Research in Astronomy and Astrophysics, 25:105006 (32pp), 2025 October
https://doi.org/10.1088/1674-4527/adf38a
CSTR: 32081.14.RAA.adf38a
© 2025. National Astronomical Observatories, CAS and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved. Printed in China.

Chasing the Formation History of the Galactic Metal-poor Disk

Xiao-Kun Hou (侯晓琨)[1,2], Rui-Zhi Zhang[1], Hai-Ning Li[1], and Gang Zhao[1,2]
[1] CAS Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; gzhao@nao.cas.cn
[2] School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China
Received 2025 April 9; revised 2025 July 19; accepted 2025 July 21; published 2025 August 28

Abstract

In our previous work, we identified ∼100,000 metal-poor stars ([Fe/H] < -1.0) from the LAMOST Survey. This work estimates their chemical abundances and explores the origin and evolution of the Galactic metal-poor disk. Our chemo-dynamical analysis reveals four main populations within the metal-poor disk: (1) a primordial disk older than 12 Gyr with [Fe/H] > -1.5; (2) debris stars from the progenitor galaxy of Gaia–Sausage–Enceladus (GSE), but now residing in the Galactic disk; (3) the metal-poor tail of the metal-rich, high-α disk formed 10–12 Gyr ago, with metallicity lower limit extending to -2.0; (4) the metal-poor tail of the metal-rich, low-α disk younger than 8 Gyr, reaching a lower metallicity limit of -1.8. These results reveal the presence of a primordial disk and show that both high-α and low-α disks reach lower metallicities than previously thought. Analysis of merger debris reveals that Wukong, with extremely low metallicity, likely originates from merger events distinct from GSE. Additionally, three new substructures are identified: ShangGu-1, characterized by unusual [Fe/H]-eccentricity correlations; ShangGu-2, possibly heated disk stars; and ShangGu-3, which can be divided into four subgroups based on differing orbital directions, with two aligning with the previously known Nyx and Nyx-2.

Key words: catalogs – methods: data analysis – Galaxy: disk – Galaxy: formation – stars: abundances

1. Introduction

According to galaxy formation models, the Milky Way grew through mergers with dwarf galaxies, living behind many merger remnants. Metal-poor stars ([Fe/H] < -1.0), typically old and formed in chemically primitive environments, serve as excellent tracers of this process. However, their rarity makes their identification both challenging and essential. The LAMOST Survey (Zhao et al. 2006, 2012) has collected tens of millions of low-resolution stellar spectra, providing an unprecedented opportunity to search for metal-poor stars. Li et al. (2018) identified 10,000 very metal-poor stars ([Fe/H] < -2.0) in LAMOST DR3, and we further expanded this number to over 30,000 (Hou et al. 2024) utilizing the more updated DR9. During our search for very metal-poor stars, we also discovered approximately 100,000 metal-poor stars in total, with most of them having reliable atmospheric parameters and velocities, making them ideal for investigating the assembly process of the Milky Way. Estimating elemental abundances is crucial, as chemical patterns reveal a star's formation environment and help distinguish whether it formed in situ or was accreted from dwarf galaxies. To achieve this, we use machine learning and template fitting to determine the abundances of various elements for these metal-poor stars.

The bimodal metallicity distribution between the Galactic halo and disk reflects their distinct formation histories: the metal-poor halo comprises primarily ancient stars formed during the early chaotic assembly phase, while the metal-rich disk ([Fe/H] > -1.0) is dominated by younger stars shaped by post-merger secular evolution (Bland-Hawthorn & Gerhard 2016). Recent discoveries of a substantial metal-poor disk population (e.g., Sestito et al. 2020; Bellazzini et al. 2024; Nepal et al. 2024) challenge this view. Three principal hypotheses attempt to resolve this paradox. One hypothesis suggests that metal-poor disk stars are remnants of dwarf galaxies that merged with the Milky Way (Malhan & Rix 2024). Another proposes that they are merely the metal-poor tail of the disk population, which has been overlooked due to their smaller numbers compared to the metal-rich disk stars (Naidu et al. 2020). A more novel perspective argues that these stars originate from the Milky Way's ancient primordial disk, distinct from the commonly recognized high-α and low-α disks (Xiang et al. 2024). Observations of high-redshift galaxies also suggest that disk-like structures can emerge even in the early stages of galaxy evolution (e.g., Fujimoto et al. 2024; Rowland et al. 2024; Baker et al. 2025). However, some researchers argue that the concept of a primordial disk is unnecessary. For instance, Dillamore et al. (2023) proposed that resonances with the Galactic bar could lead to some halo stars acquiring disk-like orbits. Similarly, Li et al. (2024) suggested that bulge stars could also evolve into disk-like orbits through the influence of the decelerating Galactic bar.

Beyond the origin of the Galactic metal-poor disk, many substructures in the Galactic stellar halo trace past mergers that shaped the Milky Way. One of the earliest and most prominent examples is the Sagittarius dwarf galaxy, discovered by Ibata et al. (1994), which provided definitive evidence of an ongoing merger event. Subsequently, Helmi et al. (1999) identified a kinematically coherent stream, suggesting the remnants of a disrupted dwarf galaxy. With the advent of Gaia data, more relics came to light: Belokurov et al. (2018), Helmi et al. (2018) pinpointed the Gaia–Sausage–Enceladus (GSE) merger as a major contributor to the inner halo. Numerous other substructures, such as Sequoia (Myeong et al. 2019), Thamnos (Koppelman et al. 2019), and Wukong (Naidu et al. 2020), have also been discovered. However, their precise origins often remain controversial. For instance, Horta et al. (2023) argued that Wukong might share a common progenitor with GSE, while other studies dispute that connection. These findings underscore the Milky Way's complex merger history and the need for more detailed chemical and kinematic data to distinguish in situ from accreted populations, refine our understanding of known substructures, and discover any yet-undetected mergers.

The paper is structured as follows: Section 2 describes the data and methods used in this study. Section 3 presents the main results. Section 4 analyzes these results in detail, discussing the formation process of the Galactic metal-poor disk, the possible origins of merger remnants, and properties of three newly discovered substructures, while comparing our findings with previous studies. Finally, Section 5 summarizes the key discoveries.

2. Data and Methods

In our previous work (Hou et al. 2024), we identified ~100,000 metal-poor ([Fe/H] < -1.0) stars from LAMOST low-resolution spectra, which are valuable for studying the Milky Way's formation history. In this study, we supplemented these stars with chemical abundances and kinematic properties to explore the formation process of the Galactic metal-poor disk. We estimated multi-element abundances using two methods: machine learning and template fitting. Machine learning provided abundances for 17 elements (C, Na, Mg, Si, Ca, Sc, Ti, Cr, Mn, Fe, Co, Ni, Zn, Sr, Ba, La and Eu) and three atmospheric parameters: the effective temperature ($T_{\rm eff}$), surface gravity (log g), and microturbulence velocity ($v_{\rm mic}$), while template fitting estimated abundances for four elements (Fe, Mg, Ca and C). Using a subset of these stars, we explored some key questions about the Milky Way. This section begins by detailing the process of estimating abundances for all metal-poor stars, followed by the methods used to calculate their kinematic parameters. Next, we describe the criteria for selecting sample stars for further study. Finally, we outline the classification of these stars into the kinematically cold disk, kinematically hot disk, and various well-known merger remnants.

2.1. Estimating Abundances

2.1.1. Machine Learning

We developed a neural network to predict chemical abundances and atmospheric parameters. The model is based on the Bayesian convolutional neural network framework from astroNN (Leung & Bovy 2019). To account for the influence of model structure on predicted values, this framework incorporates a dropout mechanism during prediction, which randomly deactivates some nodes. The loss function used in the model is defined as

$$\mathrm{Loss}_i = \begin{cases} \frac{1}{2}(\hat{y_i} - y_i)^2 e^{-s_i} + \frac{1}{2}s_i & \text{for } y_i \neq \mathrm{Nan} \\ 0 & \text{for } y_i = \mathrm{Nan} \end{cases}. \quad (1)$$

Here y_i and $\hat{y_i}$ are the label and predicted value for ith reference star, respectively, and $s_i = \ln[\sigma^2_{\mathrm{known},i} + \sigma^2_{\mathrm{predicted},i}]$, where $\sigma^2_{\mathrm{known},i}$ and $\sigma^2_{\mathrm{predicted},i}$ are the uncertainties of the label and prediction. This loss function serves two purposes: it handles missing labels during training by ignoring them and incorporates both label and prediction uncertainties, enabling the model to output predictions along with their associated errors. The model input comprises normalized LAMOST low-resolution spectra within the wavelength range of 3900–5500 Å. This range was chosen because it includes key spectral features such as the Ca II H and K lines (3933 and 3968 Å), the CH G-band (4300 Å), and the Mg I triplet (5167–5184 Å). Notably, this differs from the wavelength range used in our previous work (Hou et al. 2024).

The reference set for the model includes 385 stars from Li et al. (2022a) and 293 stars from the overlap between JINAbase (Abohalima & Frebel 2018) and LAMOST. The stars from Li et al. (2022a) are metal-poor, initially identified using LAMOST low-resolution spectra and later observed with high-resolution spectroscopy for precise chemical abundances. JINAbase is a database compiling metal-poor stars and their chemical abundances from various studies. We incorporated metal-poor stars from these two distinct sources into our reference set to ensure that the metallicity coverage closely aligns with that of our target stars. Among the target stars, approximately 23.3% have metallicities [Fe/H] > -1.5. However, within the 385 stars from Li et al. (2022a), only two stars meet this criterion, representing just about 0.5% of the sample. In contrast, among the 293 overlapping stars from JINAbase and LAMOST, 57 stars (approximately 19.5%) have [Fe/H] > -1.5. Moreover, the highest metallicity recorded in Li et al. (2022a) is -0.63, which still does not sufficiently cover the metallicity range of the target stars. These statistics highlight the necessity of supplementing Li et al. (2022a) with

stars from the JINAbase-LAMOST overlap to achieve a more comprehensive coverage. To further illustrate this point, we included histograms (Figure A1) showing the $T_{\rm eff}$, log g and metallicity distributions of the target sample, as well as those of Li et al. (2022a), the JINAbase-LAMOST overlap, and their combined reference set. While the overlapping stars from JINAbase and LAMOST span a broader metallicity range, their abundances and atmospheric parameters are drawn from diverse studies, which introduce some inconsistency. Therefore, we integrated these stars with those from Li et al. (2022a) to build a reference set that balances both breadth of metallicity coverage and data uniformity.

The comparison of common stars between Li et al. (2022a) and JINAbase, shown in Figure A2, revealed systematic differences in log g and the abundances of Na, Mg, C, Si, Sc, Ti, Sr, and Ba, as their values deviate from a strict 1:1 relation. For each parameter (log g and the abundance of each element), we fitted a linear relationship between values from JINAbase and Li et al. (2022a)

$$y = a \times x + b. \tag{2}$$

Here, x represents the JINAbase value for a given parameter (log g or [X/H], with X representing Na, Mg, C, Si, Sc, Ti, Sr, or Ba), and y corresponds to the respective value from Li et al. (2022a). The red dotted lines in Figure A2 show these linear fits, and the corresponding parameters a and b for each relationship are provided in Table A1. These linear functions were then used to calibrate JINAbase values to the scale of Li et al. (2022a), which provides a more consistent data set with uniformly analyzed abundances, while log g is derived from precise Gaia parallaxes. After this calibration, the reference stars were then divided into a training set and a test set in a 9:1 ratio. The model was trained on the training set, and its reliability was evaluated by comparing predicted values with labels in both sets, as shown in Figure A3. To assess consistency, we calculated the mean and standard deviation of the differences between predicted values and labels. Overall, the predictions show good agreement with the labels for all atmospheric parameters and chemical abundances. We further analyzed the relationship between the absolute difference between predicted values and labels and the prediction errors, as shown in Figure A4. The results indicate a positive correlation for most parameters and abundances, with larger prediction errors aligning with greater discrepancies between predicted and label values.

Beyond comparing predictions with labels within the reference set, we also performed an external comparison. Li et al. (2022b) used a machine learning approach to estimate atmospheric parameters and chemical abundances for a large subset of stars from LAMOST DR8. By cross-matching our catalog with theirs, we identified ∼20,000 common stars. Figure A5 compares our estimations of $T_{\rm eff}$, log g, [Fe/H], and the abundances of C, Mg, Si, Ca, Mn, and Ni with their corresponding values. The results demonstrate strong consistency for $T_{\rm eff}$ and log g. For [Fe/H] and the abundances of C, Mg, Ca, Mn, and Ni, our values are systematically lower but still exhibit overall agreement. These systematic differences likely arise from the use of different reference stars, as Li et al. (2022b) adopted stars common to LAMOST and APOGEE (Majewski et al. 2017) as their reference set. Overall, these findings indicate the robustness of our estimations.

During the spectral normalization process, we employed polynomial fitting to derive a pseudo-continuum and obtained the normalized spectrum by dividing the observed spectrum by this pseudo-continuum. However, our analysis identified several poor-quality spectra, including cases with lost flux or significant cosmic-ray contamination, which adversely affected the quality of the normalized spectra. To evaluate the normalization quality, we calculated a new pseudo-continuum from the normalized spectrum. If this pseudo-continuum was smooth and had a mean value close to 1.0, the normalized spectrum was considered high quality and assigned a flag of 0; otherwise, it was flagged as 1. Finally, we applied the model to previously identified metal-poor stars, and the predicted abundances along with their associated uncertainties are compiled into a catalog, available via the link provided at the end of the article.

2.1.2. Template Fitting

Although the neural network model performed well on the reference set, we employed an independent approach to further verify our measurements. Specifically, we used template fitting to re-estimate the abundances of Fe, Mg, Ca, and C, chosen for their strong absorption lines in LAMOST low-resolution spectra. The wavelength ranges used for fitting the abundance of each element are listed in Table A2. For other elements, factors such as low abundances, weak absorption features, severe line blending and noise make template fitting impractical.

To identify the best-fitting template for each observed spectrum, we applied a Bayesian method. In this approach, the abundance of Fe, Mg, Ca, or C is assumed to follow a Gaussian prior distribution

$$P(\theta) = \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left(-\frac{(\theta - \mu_\theta)^2}{2\sigma_\theta^2}\right), \tag{3}$$

where θ represents the abundance, and μ_θ and σ_θ are the mean and standard deviation respectively derived from the neural network model's predictions and uncertainties. Next, we constructed a likelihood function to quantify the agreement between the template and observed spectrum

$$P(S_{\rm obs}|\theta) = \exp\left(-\frac{1}{2}\chi^2(\theta)\right), \tag{4}$$

where $\chi^2(\theta)$ is calculated as

$$\chi^2(\theta) = \sum_i \frac{(S_{\mathrm{obs},i} - S_{\mathrm{template},i}(\theta))^2}{\sigma_i^2}. \quad (5)$$

Here, $S_{\mathrm{obs},i}$ and $S_{\mathrm{template},i}$ are the normalized fluxes of the observed spectrum and template at the ith pixel, and σ_i is the standard deviation of the observed flux. Smaller $\chi^2(\theta)$ values indicate better agreement between the template and observed spectrum, leading to a higher likelihood. Finally, the posterior distribution is obtained by combining the prior distribution and the likelihood function

$$P(\theta|S_{\mathrm{obs}}) = \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left(-\frac{(\theta - \mu_0)^2}{2\sigma_\theta^2}\right) \cdot \exp\left(-\frac{1}{2}\chi^2(\theta)\right). \quad (6)$$

The process of generating a template involves several steps. First, we use KMOD, a publicly available IDL-based code[3] widely adopted in many investigations (e.g., Mészáros & Prieto 2013; Yong et al. 2013), to interpolate the ATLAS9 atmospheric model grid (Castelli & Kurucz 2003). This interpolation is performed to obtain the target atmospheric model corresponding to the desired T_{eff}, $\log g$, and [Fe/H]. Once the atmospheric model is established, we employ the SPECTRUM (Gray & Corbally 1994) program to generate a template based on the sampled chemical abundances.

Abundances were estimated as the mean value of samples drawn from the posterior distribution. Their errors arise from two sources: (1) posterior distribution uncertainty, defined as half the difference between the 16th and 84th quantiles of the posterior samples, and (2) the propagated uncertainties from the atmospheric parameters T_{eff}, $\log g$, and [Fe/H]. Since our abundance measurements are based on template fitting, where the stellar parameters are fixed and the chemical abundance of the template is varied to match the observed spectrum, any uncertainty in the stellar parameters can directly affect the derived abundances. To account for this, we varied each parameter by its $\pm 1\sigma$ uncertainty and re-estimated the abundance. The difference in abundance caused by this variation was taken as the contribution to the abundance error from that parameter. This process was performed separately for T_{eff}, $\log g$, and [Fe/H]. The final abundance uncertainty was calculated as the square root of the sum of the squares of the posterior uncertainty and the individual contributions from the three parameters. For iron, only the posterior uncertainty and the contributions from T_{eff} and $\log g$ were considered.

We compared the Fe, Mg, Ca, and C abundances estimated from template fitting and machine learning, as shown in Figures A6, A7, A8, and A9. Each figure includes two panels: the left shows stars with flag = 0, while the right shows stars after error screening. For the error screening, we required that the errors from both machine learning and template fitting must not exceed 0.2 for Fe, Mg, and Ca, while a less stringent threshold of 0.3 was applied for C. After error filtering, the absolute mean differences between two methods for Fe, Mg, and Ca are 0.01, 0.03, and 0.1, respectively, with standard deviations of 0.21, 0.21, and 0.24, indicating excellent consistency. However, significant discrepancies were observed for C, with template fitting yielding higher abundances. We attribute this discrepancy to the neural network's reference set, which included labels from both JINAbase and Li et al. (2022b). As described in Section 2.1.1, [C/H] from these two sources does not follow a strict 1:1 relation. To address this, we corrected JINAbase values to the scale of Li et al. (2022b) using a linear function derived from stars common to both sources. However, due to the small number of the common stars, this calibration may be unreliable. To assess the accuracy of [Fe/H] and the abundances of C, Mg, and Ca obtained by template fitting, we compared our estimations to those from Li et al. (2022b), as shown in Figure A10. The comparison shows good agreement but our values are systematically lower than those reported by Li et al. (2022b). This offset aligns with the findings in Section 2.1.1, where we compared abundances obtained by machine learning to those from Li et al. (2022b). We estimated the abundances of Fe, Mg, Ca, and C for all spectra with flag = 0 using template fitting and compiled the results together with those from machine learning in a single catalog.

2.2. Computing Phase-space Quantities

The Integrals of Motion (IoMs) of stars from the same stellar system would remain coherent as the Galaxy evolved, which can be used to identify the Galactic substructures. To compute the positions, velocities, IoMs, and orbital parameters for further classification, we first selected stars with precise astrometric measurements in our sample.

We cross-matched our metal-poor stars with Gaia DR3 catalog (Gaia Collaboration et al. 2023) and the `StarHorse` photo-astrometric distance catalogs (Anders et al. 2022). Stars with relative distance uncertainties of Anders et al. (2022) greater than 20% or Gaia RUWE greater than 1.4 were removed since their astrometric measurements are unreliable (Lindegren et al. 2021). For stars with multiple observation results in LAMOST, we only kept the record with the highest signal-to-noise ratio.

We computed positions and velocities with coordinates and proper motions from Gaia DR3, photo-astrometric distance from Anders et al. (2022), and radial velocities from LAMOST DR10 or Gaia DR3. We prioritized the radial velocity measurement from LAMOST DR10; otherwise, we adopted the value from Gaia DR3. We assumed that the Sun is located 8.21 kpc from the Galactic Center (McMillan 2017) and 20.8 pc above the Galactic plane (Bennett & Bovy 2019). We adopted the motion of the local standard of rest (LSR) as $V_{\mathrm{LSR}} = 233.1$ km s^{-1} (McMillan 2017)

[3] http://leda.as.utexas.edu/stools

and the solar peculiar motion as $(U_\odot, V_\odot, W_\odot) = (11.1, 12.24, 7.25)$ km s^{-1} (Schönrich et al. 2010).

We further computed IoMs (e.g., energy, actions) and orbital parameters (e.g., eccentricity, maximum height to the Galactic plane) using the software `Agama` (Vasiliev 2019) with the axisymmetric potential model of the Milky Way (McMillan 2017). We assumed Gaussian uncertainties for the observables and ran a Monte Carlo (MC) simulation 10,000 times for each star to estimate the uncertainties of the dynamical parameters. The value and the uncertainty of each quantity are represented by the median value and half the difference between the 16th and 84th quantiles.

In total, we obtained phase-space quantities for ∼74,000 stars.

2.3. Stars for Studying the Milky Way

To study the Milky Way, we selected 46,178 stars from the ∼74,000 stars with phase-space quantities computed in Section 2.2, based on the precision of their kinematic parameters:

(i) $J_r > 0$ km s^{-1}
(ii) $J_z > 0$ km s^{-1}
(iii) $J_{\phi,\mathrm{err}} < 180$ km s^{-1}
(iv) $E_{\mathrm{err}}/E < 0.3$
(v) $J_{r,\mathrm{err}}/J_r < 0.3$
(vi) $J_{z,\mathrm{err}}/J_z < 0.3$

Here, E represents orbital energy, while J_ϕ, J_r, and J_z are the azimuthal, radial, and vertical actions respectively, with E_{err}, $J_{\phi,\mathrm{err}}$, $J_{r,\mathrm{err}}$, and $J_{z,\mathrm{err}}$ as their respective uncertainties. As shown in Figure 1, our sample spans diverse evolutionary stages, with the majority of stars concentrated at the main-sequence turnoff, and others distributed across the subgiant, giant, and asymptotic giant branch phases. Thanks to the exceptionally high observational efficiency of the LAMOST survey, we have assembled such a remarkably large sample of metal-poor stars, enabling detailed studies of the Galactic evolution, particularly regarding the Galactic metal-poor disk. Previous investigations into the Galactic metal-poor disk have predominantly relied on photometric surveys such as Gaia and Pristine (e.g., Sestito et al. 2020; Fernández-Alvar et al. 2021; Bellazzini et al. 2024; Hong et al. 2024). Although these photometric surveys offer metallicity measurements for hundreds of millions of stars, their chemical abundance information typically remains limited to iron. Consequently, researchers often supplement these data sets with high-resolution spectroscopic surveys like GALAH and APOGEE (e.g., Carter et al. 2021; Fiorentin et al. 2021; Fernández-Alvar et al. 2024; Malhan & Rix 2024; Re Fiorentin et al. 2024). However, integrating photometric and spectroscopic data substantially reduces the available sample size. Leveraging LAMOST's extensive data set, we can maintain a large stellar sample while simultaneously exploring multi-element abundance patterns, thus providing deeper insights into the composition and kinematics of the Galactic metal-poor disk.

