

The Poor Old Heart of the Milky Way

HANS-WALTER RIX,¹ VEDANT CHANDRA,² RENÉ ANDRAE,¹ ADRIAN M. PRICE-WHELAN,³ DAVID H. WEINBERG,⁴
CHARLIE CONROY,² MORGAN FOUESNEAU,¹ DAVID W HOGG,^{5,3,1} FRANCESCA DE ANGELI,⁶ ROHAN P. NAIDU,^{7,*}
MAOSHENG XIANG,⁸ AND DANIELA RUZ-MIERES⁶¹Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany²Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA 02138, USA³Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA⁴Department of Astronomy and Center for Cosmology and AstroParticle Physics, The Ohio State University, Columbus, OH 43210, USA⁵Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA⁶Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK⁷MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Ave., Cambridge, MA 02139, USA⁸National Astronomical Observatories of China, Chinese Academy of Sciences, Beijing, 100012, China

ABSTRACT

Massive disk galaxies like our Milky Way should host an ancient, metal-poor, and centrally concentrated stellar population. This population reflects the star formation and enrichment in the few most massive progenitor components that coalesced at high redshift to form the *proto-Galaxy*. While metal-poor stars are known to reside in the inner few kiloparsecs of our Galaxy, current data do not yet provide a comprehensive picture of such a metal-poor “heart” of the Milky Way. We use information from Gaia DR3, especially the XP spectra, to construct a sample of 2 million bright ($G_{BP} < 15.5$ mag) giant stars within 30° of the Galactic Center with robust $[M/H]$ estimates, $\delta[M/H] \lesssim 0.1$. For most sample members we can calculate orbits based on Gaia RVS velocities and astrometry. This sample reveals an extensive, ancient, and metal-poor population that includes $\sim 18,000$ stars with $-2.7 < [M/H] < -1.5$, representing a stellar mass of $\gtrsim 5 \times 10^7 M_\odot$. The spatial distribution of these $[M/H] < -1.5$ stars has a Gaussian extent of only $\sigma_{R_{GC}} \sim 2.7$ kpc around the Galactic center, with most of these orbits being confined to the inner Galaxy. At high orbital eccentricities, there is clear evidence for *accreted* halo stars in their pericentral orbit phase. Stars with $[M/H] < -2$ show no net rotation, whereas those with $[M/H] \sim -1$ are rotation dominated. Most of the tightly bound stars show $[\alpha/Fe]$ -enhancement and $[Al/Fe]$ – $[Mn/Fe]$ abundance patterns expected for an origin in the more massive portions of the proto-Galaxy. These central, metal-poor stars most likely predate the oldest part of the disk ($\tau_{age} \approx 12.5$ Gyr), which implies that they formed at $z \gtrsim 5$, forging the proto-Milky Way.

1. INTRODUCTION

Understanding the formation history of our own Galaxy, especially its earliest phases, has been a central goal of *Galactic Archeology* for decades (Freeman & Bland-Hawthorn 2002), with dramatic progress enabled by a suite of spectroscopic surveys (e.g., Yanny et al. 2009; De Silva et al. 2015; Majewski et al. 2017; Conroy et al. 2019) and ESA’s Gaia mission (Perryman et al. 2001; Gaia Collaboration et al. 2022). The earliest phases of the Milky Way’s star-formation and enrich-

ment history are reflected at the present epoch in the orbit- and abundance- distribution of old and metal-poor stars. In the context of the hierarchical formation of massive disk galaxies like the Milky Way, we should expect the oldest and most metal-poor stars (say $[M/H] < -1.5$) to be a mix of stars that a) formed within one of the main overdensities which coalesced early to form the proto-Galaxy, or b) formed early in distinct satellite galaxies that eventually merged with the main body (Zolotov et al. 2009). The first channel is commonly referred to as *in situ* formation, the second *accreted* (Tumlinson 2010; Pillepich et al. 2015; El-Badry et al. 2018; Renaud et al. 2021b). If the *in situ* stars formed in a considerably more massive potential well than the *accreted* stars, this difference in origin should

Corresponding author: Hans-Walter Rix
rix@mpia.de

* NASA Hubble Fellow

also be reflected in their stars’ abundance patterns, such as $[\alpha/\text{Fe}]$ vs. $[\text{Fe}/\text{H}]$ or $[\text{Mg}/\text{Mn}]$ vs. $[\text{Al}/\text{Fe}]$ (Zolotov et al. 2010; Hawkins et al. 2015; Das et al. 2020; Horta et al. 2021; Belokurov & Kravtsov 2022; Conroy et al. 2022, and others).

However, at very early epochs, this distinction between *in situ* and *accreted* may become rather blurry. High-resolution formation simulations of galaxies like the Milky Way show (e.g. Renaud et al. 2021a,b) that pieces of comparable mass may rapidly coalesce early on (say, $z > 4$) in a sequence of “major mergers”. One terminology choice is to label only the one piece that was somewhat more massive as the *in situ* component, and to label all other pieces as accreted (which put together may constitute the majority of mass). However, simulations show that it may be operationally impossible later on to differentiate what was *in situ* and what accreted at very early epochs, whether one uses orbit or abundance information (e.g. Renaud et al. 2021b; Brauer et al. 2022; Orkney et al. 2022). Alternatively, one could label collectively all the major pieces that coalesced very early on, say at $z \gtrsim 5$, as the *proto-Galaxy*. (see Conroy et al. 2022)¹, and then differentiate subsequent additions at slightly later epochs (say, $z < 3$) as either *in situ*, if the material was brought in as gas, or as *accreted*, if the stars formed in a distinct potential well that then subsequently merged. Here, we opt for the second terminology, in part in light of the results we find. Consequently, we refer to the oldest parts of the Milky Way as the *proto-Galaxy*, without trying to single out one of the contributing pieces as *in situ*.

The nature of the accreted component of our Galaxy at $R_{\text{GC}} \gtrsim 5$ kpc has come into focus due to the combination of Gaia and large ground-based spectroscopic surveys. Tidal debris from the disrupted Sagittarius satellite (Ibata et al. 1994) dominates the halo from 20 – 50 kpc with stars mostly on highly inclined orbits with substantial angular momentum (Majewski et al. 2003; Naidu et al. 2020); at 5 – 25 kpc the debris from the disrupted Gaia-Sausage-Enceladus (GSE) satellite (Helmi et al. 2018; Belokurov et al. 2018b; Naidu et al. 2020) dominates, with stars on highly eccentric orbits ($e \gtrsim 0.7$) with little (slightly retrograde) angular momentum. Beyond these dominant components, a growing number of additional, distinct accreted components of the Galaxy have been identified (e.g., Newberg et al. 2009; Kruijssen et al. 2019; Myeong et al. 2019; Naidu et al. 2020; Yuan et al. 2020; Malhan et al. 2022). Beyond their orbits, the abundance patterns of these stars

also point toward an accreted origin: their distribution in the $[\alpha/\text{M}]-[\text{M}/\text{H}]$ plane (lower $[\alpha/\text{M}]$ at a given $[\text{M}/\text{H}]$) indicates that their birth material was enriched in a potential well of lower mass, i.e. a satellite galaxy (e.g. Frebel & Norris 2015; Hawkins et al. 2015; Lee et al. 2015). In addition to presumed satellite galaxy debris, the stellar halo also appears to encompass a large number of disrupted star clusters (e.g. Malhan et al. 2018; Shipp et al. 2018; Bonaca et al. 2021).

There has also been progress in understanding the old proto-Galactic or *in situ* component of the Milky Way. The old and metal-poor Galactic *in situ* component that has been mapped best is the old, α -enhanced disk (e.g., Hayden et al. 2015; Bonaca et al. 2020; Belokurov et al. 2020; Xiang & Rix 2022): its stars seem to date back to ≥ 12.5 Gyrs (or $z \gtrsim 5$), are centrally concentrated and form a thick disk. However, the old α -enhanced disk metallicity distribution function (MDF) is truncated, or at least drops sharply, below $[\text{M}/\text{H}] = -1$. It seems inevitable that there must have been a substantive proto-Galactic stellar population responsible for enriching the old disk’s birth material to $[\text{M}/\text{H}] = -1$. Recent surveys have provided clear evidence for a so-called *in situ* halo component in the Milky Way (e.g., Bonaca et al. 2017; Haywood et al. 2018; Di Matteo et al. 2019; Naidu et al. 2020; Bonaca et al. 2020; Belokurov et al. 2020; Horta et al. 2021). Many of these studies drew on samples at the Solar radius and beyond, and those *in situ* halo stars may reflect very early disk stars that were kicked up, or splashed to hotter orbits.

While the chemodynamical structure of the Galaxy at $R_{\text{GC}} \gtrsim 5$ kpc has come into focus over the past decade, the nature of the inner Galaxy – especially at low metallicity – has proven more elusive. Fortunately, recent work has begun to piece together a picture of the metal-poor inner Galaxy with limited sample sizes (e.g., Ness et al. 2013; García Pérez et al. 2013; Schlafman & Casey 2014; Casey & Schlafman 2015; Ness et al. 2015; Koch et al. 2016; García Pérez et al. 2018; Reggiani et al. 2020; Arentsen et al. 2020a; Lucey et al. 2021). In particular, narrow-band photometric surveys have spectacularly enabled a more efficient selection of metal-poor stars towards the inner Galaxy (e.g., Arentsen et al. 2020a,b), providing metal-poor samples in the inner Galaxy of several thousand objects. Spectroscopic follow-up has shown that these stars are indeed metal poor and show very little rotation at the lowest metallicities (Arentsen et al. 2020b). Kruijssen et al. (2019, 2020) and Forbes (2020) recently used the properties of inner Galaxy globular clusters as tracers of a very old Milky Way component, which they dubbed *Kraken*

¹ Belokurov & Kravtsov (2022) use the term “young Galaxy”

and *Koala*. They considered these possible instances of early accretion.

Most recently, Belokurov & Kravtsov (2022) and Conroy et al. (2022) have been able to shed first light on the question of how the old disk emerged from an earlier, more metal-poor population of stars on dynamically hot orbits. Belokurov & Kravtsov (2022) combined APOGEE and Gaia astrometry to study the kinematics of stars in the inner Galaxy down to metallicities of $[M/H] \sim -1.5$. They found that metal-poor stars in the inner Galaxy with abundance patterns they identified with *in situ* formation (dubbed *Aurora*) show rather little net rotation, and reasoned that these stars preceded the Milky Way’s ancient disk (see also Arentsen et al. 2020a). Conroy et al. (2022) used chemistry, kinematics, and ages from H3 and Gaia data to identify likely proto-Galactic stars down to $[M/H] \sim -2.5$ with ages $\gtrsim 13$ Gyr. This recent work clearly implies that our Milky Way has some form of a proto-Galactic population. Whether or not there exists an operationally separable accreted population buried deep in the potential well of our Galaxy remains an open question (cf. Kruijssen et al. 2020; Horta et al. 2021; Myeong et al. 2022).

It is on this background that we aim to flesh out a more comprehensive picture of the abundance-orbit distribution of metal-poor stars in the inner Galaxy, $R_{GC} \lesssim 5$ kpc. Here, sensible definitions of metal-poor could be a) more metal-poor than the oldest disk stars, i.e. $[M/H] < -1$, or b) more metal poor than large published samples of stars with abundances and orbits in the inner Galaxy, i.e. $[M/H] < -1.5$.

Specifically, we carry out a comprehensive search for metal-poor stars towards the Galactic center, drawing on the immense wealth of new information that Gaia DR3 now affords. Our work relies mainly on newly published low-resolution BP/RP spectra (hereafter collectively referred to as ‘XP’, Carrasco et al. 2021; De Angeli et al. 2022) and photometry synthesized from these spectra (Montegriffo, P. et al. 2022a). In principle, various metallicity estimates exist for millions of stars in DR3. Yet, the XP-based estimates from Andrae et al. (2022) suffer from strong biases,² while the RVS-based metallicity estimates from Recio-Blanco et al. (2022) are currently limited to bright stars with high-quality Gaia RVS spectra.

For our analysis, we select giant stars within 30° of the Galactic center, derive data-driven $[M/H]$ estimates

from their XP spectra, and combine them with RVS velocities to obtain orbits. This provides a sample of 1.5 million stars in the inner Galaxy, for a chemodynamical study.

The rest of the paper is organized as follows: in Section 2 we describe the initial selection of the sample, and the subsequent data-driven derivation of $[M/H]$ estimates, using their XP spectra and AllWISE photometry (Cutri et al. 2021). In Section 3 we present the spatial, $[M/H]$ and orbit distribution of these stars, along with the abundance patterns for a small subset. In Section 4 we summarize these results, put them in the context of Galactic archaeology, and sketch a few avenues for follow-up work.

2. DATA SETS AND DERIVED ABUNDANCE AND ORBIT QUANTITIES

We aim to devise a well-defined and large sample of stars in the “inner Galaxy” that consists of stellar tracers that cover the full age and $[M/H]$ range with comparable selection effects. For that we need tracers that are luminous enough to reach the distance of the Galactic center and beyond, even in the presence of some dust extinction. For *chemo-dynamical* mapping, we need a robust and precise estimate of $[M/H]$ and an estimate of the 6D phase space coordinates (x, v) to estimate the orbits. Ideally, we would also like to have precise age estimates and individual abundances $[X/H]$, or at least an estimate of the α -enhancement, $[\alpha/M]$. Using age estimates and individual abundances for these giant stars is beyond the scope of this work, as is a full accounting of the sample selection function.

