Element Abundance Analysis of the Metal-rich Stellar Halo and High-velocity Thick Disk in the Galaxy

Haifan Zhu1, Cuihua Du2, Yepeng Yan1, Jianrong Shi2, Jun Ma2, and Heidi Jo Newberg5

1 School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China
2 College of Astronomy and Space Sciences, University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China
3 Department of Astronomy, Beijing Normal University, Beijing 100875, People’s Republic of China
4 Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, People’s Republic of China
5 Department of Physics, Applied Physics and Astronomy, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Received 2021 March 18; revised 2021 April 28; accepted 2021 May 6; published 2021 June 28

Abstract

Based on the second Gaia data release (DR2) and APOGEE (DR16) spectroscopic surveys, we define two kinds of star samples: high-velocity thick disk (HVTD) with $v_\phi > 90$ km s$^{-1}$ and metal-rich stellar halo (MRSH) with $v_\phi < 90$ km s$^{-1}$. Due to high-resolution spectra data from APOGEE (DR16), we can accurately analyze the element abundance distribution of HVTD and MRSH. These element abundances constitute a multidimensional data space, and we introduce an algorithm method for processing multidimensional data to give the result of dimensionality reduction clustering. According to chemical property analysis, we derive that some HVTD stars could originate from the thin disk and some MRSH stars could originate from dwarf galaxies, but those stars that have similar chemical abundance characteristics in both samples may form in situ.

Unified Astronomy Thesaurus concepts: Galaxy disks (589); Galaxy kinematics (602); Galaxy abundances (574); Galaxy structure (622); Galaxy stellar halos (598)

1. Introduction

The halo and thick disk are basic components of the Galaxy, and the study of their kinematics and chemical abundance could provide important clues to the Galaxy’s formation and evolution history. Recent studies (Carollo et al. 2007, 2010; Chen et al. 2011; Beers et al. 2012; Kinman et al. 2012) have shown that the Galactic halo comprises at least two stellar populations: inner halo and outer halo. They have different kinematics, spatial distribution, and chemical composition (Carollo et al. 2007, 2012; Liu et al. 2018; Conroy et al. 2019; Bird et al. 2020). For example, the inner halo is mainly distributed at distances up to 15–20 kpc from the Galactic center, and the mean metallicity of the inner halo ranges from [Fe/H] $\sim$ −1.2 dex to −1.7 dex. The outer halo is mainly distributed at distances up to 15–20 kpc from the Galactic center, and the mean metallicity of the outer halo ranges from [Fe/H] $\sim$ −1.9 dex to −2.3 dex (e.g., Carollo et al. 2007; An et al. 2013, 2015; Gu et al. 2015, 2016, 2019; Zuo et al. 2017; Liu et al. 2018). Studies of the detailed chemical abundances and ages of halo stars have sought to place further constraints on the structure and formation of the Galactic halo (Naidu et al. 2020; Şahin & Bilir 2020). Many studies have shown that there are two chemically distinct stellar populations: an older high-\(\alpha\) halo population and a younger low-\(\alpha\) halo population (e.g., Nissen & Schuster 2010, 2011; Bergemann et al. 2017; Hayes et al. 2018).

In particular, since the second release of Gaia (Gaia DR2), a significant number of works have revealed an even more complex but detailed picture of the Galactic halo. For example, Belokurov et al. (2018) showed the velocity ellipsoid becomes strongly anisotropic for the halo stars with $-1.7 < [\text{Fe/H}] < -1.0$ dex and local velocity distribution appears highly stretched in the radial direction, taking a sausage-like shape, and they suggested that such orbital configurations could show that most of the inner halo stars should be dominated by stars accreted from an ancient massive merger event. This merger event is referred to as the Gaia–Sausage merger (e.g., Deason et al. 2018; Myeong et al. 2018, 2019; Lancaster et al. 2019). Helmi et al. (2018) also demonstrated that the inner halo is dominated by debris from the merger of a dwarf galaxy that occurred 10 Gyr ago, and the dwarf galaxy is referred to as Gaia-Enceladus. Several other works also found new chemo-dynamical properties of the stellar halo, that were unknown before Gaia (e.g., Bonaca et al. 2017, 2020; Koppelman et al. 2019; Belokurov et al. 2020; Carollo & Chiba 2021; Naidu et al. 2020; Yan et al. 2020; Yuan et al. 2020). All these observation results imply that accreted stars from satellite galaxies have been suggested to be dominant inner halo components.

