



New Families in our Solar Neighborhood: Applying Gaussian Mixture Models for Objective Classification of Structures in the Milky Way and in Simulations

Farnik Nikakhtar¹ , Robyn E. Sanderson^{1,2} , Andrew Wetzel³ , Sarah Loebman^{3,4}, Sanjib Sharma^{5,6} , Rachael Beaton^{7,8} , J. Ted Mackereth^{9,10,16}, Vijith Jacob Poovelil¹¹ , Gail Zasowski¹¹ , Ana Bonaca¹² , Sarah Martell^{6,13} ,

Henrik Jönsson¹⁴ , and Claude-André Faucher-Giguère¹⁵

¹ Department of Physics & Astronomy, University of Pennsylvania, 209 S 33rd Street, Philadelphia, PA 19104, USA; farnik@sas.upenn.edu

² Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA

³ Department of Physics & Astronomy, University of California, Davis, CA 95616, USA

⁴ Department of Physics, University of California Merced, 5200 North Lake Road, Merced, CA 95343, USA

⁵ Sydney Institute for Astronomy, School of Physics, A28, The University of Sydney, NSW 2006, Australia

⁶ Center of Excellence for Astrophysics in Three Dimensions (ASTRO-3D), Australia

⁷ Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA

⁸ The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA

⁹ Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada

¹⁰ Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada

¹¹ Department of Physics & Astronomy, University of Utah, Salt Lake City, UT 84112, USA

¹² Center for Astrophysics, Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

¹³ School of Physics, UNSW, Sydney, NSW 2052, Australia

¹⁴ Materials Science and Applied Mathematics, Malmö University, SE-205 06 Malmö, Sweden

¹⁵ Department of Physics and Astronomy and CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

Received 2021 April 19; revised 2021 July 1; accepted 2021 August 1; published 2021 November 5

Abstract

The standard picture of galaxy formation motivates the decomposition of the Milky Way into 3–4 stellar populations with distinct kinematic and elemental abundance distributions: the thin disk, thick disk, bulge, and stellar halo. To test this idea, we construct a Gaussian mixture model (GMM) for both simulated and observed stars in the solar neighborhood, using measured velocities and iron abundances (i.e., an augmented Toomre diagram) as the distributions to be decomposed. We compare results for the Gaia–APOGEE DR16 crossmatch catalog of the solar neighborhood with those from a suite of synthetic Gaia–APOGEE crossmatches constructed from FIRE-2 cosmological simulations of Milky Way mass galaxies. We find that in both the synthetic and real data, the best-fit GMM uses *five* independent components, some of whose properties resemble the standard populations predicted by galaxy formation theory. Two components can be identified unambiguously as the thin disk and another as the halo. However, instead of a single counterpart to the thick disk, there are three intermediate components with different age and alpha abundance distributions (although these data are not used to construct the model). We use decompositions of the synthetic data to show that the classified components indeed correspond to stars with different origins. By analogy with the simulated data, we show that our mixture model of the real Gaia–APOGEE crossmatch distinguishes the following components: (1) a classic thin disk of young stars on circular orbits (46%), (2) thin disk stars heated by interactions with satellites (22%), (3, 4) two components representing the velocity asymmetry of the alpha-enhanced thick disk (27%), and (5) a stellar halo consistent with early, massive accretion (4%).

Unified Astronomy Thesaurus concepts: Galaxy dynamics (591); Milky Way dynamics (1051); Solar neighborhood (1509); Galactic abundances (2002); Milky Way Galaxy physics (1056)

1. Introduction

The kinematics and elemental abundances of the Milky Way’s stars are thought to contain clues to the formation history of the Galaxy we live in (e.g., Freeman & Bland-Hawthorn 2002). In the classic picture, the distribution of stars in velocity and elemental abundances has a relatively small number of distinct components linked to different formation epochs:

1. At very early times, star formation and proto galaxy merging take place in a relatively chaotic environment, leading to a roughly spheroidal distribution variously referred to as an “early spheroid” (e.g., Elmegreen et al. 2008). Stars formed very fast in this epoch, so despite

starting from gas almost free of metals, the resultant population is quite metal-rich.

2. A subsequent epoch of accretion creates a hot disk structure (still relatively metal-poor), which forms stars in the present-day *thick disk* (e.g., Forbes et al. 2012; Bird et al. 2013) and/or stars formed early on in a thin disk are heated by scattering processes and displaced by radial migration to form the thick disk (Sharma et al. 2021). At larger radii, stars above the plane are built up by flaring of mono-age populations (Minchev et al. 2015).
3. The *thin disk* is formed by colder and more gradual accretion of more metal-rich gas, regulated by feedback from relatively steady star formation, in a process that continues to the present day (e.g., Brook et al. 2012; Garrison-Kimmel et al. 2018; Ma et al. 2017; Stern et al. 2021).

¹⁶ Banting Fellow.

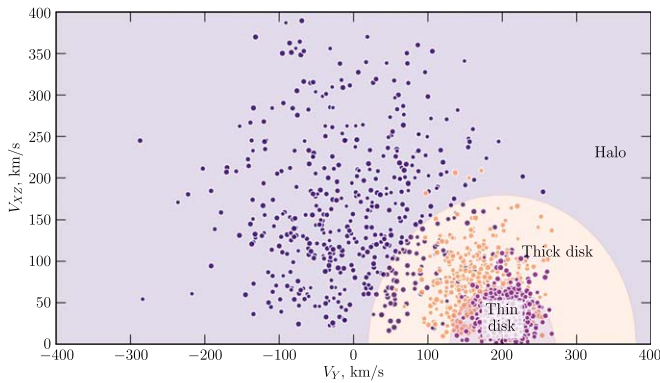


Figure 1. Typical regions of the Toomre diagram based on observation in the solar neighborhood and ascribed to different chemokinematic components following Bensby et al. (2003): the metal-poor “halo” (purple), the intermediate “thick disk” (orange), and the metal-rich “thin disk” (magenta).

4. Accretion of smaller, more metal-poor components contributes an additional, roughly spheroidal component with a larger scale radius but with substructures, commonly referred to as the “outer halo” or “*accreted halo*” (e.g., Searle & Zinn 1978).¹⁷ In studies of the solar neighborhood, this component and the first one are often jointly referred to as the “halo,” which is a suitable simplification given the short scale radius and complex kinematics of the Milky Way’s bulge-like component (Gerhard & Martinez-Valpuesta 2012).

This picture sets up the expectation of a multicomponent stellar distribution in the solar neighborhood, with old, spheroidally distributed stars at the lowest metallicities (the “halo”), young, metal-rich stars in a kinematically cold “thin disk,” and a population intermediate in age, metallicity, and kinematics, commonly referred to as the “thick disk.” Sustained star formation during the cooling of the gas reservoir, as well as gradual dynamical heating of the resulting stellar population, would predict a smooth correspondence between kinematic temperature, metallicity, and age in the disk, with kinematically hotter stars being older and more metal-poor.

Many papers prior to this one have sought to test these ideas and refine our understanding of the processes that built the solar neighborhood by selecting stars based on their kinematics and studying their abundance distributions (e.g., Bensby et al. 2003; Venn et al. 2004; Ishigaki et al. 2013; Nissen & Schuster 2010; Bonaca et al. 2017; Fernández-Trincado et al. 2019; An & Beers 2020; Hayden et al. 2020). Most often, these kinematic cuts are performed in the Toomre diagram, where the x axis has the velocity of stars in the direction of Galactic rotation V_Y , and the y axis has the perpendicular component $V_{XZ} \equiv \sqrt{V_X^2 + V_Z^2}$. Here, V_X is along the Sun–Galactic center direction, and V_Z is perpendicular to the disk plane in the direction of the total angular momentum.

Figure 1 shows the Toomre diagram of stars drawn from traditional components of the Galaxy measured by Bensby et al. (2003), with the thin disk in magenta, the thick disk in orange, and the halo in purple. These components are

overlapping, but a selection criterion based on the relative velocity with respect to the local standard of rest (LSR) can preferentially select thin disk stars comoving with the LSR (magenta shaded region in Figure 1), thick disk stars at intermediate distance from the LSR (orange shaded region), and halo stars moving at high velocity with respect to the LSR (purple shaded region).

In an alternative to such kinematic selections, other studies have selected disk stars based on their elemental abundances and from their spatial distributions tested the idea of a continuous transition between thin and thick disks (e.g., Bovy et al. 2012a, 2012b, 2016; Mackereth et al. 2017). Some cosmological simulations (e.g., Ma et al. 2017; Obreja et al. 2019; Bird et al. 2021; Buck et al. 2020) also do not predict a clean/sharp transition from the thick to the thin disk, but rather, a more gradual settling of the stellar disk (note, however, that some simulations such as the FIRE simulations analyzed in this paper suggest a sharper transition in the properties of the gas disk, as galaxies transition from highly bursty to more steady star formation rates, e.g., Stern et al. 2021). Still others were interested in searching for local interlopers from the halo to assess the Milky Way’s accretion history, simultaneously employing both kinematic and metallicity cuts to select this relatively small population from the overwhelmingly more numerous disk stars (e.g., Helmi et al. 2017; Herzog-Arbeitman et al. 2018).