To build such a sample, we first identify all likely red giant stars (RGB and RC) in the direction of the Galactic center that: a) are bright enough that we can expect a robust and precise estimate of $[M/H]$ from Gaia XP data (De Angeli et al. 2022; Montegriffo, P. et al. 2022b); and b) have radial velocities from Gaia RVS (Katz et al. 2022) to estimate the stars’ orbits in conjunction with their parallaxes, ϖ , and proper motions, $\vec{\mu}$. Then we re-derive $[M/H]$ estimates for these stars from their XP spectra with a data-driven approach that draws on the SDSS’s APOGEE survey (DR17, Abdurro’uf et al. 2022a), to address the documented shortcomings of the GSP-Phot abundance estimates in Gaia DR3 (Andrae et al. 2022). Finally, we calculate the orbits of the sample members from their (x, v) , which requires a model for the Galactic potential, $\Phi(x)$ (here Price-Whelan 2017).

2.1. Initial Gaia Query

The initial *inner Galaxy sample* was devised via the following ADQL query:

² The metallicity biases reported in Andrae et al. (2022) presumably originate from systematics in the *ab initio* SED models and from our imperfect understanding of the XP instrument used to transfer those model SEDs into XP spectra.

```

select * from gaiadr3.gaia_source
where
parallax < 100.*power(10.,0.2*(0.9 -
    (phot_g_mean_mag - 1.5*(bp_rp-1.)))
and parallax < 1.
and abs(b)<30 and (l<30 or l > 330.)
and bp_rp between 1.0 and 3.5
and phot_bp_mean_mag < 15.5
and has_xp_continuous='true'

```

The first condition on the parallax selects stars whose absolute magnitude is more luminous than $M_G = 0.9$, aimed at selecting giant stars at luminosities that just include the red clump (RC); the $-1.5*(bp_rp-1.)$ term reflects an approximately dereddened magnitude. The query is articulated in such a way that zero or negative parallaxes are accommodated. The second parallax condition only excludes the immediate “foreground” at $D < 1$ kpc. The conditions on l and b select a rectangular region of 30° around the Galactic center. The lower limit on bp_rp eliminates most luminous hot stars for which $[M/H]$ -estimates are not feasible (but not all in the presence of reddening). The condition on $phot_bp_mean_mag$ is designed to select objects bright enough in the blue (BP) part of the XP spectra for robust $[M/H]$ -estimates; it still allows to select RC stars at the distance of the Galactic center in the presence of moderate reddening. The final line in the selection ensures that continuous XP spectra are available. This query yields 2.1 million objects.

2.2. Deriving Robust $[M/H]$ Estimates

While about 75% of this sample has RVS velocities in Gaia DR3, only a small fraction of them have metallicities or abundances derived from the RVS spectra (Recio-Blanco et al. 2022). This leaves the route of estimating $[M/H]$ from the low-resolution XP spectra of these objects. Consequently, our objective is to train a machine-learning algorithm that can precisely, accurately, and robustly predict $[M/H]$ from Gaia DR3 data.

Estimating $[M/H]$ based on machine learning involves four aspects. 1) It involves the choice of the training sample, for which we adopt SDSS DR17 (APOGEE, Abdurro’uf et al. 2022b) since its abundances are well validated, it covers the inner Galaxy and contains mainly giants, many with high extinctions. 2) It involves the choice of the data features on which to train the prediction; these can be derived from the XP spectra but may also entail, as we discuss below, external data available across the sample, such as near-infrared photometry from ALLWISE (Cutri et al. 2021). 3) It involves the choice of the machine learning algorithm, for which we adopt the extreme gradient boosting algorithm (Chen

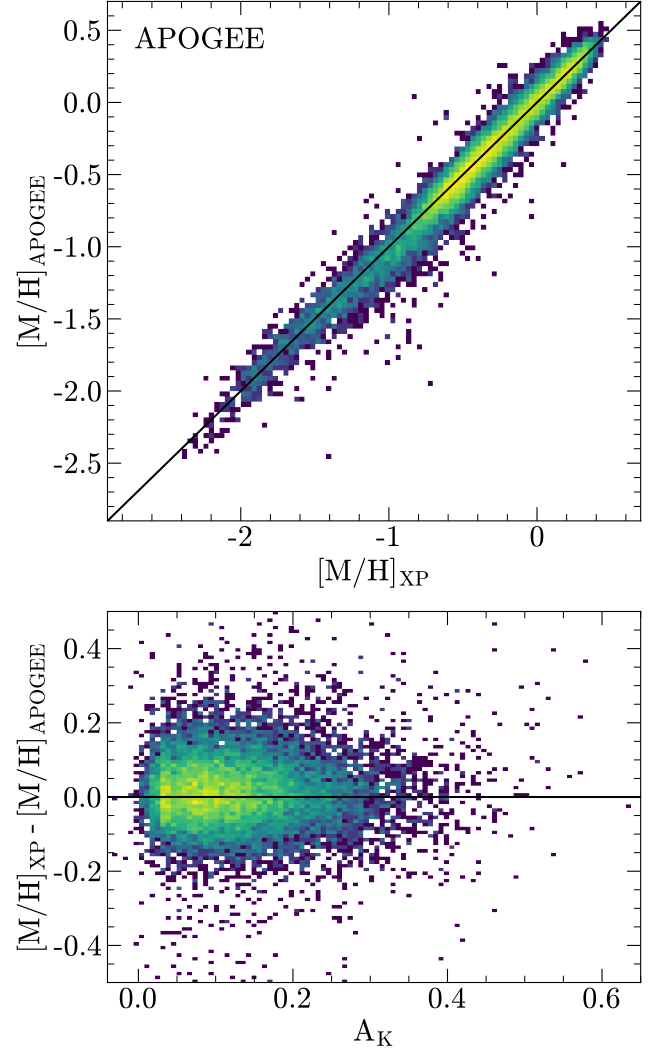


Figure 1. Top: Validation of our $[M/H]$ estimates from *Gaia* XP data as function of actual $[M/H]$ from APOGEE. Bottom: (In-)sensitivity of $[M/H]$ estimates to extinction, A_K . The Y-axis shows the difference between $[M/H]$ estimates based on XP spectra and APOGEE as a function of APOGEE’s A_K . There is no evidence for any systematic trend of $\Delta[M/H]$ with A_K to an extinction level that corresponds to $A_V \approx 3$.

& Guestrin 2016, hereafter XGBoost). XGBoost is one of the best algorithms currently available; it is straightforward and computationally inexpensive to train and it can outperform other methods such as deep learning (e.g. Grinsztajn et al. 2022). 4) and finally, we need to validate our $[M/H]$ prediction, which then speaks to the quality of the data, the training sample, and the algorithm.

The choice of the correct data features to train and predict $[M/H]$ from the XP spectra is not trivial. One obvious option is to directly use the 2×55 modified Gauss-

Hermite coefficients that describe XP spectra. But our numerical experimentation implies that other or additional data features may yield more precise and robust results with a machine learning approach³, for several pragmatic reasons. Much of the information about narrow metal line features is contained in high-order coefficients, which are often noisy. We have prior information on which parts of the spectrum are highly diagnostic in $[M/H]$ estimates and which ones are not; we need not ask an algorithm to learn this. Any estimate of $[M/H]$ requires an implicit estimate of the effective temperature T_{eff} , which is covariant with $[M/H]$; and T_{eff} is also highly covariant with reddening. Therefore, providing additional information that helps break that degeneracy is precious. The near-infrared photometry from ALLWISE (Cutri et al. 2021) exists across the sky and proves powerful in this respect.

In practice, we follow the approach of Montegriffo, P. et al. (2022a) to calculate synthetic photometry from the XP spectra in a wide range of mostly narrowband filters, in particular those filters with established utility in identifying metal-poor stars: Strömgren filters (Strömgren 1966), JPAS+ filters (Marín-Franch et al. 2012) and Pristine H & K filters (Starkenburg et al. 2017). On this basis, we use XGBoost to train the prediction of $[M/H]$ using the *entire* SDSS DR17 APOGEE of giants with $G_{\text{BP}} < 15.5$ ($\sim 230,000$ stars). Details are given in the Appendix.

Training on a particular data set and using a model that is discriminative, rather than generative, raises several issues. First, we (obviously) tie our results to the metallicity scale of SDSS DR17. Second, estimates of $[M/H]$ may be drawn into the support of the training set. In particular, XGBoost will not extrapolate $[M/H]$ significantly beyond the APOGEE DR17 range, as it is a tree-based method that segments the feature space and assigns a mean label to each such segment. Finally, generative data-driven spectral models, such as *The Cannon* (Ness et al. 2015), will presumably degrade more gracefully towards low signal to noise. Here we have addressed the last point implicitly by restricting the sample to $G_{\text{BP}} < 15.5$.

This results in $[M/H]$ estimates for about 2 million giants toward the inner Galaxy, 1.58 million of which have RVS velocities, and 1.25 million have both RVS velocities and $\varpi/\sigma_{\varpi} \geq 5$, which is our minimal condition for an orbit estimate. For subsequent analysis, we have eliminated 2% of the sample as reddened hot stars,

which can be readily recognized by their position in the m_1 -**bp_rp** color plane, where $m_1 \equiv v - 2b + y$ is the metallicity-sensitive Strömgren filter combination (see Strömgren 1966, for details); specifically, we eliminate sources with $m_1 + 0.16(\text{bp_rp} - 1) < 0$ and $\text{bp_rp} < 2.7$.

2.3. $[M/H]$ Validation

We can explore and validate the precision and accuracy of these $[M/H]$ estimates through various comparisons with analogous estimates from (higher-resolution) spectroscopic surveys. For this validation, we consider SDSS APOGEE (Abdurro'uf et al. 2022a), LAMOST (here Xiang et al. 2019), GALAH (DR3 Buder et al. 2021), and Gaia's GSP-Spec (Recio-Blanco et al. 2022).

We start by considering the $\sim 17,000$ stars in common between APOGEE and our sample, shown in Figure 1 (top panel). This panel of Figure 1 illustrates that for this sample, XGBoost provides a remarkably precise ($\lesssim 0.1$ in the median), accurate and robust $[M/H]$ estimate across $-2.5 < [M/H] < 0.5$. It should be noted that XGBoost was trained on the *full* all-sky APOGEE sample. So this is a validation on a small subset of the full training set, not a pure cross-validation. The accuracy of the $[M/H]$ prediction (i.e. lack of systematic offset) is therefore by construction. It is particularly remarkable that basically all stars predicted to be metal-poor from $[M/H]_{\text{XP}}$ are indeed metal-poor according to $[M/H]_{\text{APOGEE}}$: a selection of, e.g., metal-poor objects by $[M/H]_{\text{XP}}$ appears to be nearly pure. The bottom panel of Figure 1 shows that these $[M/H]$ estimates remain unbiased, at a level of a few percent, even in the presence of substantial dust extinction $A_K = 0.3$, corresponding to about 3 magnitudes of A_V .

Appendix A.2 details a true cross-validation of our $[M/H]_{\text{XP}}$ estimates against three external data sets: LAMOST, GALAH and GSP-Spec. Broadly speaking, the cross-validation affirms that the $[M/H]_{\text{XP}}$ estimates are precise and robust (enabling pure $[M/H]$ -based sample selection). Our estimates have a root mean square difference ≈ 0.1 dex compared to these datasets, with minimal bias. Most importantly, the performance remains unbiased toward the metal-poor end, affirming our confidence in selecting metal-poor stars based on XP metallicities. We refer to Appendix A.2 for more detailed validation of the $[M/H]$ estimates.

2.4. Orbits of the Sample Members

For sample members with suitable 6D phase-space information, we calculate their orbits. The sky positions and proper motions are available for all of them, radial velocities, and good parallax measurements only for a good fraction of the sample. In practice, we take all

³ If stellar features are to be estimated from XP spectra in a full forward-modeling approach, the situation may be different.

stars that have RVS velocities with $\delta v_{\text{RVS}} < 5 \text{ km s}^{-1}$ and $\varpi/\sigma_\varpi > 5$, about 1.25 million objects across all $[\text{M}/\text{H}]$.

To compute orbits and compute orbital properties, we use a four-component Milky Way mass model consisting of spherical Hernquist nucleus and bulge components, an (approximate) exponential disk component (Smith et al. 2015), and a spherical NFW halo component. We adopt a radial scale length $h_R = 2.6 \text{ kpc}$ and scale height $h_z = 300 \text{ pc}$ (Bland-Hawthorn & Gerhard 2016), and fit for the masses and scale radii of the nucleus, bulge, and halo components using the same compilation of Milky Way enclosed mass measurements as used to define the `MilkyWayPotential` in `gala` (Price-Whelan 2017; Price-Whelan et al. 2020), with the additional constraint of having a circular velocity at the solar position $v_c(R_0) = 229 \text{ km s}^{-1}$ (Eilers et al. 2019). We compute actions using the “Stäckel Fudge” (Binney 2012; Sanders 2012) as implemented in `galpy` (Bovy 2015). To compute the orbital eccentricity, pericenter, and apocenter values, we numerically integrate the orbits with a timestep of 0.5 Myr across four times the radial orbit period (estimated by the computed orbital frequencies from the action solver).

The $[\text{M}/\text{H}]$ estimates for 1.5 million stars towards the Galactic center, along with orbits for the RVS subsample of 1.25 million stars, are available as supplementary material to this article and are hosted online⁴ (Rix et al. 2022).

3. RESULTS

We now present the properties of the low- $[\text{M}/\text{H}]$ stellar population contained within this sample. We start with the spatial distribution, then move to the $[\text{M}/\text{H}]$ -distribution, the orbit distribution, and finally to a first exploration of the abundance patterns of the metal-poor stars, especially the level of α -enhancement. This is all done with an eye toward what we can learn about a central metal-poor *in situ* halo or bulge population.

3.1. Spatial Distribution of Metal-Poor Stars

Figure 2 compares the on-sky density distribution of metal-rich stars with $[\text{M}/\text{H}] > -0.4$ (left panel) to metal-poor ones with $[\text{M}/\text{H}] < -1.5$ (middle panel). The metal-poor stars exhibit a striking central concentration relative to the metal-rich sample, suggesting a spheroidal population in the inner Galaxy. The projected sky density of metal-poor sample members in Figure 2 (middle panel) is dramatically altered by the high

dust extinction towards the GC (right panel). The fine-scale structure in the projected $[\text{M}/\text{H}] < -1.5$ stellar distribution can mostly be explained by dust extinction, as the right-hand panel of Figure 2 shows. This is expected as we require $G_{\text{BP}} < 15.5$ for sample selection.