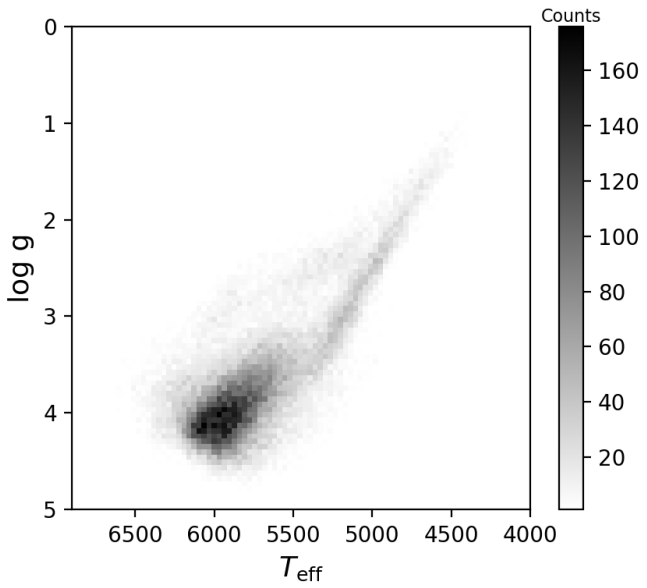


Figure 1. Kiel diagram of the selected stars for studying the Milky Way, with grayscale indicating the star count density.

For our analysis, we primarily use the abundances of Fe, Mg, and Ca obtained by machine learning, applying strict criteria to ensure the reliability: both the machine learning and template fitting uncertainties must be below 0.2, and the absolute difference between predicted values from two methods must not exceed 0.2. When these stringent criteria result in insufficient stars for study, we relax the standards to only require machine learning uncertainties below 0.2, explicitly noting when this applies. Additionally, we include the abundances of Sc, which are measured only through machine learning, with uncertainty thresholds of 0.15.

Additionally, we incorporated stellar ages from Xiang & Rix (2022), who estimated the ages of over 200,000 LAMOST subgiant stars using the isochrone fitting method. To ensure the subgiant phase, we selected stars with T_{eff} and $\log g$ errors below 100 K and 0.2, respectively. We further required ages less than 15 Gyr, age uncertainties within 1 Gyr, and relative age uncertainties under 0.2. The filtered stars were cross-matched with our sample, integrating their age data for further analysis.

In conclusion, our sample contains 46,178 stars with chemical abundances, spatial positions, orbital parameters, and IoM, including 2517 subgiant stars with precise age estimates.

2.4. Classification of Sample Stars

Based on their kinematic properties, we classified the 46,178 sample stars into two main categories: the kinematically cold disk and kinematically hot disk, as well as six well-known merger remnants: GSE, Pontus, Thamnos, Sequoia, Helmi

Table 1
Classification Criteria for Dividing Sample Stars and DTGs into the Kinematically Cold Disk, Kinematically Hot Disk, and Six Merger Remnants

Name	Criteria
Kinematically cold disk	$J_\phi > 1.66$ ($\times 10^3$ kpc km s^{-1}); $e < 0.2$; $\lvert V_z\rvert < 15$ (km s^{-1})
Kinematically hot disk	$J_\phi < 1.66$ ($\times 10^3$ kpc km s^{-1}); $0.2 < e < 0.6$; $\lvert V_z\rvert < 25$ (km s^{-1}); $-200 < V_\phi < -90$ (km s^{-1});
GSE	$e > 0.7$; $\lvert V_\phi\rvert < 100$ (km s^{-1})
Pontus	$-1.72 < E < -1.56$ ($\times 10^5$ km^2 s^{-2}); $-470 < J_\phi < 5$ (kpc km s^{-1}); $0.245 < J_r < 0.725$ ($\times 10^3$ kpc km s^{-1}); $0.115 < J_z < 0.545$ ($\times 10^3$ kpc km s^{-1}); $0.39 < L_\perp < 0.865$ ($\times 10^3$ kpc km s^{-1}); $0.5 < e < 0.8$; $8 < r_{\rm apo} < 13$ (kpc); $1 < r_{\rm peri} < 3$ (kpc);
Thamnos	$-1.8 < E < -1.6$ ($\times 10^5$ km^2 s^{-2}); $e < 0.7$; $J_\phi < 0$ (kpc km s^{-1})
Sequoia	$E > -1.6$ ($\times 10^5$ km^2 s^{-2}); $J_\phi < -0.7$ ($\times 10^3$ kpc km s^{-1}); $\eta < -0.15$
Helmi Streams	$0.75 < J_\phi < 1.7$ ($\times 10^3$ kpc km s^{-1}); $1.6 < L_\perp < 3.2$ ($\times 10^3$ kpc km s^{-1})
Wukong	$0 < J_\phi < 1$ ($\times 10^3$ kpc km s^{-1}); $E < -1.15$ ($\times 10^5$ km^2 s^{-2}); $(J_z - J_r)/J_{\rm tot} > 0.3$; $90^\circ < \arccos(L_z/L) < 120^\circ$

Note. E denotes orbital energy. J_ϕ, J_r, and J_z are the azimuthal, radial, and vertical components of action ($\boldsymbol{J}$) respectively, with $J_{\rm tot}$ being their quadrature sum. V_ϕ and V_z represent azimuthal and vertical velocities respectively. e and η stand for eccentricity and circularity, while $r_{\rm apo}$ and $r_{\rm peri}$ are the apogalacticon and perigalacticon radii respectively. L_z is the azimuthal angular momentum component and $L_\perp$ ($L_\perp = \sqrt{L_x^2 + L_y^2}$) is the perpendicular component.

Stream, and Wukong. The detailed classification criteria are provided in Table 1. However, since the Galactic disks and remnants lack clear kinematic boundaries, the initial classification may result in mutual contamination, potentially affecting subsequent analysis. To mitigate this, we applied the HDBSCAN algorithm (Campello et al. 2015) to cluster the stars in a space defined by energy (E) and actions (J_ϕ, J_r, and J_z). This approach identified 6549 stars grouped into 222 Dynamically Tagged Groups (DTGs), following the naming convention of Yuan et al. (2020b). HDBSCAN is well-suited for identifying groups in data with varying densities, and it does not require a pre-defined number of clusters. The reason why not all stars in our sample were clustered is that HDBSCAN is designed to find dense regions of data while discarding sparse regions as noise. This feature ensures that only stars with sufficient density are grouped into meaningful clusters, while stars that are too isolated or do not form a significant structure are excluded. For our analysis, we carefully selected the parameters for HDBSCAN to ensure that the clusters formed were both robust and meaningful. Specifically, we set the min_cluster_size to 15, meaning that a group had to consist of at least 15 stars to be considered a valid cluster. We also used a min_samples value of 6, which defines the number of neighboring points required to form a dense region. Additionally, the metric was set to "euclidean" to measure the similarity between stars in a straightforward manner, and the cluster_selection_method was set to "leaf," which ensures that the final clusters are chosen based on density and not by a hierarchical tree structure. These parameter choices were carefully optimized and tested to capture the true dynamical structures of the stellar populations, while minimizing the impact of statistical fluctuations and noise. Each DTG was assigned to one of the previously mentioned structures or remnants based on the mean kinematic parameters of the stars within the group. Specifically, the mean kinematic values of the member stars were used as the representative parameters for the entire DTG, ensuring that the classification reflected the overall dynamical behavior of the DTG. Based on this approach, 194 DTGs were successfully associated with known structures, while 28 DTGs did not align with the kinematic criteria of any known structure or remnant listed in Table 1. These 28 DTGs were further analyzed and grouped into six distinct categories. The criteria for this grouping, along with a detailed analysis of these six groups, are provided in Section 3.3. Due to the large number of DTGs, the kinematic parameters of all stars, along with a summary table of the DTG properties, have been made available through a link at the end of the article.

Yoshii (1982) and Gilmore & Reid (1983) found that the spatial distribution of stars in the Galactic disk can be well fitted by two exponential components. One component exhibits a larger scale height, while the other has a smaller scale height, leading to the morphological classification of the disk into "thick disk" and "thin disk." With the advancement of spectroscopic surveys, the chemical abundances of an increasing number of stars have been measured. Bovy et al. (2012) used data from the Sloan Digital Sky Survey (SDSS), proposing that the Galactic disk can also be divided into two components based on α abundance: high-α disk and low-α disk. Stars in the high-α disk are generally older and more metal-poor, while those in the low-α disk tend to be younger and more metal-rich. Furthermore, Bovy et al. (2012) and Hayden et al. (2015) found significant geometric differences between the high-α and low-α disks. High-α disk stars are predominantly located within $\sim$10 kpc from the Galactic Center and tend to lie at vertical distances greater than

0.5 kpc from the midplane. In contrast, low-α disk stars remain prominent out to $\sim$15 kpc and are mainly confined within 0.5 kpc of the midplane. Bensby et al. (2014) geometrically classified stars into the thin and thick disks and found that most thick disk stars belong to the high-α disk, although the correspondence is not perfect. In summary, the Galactic disk can be classified not only based on geometry (thick and thin disks), but also in terms of chemical abundances (high-α and low-α disks). Most studies of the Galactic disk have focused on stars with metallicities [Fe/H] > -1, where the disk populations are most prominent. As noted earlier, the high-α disk is generally associated with the geometrically thick disk, while the low-α disk corresponds roughly to the geometrically thin disk. This approximate mapping has led to the commonly held view that the Milky Way disk is composed of two distinct stellar populations, both chemically and kinematically. However, the correspondence between chemical and geometric definitions is not perfect. Recent studies (e.g., Martig et al. 2016; Lian et al. 2025) have shown that some stars with geometrically thick disk kinematics are not α-enhanced and may even be relatively young. For example, at Galactocentric distances of 10–14 kpc, Lian et al. (2025) identified stars with thick disk-like orbital properties but [α/Fe] < 0.2 and ages around 6.6 Gyr. These findings suggest that the geometrically thick disk includes a chemically and temporally diverse stellar population that challenges the traditional paradigm.

This work primarily examines the properties of the Galactic disk at the metal-poor end, demonstrating that the situation becomes increasingly complex at lower metallicities ([Fe/H] < -1.0). The kinematically hot and cold disks do not represent single stellar populations, but rather include multiple distinct components. Therefore, in this work, the kinematically defined hot and cold disks do not directly correspond to the traditional metal-rich, high-α and low-α disk populations, and they differ from the thick and thin disks typically observed at the metal-rich end of the Galactic disk. To better characterize these components in the metal-poor disk, we classify stars based on their kinematic properties and then examine the distributions of age and chemical abundances across different groups. Similarly, we classify all stars with very high eccentricities as belonging to the GSE. In this paper, GSE specifically refers to these high-eccentricity stars, rather than to the progenitor dwarf galaxy itself. It is important to note that these classifications are not intended to define physically distinct stellar populations, but rather serve as a practical framework to facilitate a clearer analysis of the structure and evolution of the Galactic metal-poor disk.

3. Results

We provide an overview of the chemical and kinematic properties of the sample stars. Figure 2 illustrates their distributions in the E–J_ϕ and V_ϕ–V_r planes. In panels (a) and (c), colors represent stellar number density, while in panels (b) and (d), they indicate metallicity. It is important to note that the stellar samples shown in the left and right panels differ slightly due to quality cuts. Panels (a) and (c) include all stars with reliable orbital parameters, as selected by the criteria described in Section 2.3. In contrast, panels (b) and (d) only include stars with high-quality [Fe/H] measurements: we require both the machine learning and template fitting estimates of [Fe/H] to have uncertainties smaller than 0.2, and the absolute difference between the two methods must also be less than 0.2. These stricter conditions reduce the number of stars shown in the metallicity panels. Panel (c) clearly highlights the GSE, characterized by nearly no net rotation ($V_\phi \approx 0$), and a wide radial velocity range ($|V_r|$ from 0 to 400 km s^{-1}). The kinematically hot disk is also evident with V_ϕ around -180 km s^{-1} and V_r near 0, while stars with V_ϕ under -200 km s^{-1} correspond to the kinematically cold disk. In panel (a), the number density shows a smooth, hook-like transition between GSE, the kinematically hot disk, and the kinematically cold disk, with the highest density in the kinematically disk region. Panels (b) and (d) reveal the kinematically cold disk has the highest metallicity, generally above -1.25. As discussed in Section 2.4, we classified these sample stars into the kinematically cold disk, hot disk, and six dwarf galaxy debris. Figure 3 illustrates their distributions in kinematic space, while Figure 4 displays the distributions of DTGs. For 28 unclassified DTGs, we grouped them into six groups (Clusters 1–6), marking them with hollow triangles, squares, diamonds, hexagons, plus signs, and crosses.

Motivated by the unclear boundaries between GSE, the kinematically hot disk, and the cold disk in the E–J_ϕ plane, we investigated the origin of the Milky Way's metal-poor disk by comparing these three structures, as detailed in Section 3.1. Also, we analyzed the properties of GSE and other merger remnants in Section 3.2. Finally, Section 3.3 details the DTGs that could not be associated.

3.1. GSE versus Galactic Disks

First, we compare the normalized metallicity histograms of the kinematically cold disk, kinematically hot disk, and GSE, along with their distributions in the [Mg/Fe]–[Fe/H] and [Ca/Fe]–[Fe/H] planes as shown in Figure 5. As stated in Section 2.4, we adopted two strategies to select the kinematically cold disk, kinematically hot disk, and GSE stars. The first classifies stars directly based on their kinematic properties, while the second identifies compact stellar groups in phase space and classifies them according to their collective kinematic features. In the comparison, samples from the first and second methods are shown in red and blue, respectively, while the entire sample is displayed in black as a background to highlight the differences among the kinematically cold disk, kinematically hot disk, and GSE. Regardless of the

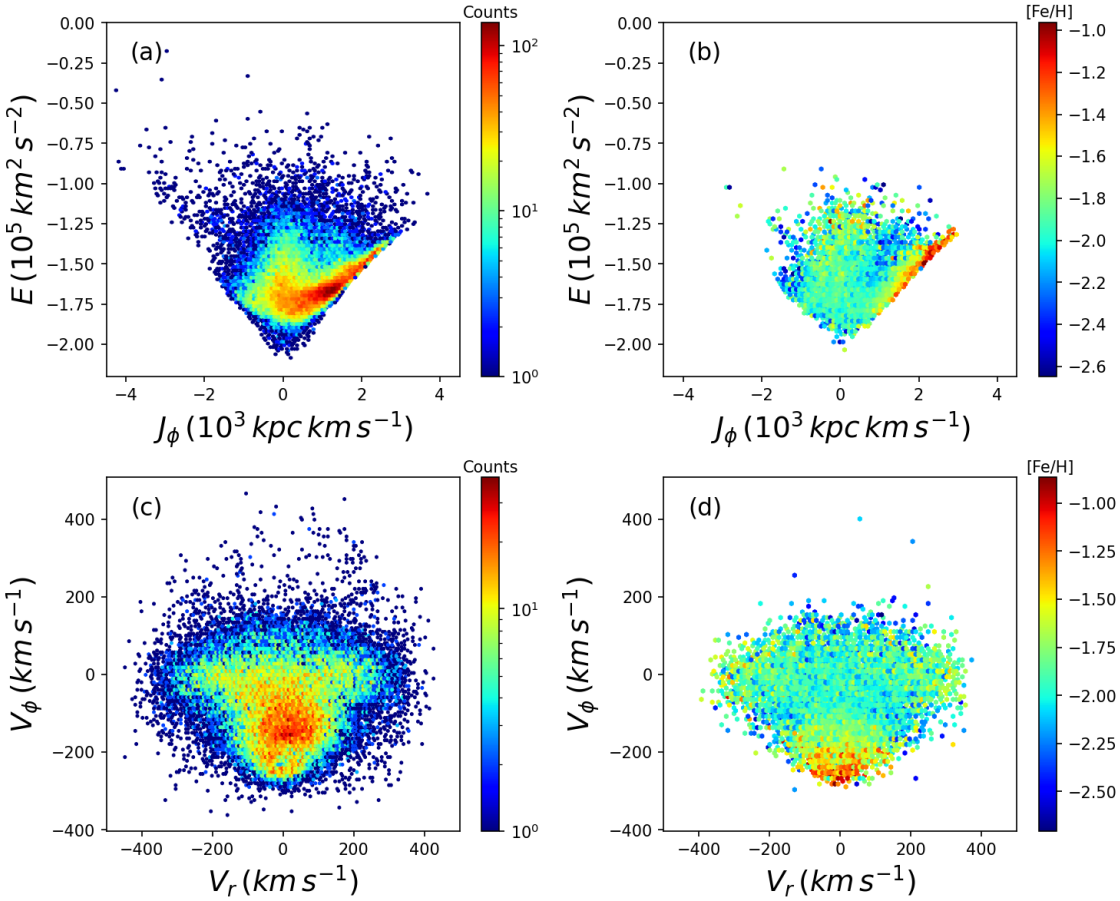


Figure 2. Panels (a) and (b) show the distributions of sample stars in the $E - J_\phi$ plane, and panels (c) and (d) those in the $V_\phi - V_r$ plane. Colors in panels (a) and (c) represent stellar number density, while in panels (b) and (d), they indicate metallicity. The number density plots ((a) and (c)) include all 46,178 sample stars, whereas the metallicity plots include only stars with sufficient [Fe/H] precision. See Section 2.3 for details.

classification methods, the kinematically cold disk shows the highest metallicity, GSE the lowest, and the kinematically hot disk lies in between. Moving from the kinematically cold disk to the kinematically hot disk and then to GSE, rotational velocity ($|V_\phi|$) decreases from over 200 km s^{-1} to nearly 0, a trend mirrored by a progressive decline in metallicity. The kinematically hot disk and GSE show unimodal metallicity histograms, peaking at approximately -1.5 and -2.0, respectively, while the kinematically cold disk exhibits three peaks at -0.8, -1.2, and -2.0, respectively.

The [α/Fe]–[Fe/H] relationship is a key indicator of the star formation and chemical enrichment history. We examine the variation of [Mg/Fe] and [Ca/Fe], two typical α-elements, with metallicity. For both the entire sample and individual structures, [α/Fe] generally decreases as [Fe/H] increases. Panels (b2), (b3), (c2), and (c3) in Figure 5 show that the kinematically hot disk and GSE share a single trend consistent with the total sample, while panels (b1) and (c1) reveal two distinct declining trends for the kinematically cold disk. One group of the kinematically cold disk stars aligns with the overall trend (dashed line), while another exhibits a steeper decline (dotted line), reflecting different formation histories. These results remain consistent even when using stars belonging to DTGs, confirming the robustness of the analysis.

Nissen & Schuster (2010) analyzed high-resolution spectra of 94 stars, including 16 thick-disk stars, and identified two distinct sequences of halo stars in the [α/Fe]–[Fe/H] plane: a high-α sequence and a low-α sequence. They found that the high-α halo stars occupy the same region as the thick-disk stars in the [α/Fe]–[Fe/H] plane. Based on this overlap, they proposed that the high-α sequence likely originates from the heated Galactic disk. In contrast, the low-α sequence was interpreted as having been accreted from external dwarf galaxies. It is important to note that all stars in their sample have [Fe/H] > -1.6, and that the separation between the high-α and low-α sequences becomes increasingly pronounced at higher metallicities. Horta et al. (2023), using APOGEE data, analyzed the [α/Fe] distribution of GSE and identified its α-knee at [Fe/H] ≈ -1.2. They reported that GSE stars have lower [α/Fe] than the disk stars for [Fe/H] > -1.2,

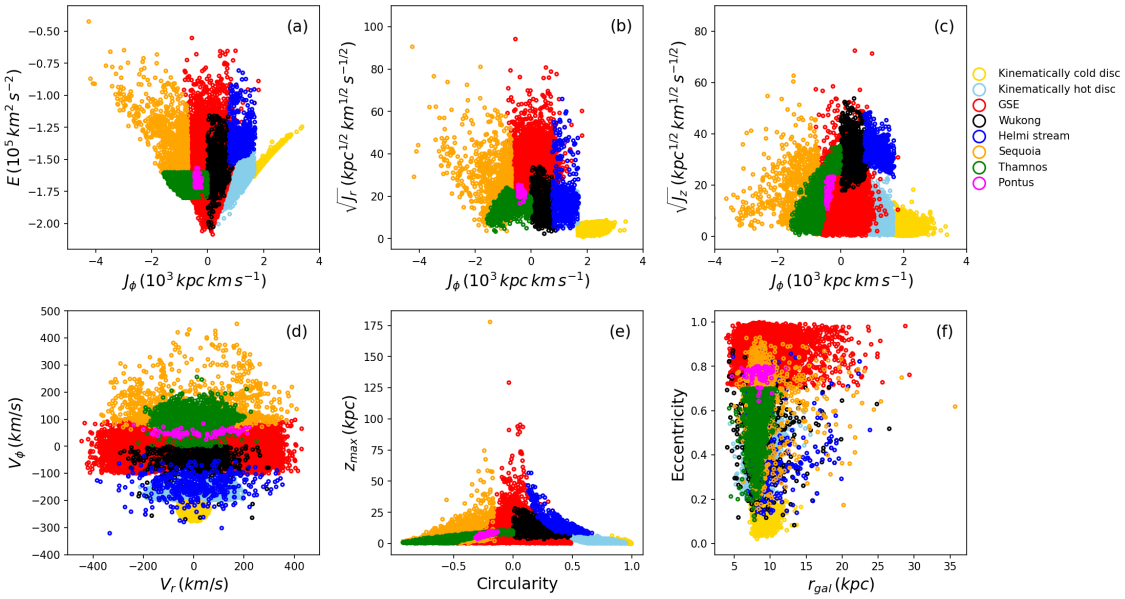


Figure 3. Kinematic space distributions of sample stars, with colors indicating different structures or remnants. See Section 2.4 for classification details.