The assumptions these approaches make about links between the kinematics and metallicity and/or alpha enhancement of stars have begun to be challenged with the advent of Gaia’s exquisite kinematic information (Gaia Collaboration et al. 2016). For example, several authors have pointed out the presence of a population that is either an intermediate component between the thick disk and the halo (Bonaca et al. 2017; Posti et al. 2018; Belokurov et al. 2018) or a sense of rotation in the halo itself (Deason et al. 2017; Kafle et al. 2017). The Gaia-Enceladus or “Sausage” structure (Belokurov et al. 2018; Helmi et al. 2018; Myeong et al. 2018; Mackereth et al. 2018) is likely a massive contributor to the local neighborhood and has 30%–50% of the halo stellar mass (Mackereth & Bovy 2020).

This influx of new information both enables and motivates the relaxation of some of the assumptions about the structure of the solar neighborhood, in favor of allowing the data themselves to tell us what the distribution looks like. In this work, we take the agnostic approach of modeling the stellar distribution as a mixture of Gaussians, with the goal of imposing as few assumptions as possible on its observed quantities.

We construct a mixture model (described briefly in Section 2) of the velocities and iron abundances of stars in the solar neighborhood, leaving the number of Gaussian components in the model free to vary and using an information criterion to pick the most suitable number. We test this approach on a new set of mock Gaia–APOGEE catalogs generated from FIRE-2 cosmological simulations¹⁸ of a Milky Way mass galaxy (Sanderson et al. 2020; described in Section 3). From these mock catalogs, we find that the best-fit model is consistent with previous arguments on the origins of solar neighborhood stars and their spatial, kinematic, and abundance distributions (Section 4). Perhaps surprisingly, we find that the optimal decomposition features five components in all of the simulations: two analogous to the thin disk and the halo, but instead of a single counterpart

¹⁷ These two terms, though sometimes used interchangeably, are not synonymous: accreted material, especially from early epochs, can certainly be found at small radii, while stars formed in outflows from regions of high star formation in the disk (Yu et al. 2020), and those kicked out by interactions with satellite galaxies (Laporte et al. 2018), can reach large radii (El-Badry et al. 2018; Starkenburg et al. 2017).

¹⁸ See the FIRE project website: <http://fire.northwestern.edu>.

to the thick disk, there are three intermediate components with distinct age—alpha distributions and formation histories, even though these data are not used to fit the model. We then apply the same strategy to stars in the solar neighborhood using the Gaia DR2 catalog crossmatched with the APOGEE DR16 (Section 3) survey and find a similar result (Section 5). We conclude in Section 6 by drawing analogies with the simulated surveys to postulate distinct origins for the five components identified in our Milky Way.

2. Gaussian Mixture Modeling

A Gaussian mixture model (GMM) describes a distribution of n_s data points (*samples*) \mathbf{x}_j using a combination of n_c Gaussian distributions with independent mean values $\bar{\mu}^i$ and covariance matrices Σ^i . The \mathbf{x}_j contain the n_f *features* (i.e., dimensions of data) used to determine the probabilities p that each of the data points belongs to each of the n_c Gaussian components.

Thus for a given sample,

$$p(\mathbf{x}_j | \bar{\tau}, \{\bar{\mu}\}, \{\Sigma\}) = \sum_{i=1}^{n_c} \tau_i f_i(\mathbf{x}_j | \bar{\mu}^i, \Sigma^i), \quad (1)$$

where f_i is a n_f -dimensional normal distribution with the given mean and covariance, and τ_i is the relative weight of each Gaussian subject to the constraint that $\sum_{i=1}^{n_c} \tau_i = 1$.

The GMM is thus specified by $n_c - 1$ free weight parameters $\bar{\tau}$, the $n_c n_f$ means $\{\bar{\mu}\}$, and the $n_c n_f(n_f + 1)/2$ components of the positive-definite, symmetric covariance matrices $\{\Sigma\}$, for a total of

$$n_p = \frac{n_c}{2}(n_f^2 + 3n_f + 2) - 1 \quad (2)$$

free parameters in the model. The probability of data point j belonging to component k in the GMM, also known as the *responsibility* \mathfrak{R} , is

$$\mathfrak{R}_k(\mathbf{x}_j | \bar{\tau}, \{\bar{\mu}\}, \{\Sigma\}) = \frac{\tau_k f_k(\mathbf{x}_j | \bar{\mu}^k, \Sigma^k)}{\sum_{i=1}^{n_c} \tau_i f_i(\mathbf{x}_j | \bar{\mu}^i, \Sigma^i)}. \quad (3)$$

The model assigns a *label* to each sample, i.e., identifies the Gaussian component to which the data point with coordinates \mathbf{x}_j is most likely to belong, by choosing the component with the highest \mathfrak{R}_k . Thus the GMM acts as both a description of the overall density distribution of the n_s samples in n_f -dimensional space, and as an unsupervised classifier that places each sample into one of n_c groups. We use the implementation provided in the Python package `scikit-learn` (Pedregosa et al. 2011).

In our case the features used for classification will be the three-dimensional velocities $\{V_X, V_Y, V_Z\}$ of the n_s stars, where V_Y is the Galactocentric velocity in the direction of the disk rotation, V_X along the Sun–Galactic center direction, and V_Z perpendicular to the disk plane in the direction of the total angular momentum. We add the iron abundances $[\text{Fe}/\text{H}]$ as a fourth dimension or feature, so for our case $n_f = 4$. This means for a model with five components $n_c = 5$, we have 74 free parameters (Equation (2)).

The reduction of dimensionality in moving from a three-dimensional velocity vector to two Toomre components of velocity $\{V_Y, V_{XZ}\}$ is motivated by the underlying symmetry of the Galaxy, but since there is a zero cutoff in V_{XZ} and our model is a mixture of Gaussians, we construct our model with the three-dimensional velocity vector plus metallicity and just

represent obtained clusters in the Toomre subspace plus metallicity. Likewise, we use Cartesian coordinates rather than cylindrical or spherical coordinates for the velocities to avoid imprinting assumptions about symmetries. Surveys of the current generation are for the most part embedded in the solar neighborhood, where the assumption of axisymmetry is appropriate. Symmetry assumptions like this are less appropriate to future surveys exploring a larger volume of the Galaxy (e.g., Beane et al. 2019), and the framework and intuition developed in this work lay ground for this transition.

To find the best-fit GMM, we start by choosing a number of components n_c . We initialize their means, covariances, and weights by preliminarily labeling each sample using k -means clustering (Steinhaus 1957) and maximizing the likelihood

$$\ln \mathcal{L}(\mathbf{x} | \bar{\tau}, \{\bar{\mu}\}, \{\Sigma\}) = \sum_{j=1}^{n_s} \sum_{i=1}^{n_c} \left[\log \tau_i - \frac{1}{2} \log |\Sigma^i| - \frac{n_f \log(2\pi)}{2} - \frac{1}{2} (\mathbf{x}_j - \bar{\mu}^i)^T \cdot (\Sigma^i)^{-1} \cdot (\mathbf{x}_j - \bar{\mu}^i) \right], \quad (4)$$

using expectation-maximization (Dempster et al. 1977), which determines the best-fit weights, means, and covariances. We then repeat this process for different values of n_c and determine the number of components that minimizes the Bayes information criterion (Schwarz 1978),

$$\text{BIC} = -2 \ln \hat{\mathcal{L}} + n_p \ln(n_s), \quad (5)$$

where $\hat{\mathcal{L}}$ is the maximum value of the likelihood function given by Equation (4) and n_p is the total number of free parameters in the model, given by Equation (2). This criterion compares the maximum likelihood values for different numbers of components (the first term) while including a penalty for introducing additional parameters into the model (the second term), to account for the fact that a model with more free parameters will always produce a better fit. The value of the penalty is derived from an asymptotic expansion of the Bayes evidence as the sample size approaches infinity, under the assumption that the data are independent samples from a distribution with an exponential form (such as a Gaussian). This agrees with the fundamental assumption of the GMM, which motivates our use of this criterion for model selection rather than an information-theory-based criterion such as the Akaike information criterion (AIC; Akaike 1974). The BIC's penalty for adding model parameters, weighted by $\ln n_s$, strongly prefers models with lower n_p relative to the AIC (Schwarz 1978).

The assumption that the data we are fitting are truly drawn from a combination of Gaussian components is not necessarily a great one; in fact, it is demonstrably not true for the Toomre coordinates, as we will discuss further in Section 4, so the penalty in the BIC for adding extra components to the mixture model is at best an approximation. Thus, although in the idealized case one would look for the minimum BIC value to select the preferred number of components in the model, in practice we do so by increasing n_c just until the BIC stops rapidly decreasing, which is called the *Elbow* rule/method (Thorndike 1953).