The Galactic disks contain a substantial fraction of their baryonic matter and angular momentum, and much of the evolutionary activity. The formation and evolution of disks are therefore very important (van der Kruit & Freeman 2011). The basic components of the Galactic disk are the thin-disk and thick-disk populations. The two components differ not only in their spatial distribution profiles but also in their kinematics, age, and metallicity (van der Kruit & Freeman 2011; Ivezić et al. 2012; Xiang et al. 2015, 2017; Jing et al. 2016; Peng et al. 2018; Gandhi & Ness 2019; Han et al. 2020; Wu et al. 2021). Compared to stars in the thin disk, stars in the thick disk are older, kinematically hotter, metal-poor, and enhanced in \(\alpha\)-elements (e.g., Chiba & Beers 2000; Prochaska et al. 2000; Bensby et al. 2005). In addition, there is evidence for the additional presence of a metal-weak thick-disk (MWTD) population that is rotationally supported, but extending to lower metallicity stars than the canonical thick disk (Morrison et al. 1990; Beers & Sommerlarsen 1995; Chiba & Beers 2000; Beers et al. 2002, 2014; Carollo et al. 2019; Yan et al. 2019). A key question now is whether these disk components can also be used to account for the chemical and kinematic measurements for the same stars. It turns out that, even with the new data collected, it is not easy to answer this question.
Despite the past three decades of thick-disk studies, there is still no consensus on models for the formation and evolution of thick disks. The proposed simulations of thick disk formation can be generally divided into four groups: (a) accretion from disrupted satellite galaxies (Abadi et al. 2003), (b) heating of the preexisting thin disk due to minor mergers (Quinn et al. 1993), (c) in situ triggered star formation during and after a gas-rich merger (Brook et al. 2004; Sales et al. 2009), and (d) in situ formation through radial migration (Sellwood & Binney 2002; Schöenrich & Binney 2009a, 2009b; Schöenrich et al. 2011).

These currently discussed models of formation mechanisms for the thick disk predict various trends between the kinematics properties and the metallicity of disk stars. For example, some simulations of accretion from disrupted satellite galaxies can help to explain why there are so many old stars in circular orbits in the outskirts of galaxies and why the specific angular momentum and radial extent of the thick disk and thin disk are comparable although their ages are significantly different (Abadi et al. 2003). Furthermore, it provides an explanation for the dynamical and evolutionary distinction between the thick and thin disk components: the thick disk is mostly tidal debris from disrupted satellites, while the young thin disk consists mostly of stars formed in situ after the merging activity abates (Abadi et al. 2003). Some simulations of thick disk formation via the accretion of satellites onto a preexisting thin disk, showed that mergers with 10%–20% mass of the host lead to the formation of a thick disk (Villalobos & Helmi 2008). Using a different numerical implementation of the radial migration scenario, Schöenrich and Binney (Schöenrich & Binney 2009a, 2009b; Schöenrich et al. 2011) showed radial migration plays an important role in the local solar neighborhood. But Liu & van de Ven (2012) argued that the old hot metal-poor high-[O/Fe] stars might have formed through early-on gas-rich mergers instead of radial migration. Feuillet et al. (2019) studied the spatial changes in the [α/M]–age relation and [M/H]–age relations of the Milky Way disk. The results are important constraints to Galactic simulations and chemical evolution models.

Kordopatis et al. (2020) used the chemo-dynamical method to give conclusions related to the formation of thick disks. They think the key to understanding the effect of the past accretions on the properties of the thick disk is the super-solar metallicity counter-rotating population. This population will provide us with a reliable sample of locally born retrograde stars in order to determine the exact time and weigh this merger (Grand et al. 2020).

Previous studies showed that most halo stars are metal-poor. With the release of more survey data, some recent works have revealed a large number of metal-rich halo stars ([Fe/H] > −1 dex) (e.g., Nissen & Schuster 2010, 2011; Bonaca et al. 2017; Posti et al. 2018; Yan et al. 2020). Bonaca et al. (2017) use the first Gaia data release, the Radial Velocity Experiment (RAVE; Steinmetz et al. 2006) and the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Eisenstein et al. 2011) to find that half of their halo sample is comprised of stars with [Fe/H] > −1 dex, and they proposed that metal-rich halo stars in the solar neighborhood actually formed in situ within the Galactic disk. It is possible that these stars have undergone radial migration that caused changes in their orbits. Yan et al. (2020) use the second Gaia data release (DR2), combined with the ongoing Large Sky Area Multi-Object Fiber Spectroscopic Telescope survey (LAMOST, also called Guoshoujing Telescope; Zhao et al. 2012) and APOGEE to show that there exists a high-velocity thick disk (HVTD) with \( v_\phi > 90 \text{ km s}^{-1} \) and a metal-rich stellar halo (MRSH) with \( v_\phi < 90 \text{ km s}^{-1} \) in the Galaxy. But the details of how HVTD and MRSH form are not yet understood.

In this work, we use the APOGEE DR16 and Gaia DR2 data to obtain samples of HVTD and MRSH to study the detailed abundance characteristics and give their possible origins. This paper is organized as follows: Section 2 introduces the data selection and the method of distance determination, and gives the division of HVTD and MRSH. Section 3 introduces the dimensionality reduction algorithm to present the element abundance analysis. In Section 4, we analyze the results and compare them with some previous works, and give the possible origin of HVTD and MRSH. The summary and conclusions are given in Section 5.

2. Data

2.1. Data Selection

The Apache Point Observatory Galactic Evolution Experiment (APOGEE), part of the Sloan Digital Sky Survey III, is a near-infrared \((H\text{-band})\) and high-resolution \((R \approx 22,500)\) spectroscopic survey (Zasowski et al. 2013). In this study, we use the data from the sixteenth SDSS Data Release (DR16), which provides accurate radial velocities, stellar parameters, and chemical abundances of about 26 chemical species for 430,000 stars covering both the northern and southern skies (Jönsson et al. 2020).