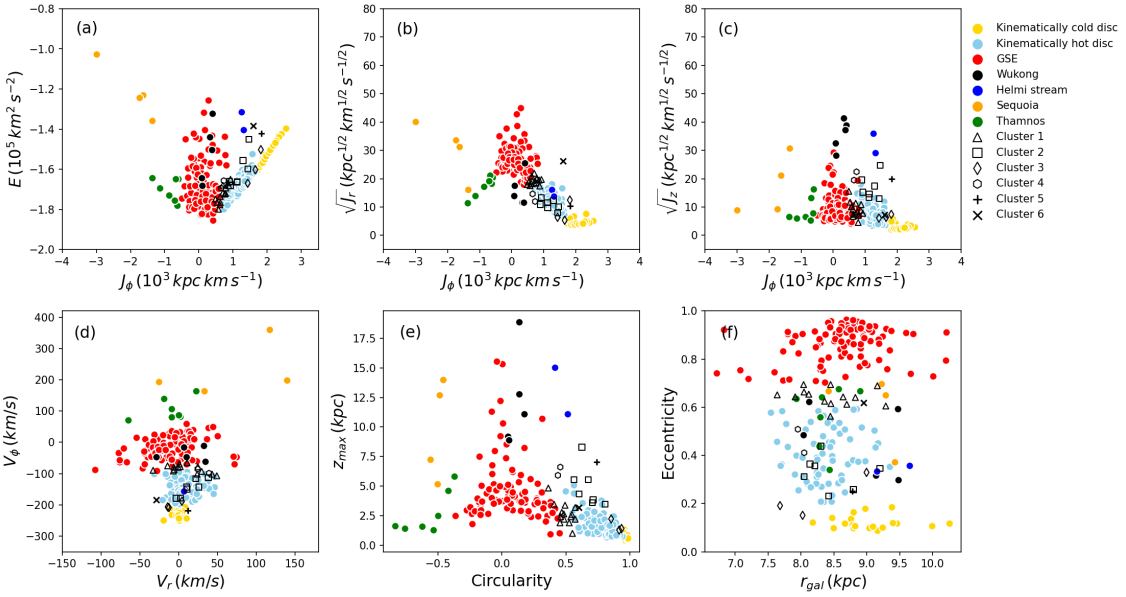


Figure 4. Similar to Figure 3, but for the DTGs. Circles represent successfully associated DTGs, while other shapes indicate those that do not align with known structures or remnants. The unclassified DTGs are divided into six distinct groups, distinguished by shapes such as squares and triangles.

corresponding to the "shin" region beyond the knee. However, their analysis does not clearly define the metal-poor disk, leaving it uncertain whether GSE and the disk are distinguishable in $[\alpha/\mathrm{Fe}]$ at lower metallicities. Naidu et al. (2020) introduced the concept of the Metal-Weak Thick Disk (MWTD), which they defined as stars with $[\mathrm{Fe/H}] < -0.8$ and $[\alpha/\mathrm{Fe}] > 0.25$. In their study, many GSE stars also show $[\alpha/\mathrm{Fe}]$ values exceeding 0.25, as seen in their Figure 16. Zhang et al. (2024) compared GSE and MWTD in the very metal-poor regime ($[\mathrm{Fe/H}] < -2$) and found no significant difference in $[\alpha/\mathrm{Fe}]$ between the two populations. Although none of these studies explicitly define a metallicity threshold where GSE and the disk diverge in $[\alpha/\mathrm{Fe}]$, they collectively suggest that a possible difference is apparent at

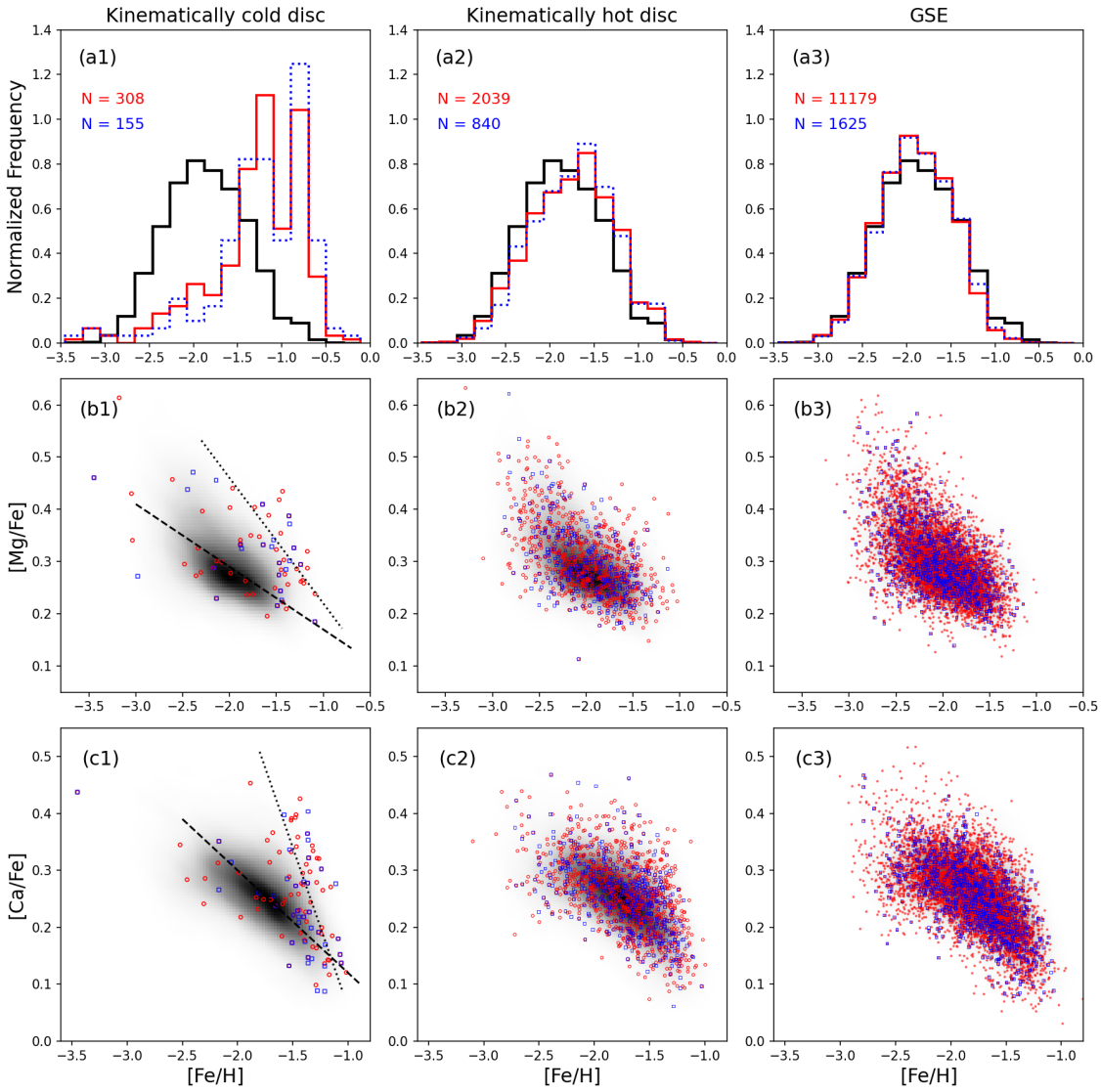


Figure 5. The first row shows the normalized metallicity histograms of the kinematically cold disk, hot disk, and GSE, while the second and third rows display [Mg/Fe]–[Fe/H] and [Ca/Fe]–[Fe/H] distributions, respectively. Black represents the entire sample, red shows stars classified directly by their kinematic properties, and blue signifies stars identified through clustering in phase space and subsequently classified based on their collective kinematic features. Panels (b1) and (c1) include dashed and dotted lines to emphasize two distinct [α/Fe]–[Fe/H] trends for the kinematically cold disk. In the histograms, red numbers indicate the count of stars selected directly, while blue numbers represent those selected indirectly.

[Fe/H] > -1.6. In our study, the initial scatter plot shown in Figure 5 displays overlapping trends between GSE and the kinematically hot disk in the [α/Fe]–[Fe/H] diagram. However, Figure 6 provides a more detailed view through a two-dimensional density map. In this figure, the color shading reflects stellar number density, with darker regions indicating higher concentrations of stars. For the GSE population, we bin stars in 0.1 intervals in [Fe/H] from -2.8 to -1.2, and plot the mean [Mg/Fe] and its 1σ dispersion as black points with vertical error bars. The same procedure is applied to the kinematically hot disk, with stars binned from [Fe/H] $= -2.8$ to -1.1. This consistent binning enables a direct comparison between the two populations across their overlapping metallicity ranges. The plot reveals that GSE exhibits a clear decrease in [Mg/Fe] at [Fe/H] > -1.6, whereas the kinematically hot disk displays a much flatter [Mg/Fe]–[Fe/H] relation over the same range. This suggests a chemical distinction that is not clearly visible in the scatter plot. It is also important to note that stars with [Fe/H] > -1.6 make up only a small fraction of our sample. This underrepresentation of higher-metallicity stars may contribute to the difficulty in identifying the

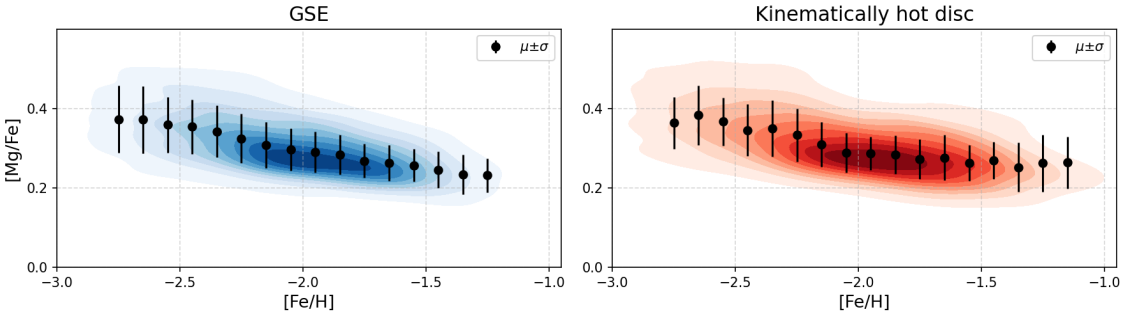


Figure 6. Two-dimensional density maps of the [Mg/Fe]–[Fe/H] distribution for GSE and the kinematically hot disk. The color shading represents stellar number density, with darker regions indicating higher densities. Stars are binned in 0.1 intervals in [Fe/H] from −2.8 to −1.2 for GSE, and from −2.8 to −1.1 for the kinematically hot disk. In each bin, the mean [Mg/Fe] and 1σ dispersion are shown as black points with vertical error bars. The stellar samples used in both panels are identical to those shown in Figure 5.

[α/Fe] difference between GSE and the disk. Finally, we emphasize that the kinematically hot disk in our study is not chemically homogeneous. As discussed in Section 4, it contains multiple components.

We then incorporated stellar ages into our analysis, but the limited number of stars with age data prevents the use of strict precision constraints on iron and magnesium abundances, as this would reduce the sample size. Therefore, we applied relatively loose constraints (see Section 2.3 for details). Figure 7 shows the distributions of the kinematically cold disk, hot disk, and GSE in the age-[Fe/H] plane. Most kinematically hot disk and GSE stars are older than 8 Gyr, while kinematically cold stars are about 50% older and 50% younger than 8 Gyr. For stars older than 8 Gyr, linear fits to the age–metallicity relationship reveal similar slopes for the kinematically cold and hot disks, both smaller than the slope for GSE. Marginal distributions further show that the metallicity peak of the kinematically hot disk corresponds to a higher metallicity than that of GSE, while the age peak of the kinematically hot disk corresponds to an older age compared to GSE.

The stars were divided into four age groups: <8, 8–10, 10–12, and > 12 Gyr. The normalized metallicity histograms of the kinematically cold disk, kinematically hot disk, and GSE were re-examined, as shown in Figure 8, revealing the following patterns:

1. For all age groups, metallicity generally follows the order: kinematically cold disk > kinematically hot disk > GSE. This pattern is most pronounced for stars older than 12 Gyr, where the kinematically hot disk's metallicity histogram appears to combine characteristics of both the kinematically cold disk and GSE in specific proportions. For stars aged 10–12 Gyr, the differences between the three structures become less distinct, with all histograms peaking around −1.2.

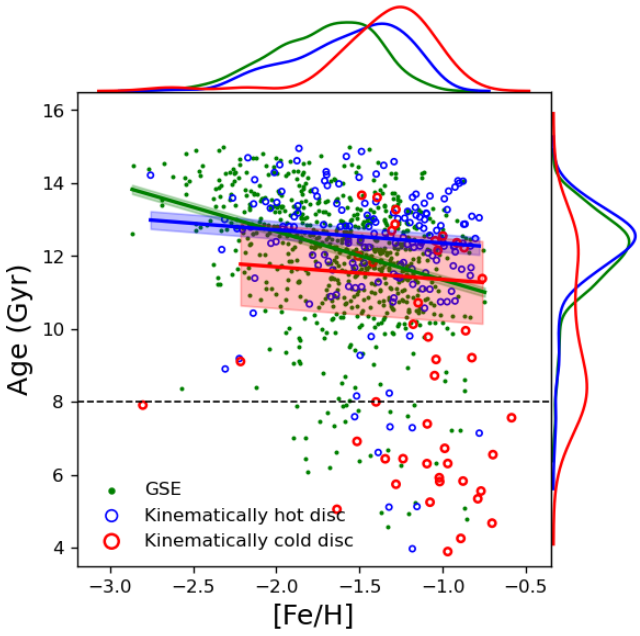


Figure 7. Stellar age–metallicity relation for the kinematically cold disk, hot disk, and GSE, represented by red, blue, and green, respectively. Solid lines indicate linear fits of this relationship, with shaded regions showing 1σ dispersion. All fits utilize stars older than 8 Gyr. A dashed line marks an age of 8 Gyr, and marginal distributions for age and [Fe/H] are also included.

2. In the kinematically cold disk, stars older than 12 Gyr generally have [Fe/H] > −1.5, while stars with [Fe/H] < −1.5 first appear in the 10–12 Gyr group, and some in the 8–10 Gyr group even have [Fe/H] < −2.0. For stars younger than 8 Gyr, nearly all have [Fe/H] > −1.7 (except for one star with [Fe/H] = −2.8), and the histogram peaks at a [Fe/H] value around −1.0. Although the sample size is limited, this observation suggests a pattern of metallicity change in the kinematically cold disk.

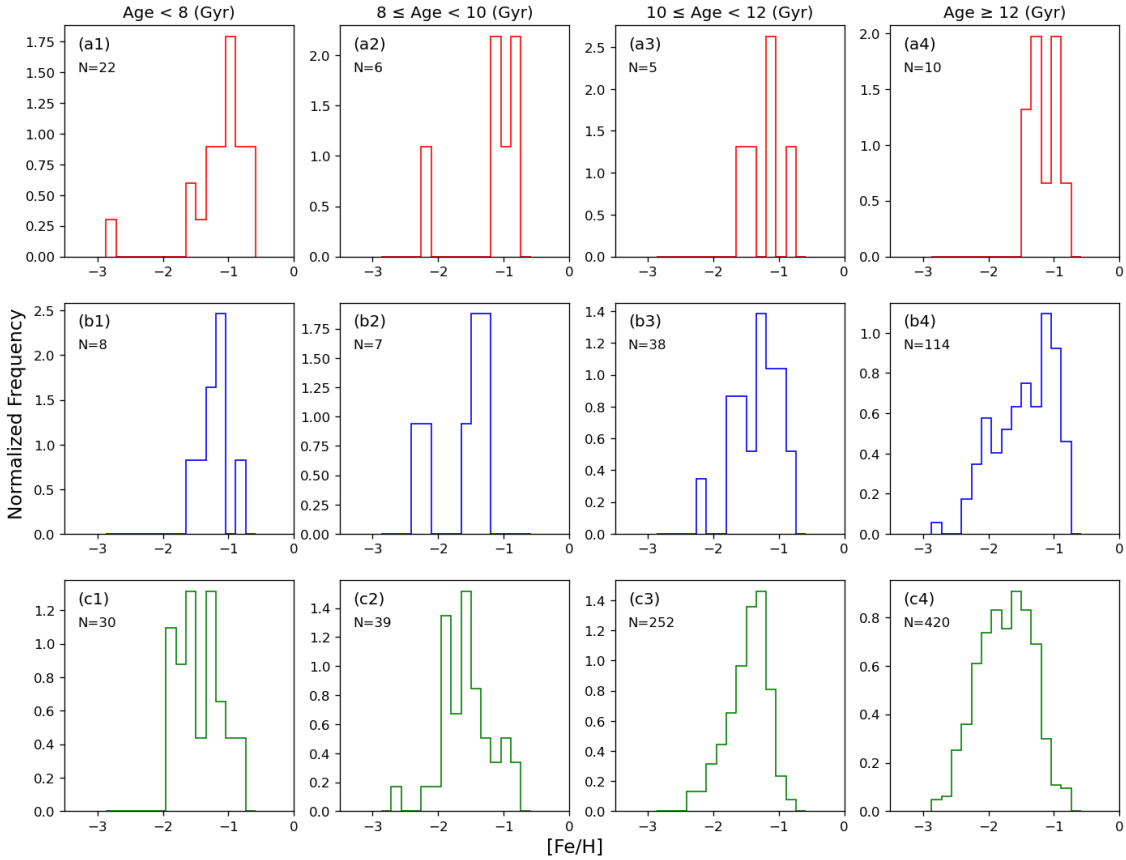


Figure 8. Normalized metallicity histograms for the kinematically cold disk, hot disk, and GSE across four age groups: <8, 8–10, 10–12, and >12 Gyr. The three rows from top to bottom represent the kinematically cold disk, hot disk, and GSE, respectively.

3. The kinematically hot disk shows a clear trend in metallicity evolution over time. For stars older than 12 Gyr, the proportion of stars with $[Fe/H] < -2.0$ is relatively high, approximately 13.2%. In the 10–12 Gyr group, this proportion drops significantly to only 5.3%. From the older than 12 Gyr group to stars younger than 8 Gyr, the median metallicities are −1.4, −1.27, −1.5, and −1.25, while the mean metallicities are −1.45, −1.32, −1.64, and −1.23, respectively.
4. For GSE stars older than 12 Gyr, the metallicity histogram peaks around −1.5, with a significant proportion having $[Fe/H] < -2.0$. Specifically, out of the 420 stars in the older than 12 Gyr group, 112 stars have $[Fe/H] < -2$, accounting for approximately 26.7%. In the 10–12 Gyr group, the peak [Fe/H] value shifts to −1.2, with only 17 stars out of 252 having $[Fe/H] < -2.0$, which is about 6.7%. In the 8–10 Gyr group, the peak [Fe/H] value decreases again to −1.5. The median metallicities for these groups are −1.75, −1.42, and −1.56, while the mean metallicities are −1.74, −1.47, and −1.56, respectively. This trend suggests that the metallicity of GSE initially increased after its formation and then decreased, with this process concluding around 8 Gyr ago. The progenitor of GSE is believed to have merged with the Milky Way around 10 Gyr ago, making the observation of GSE stars younger than 8 Gyr unusual. These younger stars, which exhibit significant radial velocities and high eccentricities, may not originate from the GSE merger itself. They were likely included in our sample because they were selected using kinematic criteria associated with GSE membership.
5. Across GSE and the kinematically hot disk samples, stars aged 8–10 Gyr are less numerous compared to those aged 10–12 Gyr and over 12 Gyr.