We do not have precise parallaxes for the entire sample, which complicates the analysis of the spatial structure. For the overall sample, the median parallax precision is $\varpi/\sigma_\varpi \sim 12$, with 10% of the sample having $\varpi/\sigma_\varpi \lesssim 4$, and 1% of the sample with negative parallaxes. For the metal-poor sample members with $[\text{M}/\text{H}] < -1.5$, which turn out to be more distant than the sample mean, the median precision is only $\varpi/\sigma_\varpi \sim 5$. Therefore, the most robust way to present the distance distribution of objects in a directional cone may be to consider their parallaxes. Given that the objects of particular interest are at $\sim 8 \text{ kpc}$ distance, it matters that we apply a Gaia parallax zero-point correction, for which we adopt 0.02 mas (Lindgren et al. 2021).

Figure 3 shows $n_*(\varpi | [\text{M}/\text{H}])$ in five $[\text{M}/\text{H}]$ -bins, with the minimal distance (maximal parallax) of 1 kpc at the right edge, and the parallax expected for the Galactic Center indicated by the dotted line (1/8.2 kpc, or 0.12 mas; Gravity Collaboration et al. 2021). This figure shows that the distance distribution of giants towards the Galactic center depends dramatically on $[\text{M}/\text{H}]$. Metal-rich stars ($[\text{M}/\text{H}] > -0.4$) are distributed throughout the disk with heliocentric distances $D_\odot = 1 - 5 \text{ kpc}$ (or 0.2-1.0 mas). In contrast, metal-poor populations become increasingly more concentrated toward the Galactic Center (GC). Indeed, the most metal-poor stars are predominantly clustered around the GC’s parallax, revealing that there is a centrally concentrated very metal-poor population in the Milky Way.

Figures 2 and 3 qualitatively confirm the picture of a metal-poor population that is very concentrated toward the Galactic center. The fact that there are almost no $[\text{M}/\text{H}] < -1.5$ stars at very low latitudes shows that almost all lie behind the strong dust extinction 2-3 kpc inward of the Sun. Rigorous modelling of these stars’ spatial distribution of these stars density distribution (e.g. Rix et al. 2021) is beyond the scope of this paper, with the severe dust extinction and crowding effects presenting formidable challenges. However, we can model the parallax distribution (Fig. 3) in the central regions less affected by dust to quantify the radial extent of this metal-poor population.

We assume that the 3D distribution is a spherical Gaussian as a function of Galactocentric radius R $\rho_{\text{Gauss}}(R, \sigma_{R_{\text{GC}}})$, with a width of $\sigma_{R_{\text{GC}}}$. As long as we can see to the far side of the Galactic center (see Fig. 3),

⁴ <https://doi.org/10.5281/zenodo.7035809>

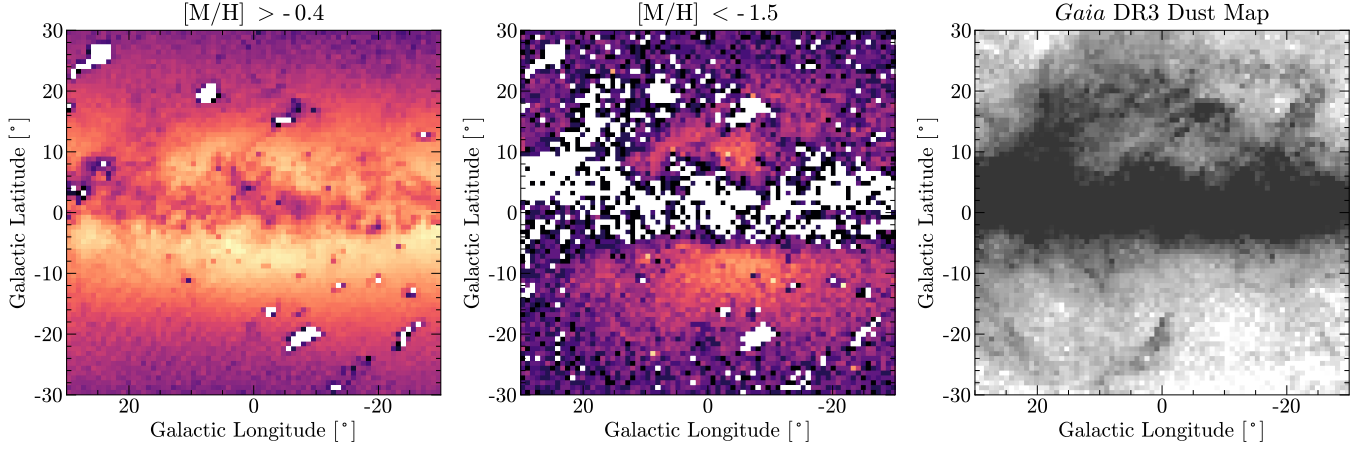


Figure 2. On-sky logarithmic density distribution of stars in our sample for metal-rich (left panel) and metal-poor (middle panel) selections. The distribution of metal-poor stars appears very centrally concentrated, but the morphology is dramatically modulated by the foreground dust extinction, as illustrated in the right panel. The right panel demonstrates that the dearth of stars is highly correlated with the distribution of dust.

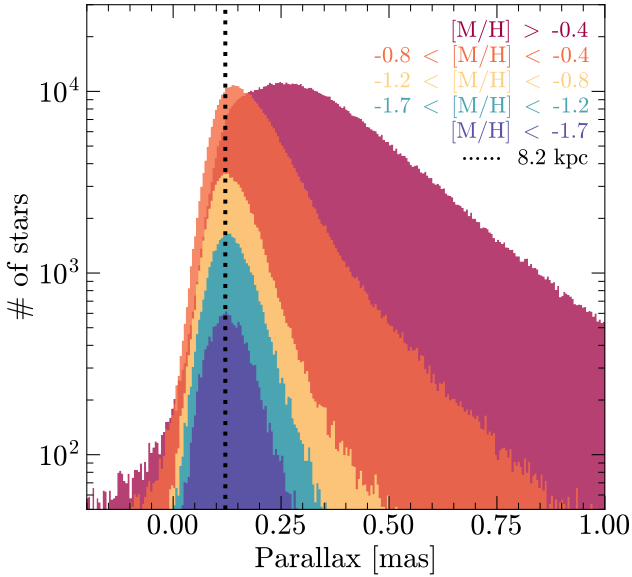


Figure 3. Parallax distribution of the different sub-samples, selected by $[M/H]$, taking ϖ as a proxy for distance. It is apparent that the metal-poor samples are centered around the expected parallax of the Galactic center (dotted black line; Gravity Collaboration et al. 2021), i.e. they are very concentrated near the Galactic Center (GC), increasingly so towards lower $[M/H]$.

we should expect the parallax (ϖ) distribution in a cone of $d\Omega$ in the direction (l, b) to be

$$n(\varpi) = d\Omega d\varpi \varpi^{-4} \rho_{Gauss}(R(\varpi, l, b), \sigma_{RGC}). \quad (1)$$

Phrasing this model in terms of the ϖ allows the model to deal with vanishing (or even negative) observed parallaxes. We find that the maximum likelihood of the observed parallaxes (and their uncertainties) for the

$[M/H] < -1.5$ stars implies an extent of $\sigma_{RGC} \sim 2.7$ kpc. Interestingly, this radius corresponds to an angle of 18° at the distance of the Galactic center, which seems very plausible in light of the projected density distribution (Figure 2; middle panel).

3.2. The $[M/H]$ -Distribution in the Inner Galaxy

We now turn our attention to the $[M/H]$ -distribution, $n_*([M/H])$, in the inner Galaxy, which is illustrated in Figure 4. The first remarkable feature of the distribution is the sheer number of objects, most apparent from the cumulative distribution in Figure 4: there are $> 4,000$ stars with an estimated $[M/H] < -2$, $\sim 18,000$ with $[M/H] < -1.5$, and almost 100,000 stars with $[M/H] < -1$, i.e. below the abundance floor of the old, α -enhanced disk. This is an order of magnitude more objects than previously published metal-poor samples ($[M/H] < -1$) of the inner Galaxy (e.g., Arentsen et al. 2020a,b).

A rough estimate shows that this corresponds to $\gtrsim 5 \times 10^7 M_\odot$ in stars at $[M/H] < -1.5$ in the inner few kpc of the Galaxy. Taking the nearest globular cluster M4 as a template for an old, $[M/H] < -1$ population, we find that there are about 35 giants with $M_G < 0.5$ in the Gaia DR3 catalog beyond M4's half mass radius of $6'$ (Richer et al. 2004, where the Gaia catalog should be approximately complete). Given M4's total mass of $7 \times 10^4 M_\odot$ (Marks & Kroupa 2010), this implies one giant (at $M_G < 0.5$) per $800 M_\odot$ of total stellar population mass. At face value, this implies $M_{tot}([M/H] < -1.5) \sim 2 \times 10^7 M_\odot$. But this does not account for the incompleteness of XP spectra in very crowded fields, not for the far-reaching effects of dust extinction. A conservative estimate of these effects is a

factor of 2.5, leading us to a lower limit on the mass of $M_{tot}([M/H] < -1.5) \gtrsim 5 \times 10^7 M_\odot$.

We have already presented two arguments that the low- $[M/H]$ portion of $n_*([M/H])$ is not severely contaminated by spurious $[M/H]$ estimates from some of the far more numerous stars of higher (true) metallicity: the validation with APOGEE and the parallax, or distance, distribution, which is most peaked for the most metal-poor objects. The $[M/H]$ distribution itself provides additional evidence. It rises very steeply from $[M/H] = -2.5$ to -2.2 . The distribution then follows a power law of slope $d(\log n_*)/d[M/H] \approx 1$ over a wide range of metallicities to $[M/H] \approx -0.9$, where it steepens towards a peak at $[M/H] = -0.6$, beyond which $n_*([M/H])$ starts dropping towards the highest metallicities present, $[M/H] \approx 0.5$.

The bottom panel of Figure 4 shows the metallicity distribution with stars of an estimated Galactocentric distance $\hat{R}_{GC} < 4$ kpc in green, and the stars of the intervening disk in gray. As giants, in particular red clump stars, are of comparable luminosity and color over a wide range of metallicities, the *differential* selection effects across different $[M/H]$ should be modest. Hence, the $[M/H]$ -distribution of stars $R_{GC} < 4$ kpc in Figure 4 should be relatively unbiased. The $[M/H]$ distribution of course varies with position, in the Galaxy (both $|z|$ and R_{GC}), as Figure 4 illustrates.

The decrease in $n_*([M/H])$ for stars at the metal-rich end ($[M/H] > 0$) with $R_{GC} < 4$ kpc (green) may seem surprising, because we know from APOGEE (e.g. Eilers et al. 2022) that the inner Galaxy is teeming with high-metallicity stars. This decrease in $n_*([M/H])$ is straightforwardly explained by Figure 2, which shows that in the inner Galaxy, we are almost completely lacking sample members with $|Z| < 700$ pc; any thinner population will inevitably be absent from our sample due to dust extinction. The APOGEE sample has a different MDF in part because its 2MASS-selected stars better sample the low- $|Z|$, high- $[M/H]$ population.

The bulk of the metal-poor population, below the minimal $[M/H]$ of the old disk at $[M/H] \sim -1$, follows $d(\log n_*)/d[M/H] \approx 1$: the cumulative number of stars $N_*(< Z)$ below a (linear) metallicity Z grows linearly with Z . This seems to be a natural slope for $n_*([M/H])$ in the early phases of a self-enriching, *in situ* system, as shown by the following argument, which builds on the models of Weinberg et al. (2017).