The second Gaia data release, Gaia DR2, provides high-precision positions, parallaxes, and proper motions for 1.3 billion sources brighter than \( G \sim 21 \text{ mag} \). More detailed information about Gaia can be found in Gaia Collaboration et al. (2018a, 2018b). In this study, we get the high-resolution sample stars by cross-matching between APOGEE DR16 and the Gaia DR2 catalog. Stellar parameters (metallicity abundances, radial velocity, effective temperature, and surface gravity) are from APOGEE DR16. Other necessary parameters (position, proper motion, and parallax) are from the Gaia DR2 catalog.

In order to obtain reliable results, we utilize the following selection criteria:

1. parallax uncertainties < 20%,
2. proper motion error < 0.2 mas yr\(^{-1}\),
3. radial velocity uncertainties < 10 km s\(^{-1}\),
4. [Fe/H] error < 0.2 dex, and
5. S/N > 20 in the \( G \) band.

2.2. Distance and Velocity Determination

Bailer-Jones (2015) discussed that the inversion of the parallax to obtain distance is not appropriate when the relative parallax error is more than 20%. We use the Bayesian approach (Bailer-Jones 2015; Astraatmadja & Bailer-Jones 2016a, 2016b; Luri et al. 2018; Yan et al. 2019) to derive stellar distance.

According to the Bayesian formula, the posterior probability can be written as follows:

\[
P(\theta|x) \propto P(x|\theta)P(\theta)
\]

\[
= \exp \left[ - \frac{1}{2} (x - m(\theta))\Sigma^{-1}(x - m(\theta)) \right] P(\theta),
\]

\( P(\theta) \) is the prior probability distribution of the parameter \( \theta \), \( P(x|\theta) \) is the likelihood function, and \( x \) is the observable data.
where \( P(x|\theta) \) is the likelihood, the symbol \( \theta \) represents the parameter vectors, which consists of heliocentric distance \( (d) \), tangential speed \( (v) \), and travel direction \( (\phi, \text{ increasing counterclockwise from north}) \), written as
\[
\theta = (d, v, \phi)^T.
\]
x is the observed data vector, which consists of the parallax \( (\pi) \), proper motion in R.A. \( (\mu_\alpha) \), and decl. \( (\mu_\delta) \), written as
\[
x = (\pi, \mu_\alpha, \mu_\delta)^T.
\]
The likelihood probability is a multidimensional Gaussian distribution centered on \( m \) as we can see in formula (1). \( \Sigma \) is covariance matrix
\[
\Sigma = \begin{pmatrix} 
\sigma_\pi^2 & \sigma_\pi \sigma_\mu \rho(\pi, \mu_\alpha) & \sigma_\pi \sigma_\mu \rho(\pi, \mu_\delta) \\
\sigma_\mu \sigma_\pi \rho(\mu_\alpha, \mu_\delta) & \sigma_\mu^2 & \sigma_\mu \rho(\mu_\alpha, \mu_\delta) \\
\sigma_\mu \sigma_\pi \rho(\mu_\alpha, \mu_\delta) & \sigma_\mu \rho(\mu_\alpha, \mu_\delta) & \sigma_\mu^2 
\end{pmatrix},
\]
where \( \rho(i, j) \) denotes the correlation coefficient between \( i \) and \( j \) and \( \sigma_\phi \) denotes the standard deviation of parameters \( k, m \) represents a set of theoretical values predicted by our model, written as
\[
m = \left( \frac{10^3}{d}, \frac{10^3 \sin \phi}{d}, \frac{10^3 \cos \phi}{d} \right)^T,
\]
where \( c = (\text{pc mas yr}^{-1})/(4.74 \text{ km s}^{-1}) \).

\( P(\theta) \) is the prior distribution given by Luri et al. (2018)
\[
P(\theta) = P(d)P(v)P(\phi),
\]
with
\[
P(d) \propto \begin{cases} 
d^2 e^{-d/L(a,b)} & d > 0 \\
0 & d \leq 0
\end{cases}
\]
and
\[
P(v) \propto \begin{cases} 
\left( v/v_{\text{max}} \right)^{\alpha-1} \left( 1 - \frac{v}{v_{\text{max}}} \right)^{\beta-1} & 0 \leq v \leq v_{\text{max}} \\
0 & \text{otherwise}
\end{cases}
\]
and
\[
P(\phi) \propto \frac{1}{2\pi}.
\]
The prior distribution of distance is the exponentially decreasing and \( L(a, b) \) is length scale of Galactic longitude- and latitude-dependent (Bailer-Jones 2015, 2018). Here we take \( \alpha = 2, \beta = 3 \), and \( v_{\text{max}} = 750 \text{ km s}^{-1} \). Then we get the posterior distribution through the Markov Chain Monte Carlo (MCMC) sampler EMCEE (Foreman-Mackey et al. 2013). We run each chain using 100 walkers and 100 steps, for a total of 10,000 random samples drawn from the posterior distribution. Finally, we adopt the data to calculate the kinematic parameters by using Astropy (Astropy Collaboration et al. 2013, 2018) and galpy (Bovy 2015). Astropy provides many functions to calculate the parameters and coordinate transformations. Galpy is a Python package for galactic dynamics and it supports orbit integration in various potentials.