We also re-examined the [Mg/Fe]–[Fe/H] distributions under the same age groupings with relaxed error constraints on Mg and Fe abundances. Figure 9 presents these distributions, with the kinematically cold disk, hot disk, and GSE displayed from top to bottom. Each panel in Figure 9 includes one auxiliary line, except panel (c1), which contains two. The auxiliary lines are of two types: dashed and dotted. Dashed lines correspond to the linear fit of the [Mg/Fe]–[Fe/H]

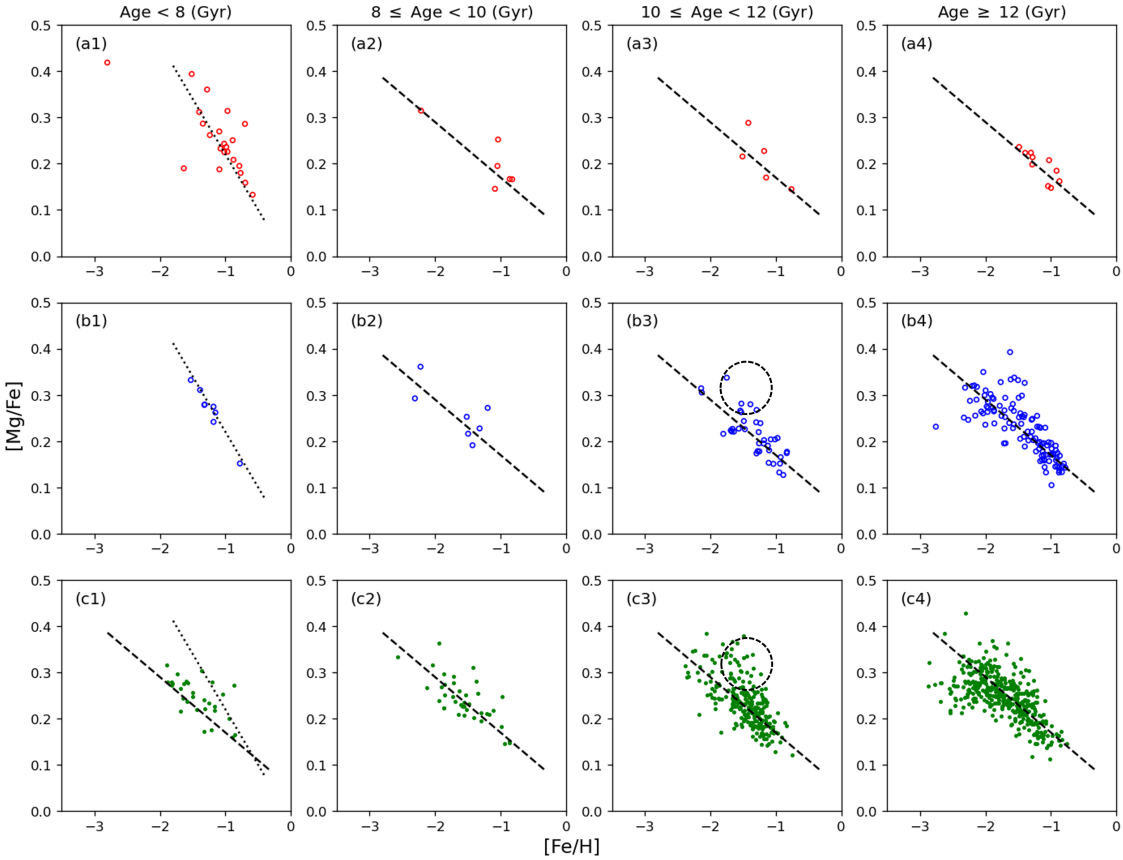


Figure 9. [Mg/Fe]–[Fe/H] distributions for the kinematically cold disk, hot disk, and GSE (shown in top, middle, and bottom rows) across four age groups: <8, 8–10, 10–12, and >12 Gyr. Each panel contains one or two auxiliary lines of two types: dashed lines derived from fitting kinematically cold disk stars older than 12 Gyr, and dotted lines from kinematically cold disk stars younger than 8 Gyr. Panels (b3) and (c3) include dashed circles highlighting Mg-enhanced stars, which are defined as stars with [Mg/Fe] values significantly higher than predicted by the dashed line relationship at a given [Fe/H].

relationship for kinematically cold stars older than 12 Gyr, while dotted lines represent the fit for kinematically cold stars younger than 8 Gyr. The equations for the lines are:

$$[\mathrm{Mg/Fe}] = -0.12 \times [\mathrm{Fe/H}] + 0.05, \quad (7)$$

$$[\mathrm{Mg/Fe}] = -0.24 \times [\mathrm{Fe/H}] - 0.02. \quad (8)$$

From Figure 9, we can observe the following patterns:

1. For the kinematically cold and hot disks, stars older than 8 Gyr show significantly different [Mg/Fe]–[Fe/H] relationships compared to younger stars. Stars younger than 8 Gyr exhibit a steeper slope, with [Mg/Fe] decreasing more rapidly as [Fe/H] increases.
2. The kinematically hot disk and GSE exhibit similar distributions across three age groups: 8–10, 10–12, and >12 Gyr. In the oldest group (>12 Gyr), both exhibit an α-knee at [Fe/H] ≈ -1.3. Dashed lines represent the expected [Mg/Fe]–[Fe/H] relationship, with Mg-enhanced stars defined as those showing significantly higher [Mg/Fe] values than predicted at a given [Fe/H]. In the 10–12 Gyr group, many stars in both the kinematically hot disk and GSE show such Mg enhancements, as shown in panels (b3) and (c3) of Figure 9. This enhancement is less pronounced in the other age groups.
3. For stars younger than 8 Gyr, the number of stars following the dotted [Mg/Fe]–[Fe/H] relationship decreases progressively from kinematically cold disk to kinematically hot disk to GSE. While the kinematically cold disk contains the most stars conforming to this relationship, such stars are nearly absent in GSE.

In Figure 7, we find that both the kinematically hot and cold disk samples contain a subset of metal-poor stars with ages younger than 8 Gyr. This is unexpected, given that younger stars are generally associated with higher [Fe/H] in the context of Galactic chemical evolution. Building on this, Figure 5 reveals that within the kinematically cold disk sample, there are two distinct sequences in the [α/Fe]–[Fe/H] plane: a high-

Table 2
Line-of-sight Velocities and Corresponding Uncertainties for 14 Metal-poor, High-α Sequence Stars Younger than 8 Gyr, Each Observed Multiple Times by the LAMOST Survey

Uid	Velocity List	Velocity Error List	χ^2	p-value
G14028202809079	8.31; 16.96; 8.94; 13.68	3.48; 5.28; 6.47; 5.49	2.316	0.509
G15971821115630	45.87; 37.86	7.63; 2.39	3.084	0.079
G16220243843745	−7.20; −7.12	4.04; 3.84	0.000	0.989
G16225978524450	−1.78; −3.19	3.83; 4.69	0.056	0.812
G16237508594175	2.23; −4.19	5.37; 6.58	0.595	0.440
G16263861478746	−10.79; −5.38	3.30; 4.74	0.998	0.317
G16310961074871	−26.86; −29.30; −32.14; −28.23	5.45; 5.99; 3.12; 4.38	1.146	0.766
G16424162144631	−24.13; −26.35; −16.38; −21.11; −22.34; −14.24	4.36; 4.76; 4.79 2.89; 4.52; 15.19	3.135	0.679
G16496513210296	40.96; 33.03	2.83; 5.04	2.582	0.108
G16498871828495	−24.25; −18.38	7.05; 6.30	0.390	0.532
G17297423527435	11.77; 13.41	3.25; 6.40	0.080	0.777
G17567156788740	−6.75; −20.85	4.50; 8.77	3.101	0.078
G15215012847084	−129.00; −124.24	4.74; 7.98	0.341	0.559
G17266455315544	−28.94; −34.52	3.97; 3.72	1.056	0.304

Note. Uid is the unique source ID in LAMOST. The "velocity list" and "velocity error list" entries for each star are paired in order.

[α/Fe] sequence and a low-[α/Fe] sequence. When stellar ages are incorporated, as shown in panels (a1) and (b1) of Figure 9, we find that disk stars on the high-[α/Fe] sequence are precisely those younger than 8 Gyr. This finding is particularly surprising, as it contradicts the widely accepted trend that [α/Fe] decreases with time in the Galactic disk. One possible explanation is that binary stellar evolution might lead to underestimated ages for some genuinely old stars. To investigate this possibility, we examined whether stars shown in (a1) and (b1) of Figure 9 might be unresolved binaries. Penoyre et al. (2022) noted that Gaia astrometric measurements with RUWE larger than 1.4 could indicate binarity. However, in our study, stars with RUWE > 1.4 were already excluded in Section 2.2. Additionally, Mowlavi et al. (2023) released a catalog of eclipsing-binary candidates based on Gaia DR3, and after cross-matching, we found no overlap between stars in our high-[α/Fe] sequence and their eclipsing-binary catalog. Furthermore, 14 stars from (a1) and (b1) in Figure 9 have been repeatedly observed by the LAMOST low-resolution survey. The multiple line-of-sight velocity measurements and their corresponding uncertainties for these stars are listed in Table 2. We used these line-of-sight velocity variations to statistically evaluate the possibility of binarity. To quantify whether a star exhibits significant radial velocity variations, indicative of a binary system, we employed the χ^2 test. Specifically, we computed

$$\chi^2 = \sum_{i=1}^{n} \frac{(v_i - \bar{v})^2}{\sigma_i^2}, \tag{9}$$

where v_i is the ith line-of-sight velocity measurement, $\bar{v}$ is the mean velocity, σ_i is the uncertainty of the ith measurement, and n is the number of observations. Under the null hypothesis-that the star is single and maintains a constant line-of-sight velocity, the calculated χ^2 statistic follows a chi-square distribution with $n - 1$ degrees of freedom. For each star, we calculated the associated p-value, which represents the probability of observing a χ^2 value as large as (or larger than) the measured one under the null hypothesis. A low p-value (typically $p < 0.05$) indicates statistically significant velocity variations and suggests the presence of binarity. The computed χ^2 values and their corresponding p-values are listed in Table 2. Based on these p-values, none of the stars exhibit velocity variations significant enough to be identified as binaries. In summary, current evidence does not support the hypothesis that these high-α, young stars are unresolved binaries.

As previously discussed, our kinematically cold disk sample reveals two distinct populations: metal-poor, high-α and metal-poor, low-α stars (Figure 5). When incorporating stellar age information, we find that the metal-poor, high-α population predominantly comprises stars younger than 8 Gyr (Figure 9). Notably, age information also reveals the presence of metal-poor, high-α stars within the kinematically hot disk, although they constitute a smaller fraction compared to the kinematically cold disk. From the 2517 stars with reliable age estimates, we applied two selection criteria to isolate disk stars. First, we excluded stars with orbital eccentricities greater than 0.6 to eliminate those with halo-like orbits. Second, we removed stars exhibiting retrograde motion by filtering out those with $V_\phi > 0$. After these selections, 1396 stars remained. Among them, 329 stars with ages younger than 8 Gyr were classified as high-α population, with the remainder designated were as low-α population. We examined the abundance trends of the two groups. Figure 10 shows the [Mg/Fe]–[Fe/H] and [Ca/Fe]–[Fe/H] distributions, including only stars with machine learning-estimated [Mg/H] and [Fe/H] having uncertainties less than 0.2. As shown in Figure 10, the younger group clearly exhibits higher [Mg/Fe] and [Ca/Fe] ratios compared to the older group. A detailed comparison between the high-α and low-α populations was subsequently performed, focusing on their spatial distribution and kinematic properties. As shown in Figure 11, the normalized histograms highlight clear and significant differences between the two groups:

(i) High-α stars are preferentially found at larger Galactocentric radii, with a distribution peak near 9.5 kpc and a substantial fraction extending beyond 10 kpc. In contrast,

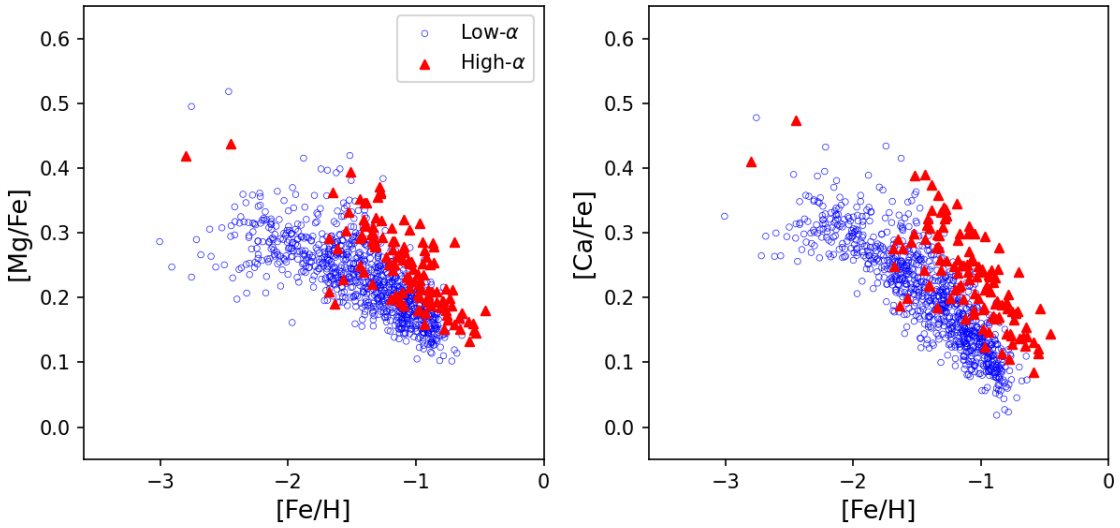


Figure 10. Distributions of [Mg/Fe] and [Ca/Fe] as a function of [Fe/H] for two stellar populations: high-α (red triangles) and low-α (blue circles). These groups are defined based on stellar age, with high-α stars being younger than 8 Gyr and low-α stars older. Only stars with [Mg/H] and [Fe/H] uncertainties less than 0.2 (from machine learning estimates) are included.

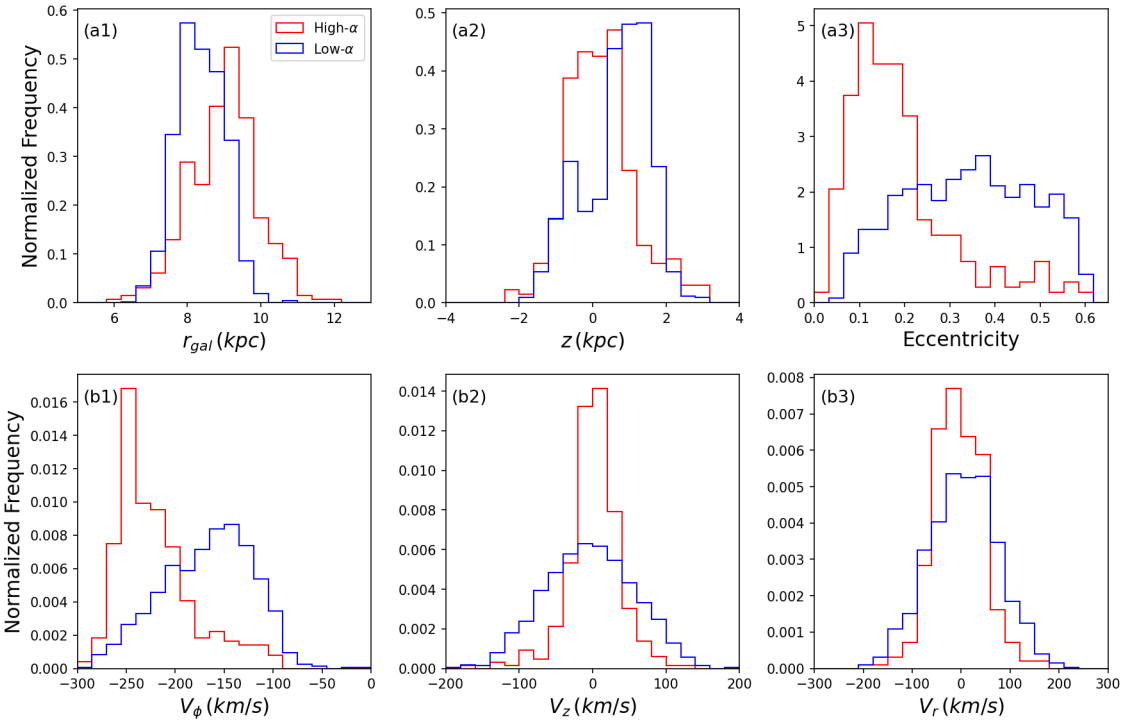


Figure 11. Normalized histograms showing spatial and kinematic properties of high-α (red) and low-α (blue) disk stars. Top panels: galactocentric radius $r_{\rm gal}$ (a1), vertical distance from the midplane z (a2), and orbital eccentricity (a3). Bottom panels: azimuthal velocity V_ϕ (b1), vertical velocity V_z (b2), and radial velocity V_r (b3). The two populations are selected from stars with reliable age estimates, excluding those with $e > 0.6$ or retrograde motion ($V_\phi > 0$). High-α stars are defined as those younger than 8 Gyr; the rest are classified as low-α.

low-α stars are more centrally concentrated, peaking around 8.2 kpc, with almost no stars located beyond 10 kpc (Figure 11(a1)).

(ii) High-α stars are more tightly confined to the Galactic midplane, with the majority located within $\pm$1.5 kpc and a distribution peak near $z \approx 0.5$ kpc. On the other hand, low-α cases exhibit a broader vertical distribution, peaking slightly higher around $z \approx 1.0$ kpc (Figure 11(a2)).

(iii) The eccentricity distribution of high-α stars is relatively narrow, peaking at $e \approx 0.12$, and more than 90% of the stars have $e < 0.4$, indicating nearly-circular orbits. In contrast, low-α stars show a flatter and broader distribution peaking around $e \approx 0.4$, with a notable fraction extending to $e > 0.5$ (Figure 11(a3)).

(iv) High-α stars demonstrate higher absolute azimuthal velocities, with a sharp distribution peak at $V_\phi \approx -250\,\mathrm{km\,s^{-1}}$, suggesting faster rotation. In contrast, low-α stars peak around $-150\,\mathrm{km\,s^{-1}}$, indicating slower rotational support and a kinematically hotter component (Figure 11(b1)).

(v) In both radial and vertical velocities, high-α stars exhibit narrower and more centrally concentrated distributions. Their V_r and V_z values are tightly clustered around $0\,\mathrm{km\,s^{-1}}$, with standard deviations significantly smaller than those of low-α stars. Low-α stars show broader velocity spreads in both components, reaching up to $\pm 200\ \mathrm{km\,s^{-1}}$, further reinforcing their dynamically hotter nature (Figure 11(b2) and (b3)).

Han et al. (2020) investigated the Galactic disk using a sample of G- and K-type dwarfs, classifying stars into thin and thick disk populations based on their [α/Fe]. Their sample primarily consists of stars with [Fe/H] > -1.0. In their classification, stars with lower [α/Fe] are assigned to the thin disk, while those with higher [α/Fe] are identified as thick disk members. They found that thin disk stars (metal-rich and low-α) generally have orbital eccentricities below 0.4, with a peak near 0.2, whereas thick disk stars (metal-rich and high-α) exhibit a broader eccentricity distribution, extending up to 0.8 (see their Figure 8). In terms of eccentricity, our sample of young, metal-poor, high-α disk stars shows kinematic behavior similar to the thin disk population (metal-rich and low-α) described by Han et al. (2020), while our old, metal-poor, low-α stars more closely resemble their thick disk (metal-rich and high-α). A similar chemo-dynamic comparison was presented by Lee et al. (2011), who analyzed G-type dwarfs from the SEGUE survey to characterize the Galactic thin and thick disks. In their study, thin disk stars (metal-rich and low-α) predominantly exhibit azimuthal velocities $|V_\phi| > 200\,\mathrm{km\,s^{-1}}$, while the thick disk stars (metal-rich and high-α) tend to show lower $|V_\phi|$ values. Spatially, their thin disk population (metal-rich and low-α) is more radially extended and located farther from the Galactic Center. Vertically, their thin disk stars (metal-rich and low-α) are primarily confined within 1.5 kpc of the Galactic midplane, whereas thick disk (metal-rich and high-α) stars are concentrated between 0.5 and 2 kpc, peaking around 1 kpc. Based on these characteristics, our old, metal-poor, low-α population shows similarities with the thick disk (metal-rich and high-α) defined by Lee et al. (2011), while our young, metal-poor and high-α stars resemble their thin disk (metal-rich and low-α). However, we note an important distinction: a subset of our young, metal-poor, high-α stars are located at significant vertical distances (Figure 11(a2)), with several exceeding 1.5 kpc and some reaching up to 3 kpc in height. In summary, combining insights from our sample and prior studies (e.g., Lee et al. 2011; Han et al. 2020), we find that the chemo-dynamic properties of disk stars vary markedly across the metallicity regime. At higher metallicities, high-α stars are dynamically hotter and they are found closer to the Galactic Center, possess higher orbital eccentricities, lower rotational velocities, and extend to greater vertical heights. In contrast, low-α stars at high metallicity tend to be dynamically colder, with lower eccentricities and higher azimuthal velocities. Interestingly, this trend reverses at lower metallicities: high-α stars in this regime exhibit kinematically colder properties, while low-α stars appear dynamically hotter.