3. Observational and Mock Gaia–APOGEE Catalogs

To apply the concept of separating populations of stars using mixture modeling, we created a suite of mock catalogs mimicking the crossmatch between the Gaia Data Release 2

(Gaia Collaboration et al. 2016, 2018) and the 16th data release (DR16) of the Apache Point Observatory Galactic Evolution Experiment (APOGEE; Ahumada et al. 2020; Jönsson et al. 2020).

3.1. Observed Catalog

APOGEE-2 (Majewski et al. 2017) is a dual hemisphere survey that uses cloned spectrographs (Wilson et al. 2019) operating each at the Apache Point Observatory on the Sloan Foundation 2.5 m telescope (Gunn et al. 2006) and at Las Campanas Observatory on the duPont Telescope (Bowen & Vaughan 1973). APOGEE targets primarily red giant stars in all components of the Milky Way (MW), with substantial additional numbers of main-sequence and massive evolved stars, which are selected using a simple set of dereddened-color and magnitude criteria (Zasowski et al. 2013, 2017, R. Beaton et al. 2021 in preparation, F. Santana et al. 2021 in preparation). The exact criteria vary by location in the Galaxy and the length of time a given field will be observed.

A custom processing pipeline (Nidever et al. 2015) reduces the data and calculates heliocentric radial velocities (RVs), and the APOGEE Stellar Parameters and Abundances Pipeline (ASPCAP, García Pérez et al. 2016) produces fundamental stellar parameters (e.g., $\log(g)$ and T_{eff}) and elemental abundances for up to 26 species. The DR16 catalog contains measurements for 430,000 stars. We also make some quality selections on the data to remove stars with very small “observed” parallaxes (i.e., spuriously large distances, $\pi < 0.1 \mu\text{as}$ or $d > 100 \text{ kpc}$) and/or large measurement errors on the parallax ($\Delta\pi/\pi > 0.1$), metallicity, and radial velocity. After applying these quality cuts around 150,000 stars remained, which give us a radial coverage of 4–12 kpc Galactocentric distances.

3.2. Mock Catalog

To create the mock Gaia–APOGEE crossmatches, we start from the synthetic Gaia surveys Sanderson et al. (2020) created from three MW-mass galaxies in the Latte suite (first introduced in Wetzel et al. 2016) of FIRE-2 cosmological simulations (Hopkins et al. 2018), which feature self-consistent clustering of star formation in dense molecular clouds and thin stellar/gaseous disks in live cosmological halos with satellite dwarf galaxies and stellar halos. In each of those simulations, there are three solar viewpoints that generate nine synthetic Gaia-like surveys (3 galaxies \times 3 neighborhoods). The synthetic stars are sampled by assuming each star particle represents a single stellar population. The simulations have initial star particle masses of $7070 M_{\odot}$, but because of stellar mass loss, a typical star particle, at $z = 0$, has a mass of $\sim 5000 M_{\odot}$. At each neighborhood, dust extinction is computed from the simulated gas metallicity distribution.

Regarding the iron abundance, it is important to bear in mind that supernovae (core collapse and type Ia) and stellar winds generate and disperse metals, which are then deposited into gas particles. For supernovae Ia, the stellar nucleosynthesis yields are adopted from Iwamoto et al. (1999), where the rates follow Mannucci et al. (2006), including both prompt and delayed populations. For core collapse supernovae, yields are from Nomoto et al. (2006); for stellar winds (AGB and O/B stars), yields are from a compilation of van den Hoek & Groenewegen (1997), Marigo (2001), and Izzard et al. (2004).

These simulations also include an explicit treatment for unresolved turbulent diffusion of metals in gas, which produces

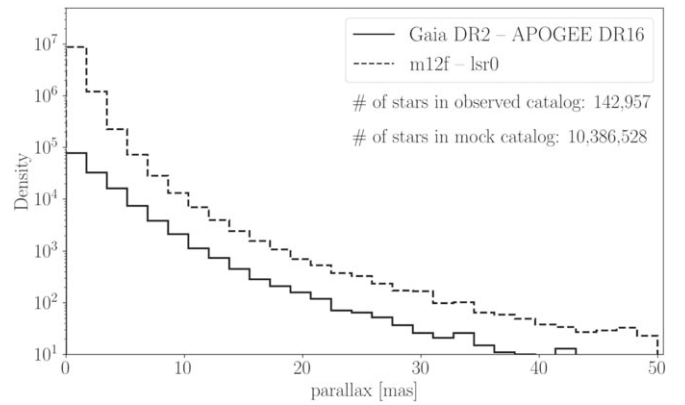


Figure 2. Parallax distribution of stars in the observed catalog and the m12f-LSR0 mock catalog after applying the quality cuts on parallax, metallicity, and radial velocity.

more realistic abundance distributions in both the MW-like galaxies and in their satellite dwarf galaxies (Su et al. 2017; Hopkins et al. 2018; Escala et al. 2018).

We also add an APOGEE-like error model for the elemental abundances [Fe/H], [Mg/Fe], [C/Fe], [N/Fe], [S/Fe], [O/Fe], [Si/Fe], and [Ca/Fe] as determined in V. J. Poovelil et al. (2021 in preparation)¹⁹ where we assume a signal-to-noise ratio of 100 for every star. Although we only use [Fe/H] for constructing the GMM, we need to track individual abundances (which are not used to fit the model) to see the physical origin and formation histories of each component. In this paper, [Mg/Fe] is used for this purpose, but other abundances will be explored in forthcoming works.

For each star the error-convolved abundances and radial velocities are each drawn from a one-dimensional normal distribution around the “true” value generated from the simulation, with the width of the normal distribution equivalent to the error on each property (qualitatively the same approach as used in Sanderson et al. (2020) to sample the Gaia error model). Our mock catalog is thus essentially an all-sky version of the Gaia–APOGEE crossmatch. Although they are based on the synthetic Gaia surveys constructed in Sanderson et al. (2020), we refer to these simulated observations as *mock catalogs* rather than *synthetic surveys* since we do not overlay the APOGEE selection function. However, this will be done in future work to produce true synthetic surveys for the Gaia–APOGEE crossmatch. Finally, before running the mixture model we make the same quality selections on the mock catalogs as the observed catalog. These selections are all based on the simulated Gaia observations, not on our rudimentary APOGEE error modeling.

To compare the spatial extent of the observed and mock catalogs, the parallax distribution of stars in both real data and one of the mocks (see Section 4 below) is shown in Figure 2. Our mock catalogs in general contain more stars than the real catalog for a few reasons. First, the total stellar mass of our simulations ranges over a factor of a few, some likely larger than the MW (Sanderson et al. 2020). Second, these are mock catalogs and not true synthetic surveys, so for example the APOGEE sky footprint has not been applied, and crowding effects are not modeled, so that even in the densest regions every star enters our mock catalogs. These effects also serve to increase the number of stars relative to the real survey.

¹⁹ A similar process was ultimately adopted for DR16 as is described in Jönsson et al. (2020).

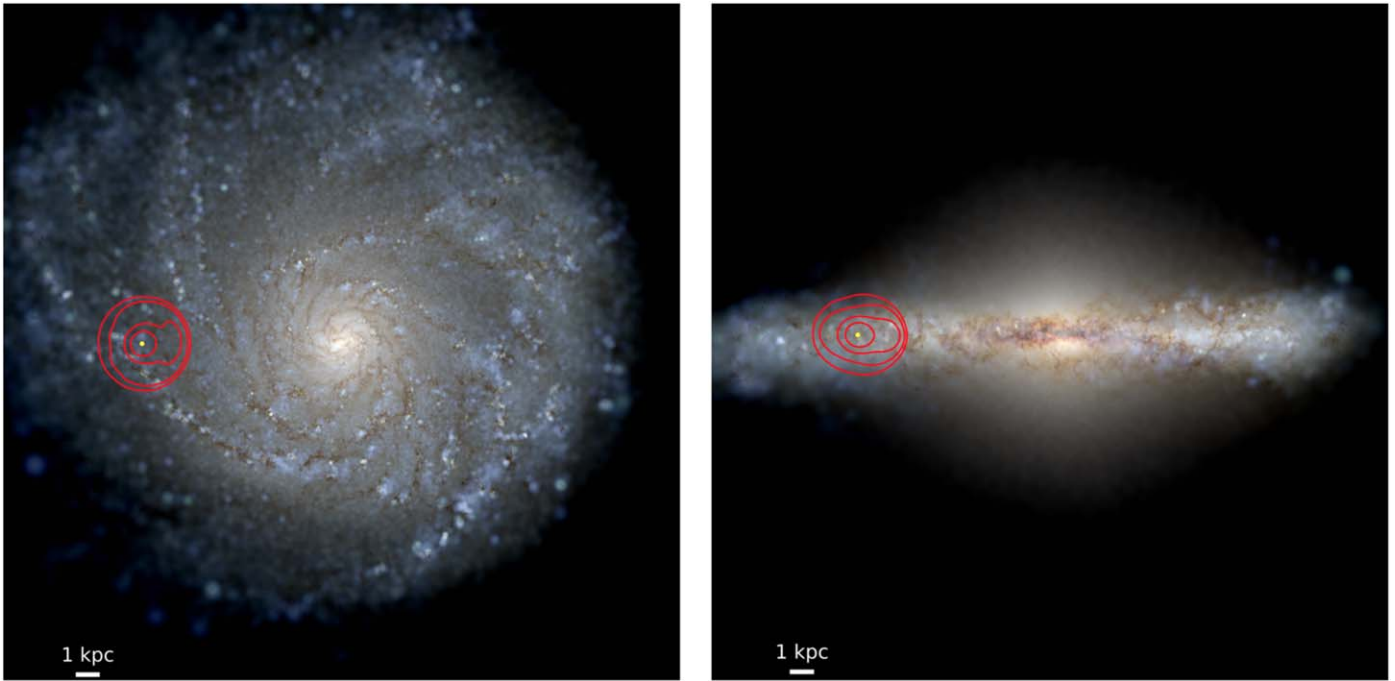


Figure 3. Face-on (left) and edge-on (right) views of m12i, one of the simulated galaxies used to generate the mock catalog. The red contours represent density contours of the sample stars around the solar position that are used for constructing the Gaussian mixture model.