Under fairly general conditions, the evolution of the ISM metallicity in a simple, fully mixed system can be described by

$$\frac{dM_Z}{dt} = y_Z \dot{M}_* - (1 + \eta - r)\dot{M}_* Z, \quad (2)$$

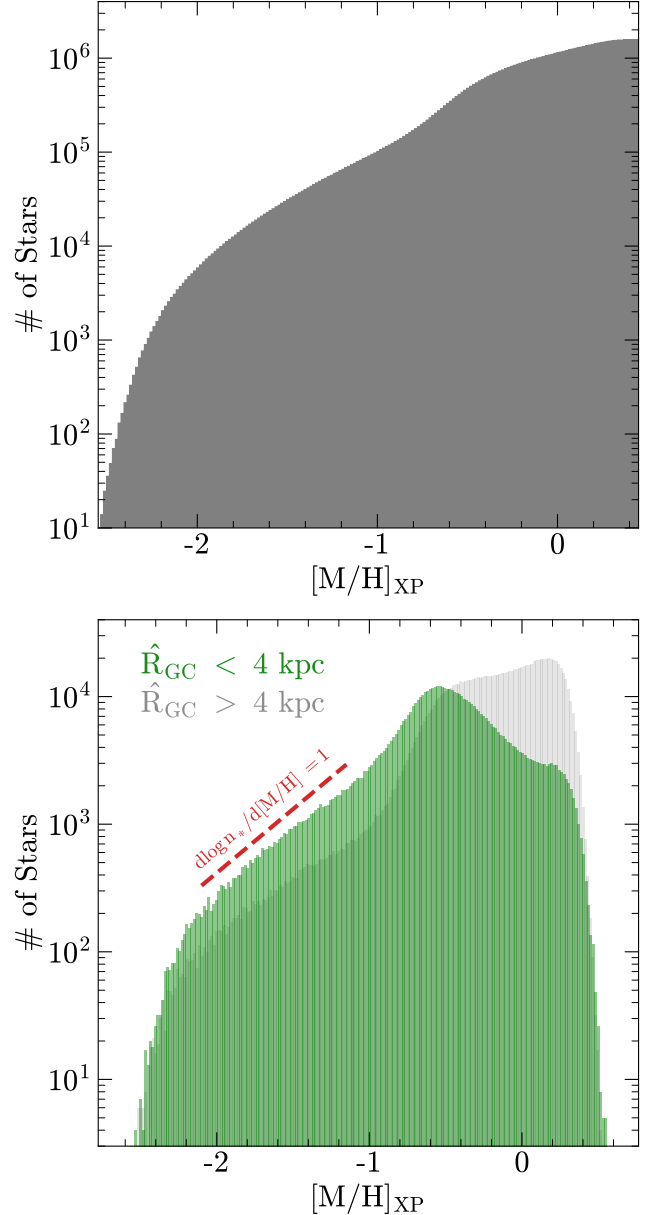


Figure 4. Top: Cumulative distribution of metallicities $[M/H]$ for the entire sample of ~ 1.7 million objects, estimated from XP spectra and broad-band photometry via XGBoost prediction trained on the APOGEE DR17: there are $\sim 5,000$ stars with $[M/H] < -2$, $\sim 20,000$ stars with $[M/H] < -1.5$, and nearly $100,000$ stars with $[M/H] < -1$. Bottom: Distribution of $[M/H]$ for our sample, split into sources with presumed Galactocentric radii $R_{GC} < 4$ kpc (in green), which are the focus of our analysis, and those at larger R_{GC} . The (green) distribution shows a steep rise to $[M/H] \sim -2$, then a power-law of slope $d(\log n_*)/d[M/H] \approx 1$ to $[M/H] \sim -1$, and finally a yet again steeper rise to $[M/H] \sim -0.6$. A possible interpretation of these slopes in the context of a simple chemical evolution model is given in Section 3.2.

where M_Z is the mass in metals, y_Z is the IMF-averaged yield, \dot{M}_* is the star formation rate (SFR), $Z \equiv M_Z/M_g$ is the metal mass fraction of the star-forming gas, $\eta \equiv \dot{M}_{\text{out}}/\dot{M}_*$ is the mass loading factor of a gaseous outflow, and $r \approx 0.4$ represents the recycling of metals that are formed into stars but quickly returned when the stars die. Equation (2) treats enrichment as instantaneous, which should be a good approximation at low metallicities where core collapse supernovae and massive star winds are the dominant sources. It also assumes that outflows have the same metallicity as the ambient ISM; if outflows instead consist of a fraction f of the supernova ejecta plus entrained ISM, then the effect is to reduce y_Z by a factor $(1 - f)$ so that it reflects only metals retained by the ISM. Importantly for the case at hand, if Z is well below the equilibrium metallicity $Z_{\text{eq}} \equiv y_Z/(1 + \eta - r)$ then the entire second term of equation (2) can be neglected.

In this low metallicity limit, equation (2) is simply $\dot{M}_Z = y_Z \dot{M}_*$, and for a metallicity-independent yield, its time integral implies $M_Z = y_Z M_*$ and thus $Z = y_Z M_*/M_g$. The log-slope of the MDF is

$$\frac{d \log n_*}{d[\text{M}/\text{H}]} = \frac{d \log M_*}{d \log Z} = \frac{Z}{M_*} \cdot \frac{\dot{M}_*}{\dot{Z}} = \frac{y_Z}{\dot{Z} \tau_*}, \quad (3)$$

where the last equality introduces the star formation efficiency (SFE) timescale $\tau_* \equiv M_g/\dot{M}_*$. Using

$$\dot{Z} = \frac{d}{dt}(M_Z/M_g) = y_Z \frac{\dot{M}_*}{M_g} - y_Z \frac{M_* \dot{M}_g}{M_g^2} \quad (4)$$

gives the end result

$$\frac{d \log n_*}{d[\text{M}/\text{H}]} = \left(\frac{\dot{Z} \tau_*}{y_Z} \right)^{-1} = \left[1 - \left(\frac{M_*}{M_g} \right) \left(\frac{\tau_* \dot{M}_g}{M_g} \right) \right]^{-1}. \quad (5)$$

The combination $(\tau_* \dot{M}_g/M_g)$ corresponds to the fractional change of the gas reservoir mass over one SFE timescale.

Equation (5) shows that an MDF log-slope ≈ 1 is a generic result in the low metallicity regime ($Z \ll Z_{\text{eq}}$), arising whenever the gas mass is constant ($\dot{M}_g = 0$) or more generally when the star-to-gas ratio M_*/M_g is small enough that the second term in Equation (4) can be neglected. Figure 4 exhibits this generic slope over the range $-2 < [\text{M}/\text{H}] < -1$. The steeper slope at $[\text{M}/\text{H}] < -2$ could plausibly arise because the gas reservoir is small but rapidly growing, with $\tau_* \dot{M}_g/M_g > 1$. The steeper slope at $-1 < [\text{M}/\text{H}] < -0.6$ may reflect rapid gas accretion (high \dot{M}_g/M_g) coinciding with the onset of the old α -enhanced disk (e.g. Belokurov & Kravtsov 2022; Xiang & Rix 2022; Conroy et al. 2022). At still higher metallicities our neglect of sink terms in equation (2) becomes a poor approximation.

3.3. The Orbit Distribution of low-[M/H] Stars in the Inner Galaxy

We now turn to the orbit distribution of the sample at hand (see Section 2.4), which can tell us which of the metal-poor stars *selected* in the inner Galaxy remain confined to the central regions of the Milky Way. The orbit distribution can tell us which stars are just passers-through near their pericenter, but on orbits that take them into the outer halo. In particular, we might expect to see the pericenter members of accreted halo components like GSE, with stars on highly eccentric orbits that take them to $R_{\text{apo}} > 10$ kpc. We can also delineate how early – or how metal-poor, if we use abundances to estimate relative stellar ages – the population develops net rotation and how rapidly the kinematics change with $[\text{M}/\text{H}]$ towards near-circular motion in the plane of the Galaxy (see Arentsen et al. 2020a; Belokurov & Kravtsov 2022; Conroy et al. 2022).

We start by considering the orbit distribution, $n_*(e, R_{\text{apo}})$, in terms of the orbital eccentricity, e and the apocenter R_{apo} . We initially focus on the $[\text{M}/\text{H}]$ -range $-1.9 < [\text{M}/\text{H}] < -1.2$ since this encompasses the bulk of known GSE stars. This distribution of 28,000 stars is shown in Figure 5. The density in the (e, R_{apo}) -plane is affected and limited by two experimental aspects. First, our initial Gaia query selects almost exclusively stars with $R_{\text{GC}} < 8$ kpc, and hence all stars whose pericenter distance exceeds 8 kpc are excluded; this boundary is indicated by the thick gray line. Except at $e < 0.2$ the $n_*(e, R_{\text{apo}})$ distribution falls off toward large R_{apo} well before this limit. There may be two reasons for the seeming dearth of stars with $R_{\text{apo}} < 2$ kpc near the center. First, severe dust extinction simply obscures much of the central kiloparsec (Fig. 2). Second, any uncertainties in the 6D phases-space coordinates, especially uncertainties in the distance, are far more likely to increase the inferred R_{apo} than decrease it, as it is a positive definite quantity. These aspects deserve careful modeling, which is beyond the scope of this paper.

The distribution in Figure 5 can be characterized by two regimes: most stars form a nearly flat distribution in eccentricity from $e = 0.1$ to $e = 0.8$, but with a narrow range in R_{apo} , with $\langle R_{\text{apo}} \rangle \approx 4$ kpc. This implies that most stars form an approximately isotropic distribution that stays confined to the inner Galaxy and hence will not appear in surveys of halo stars that focussed on larger R_{GC} and off the Galactic plane. The second regime is that of $e \geq 0.75$ and $R_{\text{apo}} \gtrsim 10$ kpc. The orbits of these stars fully match the expectations of the pericenter members of the GSE stars, the accreted component that dominates the halo population between 10 and 30 kpc (Helmi et al. 2018; Belokurov

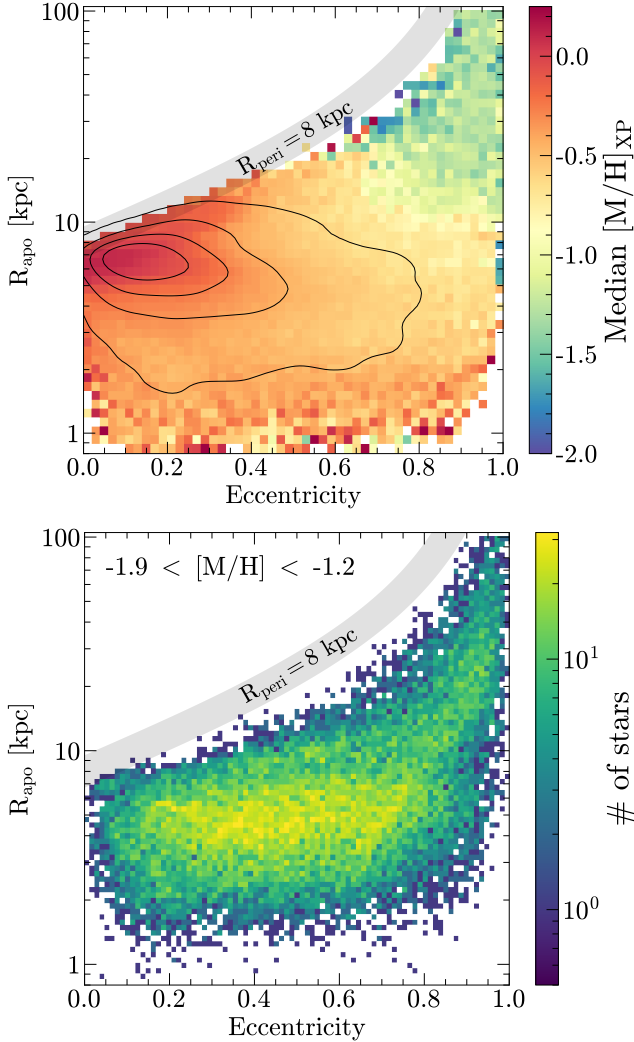


Figure 5. Top: R_{apo} vs. eccentricity space colored by the median XP metallicity. We overlay contours of linear density using a Gaussian kernel density estimate. Bottom: Distribution of the $-1.9 < [\text{M}/\text{H}] < -1.2$ sub-sample in R_{apo} vs. eccentricity space. This figure shows that most of the low-metallicity sample members have orbits that remain confined to the inner Galaxy ($R_{\text{GC}} < 5$ kpc), with a broad eccentricity distribution. At high eccentricities, there are many stars with large apocenters, potentially representing a population of accreted stars on highly radial orbits.

et al. 2018a; Naidu et al. 2021). Indeed, recent work by Belokurov et al. (2022) has shown quite clearly that there are likely GSE members with pericenters within 3 kpc. There is tantalizing but inconclusive evidence from Figure 5 that the distribution of GSE stars might extend to small R_{apo} and lower e : Figure 5 indicates an excess of stars near $(e, R_{\text{apo}}) \approx (0.7, 4 \text{ kpc})$. Here it would become important to disentangle GSE from other

proposed accreted structures in the inner Galaxy (e.g., Kruijssen et al. 2019; Horta et al. 2021).

We now ask to what extent different metal-poor stars in the inner Galaxy have net angular momentum. In general, more metal-rich stars are generally expected and observed to have a more coherent sense of angular momentum. We explore this in the two panels Figure 6, which show how close, or far, stars at a given $[\text{M}/\text{H}]$ and R_{apo} are from being an ensemble of stars on circular, in-plane orbits. This is quantified via the ratio of the angular momentum, which is the azimuthal action J_ϕ , to the total action $J_{\text{tot}} = \sqrt{J_\phi^2 + J_R^2 + J_z^2}$. The top panel shows that the increase of J_ϕ/J_{tot} with $[\text{M}/\text{H}]$ depends somewhat on the orbit size: At a given $[\text{M}/\text{H}]$ orbits with smaller R_{apo} are farther away from the coplanar quasi-circular orbits than at $R_{\text{apo}} \sim 5$ kpc.

The bottom panel of Figure 6 shows $\overline{J_\phi/J_{\text{tot}}}([\text{M}/\text{H}])$, marginalized over $3 < R_{\text{apo}}/\text{kpc} < 7$: starting already at $[\text{M}/\text{H}] < -2$ there is some, albeit small, net $\overline{J_\phi/J_{\text{tot}}}([\text{M}/\text{H}])$, which increases gradually to ~ 0.3 at $[\text{M}/\text{H}] = -1.3$. Among more metal-rich populations, $\overline{J_\phi/J_{\text{tot}}}([\text{M}/\text{H}])$ steeply rise with rising $[\text{M}/\text{H}]$ to $[\text{M}/\text{H}] \sim -0.7$, after which it levels off in the regime dominated by cold rotation. This analysis implies that some modest net rotation is present in the population well below $[\text{M}/\text{H}] < -1.5$, affirming and extending the findings of Arentsen et al. (2020a), Belokurov & Kravtsov (2022) and Conroy et al. (2022).

3.4. Abundance Patterns in the Inner Galaxy

We now enlist element abundances for our sample stars that overlap with APOGEE to explore which parts of the sample appear accreted from a chemical enrichment perspective (as opposed to *proto-Galactic*, or *in situ*). Such abundance-based arguments for the origin of stars are based on the idea that enrichment works differently in (sub-)halos of different mass (e.g., Hawkins et al. 2015; Das et al. 2020; Horta et al. 2021; Belokurov & Kravtsov 2022): the production of α -elements and Al aided by high star-formation intensities in the deeper potential wells of the proto-Galaxy, compared to the shallower potential of lower-mass satellite galaxies. However, these abundance diagnostics of the stars' formation environment may be less discriminating at lower metallicities (e.g., Conroy et al. 2022).