A number of recent studies have identified a peculiar population of stars exhibiting relatively high [α/Fe] despite their young ages. Using asteroseismic data to determine stellar ages and spectroscopic measurements to derive chemical abundances, Chiappini et al. (2015), Martig et al. (2015), Wu et al. (2018) independently reported the existence of such stars. Chiappini et al. (2015) found that these high-α stars are younger than 8 Gyr, while Martig et al. (2015), even under conservative assumptions intended to overestimate stellar ages, identified high-α stars with ages below 6 Gyr. Both studies showed that these young high-α stars occupy the same region in the [Mg/Fe]–[Fe/H] diagram as the old, metal-rich, high-α disk stars. In terms of metallicity, the young high-α stars studied by Martig et al. (2015) have [Fe/H] > -0.5, while those in Chiappini et al. (2015) have [Fe/H] > -0.7. Chiappini et al. (2015) further noted that these stars are primarily located within 6 kpc of the Galactic Center and proposed that their formation may be influenced by the Galactic bar. In contrast, Martig et al. (2015) found that the guiding radii and radial velocity distributions of these stars are similar to those of the canonical thick disk, implying that their spatial and kinematic properties are consistent with those old, high-α stars. In our study, we also identify a population of high-α stars with ages younger than 8 Gyr. However, in contrast to the young, high-α stars reported by Chiappini et al. (2015), Martig et al. (2015), our high-α stars are metal-poor. Furthermore, our analysis reveals that our young, metal-poor, high-α stars exhibit kinematic properties more closely aligned with the thin disk: they are found at larger Galactocentric radii and exhibit dynamically colder motions. These results suggest that our young, metal-poor, high-α stars are distinct from

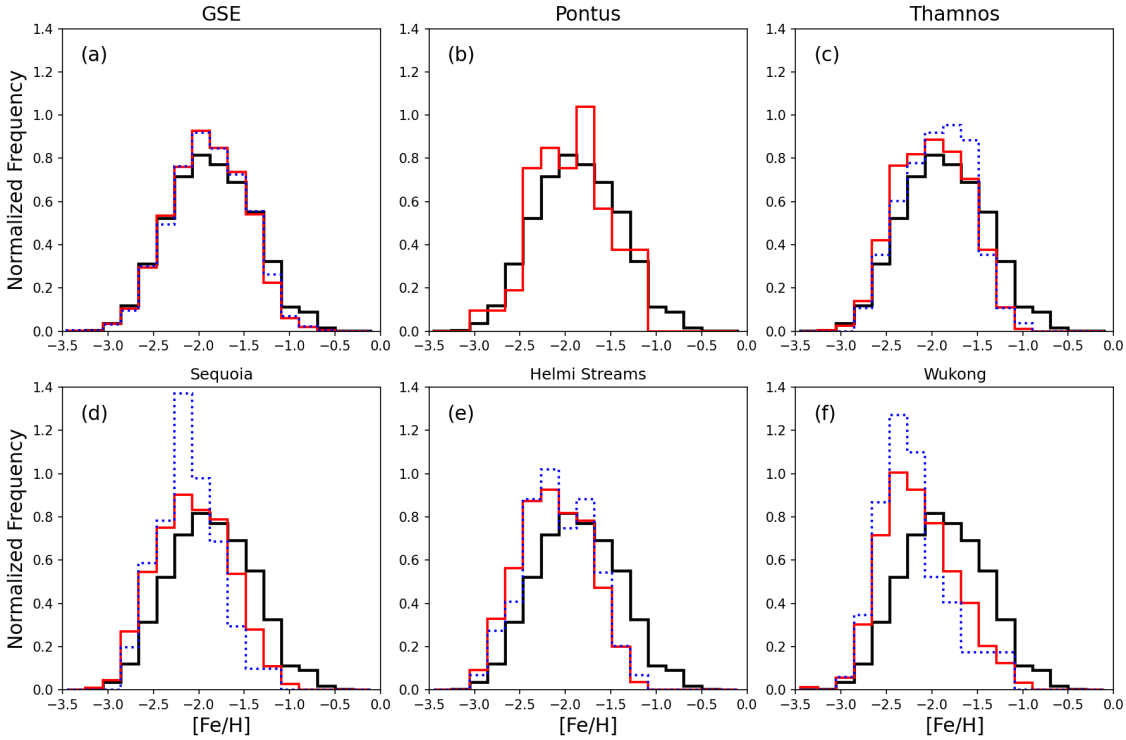


Figure 12. Normalized metallicity histograms of GSE, Pontus, Thamnos, Sequoia, Helmi Streams, and Wukong. As in Figure 5, each panel shows: the black solid line for all sample stars, the red solid line for stars directly classified by their kinematic properties, and the blue dotted line for stars identified through clustering in phase space and subsequently attributed to that remnant.

previously identified metal-rich, young, high-α populations, both chemically and kinematically, and may represent a different evolutionary pathway or formation environment.

3.2. GSE versus Other Merger Remnants

Gaia's observations have uncovered numerous dwarf galaxy remnants in the Galactic halo, with stars in each remnant typically exhibiting similar kinematic and chemical properties. However, the origins of these remnants remain unclear. Some researchers suggest that certain remnants share a common progenitor, while others argue they originate from different merger events. This study aims to provide new evidence on this topic, focusing on GSE, Pontus, Thamnos, Sequoia, Helmi Streams, and Wukong. We excluded other remnants like Sagittarius, Cetus, and Heracles. Despite its size, Sagittarius' great distance from the Sun limits the number of its stars in our sample. Cetus is mainly located in the southern celestial hemisphere, while Heracles lies near the Galactic Center, where LAMOST lacks observational coverage.

We compared the normalized metallicity histograms of GSE, Pontus, Thamnos, Sequoia, Helmi Streams, and Wukong, as shown in Figure 12. The key findings are as follows:

1. Pontus contains proportionally fewer stars with [Fe/H] > -1.7, and more with $-2.5 <$ [Fe/H] < -2.0 compared to GSE.
2. Thamnos peaks at [Fe/H] ≈ -2.0, and contains proportionally fewer stars with [Fe/H] > -1.5 and a higher fraction of stars in the range $-2.7 <$ [Fe/H] < -2.0 compared to GSE.
3. Sequoia, Helmi Streams, and Wukong have metallicity histograms peaking at lower [Fe/H] values than GSE and contain more stars with [Fe/H] < -2.0, with Wukong exhibiting the lowest metallicities.

3.3. Unclassified DTGs

According to our classification criteria for the Galactic disks and dwarf galaxy remnants, 28 DTGs could not be assigned to either category. These 28 DTGs were further analyzed and divided into six groups based on their kinematic characteristics, labeled Clusters 1–6. Stars within each group exhibit similar kinematic properties, as shown by their distributions in the kinematic space in Figure 4. Specifically, Clusters 1–3 likely correspond to known structures but were excluded from the main classification due to the strict criteria applied, while

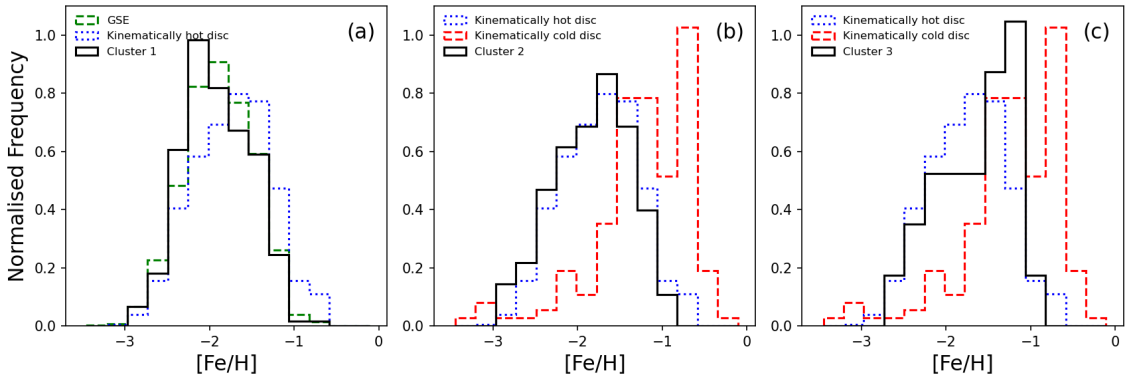


Figure 13. (a) Comparison of normalized metallicity histograms between GSE, the kinematically hot disk, and Cluster 1. (b) Comparison of normalized metallicity histograms between the kinematically hot disk, the kinematically cold disk, and Cluster 2. (c) Comparison of normalized metallicity histograms between the kinematically hot disk, the kinematically cold disk, and Cluster 3. GSE, kinematically hot disk, and kinematically cold disk stars shown here were first identified through phase-space clustering and then classified accordingly.

Clusters 4–6 likely represent unknown substructures. The detailed kinematic characteristics of each group are provided in the following paragraphs. Additionally, a table listing the specific DTGs within each group can be found at the end of the article via a provided link.

Cluster 1, containing 14 DTGs, exhibits consistent kinematic properties, such as V_ϕ around $-100\,\mathrm{km\,s^{-1}}$ and eccentricities near 0.65. These DTGs are located at the periphery of the region occupied by GSE and share a similar metallicity histogram, as shown in Figure 13. However, they were excluded from GSE due to our arbitrary eccentricity criterion ($e > 0.7$). Given their kinematics and metallicity, we conclude that Cluster 1 is part of GSE. Cluster 2, comprising seven DTGs, overlaps with the kinematically hot disk in kinematic space. The DTGs in this cluster have $|V_z|$ values slightly above $25\,\mathrm{km\,s^{-1}}$, exceeding the kinematically hot disk criterion of $|V_z| < 25\,\mathrm{km\,s^{-1}}$. Nevertheless, their metallicity histogram, illustrated in panel (b) of Figure 13, closely resembles that of the kinematically hot disk. Therefore, we conclude that Cluster 2 is part of the kinematically hot disk. Cluster 3 consists of three DTGs with kinematic properties intermediate between those of the kinematically hot and cold disks. Two DTGs were excluded from the kinematically cold disk because, despite them having $J_\phi > 1660\,\mathrm{kpc\,km\,s^{-1}}$ and $|V_z| < 15\,\mathrm{km\,s^{-1}}$, their eccentricities are greater than 0.3. The third DTG, on the other hand, was excluded from the kinematically hot disk because, while its $J_\phi < 1660\,\mathrm{kpc\,km\,s^{-1}}$, its eccentricity is lower than 0.2. These DTGs exhibit kinematic characteristics partially aligning with both disks, suggesting they belong to the transition region between the two. Their metallicity histogram, shown in panel (c) of Figure 13, lies between that of the kinematically cold and hot disks, supporting this conclusion. Therefore, we infer that Cluster 3 represents the transition from the kinematically cold disk to the kinematically hot disk. It should be noted that the stars shown in Figure 13 for the kinematically hot disk, cold disk, and GSE were first identified through phase-space clustering, and subsequently assigned to each component based on their kinematic properties. For plotting purposes, we only include stars with [Fe/H] uncertainties less than 0.2 from both machine learning and template fitting methods, and with a difference between the two methods also smaller than 0.2.

For Clusters 4–6, Figure 14 shows the distribution of their stars in the V_ϕ–V_r and V_ϕ–V_z planes, alongside stars belonging to GSE, the kinematically hot disk, and the kinematically cold disk. As in Figure 13, the disk and GSE samples were first identified through phase-space clustering and then assigned to their respective components based on their kinematic properties. The following observations can be made:

1. Stars in Cluster 4 have V_ϕ near $-100\,\mathrm{km\,s^{-1}}$, between GSE and the kinematically hot disk. Their V_r dispersion is similar to the kinematically hot disk, but their $|V_z|$ values are significantly higher, mostly exceeding $100\,\mathrm{km\,s^{-1}}$.
2. Stars in Cluster 5 have V_ϕ and V_r similar to the kinematically cold disk, with V_ϕ mostly below $-200\,\mathrm{km\,s^{-1}}$ and $|V_r|$ below $100\,\mathrm{km\,s^{-1}}$. However, their $|V_z|$ values are much higher than those of the kinematically cold disk, mostly exceeding $100\,\mathrm{km\,s^{-1}}$.
3. Stars in Cluster 6 exhibit disk-like V_z and V_ϕ. Specifically, $|V_z|$ is mostly below $100\,\mathrm{km\,s^{-1}}$, consistent with the kinematically hot disk, while V_ϕ ranges from -240 to $-130\,\mathrm{km\,s^{-1}}$, showing features of both the kinematically hot and cold disks. However, their $|V_r|$ is unusually large, mostly around $200\,\mathrm{km\,s^{-1}}$.

The kinematics of stars in Cluster 4 are intermediate between the GSE sample and the kinematically hot disk sample, but their metallicity is generally lower than both, as shown in Figure 15(a). It is important to note that the metallicity histogram for Cluster 4 stars is based on a sample of 24 stars. The mean and median metallicities for Cluster 4 are [Fe/H] = −1.93 and −2.06, respectively, while

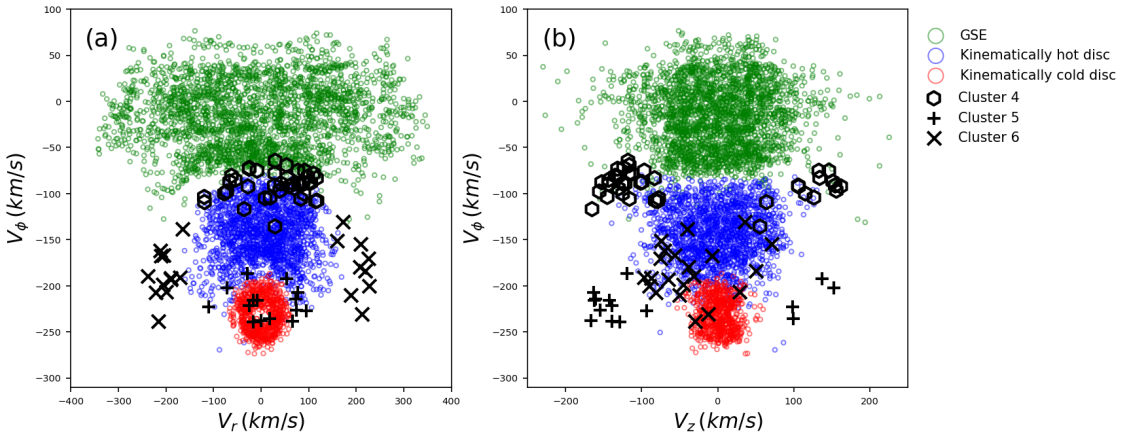


Figure 14. V_ϕ–V_r and V_ϕ–V_z distributions of stars in Clusters 4, 5, and 6, represented by hexagons, plus signs and crosses, respectively. The distributions of GSE, the kinematically hot disk, and the kinematically cold disk are shown by green, blue, and red circles, respectively.

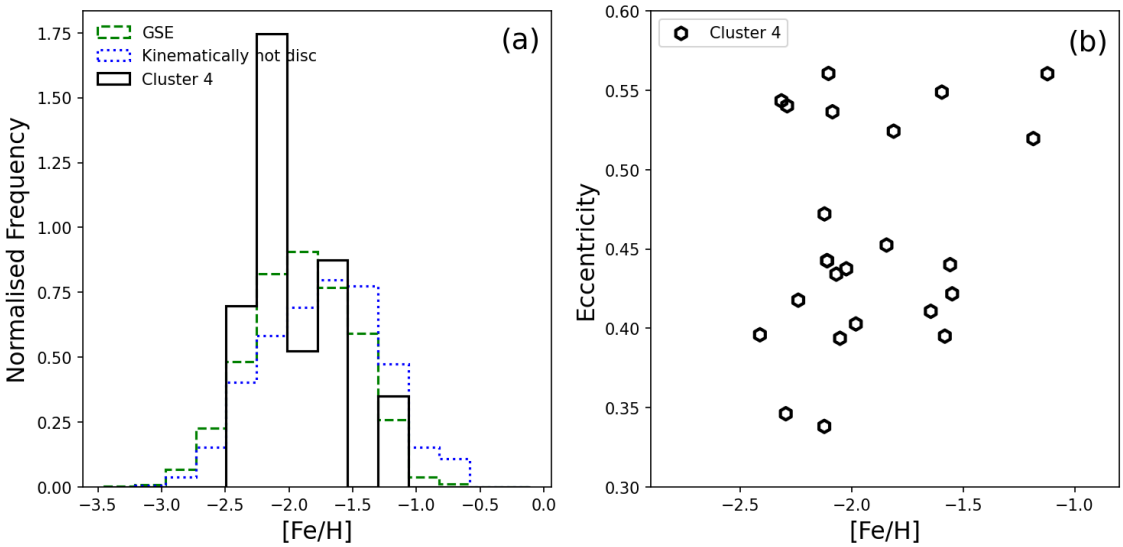


Figure 15. (a) Comparison of normalized metallicity histograms between GSE, the kinematically hot disk, and Cluster 4. (b) e–[Fe/H] distribution of stars in Cluster 4. Stars associated with GSE and the kinematically hot disk were identified through phase-space clustering and subsequently assigned to their respective components based on kinematic criteria. For both panels, only stars with [Fe/H] uncertainties smaller than 0.2 from both the machine learning and template fitting methods, and with a discrepancy between the two methods less than 0.2 are included.

the mean and median metallicities for GSE are [Fe/H] = −1.89 and −1.89, and for the kinematically hot disk, they are [Fe/H] = −1.73 and −1.7, respectively. Additionally, Figure 15(b) reveals a positive correlation between eccentricity and [Fe/H] for Cluster 4 stars. For Cluster 5, [Mg/Fe]–[Fe/H] distributions reveal two distinct sequences resembling those of the kinematically cold disk sample, as shown in Figure 16. Stars in Cluster 6 show large $|V_r|$, similar to GSE, prompting a comparison of their [Mg/Fe]–[Fe/H] and [Sc/Fe]–[Fe/H] distributions with GSE in Figure 17. Cluster 6 shows an α-knee at [Fe/H] ≈ -1.3, consistent with GSE. In the [Sc/Fe]–[Fe/H] plane, both Cluster 6 and GSE exhibit linear trends, but Cluster 6 has smaller [Sc/Fe] dispersion at a given [Fe/H]. In Figures 16 and 17, the disk and GSE samples are drawn from their respective DTGs. A relaxed error screening is applied to Mg, Fe, and Sc abundances in these figures: stars are retained only if the machine learning-derived uncertainties in [Mg/H] and [Fe/H] are less than 0.2, and the uncertainty in [Sc/H] is less than 0.15.

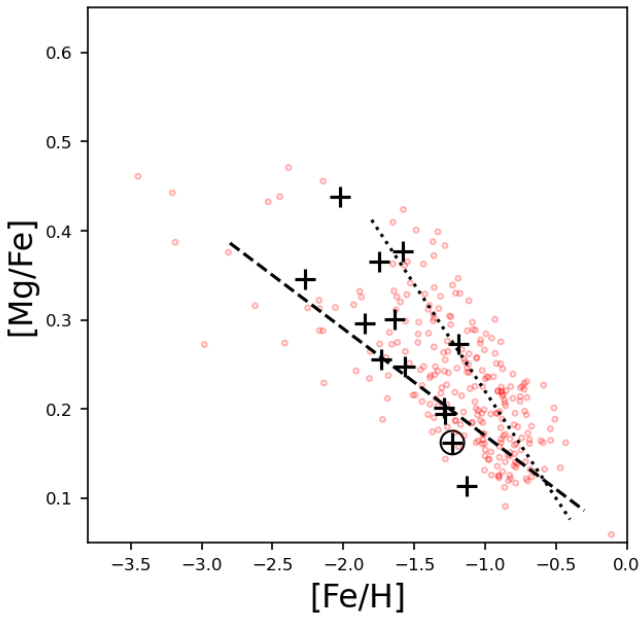


Figure 16. [Mg/Fe]–[Fe/H] distribution of stars in Cluster 5 and the kinematically cold disk, represented by plus signs and circles, respectively. Auxiliary lines from Figure 5(b1) are also included to illustrate the two distinct trends. One star in Cluster 5 with available age information is marked by a black circle.

4. Discussion

4.1. Formation History of the Galactic Metal-poor Disk

4.1.1. Formation Process

We suggest a possible formation scenario that can account for the above observed phenomena, and we summarize it in four stages:

1. *Primordial phase*. Figures 7 and 8 indicate that the kinematically cold disk sample contains stars older than 12 Gyr, which we identify as the Galactic primordial disk. Although we believe the merger between the progenitor of GSE and the Milky Way caused significant contamination of the disk by stars from GSE's progenitor, we argue that the kinematically cold disk sample older than 12 Gyr is not the result of this contamination, since these stars have higher metallicity than GSE stars of the same age, and lack the distinct α-knee at [Fe/H] ≈ -1.3. Helmi et al. (2018) noted that the mass of GSE's progenitor is about one-fourth of the Milky Way's, explaining why the primordial disk has higher metallicity, and smaller [Mg/Fe] dispersion at a given [Fe/H]. The Milky Way's larger mass led to a higher star formation rate and faster metallicity enrichment at early stage, whereas the smaller mass of GSE's progenitor and slower star formation caused its metallicity to enrich more slowly. Additionally, the Milky Way's greater mass ensured more uniform gas distribution, whereas the progenitor of GSE likely experienced multiple star formation events, resulting in less uniform gas and greater [Mg/Fe] dispersion. In fact, based on our sample, there should also be primordial Galactic disk stars present in our kinematically hot disk sample. As shown in Figure 7, the kinematically hot disk contains some members older than 12 Gyr. However, these kinematically hot disk stars with ages greater than 12 Gyr span a wide range of metallicities, some have [Fe/H] > -1.5, while others are more metal-poor, even with [Fe/H] < -2.5. In Figure 8, if we consider the metallicity histogram in panel (a4) as representative of primordial disk stars, and panel (c4) as that of stars from the progenitor of GSE, then the histogram in panel (b4) can be well described as a combination of (a4) and (c4) in certain proportions (see Section 4.1.3 for details). This indicates the presence of primordial disk stars in our kinematically hot disk sample. Therefore, we suggest that the primordial Galactic disk may not have been "thin."
2. *Merger phase.* We infer that the merger between the GSE progenitor and the Milky Way occurred approximately 10–12 Gyr ago. This inference is mainly based on the evolution of metallicity and [Mg/Fe] in the GSE sample, along with the metallicity trends observed in the kinematically hot disk. Figure 18 presents the metallicity histograms for both the kinematically hot disk and GSE samples in the 10–12 Gyr and >12 Gyr groups. Unlike in Figure 8, Figure 18 includes some key statistical indicators such as the mean and median metallicities. Furthermore, we apply kernel density estimation (KDE) to model the metallicity distribution of the four subsamples. The KDE fits are shown as dashed lines, with the corresponding peak [Fe/H] values annotated.

   The first line of observational evidence is that the 10–12 Gyr GSE sample exhibits higher metallicities compared to the older (>12 Gyr) GSE population. This trend is consistently reflected in the median, mean, and peak values of the [Fe/H] distributions. Moreover, the >12 Gyr GSE sample contains a significantly larger fraction of very metal-poor stars ([Fe/H] < -2.0), whereas such stars are rare in the 10–12 Gyr age group.