When calculating with galpy, we choose the potential model from MWPotential2014 (Bovy 2015), the distance from the Sun to the Galactic center is 8.2 kpc and the height above the plane is about 15 pc (Bland-Hawthorn \\& Gerhard 2016). When deriving velocity, we adopt \( v_{LSR} = 223.8 \text{ km s}^{-1} \) (McMillan 2017). The velocity of the Sun with respect to the the local standard of rest (LSR) is \((U, V, W) = (10.0, 11.0, 7.0) \text{ km s}^{-1}\) (Bland-Hawthorn \\& Gerhard 2016). Using the above, we can get the coordinate position, speed, and orbit parameters.

2.3. Determination of HVTD and MRSH

In this study, we first get 4115 high-velocity sample stars with \( v_{tot} > 220 \text{ km s}^{-1} \). The sample star distribution in the Toomre diagram is presented in Figure 1. In order to determine the HVTD component and the MRSH component, we also give the rotation velocity distribution of the sample stars with different metallicity selection in Figure 2. According to Yan et al. (2020), two-dimensional Gaussian fitting implied the existence of the HVTD and MRSH in the high-velocity sample, and they gave a definition of the division of the two components: HVTD with \( v_\phi > 90 \text{ km s}^{-1} \) and \([\text{Fe/H}] > -1 \text{ dex} \), MRSH with \( v_\phi < 90 \text{ km s}^{-1} \) and \([\text{Fe/H}] > -1 \text{ dex} \). Figure 3 gives the Toomre diagram. We can see that the distribution of HVTD and MRSH samples is quite different in the Toomre diagram. There is a small overlap near the dividing line.

The standard for division of two samples is derived from the characteristics of the bimodal distribution of the overall sample. It is not easy to obtain a clear boundary for the sample distribution. In this work, we directly adopted the previous division method.

The rotational velocity from 75 to 125 \text{ km s}^{-1} between the two peaks of the velocity distribution is suitable as the division as Figure 2. Different dividing lines will affect the number of samples of MRSH and HVTD, but there are only 204 sample stars in this range for \([\text{Fe/H}] > -1 \text{ dex} \) and 176 sample stars for \([\text{Fe/H}] > -0.8 \text{ dex} \). Furthermore, there are 65 stars in interval of 
\( v_\phi \in [75 \text{ km s}^{-1}, 90 \text{ km s}^{-1}]) \), 84 stars in interval \([90 \text{ km s}^{-1}, 110 \text{ km s}^{-1}]) \), and 55 stars in interval \([110 \text{ km s}^{-1}, 125 \text{ km s}^{-1}]) \) with \([\text{Fe/H}] > -1 \text{ dex} \). For \([\text{Fe/H}] > -0.8 \text{ dex} \), the sample is even smaller. So even if changing the classification standard in this range, only classification of a
very small sample of stars will be changed, and these changes will not have a substantial effect on the subsequent study.

2.4. Metallicity Distribution and Gradient

According to previous research, we use kinematic information to divide the sample stars into HVTD and MRSH. Figure 4 shows the metallicity distribution in the vertical distance \(|z|\) of the two samples. We note that the distribution of HVTD is relatively concentrated, mainly in the range <3 kpc, and the distribution of MRSH is relatively uniform.

We use maximum likelihood estimation to derive the metallicity gradient in the \(|z|\) direction. For a given sample with total number of stars \(N\), the data points conform to the linear form \([\text{Fe}/\text{H}] = kz_i + b\). In this model, if position \(z_i\), \([\text{Fe}/\text{H}]\) uncertainty \(\sigma_i\), slope \(k\) (also \(d[\text{Fe}/\text{H}]/dz\)), and intercept \(b\) are given, the distribution \(P([\text{Fe}/\text{H}]|z_i, \sigma_i, k, b)\) for \([\text{Fe}/\text{H}]\) is

\[
P([\text{Fe}/\text{H}]|z_i, \sigma_i, k, b) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{(\text{[Fe}/\text{H}] - kz_i - b)^2}{2\sigma_i^2}\right)
\]

and the likelihood \(L\) is

\[
L = \prod_{i=1}^{N} P([\text{Fe}/\text{H}]|z_i, \sigma_i, k, b).
\]

We aim to derive the parameter value \((k, b)\) that maximizes the likelihood function. Here, we use the Bayesian approach to derive the uncertainty of the metallicity vertical gradient. The metallicity vertical gradient of the HVTD is \(d[\text{Fe}/\text{H}]/dz|_{z=3 \text{ kpc}} = -0.11 \pm 0.0004 \text{ dex kpc}^{-1}\) as shown in the bottom left panel of Figure 4. The HVTD is mainly distributed in \(z < 3 \text{ kpc}\), we calculated the metallicity vertical gradient in this range, \(d[\text{Fe}/\text{H}]/dz|_{z=3 \text{ kpc}} = -0.16 \pm 0.0008 \text{ dex kpc}^{-1}\).