We remove stars belonging to globular clusters and apply a few basic quality cuts to the APOGEE abundance data based on their uncertainties and flags. We plot the resulting chemical distributions, $[\text{X}/\text{Fe}]$ vs. $[\text{Fe}/\text{H}]$ with $\text{X} = \{\alpha, \text{Al}, \text{Mn}\}$, as a function of eccentricity in Figure 7, by median apocenter radius. The presumed

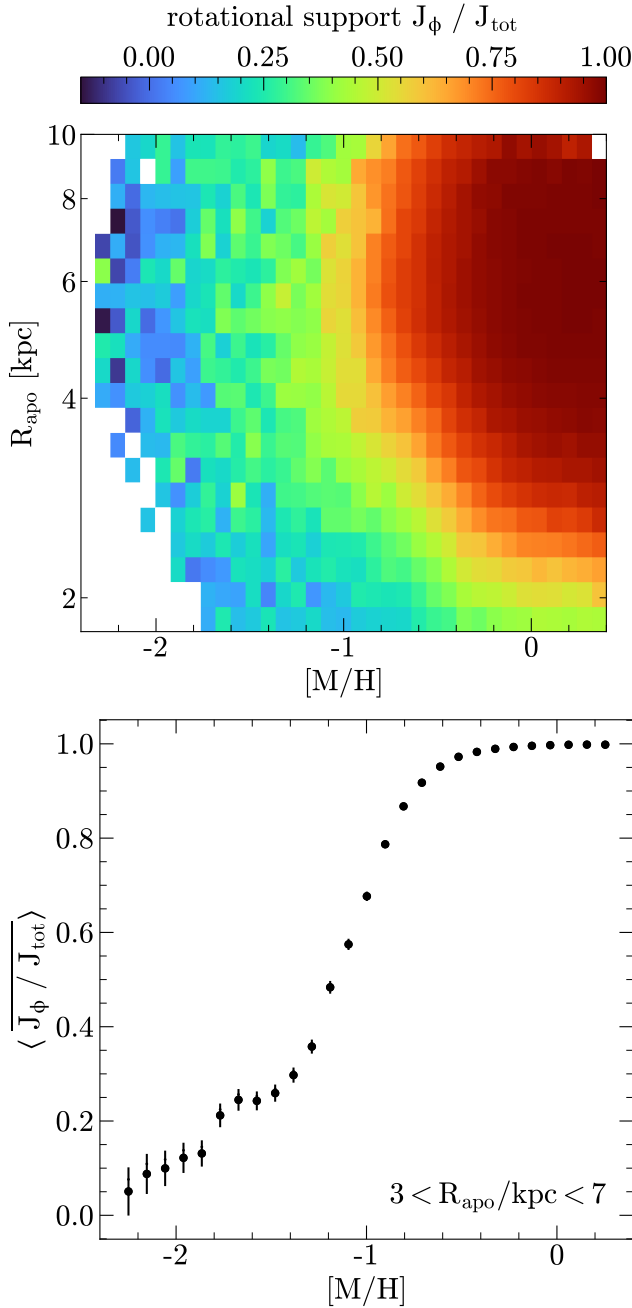


Figure 6. Dependence of the orbits’ level of rotation support, as a function of $[M/H]$ and R_{apo} . The rotation support is defined as $\langle J_\phi / J_{\text{tot}} \rangle$, where 1 is an ensemble of circular, in-plane orbits, covering both an increase in net rotation and the approach towards circular, co-planar orbits. The top panel shows $\langle J_\phi / J_{\text{tot}} \rangle$ as a function of $[M/H]$ and R_{apo} . The bottom panel shows $J_\phi / J_{\text{tot}}([M/H])$, marginalized over $3 \text{ kpc} < R_{\text{apo}} < 7 \text{ kpc}$: starting already at $[M/H] < -2$ there is some net $J_\phi / J_{\text{tot}}([M/H])$, which increases to ~ 0.3 at $[M/H] = -1.3$, then steeply towards $[M/H] \sim -0.7$, after which it levels off in the cold rotation dominated regime.

boundaries between *in situ* and accreted are indicated by the yellow dashed lines.

Figure 7 shows clearly that at modest eccentricities (left and middle columns), most stars belong to the α -enhanced sequence, implying a proto-Galactic origin. At high eccentricities (right columns), a prominent low- α sequence with large apocenters appears. These stars on eccentric orbits with large apocenter appear to be remnants of the accreted GSE merger based on their chemistry and kinematics (e.g., Mackereth et al. 2019; Naidu et al. 2020; Bonaca et al. 2020; Hasselquist et al. 2021; Horta et al. 2022). Note that among the stars of highly eccentric orbits there is a portion whose abundances point towards proto-Galactic origin (right column). Remarkably, they almost exclusively have apocenters $\lesssim 5$ kpc, very much like their cousins on less eccentric orbits, but very unlike the chemically identified GSE debris. The population of eccentric stars attributable to the proto-Galaxy has been seen before. It has been dubbed ‘Aurora’ by Belokurov & Kravtsov (2022) and has been explored at even lower metallicities by Conroy et al. (2022). Our sample probes to considerably lower R_{GC} than either of these works, and reveals the bulk and full extent of the metal-poor proto-Galaxy at the heart of the Milky Way.

It is apparent from Figures 5 and 7 that both orbits and abundances can help to differentiate between presumed proto-Galactic stars and stars accreted later from a lower-mass satellite. A differentiation by chemistry alone appears increasingly difficult at low $[M/H]$ (Conroy et al. 2022). But the combination of both aspects, summarized in Figure 7, can make such a differentiation more convincing.

Horta et al. (2021) selected “accreted” stars in these APOGEE chemical spaces as evidence of a past ‘Heracles’ merger buried deep within the MW, distinguished from GSE on the basis of a bimodal distribution in total orbital energy. This claimed galaxy bears resemblance to the previously proposed Kraken/Koala mergers (Kruijssen et al. 2019, 2020; Forbes 2020), and these names may well refer to the same event. However, Lane et al. (2022) pointed out that the apparent bimodality in orbital energy used by Horta et al. (2021) to select Heracles/Kraken/Koala might be an artifact of the spatial selection function of APOGEE. Furthermore, Heracles/Kraken/Koala and the proto-Galactic *Aurora* are virtually indistinguishable in all chemical spaces (Horta et al. 2022; Naidu et al. 2022; Myeong et al. 2022).

Taken together, these lines of evidence suggest that the stars attributed to *Heracles*, *Kraken*, *Koala* and *Aurora*, along with the numerous metal-poor stars identified here are different orbit and metallicity regimes

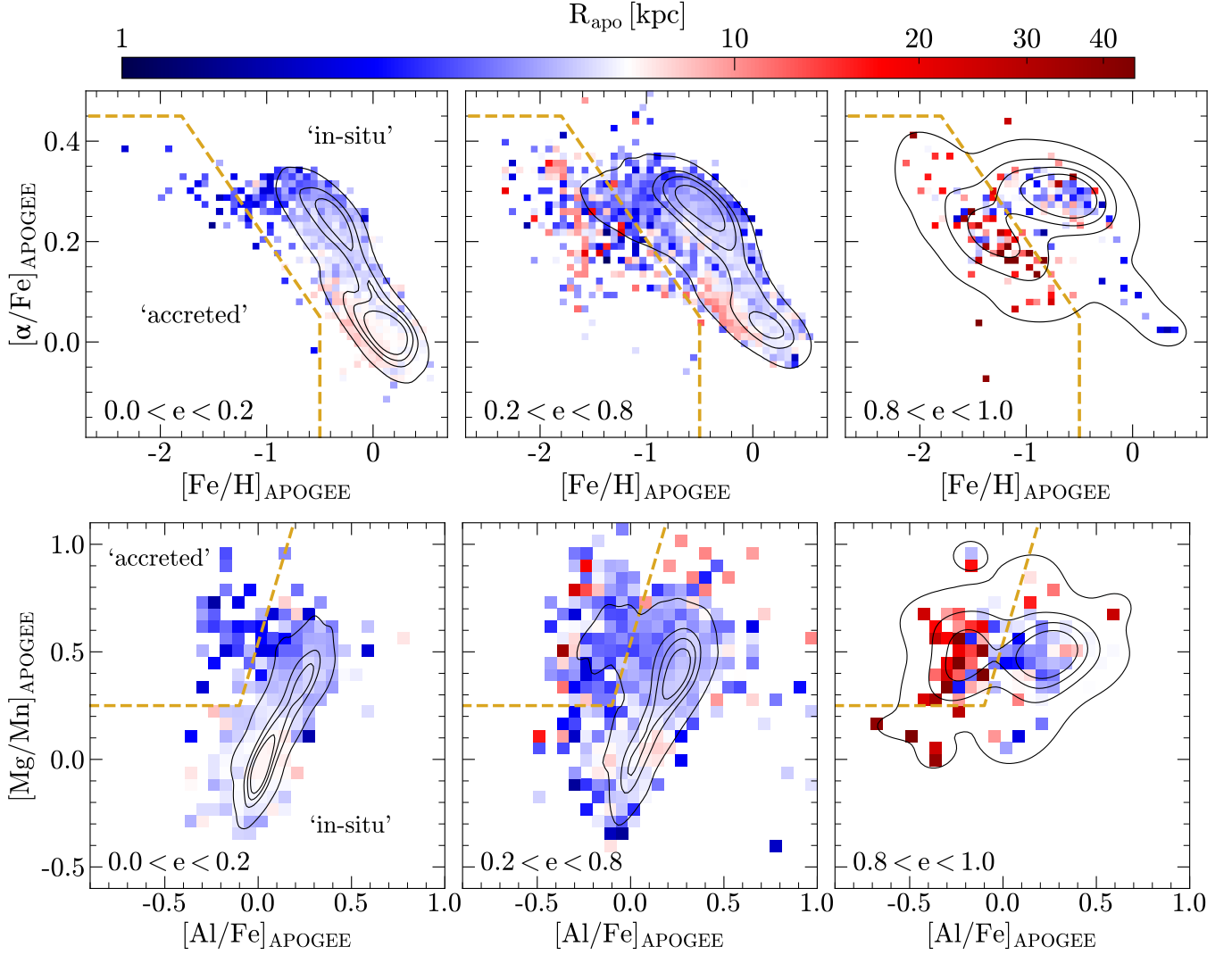


Figure 7. Abundance diagnostics for stars in the inner Galaxy, and their relation to these stars’ orbital apocenters. We show $[\text{Fe}/\text{H}]$, $[\alpha/\text{Fe}]$ (top row) and $[\text{Mg}/\text{Mn}]$ and $[\text{Al}/\text{Fe}]$ (bottom row) from SDSS/APOGEE (DR17) for stars with $\varpi < 0.25$ mas. The three columns of panels show slices of increasing orbital eccentricity from left to right, with stars colored by their median R_{apo} . The color map transitions from blue to red for apocenters at the Solar Galactocentric radius; linear density contours of the underlying sample are overlaid in black. The gold dashed lines indicate boundaries that have been used for selecting ‘accreted’ stars on the basis of their abundance patterns (e.g., [Hawkins et al. 2015](#); [Das et al. 2020](#); [Horta et al. 2021](#); [Belokurov & Kravtsov 2022](#); [Conroy et al. 2022](#)). Note, however, that the $[\text{Mg}/\text{Mn}]$ – $[\text{Al}/\text{Fe}]$ diagnostic has been found to be more ambiguous at low metallicity (see [Belokurov & Kravtsov 2022](#); [Conroy et al. 2022](#)). Overall, the figure shows that there is a high degree of correlation between the orbital and the abundance properties. At eccentricities $e < 0.8$ most stars with $[\text{M}/\text{H}] < -1$ have both $R_{\text{apo}} \lesssim R_{\text{GC}}(\text{Sun})$ and abundance patterns attributed to *in situ* formation. Only at the highest eccentricities, $e \geq 0.8$, is there a good fraction of stars that have large apocenters and at the same time abundance patterns characteristic of accreted stars. Both of these latter properties are expected for members of GSE near pericenter.

of one proto-Galactic stellar population, now forming the metal-poor, old heart of the Milky Way. Whereas past studies have mainly identified the eccentric tail of this population that traverses the solar neighborhood or off the Galactic plane (Belokurov & Kravtsov 2022; Conroy et al. 2022; Myeong et al. 2022), we now show more directly that the bulk of this population resides in the bulge itself, mostly within a few kpc of the Galactic center. Indeed, this matches theoretical predictions from cosmological zoom-in simulations, which predict a spheroidal distribution of *proto-Galaxy* stars predominantly concentrated around the Galactic center (e.g., El-Badry et al. 2018, Wetzel et al. 2022, Figure 13 in Belokurov & Kravtsov 2022).

The picture presented here appears to obviate the need for an *ex-situ* merger like Heracles/Kraken/Koala, at least based on the orbits and chemistry of field stars. This does not necessarily imply that these mergers did not occur; rather, it emphasizes the – conceptual and practical – ambiguity between *in situ* and *accreted* stars during the earliest phases of the Milky Way, which led us to use the collective term proto-Galaxy. This view is supported by cosmological zoom-in simulations, which affirm the difficulty in distinguishing accreted galaxies from the MW proto-galaxy at early times (e.g., Renaud et al. 2021a,b; Orkney et al. 2022).

4. SUMMARY, DISCUSSION AND OUTLOOK

We have presented an exploration of the metal-poor stellar population in the inner Galaxy ($R_{GC} \lesssim 5$ kpc), drawing on the newly available low-resolution XP spectroscopy from Gaia’s DR3. The primary goal was to see whether there is an extensive ancient and metal-poor stellar population at the “heart” of the Milky Way. We wanted to learn whether the bulk of these stars are a tightly bound proto-Galactic population, rather than stars accreted from more distant satellites, whose location in the inner Galaxy just reflects the pericenter phase of orbits that take them out to the “classical” halo at $R_{GC} \gtrsim R_{\odot}$. As simulations imply that the earliest heart of galaxies often arises from the high-redshift coalescence of clumps with comparable mass, we prefer the term *proto-Galactic* for the resulting population, rather than *in situ*, a term that singles out one of these clumps even if it has only modestly larger mass than other nearby clumps.