   The second piece of evidence is that stars in the 10–12 Gyr GSE sample show higher [Mg/Fe] compared to those in the >12 Gyr sample. As shown in Figure 9, [Mg/Fe] decreases with increasing [Fe/H] among stars older than 12 Gyr, with a noticeable α-knee appearing at [Fe/H] ≈ -1.3. In contrast, the 10–12 Gyr GSE group does not show a distinct α-knee. Instead, for [Fe/H] < -1.5, these younger stars display elevated [Mg/Fe] relative to their older counterparts. In panels (c3) and

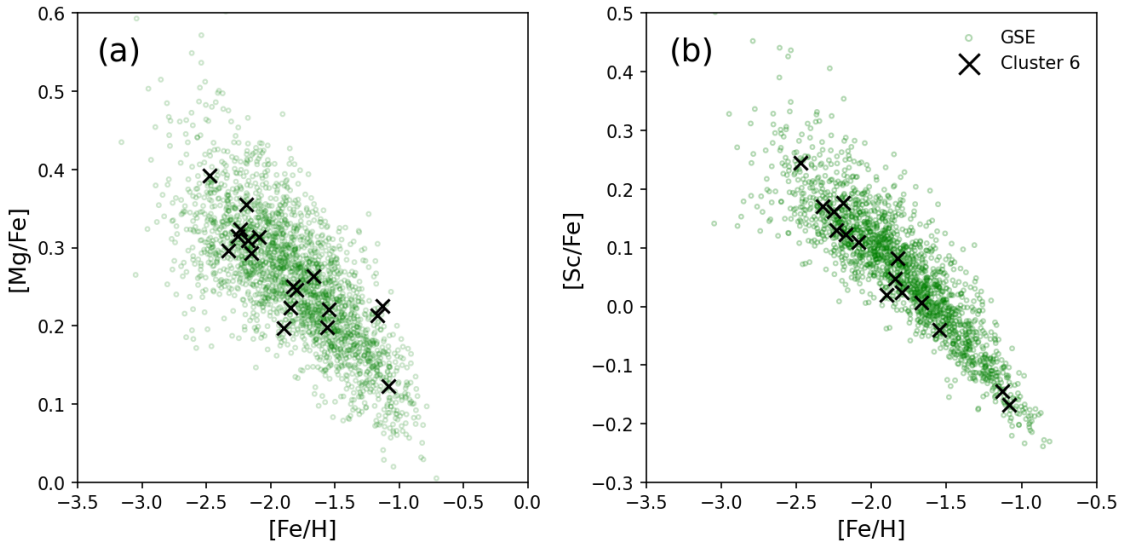


Figure 17. (a) [Mg/Fe]–[Fe/H] distribution of Cluster 6 and GSE. (b) [Sc/Fe]–[Fe/H] distribution of Cluster 6 and GSE. Each cross or circle represents a star.

(c4) of Figure 9, the >12 Gyr stars are roughly symmetrically distributed around the dashed reference line, whereas the majority of stars in the 10–12 Gyr group lie above the line.

The third line of evidence comes from the kinematically hot disk. As discussed in Section 3.1, it superficially appears that the metallicity of the kinematically hot disk increases from the >12 Gyr group to the 10–12 Gyr group. This impression arises because the older group contains a substantial fraction of very metal-poor stars, while such stars are nearly absent in the 10–12 Gyr group. Additionally, as seen in Figure 18, both the median and mean [Fe/H] values increase from the >12 Gyr group to the 10–12 Gyr group. However, we argue that this apparent metallicity increase does not reflect genuine chemical evolution. Instead, the >12 Gyr kinematically hot disk is likely a composite population, consisting of (1) stars formed in the early Milky Way, and (2) stars originally from the GSE progenitor that now exhibit disk-like kinematics. In Section 4.1.3, we demonstrate that the metallicity distribution of the >12 Gyr hot disk can be reproduced as a mixture of the >12 Gyr GSE and cold disk samples. This suggests that the >12 Gyr hot disk is not a chemically homogeneous population, but rather a blend of native and accreted stars. Because GSE progenitor was less massive than the Milky Way, its stars are generally more metal-poor than those formed in situ. As a result, their presence biases interpretations of metallicity evolution in the native disk. Based on the KDE peak values in Figure 18, the peak [Fe/H] for the >12 Gyr hot disk is −1.1, while it shifts to −1.24 for the 10–12 Gyr group. This implies that, contrary to the initial impression, the metallicity of the native disk decreases over time. In Figure 19, we compare the KDE-fitted metallicity distributions of the 10–12 Gyr hot disk with the two decomposed components of the >12 Gyr hot disk. Relative to the native component of the older hot disk, the 10–12 Gyr population exhibits lower median, mean, and peak metallicities, further supporting a decline in metallicity within the in situ disk over time. This conclusion hinges on the population decomposition of the >12 Gyr kinematically hot disk, based on a comparative analysis of the metallicity distributions among the hot disk, cold disk, and GSE samples, as detailed in Section 4.1.3. Previous studies suggest that stars formed in the GSE progenitor typically have lower [Al/Fe], while stars born in the Milky Way show higher [Al/Fe] (e.g., Feuillet et al. 2021; Belokurov & Kravtsov 2022). However, due to the limited spectral resolution of LAMOST, we are unable to measure [Al/Fe] for the stars in our sample and thus cannot provide direct chemical evidence for this decomposition. Nevertheless, other studies have identified metal-poor stars in the Galactic disk with [Al/Fe] < 0 (e.g., Malhan & Rix 2024; Re Fiorentin et al. 2024). It is important to note that the metallicity statistics presented in Figures 18 and 19 may show slight discrepancies. This is because the values in Figure 18 are derived from the observed stellar samples, whereas those in Figure 19 are based on KDE-fitted functions.

These phenomena can be explained by the interaction between the GSE progenitor and the Milky Way. Since the GSE progenitor had a lower mass than the Milky Way, its gas should been correspondingly more metal-poor. During their interaction, gas from both systems likely mixed, producing an interstellar medium with intermediate metallicity, higher than that of the GSE progenitor, but

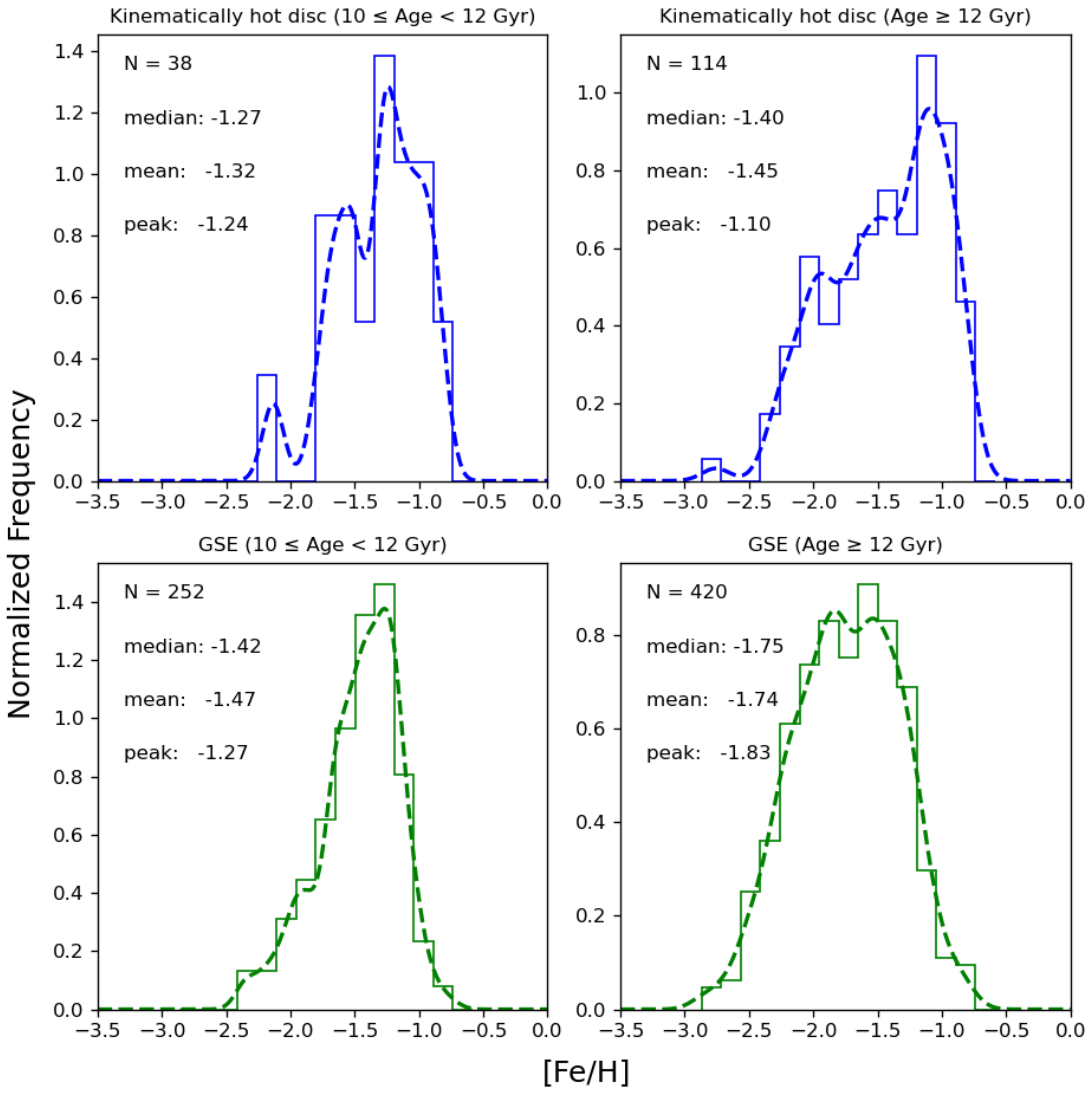


Figure 18. Normalized [Fe/H] histograms for the kinematically hot disk and GSE subsamples, shown for the 10–12 Gyr and >12 Gyr age groups only. Each panel lists the number of stars (*N*), along with the median, mean, and KDE peak [Fe/H] values. Solid lines show the histograms; dashed lines indicate KDE fits.

lower than that of the Milky Way. Stars formed from this mixed gas would reflect these metallicity properties. In particular, stars with highly eccentric orbits would appear more metal-rich than older GSE stars. At the same time, disk stars formed from this mixed gas would be more metal-poor than the older, in situ disk stars. The interaction likely triggered a burst of star formation, leading to creation of many massive stars. These massive stars, with their short lifespans, quickly ended their lives as type II supernovae, releasing large amounts of α elements such as magnesium into the surrounding interstellar medium. This rapid enrichment would result in elevated [Mg/Fe] in stars formed during the merger event. It is important to note that [Mg/Fe] is also influenced by iron production, which predominantly comes from type Ia supernovae. In the early stages of galaxy evolution, before type Ia supernovae had significantly enriched the environment with iron, [Mg/Fe] would naturally remain high. However, in our case, we observe that the 10–12 Gyr GSE stars are not only more metal-rich but also show higher [Mg/Fe] compared to the older (>12 Gyr) GSE population. If no starburst had occurred to enhance magnesium, one would expect the younger, more metal-rich GSE stars to exhibit lower [Mg/Fe] than their older counterparts. The fact that this is not observed indicates the need for an additional source of magnesium. The most plausible explanation is a burst of star formation triggered by the merger between GSE and the Milky Way.

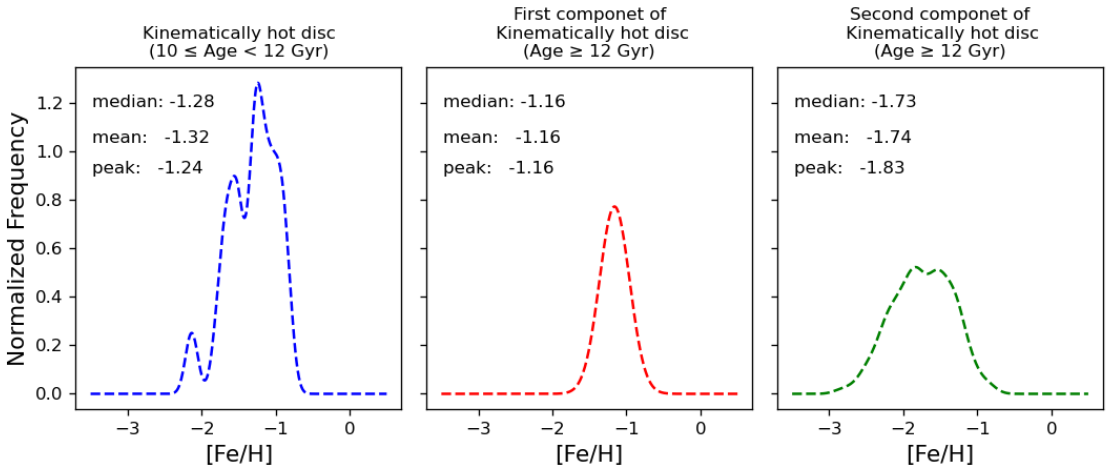


Figure 19. KDE-fitted [Fe/H] distributions for the kinematically hot disk with ages 10–12 Gyr (left) and the two decomposed components of the >12 Gyr kinematically hot disk (middle and right). The >12 Gyr hot disk is modeled as a composite population consisting of stars formed in the early Milky Way (middle panel) and stars originally from the GSE progenitor (right panel). The decomposition process is described in Section 4.1.3. Each panel reports the median, mean, and peak [Fe/H] values derived from the fitted KDE distributions.

This timescale is also consistent with previous studies. Helmi et al. (2018) estimated the GSE merger at 10 Gyr ago, and Belokurov et al. (2018) suggested a range of 8–11 Gyr. Similarly, Ciucă et al. (2023) identified a population of stars aged 10–12 Gyr that are more [Mg/Fe]-rich than older stars, attributing this to starbust by GSE merger. Aguado-Agelet et al. (2025) found two groups of GSE-associated globular cluster: one around ∼13 Gyr and [Fe/H] ≈ −1.6, likely formed before the merger, and another around ∼11.5 Gyr and [Fe/H] ≈ –1.1, likely formed during the merger process. Vincenzo et al. (2019) also concluded, via chemical evolution modeling, that the merger occurred about 10 Gyr ago.

In conclusion, the GSE merger had a profound impact on the Milky Way. First, it introduced stars from the GSE progenitor into the Galactic disk. Our previous analysis shows that kinematically cold disk stars older than 12 Gyr are largely uncontaminated by GSE debris, as their [Fe/H] is much higher than that of GSE stars of the same age. Thus, the kinematically cold disk stars in this age range can represent the metallicity of the Galactic primordial disk. In contrast, the stars from the GSE progenitor that polluted the primordial disk are found mainly in our kinematically hot disk sample. Therefore, kinematically hot disk stars older than 12 Gyr include both primordial disk stars and GSE-origin stars. This explains why the metallicity distribution of kinematically hot disk stars older than 12 Gyr falls between that of the cold disk and the GSE at the same age. Second, the merger brought metal-poor gas into the Milky Way, triggering a starburst that formed a new generation of metal-poor disk stars. Because this mixed gas was more metal-poor than the original Galactic gas, kinematically hot disk stars formed 10–12 Gyr are more metal-poor than older in situ disk stars.

3. *Quiescent phase.* Following the GSE merger, the Milky Way's disk entered a quiescent phase 8–10 Gyr ago, marked by slow gas accretion and reduced star formation. This is evidenced by the significantly fewer stars aged 8–10 Gyr compared to 10–12 Gyr, the absence of Mg-enhanced stars in the 8–10 Gyr group.
4. Latest formation phase: From 8 Gyr ago to the present, the formation of the Milky Way's disk proceeded independently of the GSE merger. Stars younger than 8 Gyr in the kinematically cold disk and hot disk exhibit a steep decline in [Mg/Fe] with increasing [Fe/H], more pronounced than in other age groups and lacking an α-knee. This pattern suggests an early phase of α-element enhancement, followed by a gradual enrichment in iron over time. The missing α-knee might occur at very low [Fe/H] (below −2.0). An outlier star in Figure 9(a1) with [Fe/H] = −2.8 and [Mg/Fe] = 0.42 could suggest a knee between [Fe/H] = −3 and −2, but this single star might be unrelated to this latest formed disk.

In summary, the metal-poor disk of the Milky Way primarily consists of four components: the primordial Galactic disk, stars originating from GSE's progenitor, stars formed from the mixed gas of GSE's progenitor and the ancient Milky Way, and stars formed within the last 8 Gyr. We believe these recently formed

stars represent the metal-poor tail of the otherwise metal-rich, low-α disk, and we will explain this interpretation in detail in Section 4.1.4. This formation scenario explains the metallicity histogram of the kinematically cold disk shown in panel (a1) of Figure 5, where three distinct peaks are observed. We attribute the prominent peak at [Fe/H] = −0.8 to young, metal-poor, high-α stars, which we identify as the metal-poor tail of the traditional metal-rich, low-α disk, while the peak at [Fe/H] = −1.2 likely originates from the Galactic primordial disk. The less prominent peak at [Fe/H] = −2.0 suggests stars related to GSE merger.

4.1.2. The Primordial Disk

The primordial disk formed over 12 Gyr ago with a metallicity lower limit of about −1.5. Xiang et al. (2024) also identified a primordial disk but suggested a higher metallicity limit ([Fe/H] > −0.5). This discrepancy likely stems from methodological differences: we directly applied a kinematic criterion to identify disk stars, while Xiang et al. (2024) did not distinguish between disk and halo stars. Instead, they grouped stars by age and studied the relationship between age and morphology, defining morphology as the ratio of scale length to scale height. In their analysis, the primordial disk at [Fe/H] < −0.5 is obscured by numerous GSE stars, which typically have [Fe/H] < −1.0 and exhibit a halo-like geometry. Xiang et al. (2024) observed disk-like structures in stars as old as 13 Gyr, with these features becoming more prominent in stars younger than 12.5 Gyr. This can be explained by our findings: disks older than 12 Gyr are primarily composed of the primordial disk stars and stars from GSE's progenitor, and Figure 7 suggests that our kinematically hot disk sample is slightly older than the GSE sample. Based on the above analysis, our kinematically hot disk sample consists of stars originating from both the progenitor of GSE and the Galactic primordial disk. Therefore, the Galactic primordial disk itself should be older than the GSE progenitor. At 13 Gyr, the relatively low presence of GSE stars allows for detection of weak disk-like features. However, between 12.5 and 13 Gyr, the increasing number of GSE stars significantly obscures the primordial disk, making the disk less apparent; 12.5 Gyr ago, the GSE merger triggered the growth of the disk, leading to the prominent disk-like structures observed by Xiang et al. (2024). These findings demonstrate consistency between our results and their observations.

Several studies have provided indirect evidence for an ancient primordial disk in the Milky Way. Using common sources between APOGEE DR17 and Gaia DR3, Belokurov & Kravtsov (2022) identified Galactic in situ stars with [Al/Fe] > −0.1. These stars have [Fe/H] between −1.5 and −0.5. Stars with [Fe/H] < −1.3 exhibit isotropic velocity distributions and large chemical abundance dispersion, characteristics of a local primordial halo they named "Aurora." Stars with −1.3 < [Fe/H] < −0.9 show increasing azimuthal velocity ($|V_\phi|$) and decreasing V_ϕ dispersion, indicating a transition from chaotic to ordered motion around redshift 2.5. Their finding of the lower [Fe/H] limit of −1.5 for early in situ stars matches our lower limit for the primordial disk. While they observed disk-like features emerging at [Fe/H] = −1.3, slightly higher than −1.5, this difference likely stems from methodology. Their approach might ignore sparse disk-like stars between [Fe/H] −1.5 and −1.3. Furthermore, Belokurov et al. (2020) identified "Splash," a local halo component with [Fe/H] > −0.7, suggesting it originated from the heating of the Galactic primordial disk due to GSE merger. They believed Splash extends to [Fe/H] below −0.7, but most observed Splash stars have [Fe/H] > −0.7, because at lower [Fe/H], Splash becomes difficult to distinguish from other halo populations such as debris from dwarf galaxies. In other words, the metallicity of the primordial disk could extend below −0.7, which is also consistent with our findings. Gallart et al. (2019) discovered two distinct sequences in the color–magnitude diagram of the Galactic halo stars. Through analysis, they identified one sequence as GSE and the other as in situ halo stars. Similar to the Splash (Belokurov et al. 2020), they proposed these in situ stars come from the heated Galactic primordial disk, due to the merger with GSE's progenitor. Their results show that both in situ stars and GSE stars have ages primarily between 10 and 14 Gyr, consistent with our findings. Additionally, they found that in situ stars generally have higher metallicities than GSE stars, with a lower metallicity limit of −1.7 for in situ stars, which broadly agrees with our −1.5 lower limit. Although they did not directly identify the primordial disk, they suggested that the primordial disk was not completely heated, with some stars maintaining their disk-like motions. These studies indicate that ancient Galactic stars generally have [Fe/H] above −1.5, and suggest a possible primordial disk. Regarding the in situ halo, some propose it contains both heated primordial disk stars and primordial halo stars (Belokurov & Kravtsov 2022), while others argue it formed entirely from the heated early disk (Matteo et al. 2019). Whether a primordial halo exists remains an open question.