This result is roughly consistent with some previous results of the thick disk. For example, Chen et al. (2011) used RHB stars in the range of 0.5–3 kpc to give the gradient for the thick disk \(d[\text{Fe}/\text{H}]/dz|_{z=3 \text{ kpc}} = -0.12 \pm 0.01 \text{ dex kpc}^{-1}\). Bilir et al. (2012) used red clump stars to derive that the vertical metallicity gradient for the thick disk is close to zero. Mikolaitis et al. (2014) gave \(d[\text{Fe}/\text{H}]/dz|_{z=3 \text{ kpc}} = -0.072 \pm 0.006 \text{ dex kpc}^{-1}\). Li & Zhao (2017) gave the result \(d[\text{Fe}/\text{H}]/dz|_{z=3 \text{ kpc}} = -0.164 \pm 0.010 \text{ dex kpc}^{-1}\). Tunçel Güçtekin et al. (2019) reported \(d[\text{Fe}/\text{H}]/dz|_{z=3 \text{ kpc}} = -0.164 \pm 0.014 \text{ dex kpc}^{-1}\) with the range of 6 < \(R\) < 10 dex kpc\(^{-1}\) and 2 < \(|z|\) < 5 dex kpc\(^{-1}\). Yan et al. (2019) reported \(d[\text{Fe}/\text{H}]/dz|_{z=3 \text{ kpc}} = -0.074 \pm 0.0009 \text{ dex kpc}^{-1}\) for the thick disk. According to these studies, the gradient of our HVTD sample is very similar to the thick disk, which implies the HVTD sample has very obvious characteristics of the thick disk.

In the top right panel of Figure 4, the metallicity vertical gradient of the MRSH is \(d[\text{Fe}/\text{H}]/dz|_{z=3 \text{ kpc}} = -0.02 \pm 0.0003 \text{ dex kpc}^{-1}\), which is a very flat gradient. This result is roughly consistent with previous results of the
halo, such as those of Tunçel Güçtekin et al. (2019), who reported the vertical metallicity gradients \( d[\text{Fe}/\text{H}]/dz = -0.023 \pm 0.006 \text{dex kpc}^{-1} \) in the interval \( 6 < R < 10 \text{kpc} \). Peng et al. (2012, 2013) derived that the vertical gradient \( d[\text{Fe}/\text{H}]/dz = -0.05 \pm 0.04 \text{dex kpc}^{-1} \) and \( d[\text{Fe}/\text{H}]/dz = -0.03 \pm 0.02 \text{dex kpc}^{-1} \) for distance \( 5 < z < 14 \text{kpc} \), and they concluded that there is little or no gradient in the halo. Other studies also report the gradient close to zero at larger galactocentric distances for the halo.

3. The Chemical Element Abundance Analysis

3.1. Stellar Elements Abundance Distribution

APOGEE DR16 provides a large number of chemical element abundances of stars. Jönsson et al. (2020) point out that some parameters are not accurate enough and should be selected according to the effective temperature. We selected 11 elements and \( \alpha \)-elements as a function of \([\text{Fe}/\text{H}]\) to study the chemical properties and origins of sample stars. These element abundance distributions are given in Figure 5. We also derived the important parameter \([\alpha/\text{Fe}]\), where \( \alpha \) refers to the average abundance of Mg, Si, Ca, and Ti. The \( \alpha \)-elements are often used as an indicator of timescale in studying the history of star formation. This is mainly due to the \( \alpha \)-elements produced in the type II supernovae explosion on a relatively short timescale (\( 10^7 \) yr), while iron is produced on a much longer timescale (\( 10^9 \) yr) by a type Ia supernovae explosion. Thus the analysis of the \( \alpha \)-elements is particularly important.

C, N, and O abundances are also important for probing the star formation history. For example, Masseron & Gilmore (2015) used the variations of carbon and nitrogen abundances of stars in the thin disk and thick disk of the Galaxy to gather information on the relative ages of stars. Martig et al. (2016) used carbon and nitrogen to infer the mass of the stars. Although the derived ages are with rms errors of 40%, these element abundances are still important to study the formation history.

As shown in each panel of Figure 5, there is a large number of overlapping areas in these distributions. In areas with rich metallicity, most sample stars belong to the HVTD. The MRSH sample stars are mainly distributed in the range with \([\text{Fe}/\text{H}] < -0.2 \) dex. In addition, from the distribution in the \([X/\text{Fe}]\) versus \([\text{Fe}/\text{H}]\) (\( X \) refers to the selected elements), we noted that some chemical elements display centralized distribution, like N, Ca, Si, and Mn for the HVTD and MRSH, as well as C, Al, and Co for the HVTD, but some elements appear to have scattered distribution, like O, Mg, Ti, and Ni for the HVTD and MRSH.