The practical foundations of our analysis were estimates of $[M/H]$ of ~ 2 million giant stars within 30° of the Galactic center, at least 1 kpc from us and sufficiently bright for good S/N in the blue part of the XP spectra, $G_{BP} \leq 15.5$. We converted the XP spectral information into a set of narrow-band fluxes in fil-

ters that are known to be highly diagnostic of $[M/H]$ in cool stars. We combined this synthetic photometry with ALLWISE photometry and then used them – along with the XP spectral coefficients – for a data-driven $[M/H]$ prediction, trained on the APOGEE DR17 data using XGBoost. We found these $[M/H]$ estimates to be precise to ~ 0.1 dex. We also found them robust and pure enough to identify metal-poor samples of stars in the inner Galaxy ($[M/H] < -1$), despite the presence of a vastly dominant population of more metal-rich stars in the inner Galaxy. For most of these stars, radial velocities exist from Gaia RVS, allowing the estimates of their orbits.

On this basis, we identified metal-poor samples of stars towards the inner Galaxy that are about two orders of magnitude larger than published ones: $> 4,000$ stars with $[M/H] < -2$, $\sim 20,000$ stars with $[M/H] < -1.5$, and $\sim 70,000$ stars with $[M/H] < -1$. The most metal-poor tail of the metallicity distribution extends to $[M/H] \approx -2.5$. Whether this constitutes a genuine metallicity floor or reflects that the APOGEE training set for XGBoost only extended to $[M/H] = -2.5$ requires follow-up data.

These samples reveal a large metal-poor population ($[M/H] < -1.5$) in the inner few kpc of our Galaxy. Three lines of argument imply that this population is centrally concentrated: first, the parallax distribution implies that most stars are within 5 kpc of the Galactic center (Figure 3). If we model the spatial density distribution of stars with $[M/H] < -1.5$ as a Gaussian centered on the Galactic center, with an extent of $\sigma_{R_{GC}}$, this parallax distribution implies $\sigma_{R_{GC}} \sim 2.7$ kpc. Second, the number density projected onto the sky shows a centrally concentrated distribution (albeit severely modulated by intervening dust, Figure 2), which is qualitatively consistent with $\sigma_{R_{GC}} \sim 2.7$ kpc. Third, their orbits show that most stars have *apocenters* of less than 5 kpc (Figure 5). The latter property implies that most of this population could not have been found in many of the previous surveys that focused on $R_{GC} > 5$ kpc. At the same time, our analysis showed that a minority (but still a large number) of stars with $[M/H] < -1.2$ currently in the inner Galaxy are on highly eccentric orbits that can take them to > 10 kpc. These are presumably members of the accreted halo, found here near the pericenter of their orbits.

Using the members of this sample that have detailed abundances from SDSS (DR17), in particular $[\alpha/Fe]$, $[Al/Fe]$ and $[Mn/Fe]$, we explored which of these stars have abundance patterns attributable to a proto-Galactic (or, *in situ*) vs. accreted origin. We found that almost all stars that remain tightly bound within the

inner Galaxy ($R_{\text{apo}} < R_{\text{GC}}(\text{Sun})$) have abundances that point towards a proto-Galactic origin. Stars on highly eccentric orbits with $R_{\text{apo}} > 10$ kpc show abundances typical of accreted stars and are presumed pericenter members of GSE.

We quantified how much net rotation the various mono-abundance populations show and how much they resemble disk-like orbits. We did this by considering the mean ratio of the angular momentum to the total action, $\overline{J_\phi/J_{\text{tot}}}([M/H])$. We found that only the most metal-poor tail $[M/H] < -2$ of this population shows no rotation, $\overline{J_\phi/J_{\text{tot}}}([M/H] < -2) \approx 0$. Populations with higher $[M/H]$ also have higher $\overline{J_\phi/J_{\text{tot}}}([M/H])$: at $[M/H] = -1.5$, $\overline{J_\phi/J_{\text{tot}}}$ has reached 0.25, then continuously increasing to $\overline{J_\phi/J_{\text{tot}}} \approx 0.9$ in $[M/H] = -0.8$.

All of this information fits a picture in which this metal-poor heart of the Milky Way constitutes the most ancient *proto-Galactic* component of our Galaxy.

- Much of this population is confined to $R_{\text{GC}} < R_\odot$, or tightly bound at the bottom of our Galaxy’s potential well, arguing against an origin through accretion from a once-distant satellite.
- The vast majority of stars with $[M/H] < -1.5$ are distinctly α -enhanced, arguing for rapid enrichment in a deep potential well.
- The stellar mass we see at $[M/H] < -1.5$ within 5 kpc corresponds to $\sim 5 \times 10^7 M_\odot$, but manifest dust obscuration implies a large correction that suggests that the stellar mass of this metal-poor heart of the Milky Way is high, $M_* \gtrsim 10^8 M_\odot$. Such high masses of strongly α -enhanced stars at such low $[M/H]$ are most likely to occur at the center of a (eventually) quite massive halo, not at a low-mass satellite.
- We have mapped the degree of (prograde) ordered azimuthal motion, $\overline{J_\phi/J_{\text{tot}}}([M/H])$ as function of $[M/H]$. At $[M/H] \sim -2$ it is approximately zero, and increases continuously when considering sets of stars with increasing $[M/H]$. Such a relation is expected if star formation occurs – as time goes on – in gas that has increasingly settled into a disk, as recent simulations of disk galaxy formation show (Gurvich et al. 2022).
- Much of this central component has $[M/H]$ well below that of the centrally concentrated, α -enhanced, thick disk, whose oldest members (at $[M/H] \sim -1$) appear to be ≈ 12.5 Gyrs old (Xi-ang & Rix 2022). If much of this metal-poor heart of the Milky Way chemically, and therefore temporally, predates the oldest α -enhanced disk, it must

have formed at $\tau \gtrsim 12.5$ Gyrs, corresponding to $z \gtrsim 5$. This stellar population at the heart of the Milky Way should be almost exclusively ancient.

The results presented here are by no means a *new distinct* stellar component of the Milky Way. The stars in our sample seem to constitute the tightly bound part – and bulk – of a proto-Galactic spheroid. The distribution tail of stars on somewhat more extended orbits has already been recognized in recent work as *in situ* halo (Belokurov & Kravtsov 2022; Conroy et al. 2022). But our results significantly flesh out the existing picture by showing that there is indeed a tightly bound *in situ* “iceberg”, whose tips have been recognized before. The fact that there are many metal-poor stars in the inner galaxy has also already been recognized by the recent work of Arentsen et al. (2020a,b).

Our analysis here is in many ways preliminary and suggests various avenues of follow-up. The spatial distribution of this population deserves to be modeled quantitatively, including the effects of dust extinction, source crowding, parallax uncertainties, and the giant luminosity function. The abundances, and perhaps more importantly, the abundance patterns, deserve full spectroscopic follow-up to understand: How pure are the XP selected samples in this regime? Is there a floor in the $[M/H]$ distribution that might reflect the $[M/H]$ of the proto-Milky Way’s circumgalactic medium? If these are indeed among the most ancient stars in the Milky Way, do they have remarkable abundance ratios? Was the formation of this component associated with the formation of an extensive set of globular clusters? What can the degree of central concentration of this component tell us about the violence of subsequent merging events, which would scatter stars to larger radii?

On a more technical side, this analysis reflects the astounding information content of the Gaia DR3 data, particularly the XP spectra. Our data-driven approach to estimate $[M/H]$ seems to work well for the current analysis, at the price of restricting $[M/H]$ estimates to bright objects. Presumably, we are far from exploiting the information of the XP spectra, which should be unlocked by forward-modeling of the data.

VC gratefully acknowledges a Peirce Fellowship from Harvard University. The authors thank Vasily Belokurov and Andrey Kravtsov for helpful comments on the manuscript.

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.

SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics — Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observa-

tory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

This work made use of the Third Data Release of the GALAH Survey (Buder et al. 2021). The GALAH Survey is based on data acquired through the Australian Astronomical Observatory, under programs: A/2013B/13 (The GALAH pilot survey); A/2014A/25, A/2015A/19, A/2017A/18 (The GALAH survey phase 1); A/2018A/18 (Open clusters with HERMES); A/2019A/1 (Hierarchical star formation in Ori OB1); A/2019A/15 (The GALAH survey phase 2); A/2015B/19, A/2016A/22, A/2016B/10, A/2017B/16, A/2018B/15 (The HERMES-TESS program); and A/2015A/3, A/2015B/1, A/2015B/19, A/2016A/22, A/2016B/12, A/2017A/14 (The HERMES K2-follow-up program). We acknowledge the traditional owners of the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present. This paper includes data that has been provided by AAO Data Central (datacentral.org.au).

The Guoshoujing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope LAMOST) is a National Major Scientific Project built by the Chinese Academy of Sciences. Funding for the project has been provided by the National Development and Reform Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences.

This work has made use of the Python package GaiaXPy, developed and maintained by members of the Gaia Data Processing and Analysis Consortium (DPAC), and in particular, Coordination Unit 5 (CU5), and the Data Processing Centre located at the Institute of Astronomy, Cambridge, UK (DPCI).

APPENDIX

A. [M/H] ESTIMATES VIA XGBOOST TRAINED ON SDSS APOGEE ABUNDANCES

Here we expand on a few specifics of the [M/H] estimate that is summarized in Section 3.2.

A.1. Training of XGBoost model of [M/H]

We want to derive [M/H] from XP information for objects that still have significant flux in the blue ($G_{BP} < 15.5$), yet may be reddened by several magnitudes. To help break any degeneracies between extinction and T_{eff}

we include near-infrared photometry from ALLWISE (Cutri et al. 2021). We adopt SDSS DR17 APOGEE abundances (Abdurro'uf et al. 2022b) as a training sample, since it covers the inner disk and contains many giants stars with high extinctions.

For the machine learning algorithm, we choose the extreme gradient boosting algorithm (Chen & Guestrin 2016, hereafter XGBoost), as it is computationally inexpensive to train and can outperform other algorithms, e.g. Deep Learning (e.g. Grinsztajn et al. 2022). To im-

Table 1. Test errors of $[M/H]$ from XGBoost from 20-fold cross-validation using various numbers of XP coefficients. We quote the median absolute error and the root-mean-square error.

input features	median AE	RMSE
all 55 coefficients	0.061	0.145
first 40 coefficients only	0.059	0.138
first 30 coefficients only	0.060	0.137
first 20 coefficients only	0.062	0.139
first 10 coefficients only	0.066	0.144
XP colours only	0.078	0.154
all coefficients + XP colours	0.064	0.138
all coefficients + WISE	0.088	0.186
XP colours + WISE	0.067	0.136
all coefficients + XP colours + WISE	0.049	0.107

prove the $[M/H]$ predictions for our sample (consisting of red giant stars by construction), we also restrict the training sample from SDSS DR17 sample to only giants, via $\log g < 3.5$.

This leaves us with a suitable choice of data features, for which we explored a number of options. Our starting point was the set of XP coefficients (De Angeli et al. 2022; Carrasco et al. 2021), which we normalized by the Gaia apparent G flux. In Table 1, we quote the test errors for using the different numbers of XP coefficients. The performance is overall very good but depend only weakly on how many coefficients we use. This suggests that also the high-order coefficients contain useful information, and, at least for this application and the APOGEE DR17 sample, a truncation of XP coefficients is neither required nor helpful in our case.

Given the results in Montegriffo, P. et al. (2022a), we then attempted to only use photometry synthesized from XP spectra using GaiaXPpy⁵ and combine this with ALLWISE near-infrared photometry. Specifically, we synthesize photometry from the systems of Pristine, StromgrenStd, Jpas, Jplus, SkyMapper, and HstAcswfc. As we can see from Table 1, using only such colors has a similar performance as only using the XP coefficients, despite including ALLWISE photometry. However, a substantial improvement is obtained when combining all XP coefficients with these colors.

In the end, we used GaiaXPpy to compute the following synthetic photometry and combined it with all 110 XP coefficients in the XGBoost training on SDSS DR17 (APOGEE):

⁵ <https://gaia-dpci.github.io/GaiaXPpy-website/>

$$\begin{aligned}
& G - W_1 \\
& G_{BP} - G_{RP} \\
& G_{BP} - W_2 \\
& W_1 - W_2 \\
& \text{CaHK}_{\text{Pristine}} - W_1 \\
& \text{CaHK}_{\text{Pristine}} - v_{\text{StromgrenStd}} \\
& \text{CaHK}_{\text{Pristine}} - b_{\text{StromgrenStd}} \\
& \text{CaHK}_{\text{Pristine}} - y_{\text{StromgrenStd}} \\
& \text{Jplus}_g - \text{Jplus}_i \\
& \text{Jplus}_{0395} - \text{CaHK}_{\text{Pristine}} \\
& \text{Jplus}_{0515} - \text{CaHK}_{\text{Pristine}} \\
& \text{Jplus}_{0861} - \text{CaHK}_{\text{Pristine}} \\
& G_{RP} - \text{Jplus}_i \\
& v_{\text{StromgrenStd}} - 2 \cdot b_{\text{StromgrenStd}} + y_{\text{StromgrenStd}} \\
& \text{CaHK}_{\text{Pristine}} - 2 \cdot \text{Jplus}_{0410} + \text{Jplus}_{0430}
\end{aligned}$$

We emphasize that the $G_{BP} - W_2$ color has the longest wavelength leverage and therefore is highly affected by extinction, whereas the $W_1 - W_2$ color is comparatively insensitive to extinction.

As is evident from Table 1, this final configuration has slightly worse performance than the best configuration. However, the median absolute errors are similar, that is, most sources get similar results. The only noteworthy difference is in the RMS error, suggesting that the final configuration may have a few more outliers. Still, the results are better than what we can achieve from coefficients alone.