However, some works present different views. Viswanathan et al. (2024) studied the Galactic disk's formation using high-α stars from Gaia DR3, attempting to exclude GSE stars. They found stars with [Fe/H] < −1.5 have negligible rotation ($V_\phi \approx 0$), attributing them to GSE. For stars with −1.5 < [Fe/H] < −1.0, they observed an average V_ϕ of $-50\,\mathrm{km\,s^{-1}}$, which they identify as the earliest Galactic disk. When [Fe/H] exceeds −1.0, the average V_ϕ reaches $-200\,\mathrm{km\,s^{-1}}$, representing the high-$\alpha$ disk. Based on the linear relationship between age and metallicity for high-α stars, they concluded the transition from the earliest disk to high-α disk was gradual. These conclusions differ from our results in two major aspects. First, regarding the earliest disk characteristics, they suggest a modest rotation ($V_\phi \approx -50\,\mathrm{km\,s^{-1}}$), while we

find significant rotational motion ($V_\phi < -100\,\mathrm{km\,s^{-1}}$). We attribute this discrepancy to two factors: (a) their sample includes GSE stars. As shown in Figure 5, GSE stars can have [Mg/Fe] lager than 0.3 in $-1.5 < [\mathrm{Fe/H}] < -1.0$, resulting in a lower average $|V_\phi|$ due to the mixture of primordial disk and GSE stars; (b) our results show that although most stars from GSE's progenitor have highly eccentric orbits, some can exhibit disk-like orbits, making them particularly difficult to distinguish from in situ disk stars through kinematics. Second, regarding the evolution from primordial disk to high-α disk, they propose a slow, gradual transition, while we argue for an abrupt change. This second point will be discussed in detail in Section 4.1.3.

Regarding the morphology of the primordial disk, our classification criteria for kinematically hot and cold disks show that the primordial disk exists in both. This suggests that the primordial disk of the Milky Way is not a very thin, kinematically cold disk.

4.1.3. Impact of GSE Merger

GSE is a prominent merger remnant in the Galactic halo, notable for stars with nearly zero azimuthal velocity (V_ϕ) and high radial velocity (V_r) (Belokurov et al. 2018; Helmi et al. 2018). This study reveals that the merger between the Milky Way and GSE's progenitor not only contributed numerous stars into the halo but also profoundly influenced the Galactic disk's evolution. The merger effected the disk in two major ways. First, it contaminated the primordial disk, as detailed in Section 4.1.1, introducing many stars from the GSE's progenitor into the Galactic primordial disk. While GSE stars are typically associated with highly eccentric orbits, Koppelman et al. (2020), through the merger simulation, showed that 9% of stars from GSE's progenitor could exhibit eccentricities below 0.6, and some have V_ϕ reach $-170\,\mathrm{km\,s^{-1}}$. This simulation-based result indicates that some stars from GSE's progenitor may have settled into disk-like orbits. Furthermore, Vasiliev et al. (2022) demonstrated that the orbits of satellite galaxies can become more circular during mergers under specific conditions.

The term GSE is often used inconsistently in the literature, sometimes referring to the dwarf galaxy that merged with the Milky Way and other times to the resulting remnant. As most studies associate this remnant with the Galactic halo, GSE often implies halo-like kinematic characteristics. However, this study shows that the merger also contributed stars to the disk. To clarify, this work defines GSE specifically as the stars with high eccentricities.

Based on our analysis of Figure 8, we identify that our kinematically cold disk stars older than 12 Gyr are entirely composed of the in situ primordial disk stars. These stars exhibit significantly higher metallicities than the GSE stars of the same age, indicating that our kinematically cold disk could not have been contaminated by accreted stars from the GSE progenitor. In contrast, the kinematically hot disk stars older than 12 Gyr show a broad distribution in metallicity. Some stars are relatively metal-poor, with [Fe/H] values as low as −3.0, consistent with the chemical properties of the GSE sample of the same age. Meanwhile, others have comparatively higher metallicities around $[\mathrm{Fe/H}] \approx -1.0$, similar to those found in our cold disk of the same age. This wide metallicity spread suggests that the kinematically hot disk in this age range consists of two distinct stellar populations: (1) in situ stars belonging to the Galactic primordial disk, and (2) stars accreted from the GSE progenitor. In summary, the Galactic primordial disk is present in both our kinematically cold and hot disk samples older than 12 Gyr. Crucially, our kinematically hot disk older than 12 Gyr is composed exclusively of two components: the Galactic primordial disk stars and accreted stars from the GSE progenitor. No additional stellar populations are needed to explain the observed chemical abundance patterns. To evaluate how strongly the merger contaminated the Galactic primordial disk, we modeled the metallicity distribution of our kinematically hot disk sample older than 12 Gyr as a combination of the distributions of our kinematically cold disk and GSE samples of the same age

$$f_{\mathrm{hot}} = \beta \times f_{\mathrm{cold}} + (1 - \beta) \times f_{\mathrm{GSE}}. \tag{10}$$

Here, f_{hot}, f_{cold}, and f_{GSE} are the metallicity distribution functions (MDFs) for the kinematically hot disk, cold disk, and GSE, respectively. These distribution functions were obtained by fitting the observed [Fe/H] histograms: we used KDE for our kinematically hot disk sample and our GSE sample, and a Gaussian function for our kinematically cold disk sample. Figure 20 displays the observed [Fe/H] histograms alongside the corresponding fitted distribution functions. From this modeling, we derive a best-fit value of $\beta = 0.39$, suggesting that 39% of the [Fe/H] distribution of the kinematically hot disk matches that of the kinematically cold disk, while the remaining 61% aligns with our GSE sample. Given that the kinematically hot disk sample contains 114 stars, we estimate that approximately 70 of them originated from the progenitor of GSE. Assuming that these GSE-origin stars now residing in the disk share the same MDF as the GSE-origin stars currently found in the Galactic halo, and noting that our GSE sample with age older than 12 Gyr has 420 stars, we infer that 14.3% of stars from the GSE progenitor were deposited into the disk, while the remaining 85.7% stayed in the Galactic halo. Lane et al. (2023) estimated the stellar mass of the GSE progenitor to be $1.45 \times 10^8 M_\odot$ (where $M_\odot$ is the solar mass) but their analysis did not account for GSE-origin disk-residing stars. Based on our inferred fraction, we estimate the stellar mass of GSE-origin stars that ended up in the Galactic disk to be approximately $\sim 2.42 \times 10^7 M_\odot$.

More importantly, the merger triggered the formation of the so-called metal-rich, high-α disk (thick disk). Three key observations support this: (1) a significant fraction of stars in

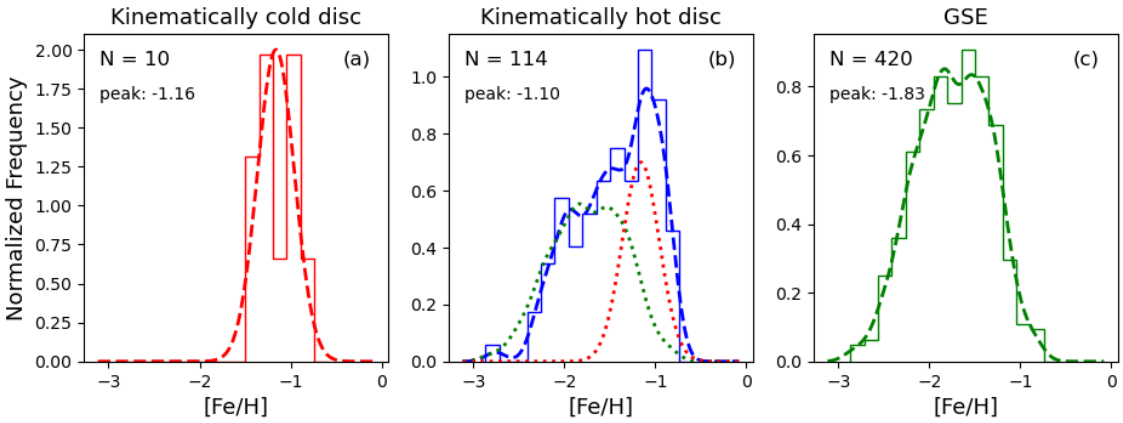


Figure 20. Solid lines show the normalized metallicity histograms for the kinematically cold disk (a), kinematically hot disk (b), and GSE (c) stars older than 12 Gyr, with the number of stars indicated in the top-left corner of each panel. Dashed lines represent the fitted distribution functions, with the peak [Fe/H] values indicated on the plots. In panel (b), the red dotted line corresponds to 0.31 times the kinematically cold disk fit, and the green dotted line corresponds to 0.69 times the GSE fit.

our sample formed 10–12 Gyr ago, consistent with the metal-rich, high-α disk formation timescale proposed by many studies (e.g., Haywood et al. 2013; Xiang & Rix 2022); (2) stars aged 10–12 Gyr show enhanced [Mg/Fe], indicating a violent starburst during this period; (3) for stars aged 10–12 Gyr, the majority exhibit kinematic properties consistent with our kinematically hot disk criterion, while the fraction meeting our kinematically cold disk criterion is relatively low. Previous studies also support this conclusion. Belokurov et al. (2018) suggested that the merger likely induced the formation of the Galactic thick disk. Similarly, Ciucă et al. (2023) identified a teardrop-shaped region in the [Fe/H]–Age plane for stars aged 9–12 Gyr, termed the "Dip." Compared to stars older than 12 Gyr, these stars show a sharp drop in [Fe/H] and a steep rise in [Mg/Fe]. They concluded that the merger between the Milky Way and the GSE's progenitor brought in metal-poor gas, triggering a starburst that formed the high-α disk. Their findings align closely with ours.

4.1.4. Metal-poor Tail of Metal-rich, Low-α Disk

We argue that disk stars with $[\mathrm{Fe/H}] < -0.8$ and age < 8 Gyr represent the metal-poor tail of the metal-poor, low-α disk based on three key observations: First, their formation timing aligns with the metal-rich, low-α disk, which began forming 8 Gyr ago (Haywood et al. 2013). Second, their spatial distribution and kinematic properties match metal-rich, low-α disk (thin disk) characteristics, as discussed in Section 3.1. Third, their [Mg/Fe]–[Fe/H] slope matches the metal-rich, low-α disk pattern. Extrapolating our fitted line yields [Mg/Fe] values of 0.11, 0.07, and 0.05 at $[\mathrm{Fe/H}] = -0.5$, -0.2, and 0, respectively, consistent with low-α characteristics at the metal-rich end. Our findings differ somewhat from those of H. Haywood et al. (2013), particularly regarding the early stages of low-α disk formation. H. Haywood et al. (2013) suggested that the inner low-α disk ($r_{\mathrm{gal}} < 10$ kpc) initially had relatively high metallicity (-0.1 to 0.1) and low α-abundance, inheriting the chemical composition of the metal-rich, high-α disk's late-stage gas, followed by a period of slow evolution. In contrast, our results indicate that the low-α disk began with much lower metallicity, possibly at least as low as -1.8, and experienced a brief period of rapid star formation, which significantly increased the [Mg/Fe]. Only after this phase did the low-α disk enter a slow evolutionary period, starting from $[\mathrm{Fe/H}] \approx -1.8$ and $[\mathrm{Mg/Fe}] \approx 0.4$. The mechanism that triggered the initial rapid star formation in the so-called metal-rich, low-α disk (thin disk) remains unclear. This suggests that there is no clear inheritance relationship between the so-called metal-rich, low-α and so-called metal-rich, high-α disks, but rather significant differences in their formation mechanisms. Some studies suggest that the thin disk's metallicity could be as low as -2.0 (Fernández-Alvar et al. 2021, 2024), but without age information, it is uncertain whether these stars are younger or older than 8 Gyr.

These stars exhibit relatively low [Fe/H] and elevated [Mg/Fe], yet they are remarkably young and display kinematically cold disk characteristics—a rather unusual combination. One possible explanation is that their ages have been underestimated due to observational or modeling uncertainties, such as unresolved binarity, inaccurate stellar parameters, or limitations in age-dating methods. In Section 3.1, we investigate whether these stars could be members of unresolved binary systems, but found insufficient evidence to support this scenario. Of course, this does not constitute definitive proof of their youth. We therefore strongly look forward to further observations, such as asteroseismic measurements and high-resolution spectroscopy, to better constrain the nature of these stars. Based on the currently available information, we propose that these stars may

represent the metal-poor tail of the metal-rich, low-α disk population. In addition, compared to older stars, they exhibit higher [Mg/Fe] at a given [Fe/H], suggesting that they have experienced α-enrichment. In our sample, the number of such young, metal-poor, high-α stars is small, and they do not appear to be associated with intense star formation activity. Consequently, the mechanism responsible for their α-enrichment remains unclear.

4.1.5. Comparative Analysis of Metal-poor Disk

The origin of the Galactic metal-poor disk has been widely debated, with various scholars offering different perspectives. Here, we compare our findings with those in the literature.

Naidu et al. (2020) identified a group of stars, named Aleph, with $V_\phi < -170\,\mathrm{km\,s^{-1}}$ and $|V_r| < 70\,\mathrm{km\,s^{-1}}$. These stars have $[\mathrm{Fe/H}] > -0.8$, $[\alpha/\mathrm{Fe}] < -0.27$, and a wide $|z|$ distribution, reaching up to 10 kpc. Horta et al. (2023) analyzed Aleph's chemical abundances, finding its properties intermediate between metal-rich, high-α and metal-rich low-α disks but closer to the low-α disk. They suggested Aleph represents disk stars perturbed by the Sagittarius. We agree with this interpretation but since the low-α disk extends to $[\mathrm{Fe/H}] \approx -1.8$, Aleph likely includes lower-metallicity members, though these may be obscured by halo stars at large distances from the midplane.

Fiorentin et al. (2021) discovered Icarus, a group of metal-poor stars with average $[\mathrm{Fe/H}] \approx -1.35$ and disk-like kinematics. Initially, Icarus was thought to be a low-α structure with an average $[\mathrm{Mg/Fe}] \approx -0.02$. However, subsequent observations revealed the average $[\mathrm{Mg/Fe}] \approx 0.27$ (Re Fiorentin et al. 2024). Their detailed comparison of chemical abundances between Icarus and GSE revealed strong similarities, suggesting Icarus is the relic of a dwarf galaxy. While they did not explicitly propose that Icarus comes from the progenitor of GSE, our work shows that some stars with disk-like kinematics belong to GSE's progenitor, supporting the idea that Icarus and GSE originated from the same merger event.

Naidu et al. (2020) defined the MWTD as stars with disk-like kinematics, $[\mathrm{Fe/H}] < -0.8$, and $[\alpha/\mathrm{Fe}] > 0.25$, suggesting it represents the metal-poor tail of the thick disk (the metal-rich high-α disk). Our analysis suggests MWTD is multi-component, including the metal-poor tail of the metal-rich low-α disk (thin disk) stars from the progenitor of GSE, and the metal-poor tail of metal-rich high-α disk (thick disk).

Malhan & Rix (2024) discovered two metal-poor structures in E–L_z space: Shakti ($L_z \approx 1000\,\mathrm{kpc\,km\,s^{-1}}$) and Shiva ($L_z \approx 300\,\mathrm{kpc\,km\,s^{-1}}$), and found that some stars in these structures have $[\mathrm{Al/Fe}] > 0$ while others have $[\mathrm{Al/Fe}] < 0$. Since Galactic stars typically have $[\mathrm{Al/Fe}] > 0$ and GSE stars $[\mathrm{Al/Fe}] < 0$, and given these stars are highly clustered in E–L_z space as well as their $[\mathrm{Fe/H}] < -1$, they concluded that Shakti and Shiva are relics of dwarf galaxies unrelated to GSE. However, we argue that these structures result from the mixing of GSE stars and the Galactic primordial disk stars. This mixing naturally explains the observed [Al/Fe] variations, with stars in both structures exhibiting $[\mathrm{Al/Fe}] > 0$ and $[\mathrm{Al/Fe}] < 0$.

Nepal et al. (2024) proposed the existence of an "old thin disk" in the Milky Way based on their study of old disk stars ($[\mathrm{Fe/H}] < -1.0$, $Z_{\max} < 1$ kpc). Unlike the commonly referred to thin disk, which is low-α and formed within last 8 Gyr, their "old thin disk" is ancient (13–14 Gyr old) and primarily consists of high-α stars. They suggested that metal-poor disk stars originated from this "old thin disk," which was later heated during the GSE merger. They also identified low-α stars older than 10 Gyr within this disk, leading to the hypothesis that it underwent different evolutionary phases. However, we disagree with their assumption that all metal-poor disk stars originate from this "old thin disk," as we believe their formation occurred in multiple stages, resulting in a diverse composition. Our proposed formation process explains several phenomena observed by Nepal et al. (2024), such as the single age peak at 13.5 Gyr for stars with $|V_\phi| > 180\,\mathrm{km\,s^{-1}}$ and two peaks at 12 and 13.5 Gyr for stars with $|V_\phi| < 180\,\mathrm{km\,s^{-1}}$. This aligns with our hypothesis that contamination from GSE merger primarily affected the warmer part of the Galactic primordial disk, leaving the colder part largely intact. The single age peak for stars with $|V_\phi| > 180\,\mathrm{km\,s^{-1}}$ corresponds to the less polluted primordial disk, while two age peaks at 12 and 13.5 Gyr for stars with $|V_\phi| < 180\,\mathrm{km\,s^{-1}}$ reflect the merger-polluted disk and the primordial disk, respectively. Additionally, Nepal et al. (2024) found that disk stars older than 13 Gyr span [Fe/H] from -2.5 to -0.5, with metal-rich stars dominating, consistent with our conclusion that the primordial disk has $[\mathrm{Fe/H}] > -1.5$ and was polluted by more metal-poor stars from GSE's progenitor. Furthermore, we did not observe low-α disk stars older than 10 Gyr and believe the formation of the low-α disk is unrelated to the ancient primordial disk.

4.2. Properties of Merger Remnants

Pontus, identified by Malhan et al. (2022), is a merger remnant with low orbital energy and slightly retrograde motion, exhibiting kinematics close to GSE and Thamnos. Their study reports that Pontus has an MDF spanning [Fe/H] from -2.3 to -1.3, with a median of $[\mathrm{Fe/H}] = -1.7$, while GSE exhibits a broader metallicity range and a significant population of stars with $[\mathrm{Fe/H}] > -1.5$. Malhan et al. (2022) also suggested that Thamnos is more metal-rich, with [Fe/H] ranging from -1.4 to -1.1, and proposed that Pontus originated from a distinct merger event compared to GSE and Thamnos. In this work, we also find that Pontus has a lower proportion of stars with [Fe/H] > -1.5 compared to GSE. However, we show that Thamnos can have [Fe/H] as low as -3.0, contradicting the claim that Thamnos is definitively more metal-rich.

Wukong (also known as LMS-1), was first identified by Yuan et al. (2020a) and later independently discovered by Naidu et al. (2020). This remnant exhibits mild prograde rotation and intermediate/high orbital energy. Its origin, whether as part of GSE or from a separate dwarf galaxy, remains uncertain. Naidu et al. (2020) and Horta et al. (2023) analyzed Wukong's chemical properties using stars from H3 and APOGEE surveys, respectively, and both constrained its metallicity to $[\mathrm{Fe/H}] < -1.45$. Naidu et al. (2020) found Wukong's [Fe/H] distribution similar to GSE's but with multiple peaks, unlike GSE's single peak. Horta et al. (2023) observed chemical similarities across multiple elements and suggested a shared origin for Wukong and GSE. In contrast, Zhang et al. (2024) and Limberg et al. (2024) found that Wukong's $[\alpha/\mathrm{Fe}]$ forms a plateau at $[\mathrm{Fe/H}] < -2.0$, unlike GSE, where $[\alpha/\mathrm{Fe}]$ decreases with rising [Fe/H]. Malhan et al. (2022) showed Wukong is the most metal-poor accreted remnant, far poorer than GSE, and hosting three most metal-poor stellar streams: C-19 ($[\mathrm{Fe/H}] = -3.38$), Sylgr ($[\mathrm{Fe/H}] = -2.92$), and Phoenix ($[\mathrm{Fe/H}] = -2.70$). Given these chemical and kinematic differences, Malhan et al. (2022) concluded that Wukong and GSE are likely products of separate merger events. Our study confirms Wukong as the most metal-poor structure among the six accreted remnants analyzed in this paper. Considering these chemical differences and their distinct kinematic properties, where GSE has most stars with eccentricities above 0.7 while Wukong's eccentricity distribution spans 0.2–0.8, and GSE has a net azimuthal velocity (V_ϕ) close to 0 whereas Wukong has $-50\,\mathrm{km\,s^{-1}}$, we tend to believe that Wukong and GSE originated from different progenitor galaxies.

Sequoia, a retrograde structure identified by Myeong et al. (2019), has slightly lower metallicity than GSE and moderate orbital eccentricity ($e \approx 0.6$). Myeong et al. (2019) concluded that Sequoia and GSE originated from different dwarf galaxies. However, Koppelman et al. (2020) and Amarante et al. (2022) suggested that a single merger event could produce both retrograde and high-eccentricity stars, implying a shared origin. Naidu et al. (2020) discovered that the metallicity distribution of Sequoia contains multiple peaks, and therefore suggested that Sequoia consists of three distinct components, with metallicity increasing in the order of I'itoi, Sequoia, and Arjuna. They believe that Arjuna belongs to GSE, while Horta et al. (2023) highlighted that different selection criteria could lead to varying conclusions. Our study found Sequoia has slightly lower metallicity than GSE, leaving the origin question unresolved.

4.3. New Substructures

We identified three new kinematic substructures in Section 3.3: Cluster 4, Cluster 5, and Cluster 6. The stars in Cluster 4 have V_ϕ of approximately $-100\,\mathrm{km\,s^{-1}}$, near the edge of the kinematically hot disk, but their $|V_z|$ is much higher, generally exceeding $100\,\mathrm{km\,s^{-1}}$. One Cluster 4 star was found to be 14.4 Gyr old. Cluster 4 is unlikely to be a disturbed remnant of the Galactic primordial disk, as its stars generally have metallicities below -1.5, whereas the primordial disk's metallicity exceeds -1.5. Furthermore, Cluster 4 stars exhibit a unique positive correlation between metallicity and orbital eccentricity. These findings suggest that Cluster 4 is likely the remnant of an undiscovered dwarf galaxy, which we designate as ShangGu-1.