3.2. Data Dimensionality Reduction

We selected 11 elements (C, N, O, Mg, Al, Si, Ca, Ti, Mn, Co, and Ni) that are accurate in APOGEE DR16. These elements with \([\text{Fe}/\text{H}]\) constitute a high-dimensional data space, and comprehensive understanding and interpreting of multidimensional abundance distribution is not easy. We try to reduce dimensionality of multidimensional data. The common data dimensionality reduction method is the principal component analysis (PCA) algorithm, which is a linear algorithm. But the results of this algorithm are not reliable when the data has complex connections. Thus for highly correlated data sets (like chemical abundance distribution in this work), a nonlinear dimensionality reduction processing algorithm is necessary.

Here, we use the t-distributed stochastic neighbor embedding method (t-SNE; Van der Maaten & Hinton 2008), which is a nonlinear machine-learning algorithm that can reduce N-dimensional data to a 2D plane. In this work, each star has 11 element abundances and metallicity \([\text{Fe}/\text{H}]\). These elements constitute a high-dimensional data space (Fe, C, N, O, Mg, Al, Si, Ca, Ti, Mn, Co, and Ni), each element is a dimension. The t-SNE algorithm can reduce each sample star with 12-dimensional data characteristics to two dimensions. It represents each star with a set of coordinates and uses a set of two-dimensional plane coordinates to make the results easier to visualize. Data points (stars) with similar characteristics on this plane will gather together.

The principle of the algorithm is as follows: given \( N \) high-dimensional data points \( x_1, \ldots, x_N \) (here represent abundances), t-SNE first computes probabilities \( p_{ij} \) that are proportional to the similarity of objects \( x_i \) and \( x_j \):

\[
p_{ji} = \frac{\exp(-||x_i - x_j||^2/2\sigma_i^2)}{\sum_{k=1}^{N} \exp(-||x_i - x_k||^2/2\sigma_i^2)},
\]

and defines symmetrized similarity:

\[
p_{ij} = \frac{p_{ji} + p_{ij}}{2N},
\]

t-SNE aims to learn a \( d \)-dimensional map \( y_1, \ldots, y_N \) (with \( y_i \in \mathbb{R}^d \)) that reflects the similarities \( p_{ij} \) as well as possible.

For the low-dimensional counterparts \( y_i \) and \( y_j \) of the high-dimensional data points \( x_i \) and \( x_j \), it is possible to compute a similar conditional probability, which is denoted by \( q_{ij} \). It reflects the similarities between two points \( y_i \) and \( y_j \) in the low-dimensional map. For the low-dimensional, we choose the variance of the Gaussian to \( 1/\sqrt{2} \)

\[
q_{ij} = \frac{\exp(-||y_i - y_j||^2)}{\sum_{k=1}^{N} \exp(-||y_i - y_k||^2)}.
\]

In the case of symmetry,

\[
q_{ij} = \frac{\exp(-||y_i - y_j||^2)}{\sum_{k=1}^{N} \exp(-||y_k - y_j||^2)}.
\]

In t-SNE algorithm, we employ a Student t-distribution with one degree of freedom as the heavy-tailed distribution in the low-dimensional map. Using this distribution, the joint probabilities \( q_{ij} \) are defined as

\[
q_{ij} = \frac{(1 + ||y_i - y_j||^2)^{-1}}{\sum_{k=1}^{N} (1 + ||y_k - y_j||^2)^{-1}}.
\]

The locations of the points \( y_i \) in the map are determined by minimizing the (nonsymmetric) Kullback–Leibler divergence of the distribution \( P \) from the distribution \( Q \), that is:

\[
\text{KL}(P||Q) = \sum_{i=1}^{N} p_{ij} \log \frac{p_{ij}}{q_{ij}}.
\]

The minimization of the Kullback–Leibler divergence is performed by using gradient descent. The result of this optimization is a 2D plane that reflects the similarities between the high-dimensional data points (Van der Maaten & Hinton 2008).
We use scikit-learn (Pedregosa et al. 2011) packages in python to implement this algorithm. The method has one main parameter perplexity, \( p \), which governs the bandwidth of the Gaussian kernels \( \sigma \) and appears in the similarities \( p_{ij} \); in this work we take \( p = 50 \). Another parameter is learning rate, which is set to 1, and the early exaggeration parameter is not important, so we take the default value, and the random state in scikit-learn packages is set to 1.

We give the final result of this algorithm in Figure 6, the HVTD sample is marked by blue dots and the metal-rich stellar halo (MRSH) is marked by the red dots.

Figure 5. Element abundance distribution vs. [Fe/H]. We removed some obviously inaccurate data. The high-velocity thick disk (HVTD) is marked by the blue dots and the metal-rich stellar halo (MRSH) is marked by the red dots.

The entire sample is divided into three areas in Figure 6:
1. area A in HVTD,
2. area B in MRSH, and
3. overlapping area of HVTD and MRSH.

Areas A and B denote two independent parts in the dimensionality reduction result, which mainly contains HVTD and MRSH, respectively.