However, the parallax distributions of the real and mock surveys are similarly shaped even if the normalization is different. The reach of the mock catalogs relative to the size of the simulated galaxies is also consistent with an “all-sky” extension of APOGEE; Figure 3 shows face-on and edge-on views of the m12i simulation with density contours of the sample stars from m12i-LSR0 to illustrate.

In addition to the APOGEE error model for elemental abundances, we added the Two Micron All Sky Survey (2MASS) J , H , and K_s magnitudes to our catalogs (Skrutskie et al. 2006), corresponding to the bands used to select APOGEE targets (see descriptions in Zasowski et al. 2013, 2017). The 2MASS photometric errors are estimated by using an exponential plus constant model and, similar to the elemental abundances, the convolved magnitudes are drawn from a zero-mean normal distribution with the corresponding variance.

4. Mixture Models of Mock Gaia–APOGEE Catalogs

To construct the mixture model, we use the three components of the space velocity in Cartesian coordinates $\{V_x, V_y, V_z\}$ computed from the error-convolved astrometry (according to the Gaia error model) and radial velocities (according to the APOGEE error model) and add the error-convolved iron abundances relative to solar, $[\text{Fe}/\text{H}]$, as a fourth dimension. Compared to the classic Toomre diagram (V_{xz} – V_y or V_{xz} – V_ϕ plane), which has historically been used to separate different stellar kinematic components in the solar neighborhood, this feature space allows for the possibility that stars in different kinematic components can have different metallicity distributions, but notably does not make any assumptions about what those distributions are. It likewise makes no assumptions about approximate symmetries in the phase space distribution (spherical, axisymmetric, or otherwise). This leaves us free to interpret the components obtained by the model in the context

of broader ideas about galaxy formation, such as the expectation that stars in the thin disk, with a velocity distribution centered most closely on the Sun’s, should also be the youngest and most metal-rich.

Figure 4 shows the results of testing different numbers of Gaussian components to model this four-dimensional “augmented Toomre diagram.” There is a clear improvement up to five components and nearly no appreciable improvement after that (Elbow rule). Moreover, the next components (sixth component and more) do not correspond to a distinct group and just split one of the pre-specified clusters. Given the preference for a low number of distinct components (consistent with the idea of a thin disk, thick disk, and halo perhaps broken into some subpopulations) we choose the five component model for further examination.

In Figure 5 the density contours of each component of the best-fit model for one of the mock catalogs are shown in three projections of Toomre+ $[\text{Fe}/\text{H}]$ space. Each set of colored contours shows the density distribution of synthetic stars for which the probability of belonging to that component is highest. We see that this model includes components that fit the standard expectations of galaxy formation and the solar neighborhood distribution: a metal-rich component with a narrow velocity distribution around the solar velocity (the “thin disk,” shown in magenta), a very metal-poor component with a broad rotational velocity distribution around zero (the “halo,” shown in dark purple), and a few progressively broader and more metal-poor components that together span the difference between these (shown in orange, black, and rose). We see a similar result, with slight variations in the relative positions of the components, for all nine mock catalogs.

The probability panel (top right) in this figure gives a sense of how well the model describes the data. It shows the distribution of probabilities of belonging to each component for the different stars in the sample. For some components, this distribution has well-defined peaks at 0 (i.e., the star is definitely not in that

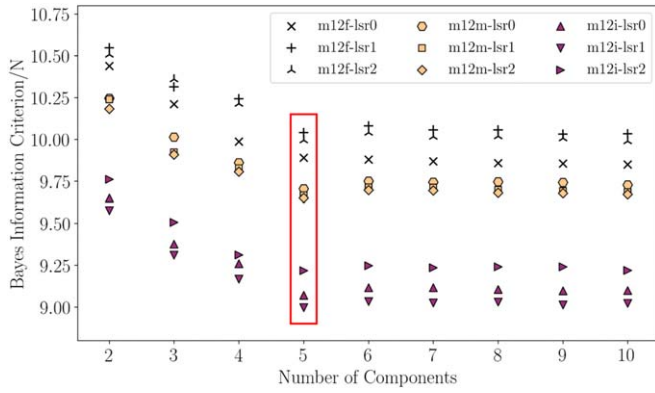


Figure 4. Bayes information criterion (BIC) over the number of objects versus the number of components of mixture models for all nine Gaia–APOGEE mocks. Black symbols are for the m12f simulation, orange symbols for the m12i simulation, and magenta symbols are for m12m. All of the 3 simulations \times 3 neighborhoods show the same BIC plateau at five components, which means the optimal decomposition features five components in all of the simulations.

component) and 1 (i.e., the star is almost certainly part of that component) with a relatively low number of probabilities at intermediate values. This means that for these components, stars are easily classified. The halo and thin disk components demonstrate this behavior.

That said, for other components, there is still a peak at zero but the distribution has more intermediate probability values and drops to zero before reaching $p = 1$. This indicates that it is more difficult for the model to securely classify stars in one of these components; in this case it is because they are relatively similar, as can be seen by examining the other panels, so many stars overlap with these. The three intermediate components demonstrate this behavior. This ambiguity between the components could come from the fact that a single Gaussian distribution in velocity is a bad description of the asymmetric velocity distribution displayed by thick disk stars, which is better modeled by a superposition of Gaussians (e.g., Schwarzschild 1907; Nordström et al. 2004). In this case one would expect that multiple otherwise similar components are being used to effectively expand the velocity distribution in a basis of Gaussians. If the GMM is simply to be used to separate the three traditional constituents of the solar neighborhood, the three intermediate components can be lumped together as the “thick disk” with no practical or conceptual difficulty. However, the preference for a small number of intermediate components, and the consistency of that number across all simulated and real data sets, also raises the intriguing possibility that the GMM is identifying subpopulations of stars with different intrinsic properties or different origins. We will evaluate this possibility, and analyze the intermediate components in detail, in Section 4.1.

We can verify the correspondence between the traditional components of the solar neighborhood and these Gaussian components by examining the distributions of various properties not used to derive the model. For example, if the most metal-rich component truly corresponds to a traditional “thin disk,” then stars in this component should have a narrow distribution of heights above the disk plane, young-to-intermediate ages, unenhanced alpha abundances relative to iron, and formation locations within the Galactic disk.

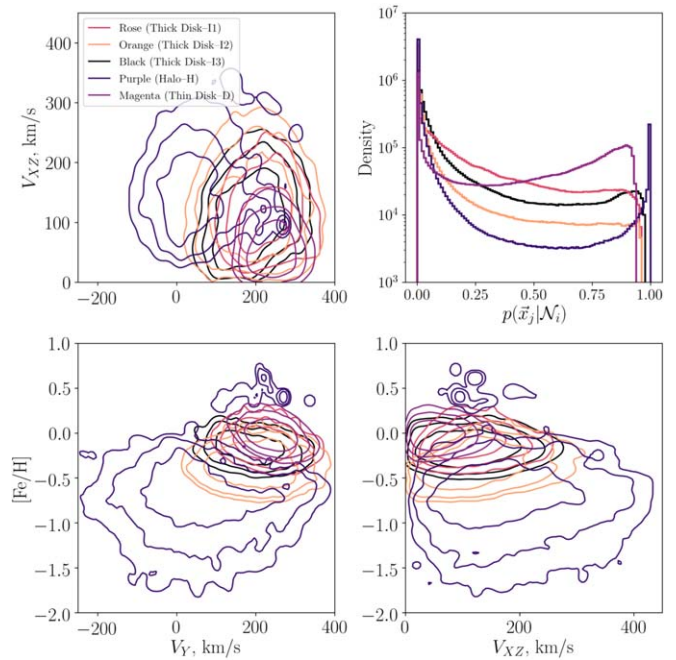


Figure 5. Augmented Toomre diagram (V_Y vs. V_{XZ} vs. $[\text{Fe}/\text{H}]$) contours of the best-fit five component mixture model to the m12f-LSR0 mock Gaia–APOGEE catalog. The top right panel shows the distribution of probabilities of belonging to each component for the different stars in the sample. In this model, two components can be identified unambiguously as the thin disk (magenta: metal-rich component with a narrow velocity distribution around the solar velocity) and another as the halo (purple: metal-poor component with a broad rotational velocity distribution around zero). Additionally, there are three progressively broader and more metal-poor components that together span the difference between these.