The stars in Cluster 5 have V_ϕ values mostly below $-200\,\mathrm{km\,s^{-1}}$, aligning with the rotational characteristics of the kinematically cold disk, but their $|V_z|$ is unusually high, generally exceeding $100\,\mathrm{km\,s^{-1}}$. Their metallicity ranges from -2.27 to -1.13, with an average of -1.58. Two stars have age information: one is 3.6 Gyr old, and the other 11.2 Gyr old. The [Fe/H] error of the star aged 3.6 Gyr is larger than 0.2, so it does not appear in Figure 16. Meanwhile, the star aged 11.2 Gyr is marked with a black circle in Figure 16. In the [Mg/Fe]–[Fe/H] distribution, Cluster 5 show two branches just like our kinematically cold disk sample. We suggest that Cluster 5 is composed of stars from the Milky Way's metal-poor disk, which have been kinematically heated recently, thereby acquiring significant vertical motions. We designate Cluster 5 as ShangGu-2.

The stars in Cluster 6 exhibit disk-like rotation but unusual high radial velocity ($|V_r| \approx 200\,\mathrm{km\,s^{-1}}$). Among the five stars with age data, the oldest is 13.8 Gyr, and the youngest 12.4 Gyr, with an average age of 13.1 Gyr and 1σ dispersion of 0.33 Gyr. The average [Fe/H] of Cluster 6 is -2.06, close to the peak metallicity value of GSE in our sample. Based on its age, metallicity, and obvious radial motion, we suggest that Cluster 6 comes from the progenitor galaxy of GSE. We designate stars with disk-like rotation and significant radial motion as ShangGu-3.

ShangGu-3 shares kinematic similarities with Nyx. Necib et al. (2020) identified Nyx with an average V_ϕ of $-130\,\mathrm{km\,s^{-1}}$, an average $|V_r|$ of $134\,\mathrm{km\,s^{-1}}$, and a narrow [Fe/H] distribution averaging -0.55. Therefore, they suggested Nyx originated from a dwarf galaxy on a prograde orbit. Horta et al. (2023) later analyzed Nyx stars in APOGEE DR17, finding most have $[\mathrm{Fe/H}] > -1.0$, though some reach as low as -2.0. Their chemical similarity to the in situ, metal-rich, high-α disk led Horta et al. (2023) to propose a local origin for Nyx. Figure 21 reveals ShangGu-3 comprises three distinct subgroups: one aligning with Nyx (positive V_x and negative V_z), marked by red arrows; a second aligning with Nyx-2 (negative V_x and V_z), marked by green arrows; and a third with negative V_x and positive V_z, marked by blue arrows. Considering that the clustering algorithm might not have captured all ShangGu-3 members, we implemented additional kinematic criteria to directly select potential members:

(i) $-240 < V_\phi < -130\,\mathrm{km\,s^{-1}}$
(ii) $150 < |V_r| < 240\,\mathrm{km\,s^{-1}}$

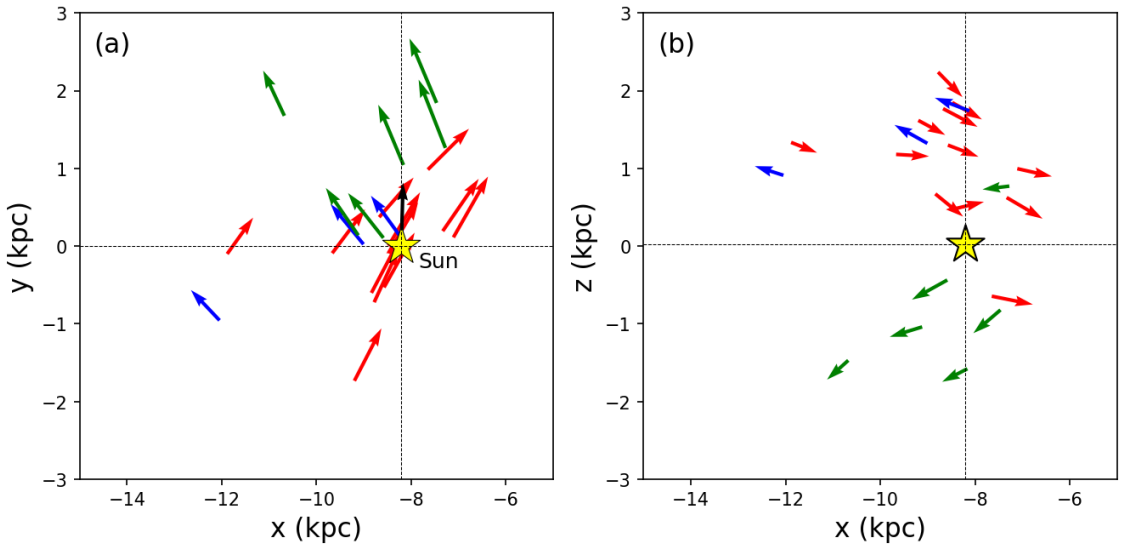


Figure 21. Velocities of ShangGu-3 stars in the Galactic x–y (a) and x–z (b) planes. The Galactic Center is at $(x, y, z) = (0, 0, 0)$ kpc while the Sun (indicated by the yellow star moving in the direction of the black arrow) is at $(x, y, z) = (-8.21, 0, 0.02)$ kpc. These stars are divided into three subgroups, shown in red, blue, and green.

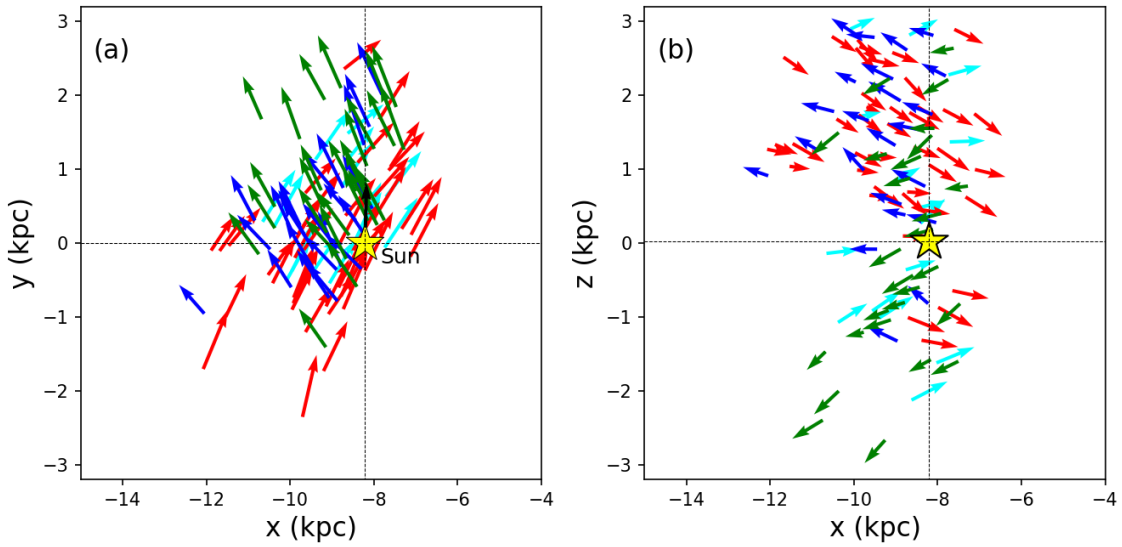


Figure 22. Similar to Figure 21, but for ShangGu-3 stars selected using kinematic criteria in Section 4.3. Compared to Figure 21, an additional subgroup with $V_x > 0$ and $V_z > 0$ is identified, marked by the cyan arrows.

(iii) $|V_z| < 100\,\mathrm{km\,s^{-1}}$
(iv) $|z| < 3$ kpc
(v) $E > -1.41 \times 10^5\,\mathrm{km^2\,s^{-2}}$

This selection yielded 106 stars, whose motions in the x–y and y–z planes are shown in Figure 22. Notably, this expanded sample reveals that ShangGu-3 actually consists of four distinct kinematic subgroups, rather than the three subgroups. Stars with positive V_x can be further divided based on their V_z values. This new finding suggests that Nyx and Nyx-2 are merely two subcomponents of a more complex structure, challenging the interpretation of Nyx as a coherent stellar stream from a merged dwarf galaxy (Necib et al. 2020).

4.4. Future Prospects

This study confirms the existence of the Galactic primordial disk, though its full scale remains uncertain due to our sample

being limited to the solar vicinity. While we analyzed the impact of the GSE merger on the evolution of the Galactic disk, the effects of other merger events remain unclear. Many merger remnants have uncertain origins and lower masses than GSE, making their influence harder to detect. Our findings reveal that Galactic structures and merger remnants lack clear kinematic boundaries. For instance, primordial disk stars are found in both our kinematically cold and hot disk samples, while stars from the GSE's progenitor appear in both the halo and disk. Understanding the distribution of Galactic structures and merger remnants across physical, kinematic, and chemical spaces is crucial for uncovering the Milky Way's assembly history. We plan to utilize deeper photometric surveys to address these issues in the future.

5. Conclusions

The LAMOST survey has collected tens of millions of low-resolution stellar spectra, from which we identified ∼100,000 metal-poor stars in our previous work. In this study, we estimated chemical abundances of these stars with two methods: machine learning and template fitting. The machine learning provided abundances for 17 elements, while template fitting estimated abundances of Fe, Mg, Ca, and C. By combining Gaia astrometry, we computed velocities, orbital parameters, and IoMs for over 70,000 stars, then selected about 46,000 with multiple-element abundances and precise kinematics to investigate the origin of the Galactic metal-poor disk and merger remnants. In addition, we supplemented age estimates for some subgiant stars in our sample.

We found that the evolution of the Galactic disk can be divided into four stages:

(i) Primordial phase (over 12 Gyr ago): The Milky Way hosted a primordial disk with metallicity above −1.5.

(ii) Merger phase (10–12 Gyr ago): The merger of the GSE progenitor with the Milky Way significantly affected the Galactic disk in two major ways. First, stars originating from the GSE progenitor, many of which have metallicities $[\mathrm{Fe/H}] < -1.5$, were deposited into the primordial disk, contaminating it with an external stellar population. Second, the merger delivered a substantial amount of metal-poor gas to the Milky Way, which fueled a rapid episode of star formation. This event gave rise to a new disk component, with some stars reaching metallicities as low as $[\mathrm{Fe/H}] \approx -2.0$. We propose that this formation process is connected to the origin of the so-called metal-rich, high-α disk (commonly referred to as the thick disk), though our findings suggest it extends to lower metallicities than traditionally assumed.

(iii) Quiescent phase (8–10 Gyr ago): After GSE merger, the Milky Way gradually accreted gas from its surroundings, leading to slow star formation.

(iv) Latest formation phase (within the last 8 Gyr): This phase corresponds to the formation of the so-called metal-rich, low-α disk (commonly referred to as the thin disk). Contrary to the assumption that the low-α disk formed with $[\mathrm{Fe/H}] > -0.7$, we propose it was initially metal-poor, with metallicity as low as −1.8. This conclusion is supported by the identification of a population of young, metal-poor, high-α stars in our study. These stars exhibit kinematic properties highly consistent with those of the metal-rich, low-α disk, suggesting a shared evolutionary origin. Although current observational data do not indicate that these stars are part of unresolved binary systems, their apparently young ages remain uncertain and may be subject to measurement biases. Furthermore, relative to other metal-poor disk stars, they show enhanced α-element abundances at a given $[\mathrm{Fe/H}]$, implying they have experienced the α-enrichment process. However, given their small numbers, it is unlikely that they are associated with episodes of intense star formation, and the underlying mechanism of their α-enrichment remains unclear. We strongly look forward to further high-resolution spectroscopic observations to better constrain the nature of these stars.

In summary, the Galactic disk formed through three major phases: the primordial phase, the GSE merger phase, and the post-merger phase. These processes resulted in four main metal-poor disk populations: stars from the Galactic primordial disk, stars from the GSE progenitor, and the metal-poor extensions of both the so-called metal-rich, high-α and low-α disks. Other populations, such as stars formed during the quiescent phase, are present but constitute a smaller fraction.

Regarding the merger relics, Wukong's much lower metallicity than GSE suggests they stemmed from separate merger events.

Additionally, we identified three new substructures: ShangGu-1, ShangGu-2, and ShangGu-3. ShangGu-1 exhibits mild rotation ($V_\phi \approx -100\,\mathrm{km\,s^{-1}}$), strong vertical motion ($|V_z| \gtrsim 100\ \mathrm{km\,s^{-1}}$), and a positive correlation between orbital eccentricity and metallicity. ShangGu-2 shows thin disk-like rotation ($V_\phi \lesssim -200\ \mathrm{km\,s^{-1}}$) combined with high vertical velocity ($|V_z| \gtrsim 100\,\mathrm{km\,s^{-1}}$). Its stellar population displays two distinct sequences in the [Mg/Fe]–[Fe/H] diagram, similar to those observed in our kinematically cold disk sample. These features suggest that ShangGu-2 is composed of metal-poor disk stars that have been dynamically heated. ShangGu-3 has prograde orbit ($-240 \lesssim V_\phi \lesssim -130\,\mathrm{km\,s^{-1}}$) with significant radial movements ($|V_r| \approx 200\,\mathrm{km\,s^{-1}}$), and a mean stellar age of 13 Gyr. It can be divided into four subgroups based on orbital directions, and we propose it originated from the progenitor of GSE.

In conclusion, this work provides a detailed analysis of the Galactic metal-poor disk, proposing a metallicity floor of −1.5 for the primordial disk, and lowering the metallicity floors for high-α and low-α disks to −2.0 and −1.8, respectively. These findings are both novel and radical, and we look forward to future high-resolution observations to verify our views. Additionally, we argue that the merger between the progenitor of GSE and the Milky Way had a profound impact on the Milky Way, not only contributing to the Galaxy's halo but also significantly influencing the evolution of the Galactic disk. The value-added catalogs of chemical abundances and kinematic parameters for metal-poor stars used in this study are available at https://github.com/tubage886/VMP-StarsII.

Acknowledgments

This study is supported by the National Natural Science Foundation of China (NSFC, grant Nos. 11988101 and 12222305), and National Key R&D Program of China No. 2024YFA1611900. 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. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences. This work has made use of data from the APOGEE (SDSS-IV) surveys. 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. This work has also made use of data from the European Space Agency (ESA) mission Gaia, processed by the Gaia Data Processing and Analysis Consortium (DPAC). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

ORCID iDs

Xiao-Kun Hou (侯晓琨) https://orcid.org/0000-0001-9430-2306

References

Abohalima, A., & Frebel, A. 2018, ApJS, 238, 36
Aguado-Agelet, F., Massari, D., Monelli, M., et al. 2025, arXiv:2502.20436
Amarante, J. A. S., Debattista, V. P., Beraldo e Silva, L., Laporte, C. F. P., & Deg, N. 2022, ApJ, 937, 12
Anders, F., Khalatyan, A., Queiroz, A. B. A., et al. 2022, A&A, 658, A91
Baker, W. M., Tacchella, S., Johnson, B. D., et al. 2025, NatAs, 9, 141
Bellazzini, M., Massari, D., Ceccarelli, E., et al. 2024, A&A, 683, A136
Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611
Belokurov, V., & Kravtsov, A. 2022, MNRAS, 514, 689
Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, MNRAS, 494, 3880
Bennett, M., & Bovy, J. 2019, MNRAS, 482, 1417
Bensby, T., Feltzing, S., & Oey, M. S. 2014, A&A, 562, A71
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Bovy, J., Rix, H.-W., Liu, C., et al. 2012, ApJ, 753, 148
Campello, R. J. G. B., Moulavi, D., Zimek, A., & Sander, J. 2015, ACM Trans. Knowl. Discov. Data, 10, 1
Carter, C., Conroy, C., Zaritsky, D., et al. 2021, ApJ, 908, 208
Castelli, F., & Kurucz, R. L. 2003, IAUS, 210, A20
Chiappini, C., Anders, F., Rodrigues, T. S., et al. 2015, A&A, 576, L12
Ciucă, I., Kawata, D., Ting, Y.-S., et al. 2023, MNRAS: Letters, 528, L122
Dillamore, A. M., Belokurov, V., Evans, N. W., & Davies, E. Y. 2023, MNRAS, 524, 3596
Fernández-Alvar, E., Kordopatis, G., Hill, V., et al. 2021, MNRAS, 508, 1509
Fernández-Alvar, E., Kordopatis, G., Hill, V., et al. 2024, A&A, 685, A151
Feuillet, D. K., Sahlholdt, C. L., Feltzing, S., & Casagrande, L. 2021, MNRAS, 508, 1489
Fiorentin, P. R., Spagna, A., Lattanzi, M. G., & Cignoni, M. 2021, ApJL, 907, L16
Fujimoto, S., Ouchi, M., Kohno, K., et al. 2024, arXiv:2402.18543
Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, A&A, 674, A1
Gallart, C., Bernard, E. J., Brook, C. B., et al. 2019, NatAs, 3, 932
Gilmore, G., & Reid, N. 1983, MNRAS, 202, 1025
Gray, R. O., & Corbally, C. J. 1994, AJ, 107, 742
Han, D., Lee, Y. S., Kim, Y. K., & Beers, T. C. 2020, ApJ, 896, 14
Hayden, D., Bovy, J., Holtzman, J. A., et al. 2015, ApJ, 808, 132
Haywood, D., Matteo, P. D., Lehnert, M. D., Katz, D., & Gómez, A. 2013, A&A, 560, A109
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Natur, 563, 85
Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, Natur, 402, 53
Hong, J., Beers, T. C., Lee, Y. S., et al. 2024, ApJS, 273, 12
Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2023, MNRAS, 520, 5671
Hou, X., Zhao, G., & Li, H. 2024, MNRAS, 532, 1099
Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, Natur, 370, 194
Koppelman, H. H., Bos, R. O. Y., & Helmi, A. 2020, A&A, 642, L18
Koppelman, H. H., Helmi, A., Massari, D., Price-Whelan, A. M., & Starkenburg, T. K. 2019, A&A, 631, L9
Lane, J. M. M., Bovy, J., & Mackereth, J. T. 2023, MNRAS, 526, 1209
Lee, Y. S., Beers, T. C., An, D., et al. 2011, ApJ, 738, 187
Leung, H. W., & Bovy, J. 2019, MNRAS, 483, 3255
Li, C., Yuan, Z., Monari, G., et al. 2024, A&A, 690, A26
Li, H., Aoki, W., Matsuno, T., et al. 2022a, ApJ, 931, 147
Li, H., Tan, K., & Zhao, G. 2018, ApJS, 238, 16
Li, Z., Zhao, G., Chen, Y., Liang, X., & Zhao, J. 2022b, MNRAS, 517, 4875
Lian, Z., Du, M., Lu, S., et al. 2025, ApJL, 984, L48
Limberg, G., Ji, A. P., Naidu, R. P., et al. 2024, MNRAS, 530, 2512
Lindegren, L., Bastian, U., Biermann, M., et al. 2021, A&A, 649, A4
Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94
Malhan, K., Ibata, R. A., Sharma, S., et al. 2022, ApJ, 926, 107
Malhan, K., & Rix, H.-W. 2024, ApJ, 964, 104
Martig, M., Fouesneau, M., Rix, H.-W., et al. 2016, MNRAS, 456, 3655
Martig, M., Rix, H.-W., Silva Aguirre, V., et al. 2015, MNRAS, 451, 2230
Matteo, P. D., Haywood, M., Lehnert, M. D., et al. 2019, A&A, 632, A4
McMillan, P. J. 2017, MNRAS, 465, 76
Mowlavi, N., Holl, B., Lecoeur-Taïbi, I., et al. 2023, A&A, 674, A16
Myeong, G. C., Vasiliev, E., Iorio, G., Evans, N. W., & Belokurov, V. 2019, MNRAS, 488, 1235
Mészáros, S., & Prieto, C. A. 2013, MNRAS, 430, 3285
Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, ApJ, 901, 48
Necib, L., Ostdiek, B., Lisanti, M., et al. 2020, NatAs, 4, 1078
Nepal, S., Chiappini, C., Queiroz, A. B., et al. 2024, A&A, 688, A167
Nissen, P. E., & Schuster, W. J. 2010, A&A, 511, L10
Penoyre, Z., Belokurov, V., & Evans, N. W. 2022, MNRAS, 513, 2437
Re Fiorentin, P., Spagna, A., Lattanzi, M. G., Cignoni, M., & Vitali, S. 2024, ApJ, 977, 278
Rowland, L. E., Hodge, J., Bouwens, R., et al. 2024, MNRAS, 535, 2068
Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829
Sestito, F., Martin, N. F., Starkenburg, E., et al. 2020, MNRAS, 497, L7
Vasiliev, E. 2019, MNRAS, 482, 1525
Vasiliev, E., Belokurov, V., & Evans, N. W. 2022, ApJ, 926, 203

Vincenzo, F., Spitoni, E., Calura, F., et al. 2019, MNRAS, 487, L47
Viswanathan, A., Horta, D., Price-Whelan, A. M., & Starkenburg, E. 2024, arXiv:2411.12165
Wu, Y., Xiang, M., Bi, S., et al. 2018, MNRAS, 475, 3633
Xiang, M., & Rix, H.-W. 2022, Natur, 603, 599
Xiang, M., Rix, H.-W., Yang, H., et al. 2025, NatAs, 9, 101
Yong, D., Norris, J. E., Bessell, M. S., et al. 2013, ApJ, 762, 26
Yoshii, Y. 1982, PASJ, 34, 365
Yuan, Z., Chang, J., Beers, T. C., & Huang, Y. 2020a, ApJL, 898, L37
Yuan, Z., Myeong, G. C., Beers, T. C., et al. 2020b, ApJ, 891, 39
Zhang, R., Matsuno, T., Li, H., et al. 2024, ApJ, 966, 174
Zhao, G., Chen, Y.-Q., Shi, J.-R., et al. 2006, ChJAA, 6, 265
Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P., & Deng, L.-C. 2012, RAA, 12, 723