A.2. Further $[M/H]$ Validation

Figure 8 illustrates a cross-validation comparison of our results to other spectroscopic surveys. Such cross-validation is easy to interpret if one can assume that the external data set represents “ground truth”. We applied a $\text{SNR}_G > 25$ to the LAMOST sample and required that there were no quality (concern) flags in any of the GALAH spectra. The left panel shows that the $[M/H]_{\text{XP}}$ estimates are accurate for LAMOST, which itself is tied to the APOGEE $[M/H]$ scale, with a slight increase in scatter towards low $[M/H]$. The comparison with GALAH (middle panel) affirms the precision and purity of the low- $[M/H]_{\text{XP}}$ estimates. Note that GALAH has a different scaling between $[M/H]$ and $[\text{Fe}/H]$ than APOGEE and LAMOST, which may explain the offset of the estimates for low- $[M/H]$ stars from the 1-to-1 line. In the parent sample without stringent GALAH quality

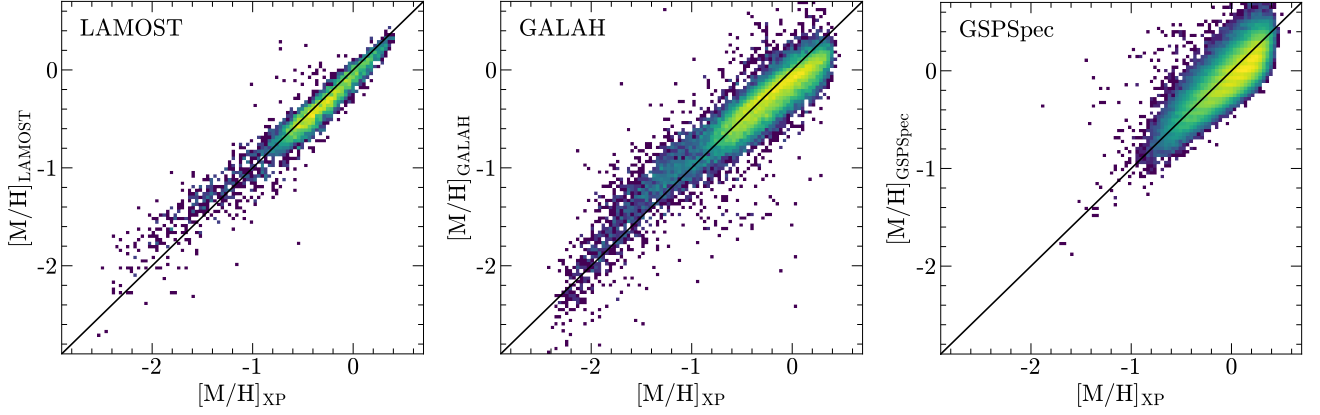


Figure 8. Cross-validation of our $[M/H]_{XP}$ estimates against three external spectroscopic data sets, LAMOST (here Xiang et al. 2019), GALAH (Buder et al. 2021) and GSP-Spec (Recio-Blanco et al. 2022). This cross-validation supports the precision and robustness of our $[M/H]_{XP}$ estimates.

cuts, there are many stars for which GALAH implies that they are $[M/H] < -1$, but the $[M/H]_{XP}$ estimate disagrees. The position of those stars in the de-reddened color-magnitude plane suggests that they indeed have $[M/H] > -1$. In other words, for these cases $[M/H]_{XP}$ appears to be the more *robust* estimate. The comparison with GSP-Spec in the right panel of Figure 8 confirms the quality of GSP-Spec’s $[M/H]$ estimates. But it also shows that the GSP-Spec values do not cover low $[M/H]$ (by design, see Recio-Blanco et al. 2022).

Because a star’s effective temperature (T_{eff}) can be very helpful in sample selection and characterization, we also estimate T_{eff} using the same XP features and XGBoost trained on SDSS DR17 APOGEE T_{eff} , as illustrated in Fig. 9. The XGBoost predictions based on XP spectra and WISE photometry match the APOGEE values within a median $\Delta T_{\text{eff}} = 32$ K, and with a mean difference of only 3 K. Figure 9 compares our $[M/H]$ estimates to APOGEE as a function of Gaia G_{BP} magnitude, showing the slight degradation towards the faintest sources in the sample.

We also compared our $[M/H]$ estimates to the extensive set of photometric estimates for metal-poor stars from Chiti et al. (2021) based on SkyMapper DR2 (SM2) photometry. We find a mean $[M/H]$ offset of 0.28 dex, and for stars with $[M/H]_{\text{SM2}} \sim -1.5$ a central 68% interval of $p([M/H]_{\text{SM2}} - [M/H]_{XP})$ of 1 dex. Clearly, the cross-validation of our XP-based $[M/H]$ estimates and SkyMapper DR2 shows far more scatter than our cross-validation with external spectroscopic data sets. This illustrates that for this sample XP-based $[M/H]$ estimates outperform those of SM2.

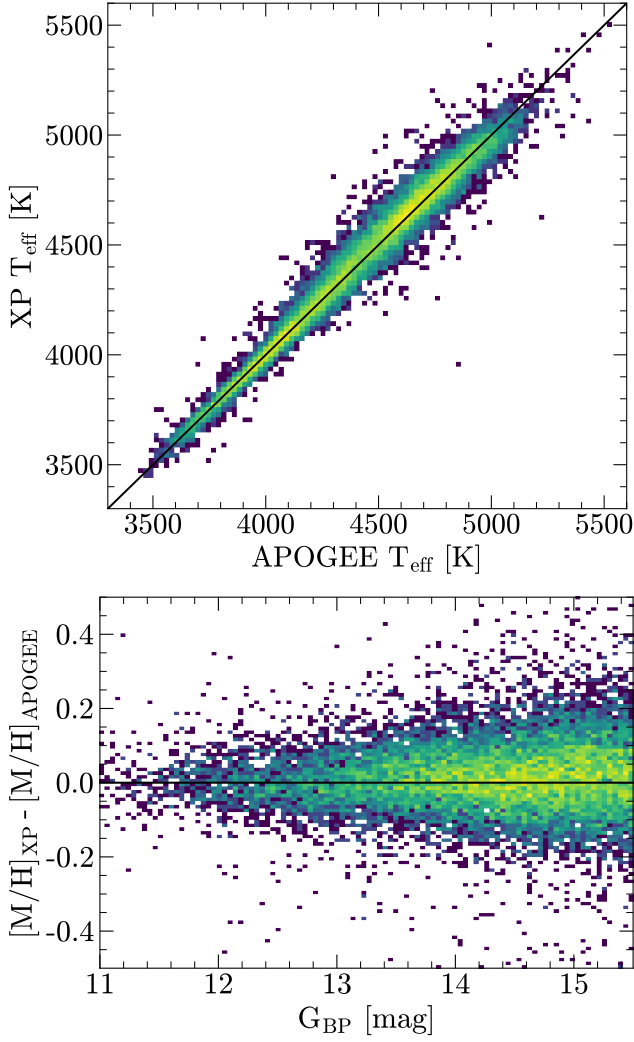


Figure 9. Top: Validation of our $T_{\text{eff}}(\text{XP})$ estimates against the T_{eff} from SDSS DR17 APOGEE. The XP predictions match the APOGEE values within a median $\Delta T_{\text{eff}} = 32$ K, and with a mean difference of only 3 K. Bottom: Validation of our $[M/H]$ estimates as a function of G_{BP} magnitude, with $\Delta[M/H] \equiv [M/H]_{\text{XP}} - [M/H]_{\text{APOGEE}}$, showing that the quality of the estimate depends somewhat on G_{BP} , as expected, but is overall unbiased.

REFERENCES

- Abdurro'uf, Accetta, K., Aerts, C., et al. 2022a, *ApJS*, 259, 35, doi: [10.3847/1538-4365/ac4414](https://doi.org/10.3847/1538-4365/ac4414)
- . 2022b, *ApJS*, 259, 35, doi: [10.3847/1538-4365/ac4414](https://doi.org/10.3847/1538-4365/ac4414)
- Andrae, R., Fouesneau, M., Sordo, R., et al. 2022, arXiv e-prints, arXiv:2206.06138.
<https://arxiv.org/abs/2206.06138>
- Arentsen, A., Starkenburg, E., Martin, N. F., et al. 2020a, *MNRAS*, 491, L11, doi: [10.1093/mnras/slz156](https://doi.org/10.1093/mnras/slz156)
- . 2020b, *MNRAS*, 496, 4964, doi: [10.1093/mnras/staa1661](https://doi.org/10.1093/mnras/staa1661)
- Belokurov, V., Deason, A. J., Koposov, S. E., et al. 2018a, *Monthly Notices of the Royal Astronomical Society*, 477, 1472, doi: [10.1093/mnras/sty615](https://doi.org/10.1093/mnras/sty615)
- Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018b, *Monthly Notices of the Royal Astronomical Society*, 478, 611, doi: [10.1093/mnras/sty982](https://doi.org/10.1093/mnras/sty982)
- Belokurov, V., & Kravtsov, A. 2022, 000.
<https://arxiv.org/abs/2203.04980>
- Belokurov, V., Vasiliev, E., Deason, A. J., et al. 2022, arXiv e-prints, arXiv:2208.11135.
<https://arxiv.org/abs/2208.11135>
- Belokurov, V., Penoyre, Z., Oh, S., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 496, 1922, doi: [10.1093/MNRAS/STAA1522](https://doi.org/10.1093/MNRAS/STAA1522)
- Binney, J. 2012, *MNRAS*, 426, 1324, doi: [10.1111/j.1365-2966.2012.21757.x](https://doi.org/10.1111/j.1365-2966.2012.21757.x)
- Bland-Hawthorn, J., & Gerhard, O. 2016, *ARA&A*, 54, 529, doi: [10.1146/annurev-astro-081915-023441](https://doi.org/10.1146/annurev-astro-081915-023441)
- Bonaca, A., Conroy, C., Wetzel, A., Hopkins, P. F., & Kereš, D. 2017, *The Astrophysical Journal*, 845, 101, doi: [10.3847/1538-4357/aa7d0c](https://doi.org/10.3847/1538-4357/aa7d0c)
- Bonaca, A., Conroy, C., Cargile, P. A., et al. 2020, *The Astrophysical Journal*, 897, L18, doi: [10.3847/2041-8213/ab9caa](https://doi.org/10.3847/2041-8213/ab9caa)
- Bonaca, A., Naidu, R. P., Conroy, C., et al. 2021, *ApJL*, 909, L26, doi: [10.3847/2041-8213/abeaa9](https://doi.org/10.3847/2041-8213/abeaa9)
- Bovy, J. 2015, *ApJS*, 216, 29, doi: [10.1088/0067-0049/216/2/29](https://doi.org/10.1088/0067-0049/216/2/29)
- Brauer, K., Andales, H. D., Ji, A. P., et al. 2022, arXiv, arXiv:2206.07057. <https://arxiv.org/abs/2206.07057>
- Buder, S., Sharma, S., Kos, J., et al. 2021, *MNRAS*, 506, 150, doi: [10.1093/mnras/stab1242](https://doi.org/10.1093/mnras/stab1242)
- Carrasco, J. M., Weiler, M., Jordi, C., et al. 2021, *A&A*, 652, A86, doi: [10.1051/0004-6361/202141249](https://doi.org/10.1051/0004-6361/202141249)
- Casey, A. R., & Schlafman, K. C. 2015, *ApJ*, 809, 110, doi: [10.1088/0004-637X/809/2/110](https://doi.org/10.1088/0004-637X/809/2/110)
- Chen, T., & Guestrin, C. 2016, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16* (New York, NY, USA: ACM), 785–794, doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)
- Chiti, A., Frebel, A., Mardini, M. K., et al. 2021, *ApJS*, 254, 31, doi: [10.3847/1538-4365/abf73d](https://doi.org/10.3847/1538-4365/abf73d)
- Conroy, C., Bonaca, A., Cargile, P., et al. 2019, *The Astrophysical Journal*, 883, 107, doi: [10.3847/1538-4357/ab38b8](https://doi.org/10.3847/1538-4357/ab38b8)
- Conroy, C., Weinberg, D. H., Naidu, R. P., et al. 2022. <https://arxiv.org/abs/2204.02989>
- Cutri, R. M., Wright, E. L., Conrow, T., et al. 2021, *VizieR Online Data Catalog*, II/328
- Das, P., Hawkins, K., & Jofré, P. 2020, *MNRAS*, 493, 5195, doi: [10.1093/mnras/stz3537](https://doi.org/10.1093/mnras/stz3537)
- De Angeli, F., Weiler, M., Montegriffo, P., et al. 2022, arXiv, arXiv:2206.06143.
<https://arxiv.org/abs/2206.06143>
- De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, *MNRAS*, 449, 2604, doi: [10.1093/mnras/stv327](https://doi.org/10.1093/mnras/stv327)
- Di Matteo, P., Haywood, M., Lehnert, M. D., et al. 2019, *A&A*, 632, A4, doi: [10.1051/0004-6361/201834929](https://doi.org/10.1051/0004-6361/201834929)
- Eilers, A.-C., Hogg, D. W., Rix, H.-W., & Ness, M. K. 2019, *ApJ*, 871, 120, doi: [10.3847/1538-4357/aaf648](https://doi.org/10.3847/1538-4357/aaf648)
- Eilers, A.-C., Hogg, D. W., Rix, H.-W., et al. 2022, *ApJ*, 928, 23, doi: [10.3847/1538-4357/ac54ad](https://doi.org/10.3847/1538-4357/ac54ad)
- El-Badry, K., Bland-Hawthorn, J., Wetzel, A., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 652, doi: [10.1093/mnras/sty1864](https://doi.org/10.1093/mnras/sty1864)
- Forbes, D. A. 2020, *MNRAS*, 493, 847, doi: [10.1093/mnras/staa245](https://doi.org/10.1093/mnras/staa245)
- Frebel, A., & Norris, J. E. 2015, *ARA&A*, 53, 631, doi: [10.1146/annurev-astro-082214-122423](https://doi.org/10.1146/annurev-astro-082214-122423)
- Freeman, K., & Bland-Hawthorn, J. 2002, *ARA&A*, 40, 487, doi: [10.1146/annurev.astro.40.060401.093840](https://doi.org/10.1146/annurev.astro.40.060401.093840)
- Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2022, arXiv e-prints, arXiv:2208.00211.
<https://arxiv.org/abs/2208.00211>
- García Pérez, A. E., Cunha, K., Shetrone, M., et al. 2013, *ApJL*, 767, L9, doi: [10.1088/2041-8205/767/1/L9](https://doi.org/10.1088/2041-8205/767/1/L9)
- García Pérez, A. E., Ness, M., Robin, A. C., et al. 2018, *ApJ*, 852, 91, doi: [10.3847/1538-4357/aa9d88](https://doi.org/10.3847/1538-4357/aa9d88)
- Gravity Collaboration, Abuter, R., Amorim, A., et al. 2021, *A&A*, 647, A59, doi: [10.1051/0004-6361/202040208](https://doi.org/10.1051/0004-6361/202040208)
- Grinsztajn, L., Oyallon, E., & Varoquaux, G. 2022, Why do tree-based models still outperform deep learning on tabular data?, arXiv, doi: [10.48550/ARXIV.2207.08815](https://doi.org/10.48550/ARXIV.2207.08815)
- Gurvich, A. B., Stern, J., Faucher-Giguère, C.-A., et al. 2022, 16, 1. <https://arxiv.org/abs/2203.04321>