4. Discussion about the Origins of MRSH and HVTD

Some previous works suggested the possible origin of those metal-rich halo stars (e.g., Bonaca et al. 2017; Belokurov et al. 2018; Haywood et al. 2018; Gallart et al. 2019; Yan et al. 2020). For example, Bonaca et al. (2017) proposed that metal-rich halo stars within 3 kpc from the Sun may have formed
in situ, rather than having been accreted from satellite systems, and these metal-rich halo stars have likely undergone substantial radial migration or heating. Yan et al. (2020) proposed that for the young stars (<9 Gyr), their formation may not be affected by the Gaia–Sausage merger. The MRSH stars were likely born in situ rather than accreted from the Gaia–Sausage merger. But for the old stars formed in situ (>9 Gyr), the Gaia–Sausage merger event may have a major effect on their formation. The MRSH stars may form in an old proto-disk, possibly dynamically heated by the Gaia–Sausage merger, and subsequently be kicked out to the halo.

In this study, based on the dimensionality reduction results, we found there are obvious regional chemical characteristics, as shown in Figure 6. The stars in zone B are mainly MRSH, the stellar metallicity in this region lies between −1 and −0.6 dex. But it still has lower α-element abundances. The [Mg/Fe], [Al/Fe], and [α/Fe] here are obviously much smaller than the overlap area. This is very similar to the chemical characteristics of dwarf galaxies. Letarte et al. (2010) showed such element abundance characteristics for field stars of dwarf galaxies (Fornax). Hawkins et al. (2015) also showed in more detail the abundance patterns of the elements as a function of metallicity for the disk stars, halo stars, and dwarf galaxy (Fornax), which clearly implied lower α-element with [Fe/H] from −1.0 to −0.5 dex for dwarf galaxies. Fattahi et al. (2019) revealed the origin of the metal-rich halo stars ([Fe/H] ~ 1 dex) with highly eccentric orbits (high orbital anisotropy, β > 0.8), by tracing their stars back to the epoch of accretion and showed that these stars could come from a single dwarf galaxy. Those sample stars in area B also have extremely high eccentricities in the last row of the middle column of Figure 7. Therefore, we conclude that those stars in area B for MRSH are chemically and kinematically consistent with dwarf galaxies.

The overlapping area in HVTD and MRSH shows the same element abundances, which means that they have a similar origin. The possible origin of these regions is in situ. After undergoing some processes, some stars become halo stars, and other stars maintain the kinematic characteristics of disk stars. The in situ population can contain stars formed in the initial gas collapse (Samland & Gerhard 2003). Cooper et al. (2015) gave two distinct origins of in situ halo stars: gas that has been stripped from satellite galaxies by tides and ram pressure, and gas that is incorporated directly into the smooth halo of the main galaxy by cosmological infall and supernova-driven outflow from the central galaxy.

In addition, the overlapping area could also include some stars that originated from within the Milky Way disk. Bonaca et al. (2017) proposed two disk heating mechanisms to form the metal-rich halo stars: runaway stars and radial migration. Runaway stars are young stars that are formed in the disk and then ejected from their birthplace. Based on the discussion of metallicity and spectral type in Bonaca et al. (2017), combined with the characteristics of our sample, we can conclude that runaway stars are a minor component of the observed metal-rich stellar halo.

Radial migration has been recognized as an important component in explaining numerous observations such as the spread in the age–metallicity relation (e.g., Sellwood & Binney 2002; Roškar et al. 2008; Minchev & Famaey 2010) and its abundance patterns (Schöenrich & Binney 2009a). El-Badry et al. (2016) proposed that stars experienced significant radial migration by two related processes. First, some stars are formed during gas outflows, so that their initial orbits can be eccentric and have large apocenters. Second, the gravitational potential of the galaxy will have strong fluctuations under the combination of the inflow gas accretion and gas outflow. This fluctuation will affect the stellar orbits, which will eventually become heated to a more isotropic distribution. Because those stars within the overlapping area have similar chemical compositions and orbital eccentricities, it indicates that they have the same origin and may have undergone similar processes. We consider radial migration as a possible formation mechanism for the stars in the overlapping area.

For the stars in area A, as shown in Figure 7, their characteristics are very obvious (e.g., high [Fe/H], low eccentricity, and low α-element). The stars in area A are mainly HVTD, and the stellar metallicity in this region lies between −0.2 and 0.4 dex. The stars in this region have characteristics of thin disk stars. According to the proposed formation mechanism of the thick disk, we infer that these stars could originate from the heating of a preexisting thin disk.

The stellar ages also provide important clues to probe the possible origins of the population. But it is difficult to obtain the accurate ages of stars. Here, we use the age range of the stars to discuss the potential origins of MRSH and HVTD. Ages of our sample stars are obtained by cross-matching with the catalog of Sanders18 (Sanders & Das 2018). Sanders & Das (2018) only estimated masses and ages for the stars metal-richer than −1.5 dex and the maximum age isochrone considered is 12.6 Gyr. The age distributions in MRSH and HVTD are shown in Figure 8, which imply that both MRSH and HVTD contain young stars (<9 Gyr) and old stars (<9 Gyr). The young MRSH stars were possibly born in situ, but the Gaia–Sausage merger event may have an important effect on the old stars. The HVTD stars also form in an old proto-disk, but these stars may be affected by the Gaia–Sausage merger event so much less than MRSH that some properties of the thick disk remain (Yan et al. 2020).