In Figure 6 we show a series of one-dimensional distributions of various stellar properties for each component in the same mock catalog. The probability panel is the same as the top right panel of Figure 5 and colors of the components are the same as in that figure; the overall distribution is shown in gray where applicable. In contrast to Figure 5, which uses the GMM as a classifier and thus includes only stars with $p_i(\mathbf{x}) > p_{j \neq i}(\mathbf{x})$ in each component, these distributions are calculated by weighting each star’s contribution by its probability of belonging to that component. This illustrates the power of mixture modeling to permit full probabilistic analysis, which is especially important given the degree of overlap between populations (illustrated by the probability distributions shown in Figure 5, which contain many intermediate values).

As seen in the top row of Figure 6, the velocity distributions of these three intermediate components bridge the gap between stars rotating in the disk plane with the Sun ($V_Y = V_{Y,\odot}$, shown with a vertical dashed line) and stars orbiting in a broader distribution centered on the Galactic center. However, the inclusion of $[\text{Fe}/\text{H}]$ as a fourth component shows that there is indeed some additional information in this distribution: components with lower mean metallicities that include metal-poor stars tend to have a broader velocity distribution. In the standard picture of galaxy formation, metal-poor stars originate predominantly from a galaxy merger and it is anticipated that they have a broader/more spheroidal velocity distribution. However, recent observations show that the Milky Way’s metal-poor stars have a strong preference to be on prograde disk orbits (Sestito et al. 2019, 2020, 2021). The similar behavior of metal-poor stars is shown for 11 of the 12 galaxies

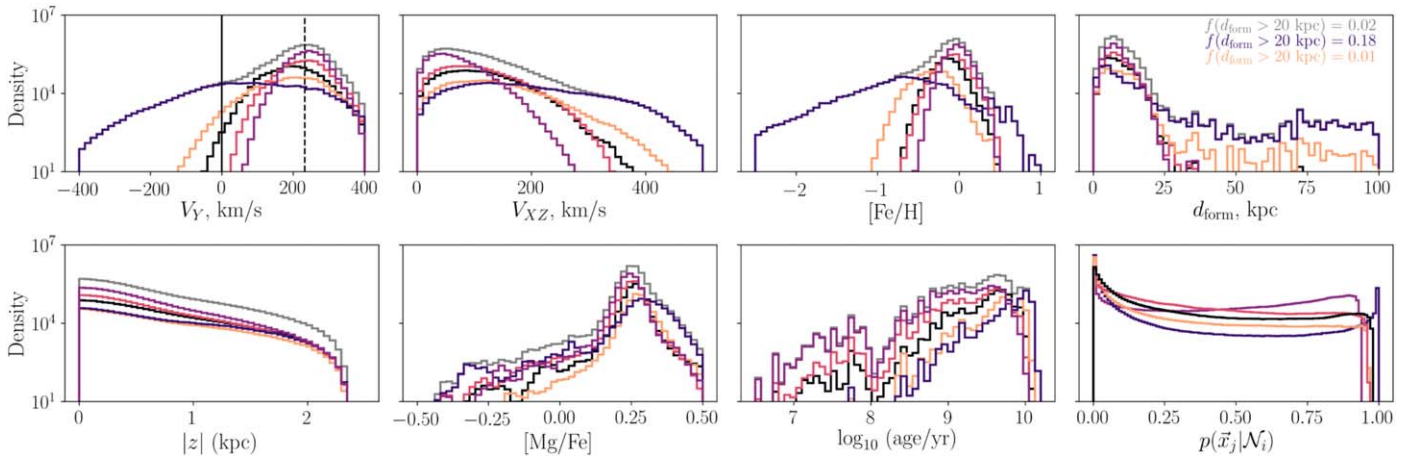


Figure 6. Distribution of various properties of stars in each component of the best-fit mixture model, for the m12f-LSR0 mock catalog. As with all our mocks, one component (dark purple) is halo-like, one (magenta) is thin-disk-like, and three (rose, orange, and black) represent intermediate populations whose properties vary with LSR and formation history. The overall distribution for all stars in this sample is shown in gray. The upper right panel shows d_{form} distributions and stars with $d_{\text{form}} > 20$ kpc are considered to be accreted. The accreted fraction of the total sample is about 2% and for the halo component it is about 20%. All five components have different age and alpha abundance distributions and these properties are not used to construct the model.

from the FIRE-2 simulation suite, including two of the simulations that we use in this work (m12f and m12m; Santistevan et al. 2021). The only galaxy among the 12 that does not show prograde orbits for metal-poor stars is m12i and the reason for this might be that all mergers occurred at about the same time, which could have distorted any coherent effect (Santistevan et al. 2021). This preference of metal-poor stars helps us to understand the velocity asymmetry that we see in Figure 6 for the halo component and the most metal-poor thick-disk-like component.

The other panels in Figure 6 show quantities that were not used in constructing the mixture model:

1. The formation distance d_{form} of the star particle from which each mock star was spawned relative to the main galaxy;
2. The height $|z|$ above the disk plane;
3. The magnesium-to-iron abundance ratio;
4. The stellar age.

These distributions can help us assess how the components identified by the model map onto ideas about the structure of the solar neighborhood outlined in Section 1.

We track the approximate formation locations of each star particle in the simulated galaxy, relative to the center of the main halo at the time of formation, by post-processing the Gizmo²⁰ snapshots saved from the simulation (Wetzel & Garrison-Kimmel 2020a). We define d_{form} as the distance of the star particle from the host galaxy center in the first snapshot after it is formed. As in Bonaca et al. (2017) and Sanderson et al. (2018), we consider stars with $d_{\text{form}} > 20$ –30 kpc to be accreted depending on the simulation (see Figure 1 of Sanderson et al. 2018), although the caveats discussed extensively in both those works apply here as well. In short, this can be considered a conservative definition of the accreted stellar component.

Overall, we see that the distribution of the magenta component closely resembles what would be considered the thin disk in the Milky Way: mostly young, metal-rich stars

formed inside the Galaxy, with alpha-to-iron ratios close to solar, orbiting with the Sun near the disk plane. The metallicity distribution of the stars in this component is not exclusively metal-rich, but extends down to at least $[\text{Fe}/\text{H}] \sim -0.5$. The component also includes some stars as old as ~ 10 Gyr that nevertheless chemically resemble the Sun. These characteristics are consistent with recent studies of the MW’s chemically-selected thin disk (e.g., Beraldo e Silva 2021). Likewise, the dark purple component pretty clearly fits our expectations for the halo: velocities consistent with a kinematically hot spheroid (slightly counter-rotating in some cases), a broad spatial distribution, and old, metal-poor, alpha-enhanced stars mostly formed at very large distances from the central galaxy. These two components are clearly identifiable in the best-fit mixture model for all nine mock catalogs. Given the clear parallels between these two components and standard interpretations of stellar populations, we will refer to them as the thin disk and halo for short in the remainder of the paper. For these two components the formation distance of the stars also supports classical theories about their origin: stars in the thin disk component all formed within 25 kpc of the Galactic center, while about 20% of the halo component was formed beyond 20 kpc, where for this simulation most material can be considered accreted rather than formed in situ. Interestingly, more than half of the stars in the halo component have an origin consistent with our picture of an early spheroid: extremely old ages and low metallicity, yet formed within the main galaxy. These stars come from early, extremely bursty epochs of star formation seen in the simulations (El-Badry et al. 2018; Yu et al. 2020; Muratov et al. 2015; Sparre et al. 2017; Faucher-Giguère 2018). The accreted fraction the halo component of m12f, m12i, and m12m simulations is about 20%, 10%, and 25%, respectively; these percentages depend on the assembly history of each simulation and these three simulations have different accretion histories. In each simulation, the accreted fraction of each component is almost the same (within 1%–2%) in the three different LSRs and also the accreted fraction of all stars is the same in all nine mocks (about 2%).

²⁰ https://bitbucket.org/awetzel/gizmo_analysis

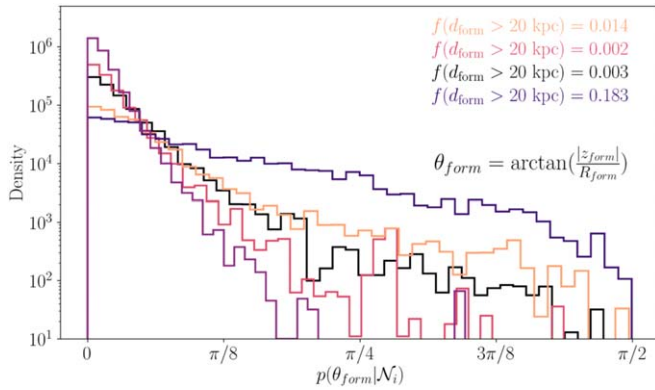


Figure 7. The distribution of the inclination angle of stars relative to the disk plane at the time of formation for all components of the m12f-LSR0 mock catalog. The orange component has a much flatter distribution than the other two intermediate components (rose and black), but it is still not as flat as the distribution of the halo component (purple).