- Hasselquist, S., Hayes, C. R., Lian, J., et al. 2021, *The Astrophysical Journal*, 923, 172, doi: [10.3847/1538-4357/ac25f9](https://doi.org/10.3847/1538-4357/ac25f9)
- Hawkins, K., Jofré, P., Masseron, T., & Gilmore, G. 2015, *MNRAS*, 453, 758, doi: [10.1093/mnras/stv1586](https://doi.org/10.1093/mnras/stv1586)
- Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, *ApJ*, 808, 132, doi: [10.1088/0004-637X/808/2/132](https://doi.org/10.1088/0004-637X/808/2/132)
- Haywood, M., Di Matteo, P., Lehnert, M. D., et al. 2018, *ApJ*, 863, 113, doi: [10.3847/1538-4357/aad235](https://doi.org/10.3847/1538-4357/aad235)
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Nature*, 563, 85, doi: [10.1038/s41586-018-0625-x](https://doi.org/10.1038/s41586-018-0625-x)
- Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2021, *MNRAS*, 500, 1385, doi: [10.1093/mnras/staa2987](https://doi.org/10.1093/mnras/staa2987)
- Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2022, 33. <https://arxiv.org/abs/2204.04233>
- Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, *Natur*, 370, 194, doi: [10.1038/370194a0](https://doi.org/10.1038/370194a0)
- Katz, D., Sartoretti, P., Guerrier, A., et al. 2022, arXiv e-prints, arXiv:2206.05902. <https://arxiv.org/abs/2206.05902>
- Koch, A., McWilliam, A., Preston, G. W., & Thompson, I. B. 2016, *A&A*, 587, A124, doi: [10.1051/0004-6361/201527413](https://doi.org/10.1051/0004-6361/201527413)
- Kruijssen, J. M. D., Pfeffer, J. L., Reina-Campos, M., Crain, R. A., & Bastian, N. 2019, *MNRAS*, 486, 3180, doi: [10.1093/mnras/sty1609](https://doi.org/10.1093/mnras/sty1609)
- Kruijssen, J. M. D., Pfeffer, J. L., Chevance, M., et al. 2020, *MNRAS*, 498, 2472, doi: [10.1093/mnras/staa2452](https://doi.org/10.1093/mnras/staa2452)
- Lane, J. M. M., Bovy, J., & Mackereth, J. T. 2022, *MNRAS*, 510, 5119, doi: [10.1093/mnras/stab3755](https://doi.org/10.1093/mnras/stab3755)
- Lee, D. M., Johnston, K. V., Sen, B., & Jessop, W. 2015, *ApJ*, 802, 48, doi: [10.1088/0004-637X/802/1/48](https://doi.org/10.1088/0004-637X/802/1/48)
- Lindgren, L., Bastian, U., Biermann, M., et al. 2021, *A&A*, 649, A4, doi: [10.1051/0004-6361/202039653](https://doi.org/10.1051/0004-6361/202039653)
- Lucey, M., Hawkins, K., Ness, M., et al. 2021, *MNRAS*, 501, 5981, doi: [10.1093/mnras/stab003](https://doi.org/10.1093/mnras/stab003)
- Mackereth, J. T., Schiavon, R. P., Pfeffer, J., et al. 2019, *MNRAS*, 482, 3426, doi: [10.1093/mnras/sty2955](https://doi.org/10.1093/mnras/sty2955)
- Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. 2003, *The Astrophysical Journal*, 599, 1082, doi: [10.1086/379504](https://doi.org/10.1086/379504)
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, 154, 94, doi: [10.3847/1538-3881/aa784d](https://doi.org/10.3847/1538-3881/aa784d)
- Malhan, K., Ibata, R. A., & Martin, N. F. 2018, *Monthly Notices of the Royal Astronomical Society*, 481, 3442, doi: [10.1093/mnras/sty2474](https://doi.org/10.1093/mnras/sty2474)
- Malhan, K., Ibata, R. A., Sharma, S., et al. 2022, *The Astrophysical Journal*, 926, 107, doi: [10.3847/1538-4357/ac4d2a](https://doi.org/10.3847/1538-4357/ac4d2a)
- Marín-Franch, A., Chueca, S., Moles, M., et al. 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8450, *Modern Technologies in Space- and Ground-based Telescopes and Instrumentation II*, ed. R. Navarro, C. R. Cunningham, & E. Prieto, 84503S, doi: [10.1117/12.925430](https://doi.org/10.1117/12.925430)
- Marks, M., & Kroupa, P. 2010, *MNRAS*, 406, 2000, doi: [10.1111/j.1365-2966.2010.16813.x](https://doi.org/10.1111/j.1365-2966.2010.16813.x)
- Montegriffo, P., Bellazzini, M., De Angeli, F., & et al. 2022a, *A&A*, doi: [10.1051/0004-6361/202243709](https://doi.org/10.1051/0004-6361/202243709)
- Montegriffo, P., De Angel, F., Andrae, R., et al. 2022b, *A&A*, doi: [10.1051/0004-6361/202243880](https://doi.org/10.1051/0004-6361/202243880)
- Myeong, G. C., Belokurov, V., Aguado, D. S., et al. 2022, arXiv, arXiv:2206.07744. <https://arxiv.org/abs/2206.07744>
- Myeong, G. C., Vasiliev, E., Iorio, G., Evans, N. W., & Belokurov, V. 2019, *Monthly Notices of the Royal Astronomical Society*, 1247, 1235, doi: [10.1093/mnras/stz1770](https://doi.org/10.1093/mnras/stz1770)
- Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, *The Astrophysical Journal*, 901, 48, doi: [10.3847/1538-4357/abaef4](https://doi.org/10.3847/1538-4357/abaef4)
- Naidu, R. P., Conroy, C., Bonaca, A., et al. 2021, *ApJ*, 923, 92, doi: [10.3847/1538-4357/ac2d2d](https://doi.org/10.3847/1538-4357/ac2d2d)
- Naidu, R. P., Ji, A. P., Conroy, C., et al. 2022, *ApJL*, 926, L36, doi: [10.3847/2041-8213/ac5589](https://doi.org/10.3847/2041-8213/ac5589)
- Ness, M., Hogg, D. W., Rix, H. W., Ho, A. Y., & Zasowski, G. 2015, *Astrophysical Journal*, 808, 16, doi: [10.1088/0004-637X/808/1/16](https://doi.org/10.1088/0004-637X/808/1/16)
- Ness, M., Freeman, K., Athanassoula, E., et al. 2013, *MNRAS*, 430, 836, doi: [10.1093/mnras/sts629](https://doi.org/10.1093/mnras/sts629)
- Newberg, H. J., Yanny, B., & Willett, B. A. 2009, *Astrophysical Journal*, 700, 10, doi: [10.1088/0004-637X/700/2/L61](https://doi.org/10.1088/0004-637X/700/2/L61)
- Orkney, M. D. A., Laporte, C. F. P., Grand, R. J. J., et al. 2022, arXiv, arXiv:2206.09246. <https://arxiv.org/abs/2206.09246>
- Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, *A&A*, 369, 339, doi: [10.1051/0004-6361:20010085](https://doi.org/10.1051/0004-6361:20010085)
- Pillepich, A., Madau, P., & Mayer, L. 2015, *ApJ*, 799, 184, doi: [10.1088/0004-637X/799/2/184](https://doi.org/10.1088/0004-637X/799/2/184)
- Price-Whelan, A., Sipőcz, B., Lenz, D., et al. 2020, *adrn/gala: v1.3*, v1.3, Zenodo, doi: [10.5281/zenodo.4159870](https://doi.org/10.5281/zenodo.4159870)
- Price-Whelan, A. M. 2017, *The Journal of Open Source Software*, 2, doi: [10.21105/joss.00388](https://doi.org/10.21105/joss.00388)
- Recio-Blanco, A., de Laverny, P., Palicio, P. A., et al. 2022, arXiv e-prints, arXiv:2206.05541. <https://arxiv.org/abs/2206.05541>

- Reggiani, H., Schlafman, K. C., Casey, A. R., & Ji, A. P. 2020, *AJ*, 160, 173, doi: [10.3847/1538-3881/aba948](https://doi.org/10.3847/1538-3881/aba948)
- Renaud, F., Agertz, O., Andersson, E. P., et al. 2021a, *MNRAS*, 503, 5868, doi: [10.1093/mnras/stab543](https://doi.org/10.1093/mnras/stab543)
- Renaud, F., Agertz, O., Read, J. I., et al. 2021b, *MNRAS*, 503, 5846, doi: [10.1093/mnras/stab250](https://doi.org/10.1093/mnras/stab250)
- Richer, H. B., Fahlman, G. G., Brewer, J., et al. 2004, *AJ*, 127, 2771, doi: [10.1086/383543](https://doi.org/10.1086/383543)
- Rix, H.-W., Chandra, V., & Andrae, R. 2022, The Poor Old Heart of the Milky Way, latest, Zenodo, doi: [10.5281/zenodo.7035809](https://doi.org/10.5281/zenodo.7035809)
- Rix, H.-W., Hogg, D. W., Boubert, D., et al. 2021, *The Astronomical Journal*, 162, 142, doi: [10.3847/1538-3881/ac0c13](https://doi.org/10.3847/1538-3881/ac0c13)
- Sanders, J. 2012, *MNRAS*, 426, 128, doi: [10.1111/j.1365-2966.2012.21698.x](https://doi.org/10.1111/j.1365-2966.2012.21698.x)
- Schlaufman, K. C., & Casey, A. R. 2014, *Astrophysical Journal*, 797, doi: [10.1088/0004-637X/797/1/13](https://doi.org/10.1088/0004-637X/797/1/13)
- Shipp, N., Drlica-Wagner, A., Balbinot, E., et al. 2018, *The Astrophysical Journal*, 862, 114, doi: [10.3847/1538-4357/aacdad](https://doi.org/10.3847/1538-4357/aacdad)
- Smith, R., Flynn, C., Candlish, G. N., Fellhauer, M., & Gibson, B. K. 2015, *MNRAS*, 448, 2934, doi: [10.1093/mnras/stv228](https://doi.org/10.1093/mnras/stv228)
- Starkenbourg, E., Martin, N., Youakim, K., et al. 2017, *MNRAS*, 471, 2587, doi: [10.1093/mnras/stx1068](https://doi.org/10.1093/mnras/stx1068)
- Strömgren, B. 1966, *ARA&A*, 4, 433, doi: [10.1146/annurev.aa.04.090166.002245](https://doi.org/10.1146/annurev.aa.04.090166.002245)
- Tumlinson, J. 2010, *ApJ*, 708, 1398, doi: [10.1088/0004-637X/708/2/1398](https://doi.org/10.1088/0004-637X/708/2/1398)
- Weinberg, D. H., Andrews, B. H., & Freudenburg, J. 2017, *ApJ*, 837, 183, doi: [10.3847/1538-4357/837/2/183](https://doi.org/10.3847/1538-4357/837/2/183)
- Wetzel, A., Hayward, C. C., Sanderson, R. E., et al. 2022, *arXiv*, arXiv:2202.06969, <https://arxiv.org/abs/2202.06969>
- Xiang, M., & Rix, H.-W. 2022, *Nature*, 603, 599, doi: [10.1038/s41586-022-04496-5](https://doi.org/10.1038/s41586-022-04496-5)
- Xiang, M., Ting, Y.-S., Rix, H.-W., et al. 2019, *The Astrophysical Journal Supplement Series*, 245, 34, doi: [10.3847/1538-4365/ab5364](https://doi.org/10.3847/1538-4365/ab5364)
- Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, *AJ*, 137, 4377, doi: [10.1088/0004-6256/137/5/4377](https://doi.org/10.1088/0004-6256/137/5/4377)
- Yuan, Z., Chang, J., Beers, T. C., & Huang, Y. 2020, *ApJL*, 898, L37, doi: [10.3847/2041-8213/aba49f](https://doi.org/10.3847/2041-8213/aba49f)
- Zolotov, A., Willman, B., Brooks, A. M., et al. 2009, *ApJ*, 702, 1058, doi: [10.1088/0004-637X/702/2/1058](https://doi.org/10.1088/0004-637X/702/2/1058)
- . 2010, *ApJ*, 721, 738, doi: [10.1088/0004-637X/721/1/738](https://doi.org/10.1088/0004-637X/721/1/738)