According to the above analysis of chemical element abundance and kinematics and age, we consider that HVTD and MRSH each have two origins. The MRSH stars are formed from the accretion of smaller galaxies and in situ formation. The HVTD stars could originate from the heating of a preexisting thin disk and in situ formation.
5. Summary and Conclusions

Based on the data of APOGEE DR16 and Gaia DR2, we first obtained high-velocity sample stars ($v_{\text{tot}} > 220$ km s$^{-1}$). According to the distribution of rotational velocity and metallicity ([Fe/H] $>-1$ dex) of these sample stars, we divided them into HVTD ($v_{\phi} > 90$ km s$^{-1}$) and MRSH ($v_{\phi} < 90$ km s$^{-1}$) and studied their element abundance distribution to confirm their origins.

We found that MRSH has relatively low metallicity ($-1.0$ dex, $-0.4$ dex), but HVTD not only has a relatively small number in the lower metallicity [Fe/H], but also has more metal-rich stars in metallicity [Fe/H] ($-0.4$ dex, $0.4$ dex). Furthermore,
we use the maximum likelihood method to estimate the vertical metallicity gradient of two samples. The metallicity vertical gradient of HVTD is $d\log(\text{Fe/H})/dz = -0.16 \pm 0.0008$ dex kpc$^{-1}$ in $z < 3$ kpc, and the metallicity vertical gradient of MRSH is $d\log(\text{Fe/H})/dz = -0.02 \pm 0.0003$ dex kpc$^{-1}$, which is a very flat gradient.

The element abundance of stars can provide important clues to probe their origin. In this study, we selected 11 elements and α-element as a function of [Fe/H] to study the chemical properties and origins of sample stars. In order to comprehensively consider various element abundance indicators, we performed dimensionality reduction processing on the data according to the t-SNE algorithm. It is clear that the sample is divided into three areas: area A in HVTD, area B in MRSH, and the overlapping area of HVTD and MRSH, which can be related to some formation mechanisms. From the chemical element abundance and kinematics and age, we can conclude that HVTD and MRSH each have two origins. The MRSH stars are formed from the accretion of smaller galaxies and in situ formation. The HVTD stars could originate from the heating of a preexisting thin disk and in situ formation.

We thank the referee for insightful comments and suggestions, which have improved the paper significantly. This work was supported by the National Natural Foundation of China (NSFC No. 11973042, No. 11973052, and No. 11873053). It was also supported by the Fundamental Research Funds for the Central Universities and the National Key R&D Program of China No. 2019YFA0405501. H.J.N. acknowledges funding from US NSF grant AST-1908653. This project was developed in part at the 2016 NYC Gaia Sprint, hosted by the Center for Computational Astrophysics at the Simons Foundation in New York City.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Aeronautics and Space Administration Observatory, National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

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

**ORCID iDs**

Cuihua Du @ https://orcid.org/0000-0002-3954-617X
Jianrong Shi @ https://orcid.org/0000-0002-0349-7839
Jun Ma @ https://orcid.org/0000-0001-6329-6644
Heidi Jo Newberg @ https://orcid.org/0000-0001-8348-0983

**References**

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, ApJ, 597, 21
An, D., Beers, T. C., Johnson, J. A., et al. 2013, ApJ, 763, 65
An, D., Beers, T. C., Santucci, R. M., et al. 2015, ApJL, 813, L28
Astraatmadja, T. L., & Bailer-Jones, C. A. L. 2016a, ApJ, 832, 137
Astraatmadja, T. L., & Bailer-Jones, C. A. L. 2016b, ApJ, 833, 119
Astropy Collaboration, Price-Whelan, A. M., Sipócz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bailer-Jones, C. A. L. 2015, PASP, 127, 994
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58
Beers, T. C., Colollo, D., Ivezic, Z., et al. 2012, ApJ, 746, 34
Beers, T. C., Drilling, J. S., Rossi, S., et al. 2002, AJ, 124, 931
Beers, T. C., Norris, J. E., Placco, V. M., et al. 2014, ApJ, 794, 58
Beers, T. C., & Sommerlarsen, J. 1995, ApJS, 96, 175
Belokurov, V., Erkal, D., Evans, N. W., Kroupov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611
Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, MNRAS, 494, 3880
Bensby, T., Feltzing, S., Lundström, I., & Ilnyi, I. 2005, A&A, 433, 185
Bergemann, M., Collet, R., Schönrich, R., et al. 2017, ApJ, 847, 16
Bilir, S., Karaa, S., Ak, S., et al. 2012, MNRAS, 421, 3362
Bird, S. A., Xue, X. B., Liu, C., et al. 2020, arXiv:2005.05980
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Bonaca, A., Conroy, C., Cargile, P. A., et al. 2020, ApJL, 897, L18
Bonaca, A., Conroy, C., Wetzel, A., et al. 2017, ApJ, 845, 101
Bovy, J. 2015, ApJS, 216, 29
Brook, C. B., Kawata, D., Gibson, B. K., & Freeman, K. C. 2004, ApJ, 612, 894
Carollo, D., Beers, T. C., Bovo, J., et al. 2012, ApJ, 744, 195
Carollo, D., Beers, T. C., Chiba, M., et al. 2010, ApJ, 712, 692
Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, Nat, 450, 1020