From the velocity and metallicity distributions, it is understandable how these two components, comprising the oldest and the youngest stars in the sample, are most easily picked out by the model with certainty. In the age distribution, which is not used in our model, they are also very well separated. They are most consistent with a Gaussian velocity distribution, though for two very different reasons: one has barely been transformed from its birth distribution at all while the other has evolved for many dynamical times. They also pick out the metal-poor tail and metal-rich peak of the total stellar distribution, respectively, making these stars particularly easy to classify.

4.1. Thick Disk Decomposition

The remaining three components, which together make up what would traditionally be referred to as the “thick disk,” lie intermediate to the thin disk and halo components in all of the characteristics we examined. The separation of this intermediate population into multiple components could indicate less consistency with the assumptions of the model, a more complex origin, or some of both. In this section we discuss these three components in detail.

The component shown in orange most resembles the halo component in terms of its velocity, metallicity, and z distributions, but differs in a few important respects. First, there are some younger stars present than in the halo component (purple); and second, although there are a few stars present formed at larger distances, the overall makeup of this sample resembles the total solar neighborhood in terms of its accreted fraction (about 1%). The stars in this component are also slightly more metal-rich than in the halo, and notably less alpha-enhanced.

Moreover, from Figure 7, which shows the distribution of the inclination angle of stars relative to the disk plane at the time of formation, we can see that this component has a much flatter distribution than the other two intermediate components, but it is still not as flat as the distribution of the halo component. The rose and black components also have distinct distributions in this view: most of the stars in the rose component have $\theta_{\text{form}} < \pi/4$ (i.e., are formed quite close to the galactic plane), while the black component has a longer tail at higher θ_{form} .

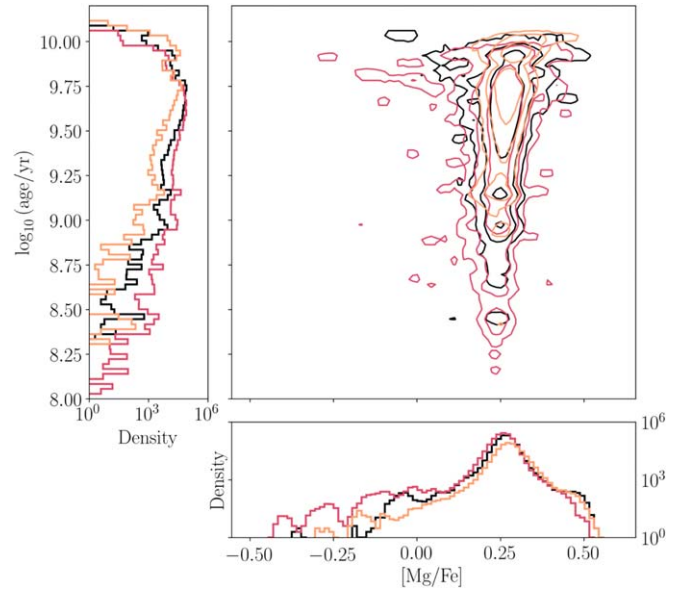


Figure 8. The distribution of age–alpha for three intermediate (thick disk) components of the m12f-LSR0 mock catalog. The rose component has younger, less alpha-enhanced stars, which is consistent with the picture given by its formation angle distribution (Figure 7).

The three components also show differences in their age and alpha abundance, two other features that were not used in classifying them. Figure 8 shows the distribution of age versus alpha for these three components, obtained by classified stars as shown in Figure 5. The rose component has younger, less alpha-enhanced stars, which is consistent with the picture given by its formation angle distribution that these stars formed after the disk plane is well established and the disk relatively cold. In these simulations, this is usually due to an influx of cold, high-angular-momentum gas at relatively late times (Garrison-Kimmel et al. 2018). Conversely, the orange component includes the oldest and most alpha-enhanced stars of the three, consistent with the picture that these stars formed earlier when the disk was kinematically hotter (often because of a wider distribution of angular momentum in the cold gas accretion).

The GMM used to classify the stars in these components is based solely on 3D velocities plus iron abundance, yet we see that they have different age and $[\text{Mg}/\text{Fe}]$ distributions and are formed in different locations relative to the galactic center (d_{form}) and to the disk plane at the time of formation (θ_{form}). Furthermore, while the thin disk and halo components are consistently identified in every mock catalog, the characteristics of intermediate components identified in different simulations (which have different assembly histories) are markedly different. If the GMM were simply decomposing the asymmetric drift in a combination of Gaussians, it is highly unlikely that all of these differences would be apparent. We thus argue that the mixture model is indeed identifying components with different physical origins within the thick disk.

4.2. Dependence on Position in the Galaxy

With mock catalogs, it is possible to study how the distribution of stars in the augmented Toomre space varies as a function of position in the Galaxy. The solar position within the simulation is a fairly arbitrary choice since the simulated galaxy does not resemble the Milky Way in its detailed

structure, such as the position and number of spiral arms, the size and orientation of a bar, or the number and location of tidal streams in the halo. To illustrate how the augmented Toomre diagram changes as a function of location, and to test the sensitivity of this approach to the choice of solar position, we generated different catalogs for three solar locations in each simulated galaxy. They have 120° displacement relative to each other on the solar circle. As shown in Figure 4, the behavior of the BIC and the number of preferred mixture components is consistent across different solar locations, implying that this technique will be successful when applied to the real data regardless of the local variations in the density or distribution of stars.

Figure 9 shows that the characteristics of the components identified by the model are not completely the same in all of the nine mock catalogs, but in all of them we have three components for the thick disk. In addition to position dependence in the galaxy, these differences can depend on the assembly history of each simulation since these three galaxies have different relative formation times and accretion histories.

4.3. Dependence on Assembly History

Many scenarios have been proposed for the origin of the thick disk. It may emerge from stars migrating outward from the hot, inner disk (e.g., Loebman et al. 2011), from a turbulent interstellar medium (ISM; e.g., Bournaud et al. 2009) or a gas-rich merger (e.g., Brook et al. 2004) at high redshift, from the nested flairs of mono-age populations (e.g. Minchev et al. 2015), from a satellite dynamically heating a preexisting stellar disk (e.g., Villalobos & Helmi 2008), or from the accretion of stars stripped from satellites (e.g., Abadi et al. 2003). In addition to each of these single-origin theories, the thick disk could also arise from various combinations of these processes at different times. Moreover, in these simulations, since there was not always a single main progenitor, the definition of in/ex situ for early-forming stars is complicated (Santistevan et al. 2020).

Figure 9 shows formation distance of stars versus their age in each component. For each simulation we can see the dependence on position in the galaxy (different LSRs) and by comparing simulations to each other, we are able to see the dependence on assembly history. In addition, the movies available here²¹ show the formation and spatial evolution of each component over time.

In m12f (top three rows), we see that the halo component is clearly accreted, but includes the very earliest star formation in the main halo as well. The oldest thick disk component (orange) is mostly composed of stars that formed very early on near the galaxy center (El-Badry et al. 2016; Ma et al. 2017; El-Badry et al. 2018). It also has some stars from a merger that comes in on a nearly co-planar orbit late in the simulation (bouncing track at high d_{form}). This leads to a far lower mean metallicity for this component than for the other two thick disk components (black and rose). These also mostly form interior to the solar circle, but at more intermediate radii, starting and finishing their star formation later than the orange component. The movie of these components shows they also have some initial diskiness. The disk in this simulation starts out perpendicular to what ends up being the disk plane at present day, and is torqued by a merger into its present configuration.

These intermediate components thus show the stages of inside-out formation in the disk, modified by merger interactions that scatter young stars formed near the solar circle, by the rotation of the disk during the merger, and by “blurring” (the selection effect that solar neighborhood stars are preferentially near apocenter). These two intermediate components are also staggered in age, especially in LSR0, and this is reflected in the systematic variation of their mean $[\text{Fe}/\text{H}]$. Closer to $z=0$, we see that the two later-forming thick disk components also contain stars that were scattered onto orbits that intersect the solar circle by the late interaction with the merging galaxy, especially in LSR1 and LSR2. The thin disk forms latest, after the merger torques all of the thick disk stars vary rapidly over into its preferred plane from a nearly 90 degree angle. The stars in the thin disk mostly come from interior to the solar circle, but the degree to which blurring can contribute to this effect is limited, since this component contains the stars on the most circular orbits (that is, the maximum eccentricity is limited by the velocity selection). This component is also scattered by the merger; it looks like perhaps some stars are removed due to the interaction, transferred to the higher-dispersion component.

Examining m12i (middle three rows), we see that its different history is reflected in the classified components, but that the same set of mechanisms is present: accretion and early star formation in the halo component, young stars and blurring in the thin disk, and radial migration and heating in the three thick disk components, which pick out stars of different ages formed in different regions: the most metal-poor component (orange) is also the oldest and most transformed by radial migration; the intermediate component (black) shows the classic combination of blurring and radial transport; the youngest and most metal-rich (rose) actually resembles the thin disk in some respects, but its stars are mostly older, have a broader velocity distribution, and show a greater alpha enhancement. Here, the different solar positions do not differ as much as in m12f, a reflection of this simulated galaxy’s overall much calmer late-time history that leads to a well-mixed system. This is also reflected in the larger difference between the halo metallicity and the disk-like components, and in the systematic variation of the mean metallicities of the disk-like components.

Finally, in m12m we find yet another superposition of the different formation channels. Here, we see stars with a high formation distance even in the thin disk component, and across all of the others. These stars come from a $\sim 1:1$ merger at $z \sim 2$ that results in a starburst across the whole galaxy; a handful of these even end up on thin-disk-like orbits.²² There are traces of this merger scattered among all of the components but the bulk of its stars are in the halo (dark purple) where it is apparent as a thick descending line at high age. We also see some variations in the makeup of the thick disk components at different solar positions: in two cases (LSR0 and LSR2) there are two components with similar average metallicities of around -0.2 that are very old stars likely related to the merger starburst, and another population at much higher metallicity (around 0.0) that looks more like a classic radially migrated distribution. In the other case (LSR1) the proportions are reversed, and so are the average metallicities (now two components have $[\text{Fe}/\text{H}] \sim 0$ and one has $[\text{Fe}/\text{H}] \sim -0.2$).

²¹ <https://web.sas.upenn.edu/dynamics/data/ananke-2a>

²² Since the merger is roughly equal mass, it is somewhat arbitrary which galaxy is the main progenitor.

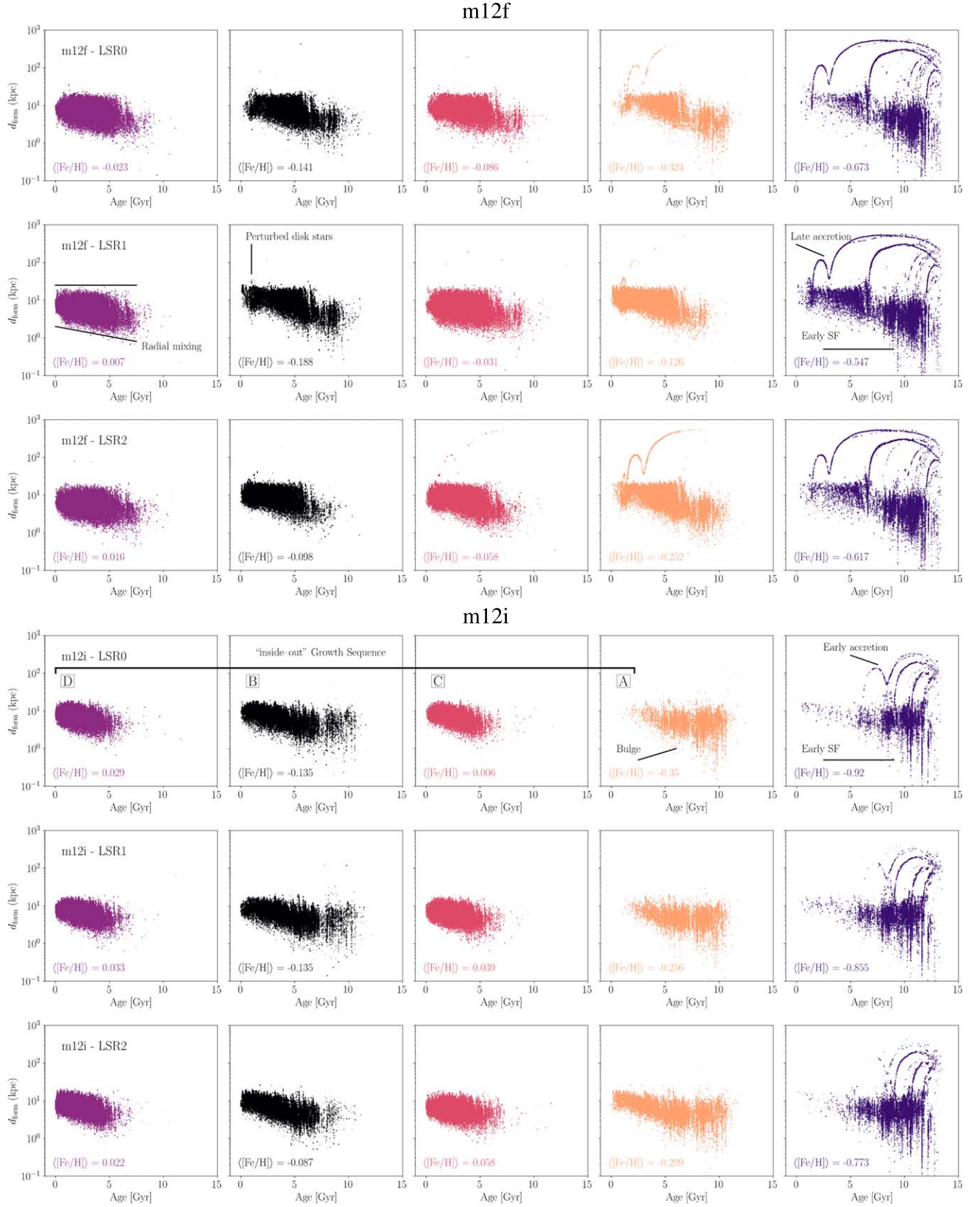


Figure 9. Formation distance–age for each component in all simulations. Some major events in each simulation are annotated in each panel (e.g., radial mixing, “inside-out” growth sequence, merger, etc.) The intermediate components show the stages of inside-out formation in the disk, modified by merger interactions that scatter young stars formed near the solar circle, by the rotation of the disk during the merger, and by “blurring” (the selection effect that solar neighborhood stars are preferentially near apocenter). In the m12f simulation the intermediate components contain stars that have been heated by interactions with satellite galaxies such as Sagittarius or a late merger. The halo components of the m12m panels support the idea that the majority of the halo component is from one or two early, massive mergers.

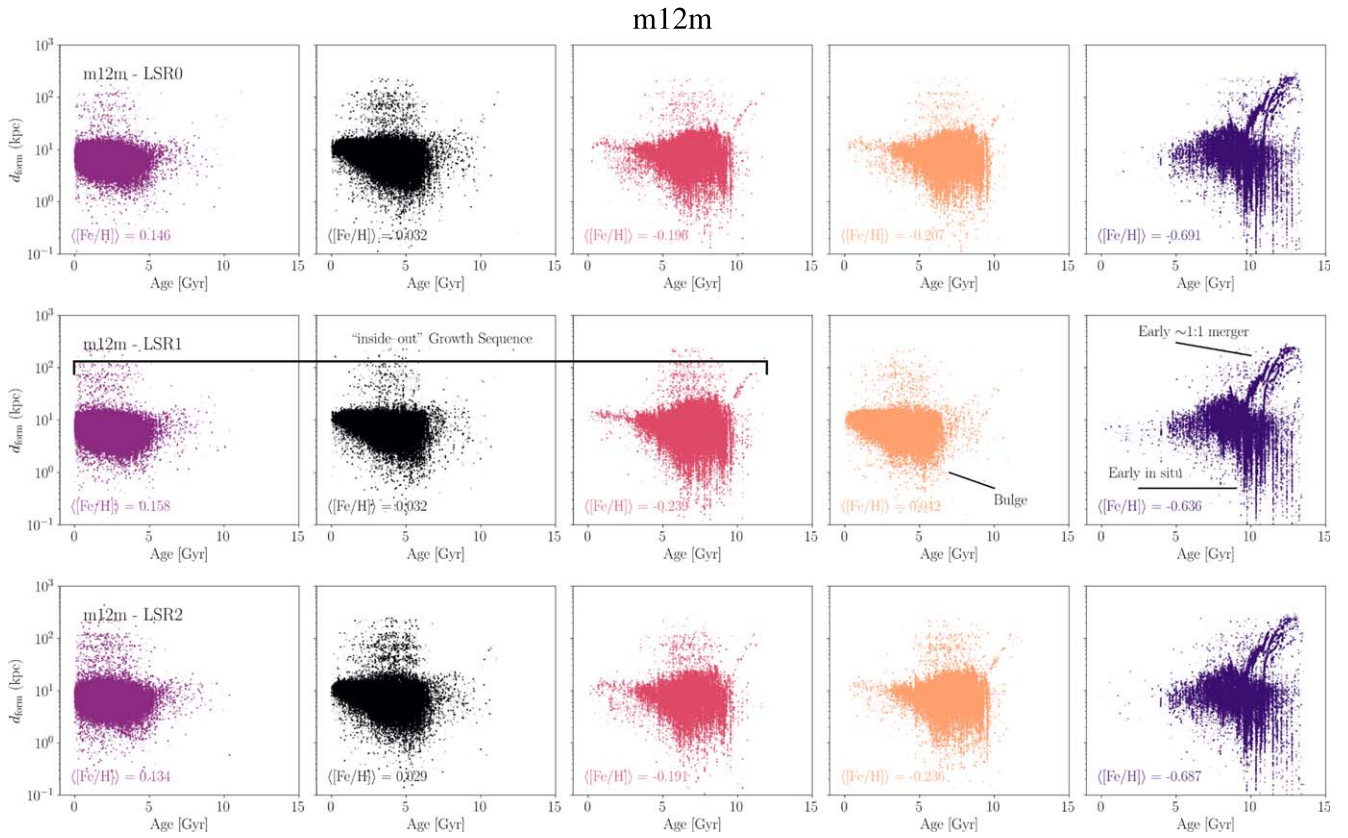


Figure 9. (Continued.)

To summarize, we find that across all simulations and solar locations, the components of our mixture models generally correspond to different formation channels for their stars, including the decomposition of the thick disk into multiple subcomponents. These can be disambiguated to some extent by the grouping of the mean metallicities of the different components, which varies based on the specific formation channels involved for each galaxy, or by examining distributions of other elemental abundances if these are available.

5. Mixture Models of the Real Gaia–APOGEE Catalog

We next constructed a mixture model of the *real* Gaia DR2–APOGEE DR16 crossmatched catalog. As before, we tried models with different numbers of Gaussian components and used the Bayes Information Criterion to choose a preferred number of components. Figure 10 shows that as with the mock catalogs, the model with five components provides a significant improvement over fewer-component models, while adding additional components improves the performance far less. We therefore proceed with five components as for the mock catalogs.

In Table 1, we present the coefficients of the five component Gaussian mixture model trained on the real Gaia–APOGEE data set (different components of the model are defined by their colors). This trained model can be used to identify members of these components in other surveys (Section 6). The correlation matrices in this table show that for the thin disk component there is a strong correlation (large off-diagonal elements) between V_X and V_Y , pointing toward a cylindrical symmetry

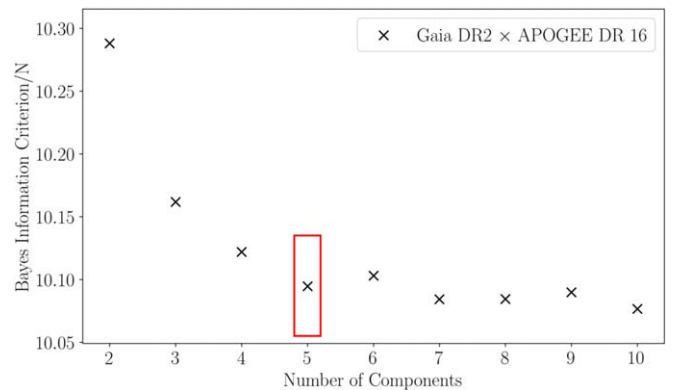


Figure 10. Bayes information criterion (BIC) over the number of stars versus the number of components of mixture models for the real Gaia DR2–APOGEE DR16 crossmatched catalog (compare to Figure 4 for the mock catalogs).

and V_R , V_ϕ as coordinates. However, for other components this is not as strong; the halo component in particular is correlated similarly strongly in V_X and V_Y and in V_X and V_Z , which suggests spherical symmetry, as expected.

In Figure 11, the left panel shows the residual plot of our GMM model in the $\{V_X, V_Y\}$ plane, and the right panel is the density plot of the stars in the solar neighborhood. In the residual plot, we compare the kernel density estimation (KDE) of the stars in that plane with the GMM probability estimation, which shows the over- and underestimation locations of the density. The positions of the centers of the stellar moving groups according to Antoja et al. (2008, 2010) are shown in the

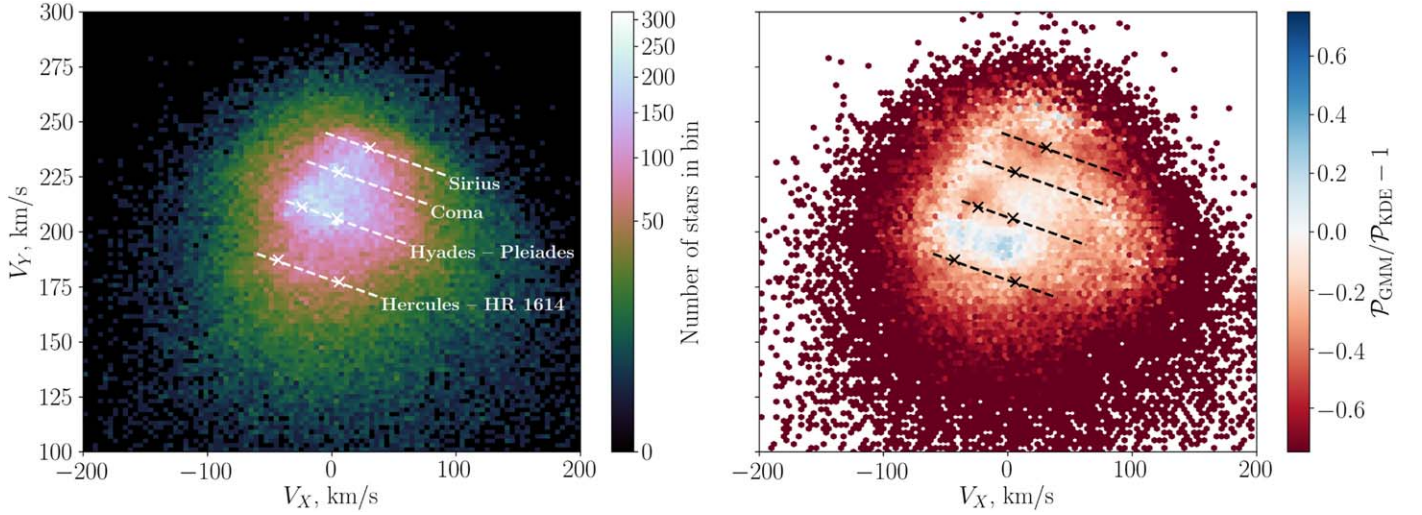


Figure 11. Left panel: 2D histogram of the velocity plane for the stars in the solar neighborhood. Right panel: residual plot for the GMM model in the velocity plane. The positions of the centers of the stellar moving groups are shown in the residual plot. The dashed lines show the approximate trace of the branches and they are at the same levels in both plots. They show that over[under]estimation of the density occurs at the locations of physical under[over]densities.

Table 1
Best-fit Gaussian Mixture Model for the Real Gaia–APOGEE Catalog

Component	Color	τ	$\vec{\mu}$ [km s ⁻¹ , km s ⁻¹ , km s ⁻¹ , dex]	$\hat{\Sigma}$
H	Dark purple	0.04	[19.01, 99.30, 3.32, -0.69]	$\begin{pmatrix} 13107.177 & -447.250 & -851.968 & 4.756 \\ -447.250 & 7498.934 & -2.700 & 15.538 \\ -851.968 & -2.700 & 4212.157 & 1.404 \\ 4.756 & 15.538 & 1.404 & 0.172 \end{pmatrix}$
I1	Rose	0.14	[-8.16, 195.53, 18.5, -0.28]	$\begin{pmatrix} 2667.817 & -81.931 & 132.633 & -1.146 \\ -81.931 & 1165.986 & -98.953 & 0.355 \\ 132.633 & -98.953 & 648.367 & -1.070 \\ -1.146 & 0.355 & -1.070 & 0.045 \end{pmatrix}$
I2	Orange	0.22	[44.25, 209.84, -2.24, -0.10]	$\begin{pmatrix} 1857.374 & -405.721 & -2.581 & 1.080 \\ -405.721 & 540.602 & 4.405 & -1.355 \\ -2.581 & 4.405 & 423.121 & 0.009 \\ 1.080 & -1.355 & 0.009 & 0.041 \end{pmatrix}$
I3	Black	0.13	[-2.16, 182.16, -19.73, -0.25]	$\begin{pmatrix} 3368.552 & -60.419 & -157.776 & -1.134 \\ -60.419 & 1164.563 & -35.883 & 0.532 \\ -157.776 & -35.883 & 711.881 & 1.784 \\ -1.134 & 0.532 & 1.784 & 0.059 \end{pmatrix}$
D	Magenta	0.46	[-1.28, 211.78, -0.37, -0.01]	$\begin{pmatrix} 1019.530 & 142.882 & 1.035 & 0.200 \\ 142.882 & 441.262 & 14.896 & -1.027 \\ 1.035 & 14.896 & 201.201 & -0.012 \\ 0.200 & -1.027 & -0.012 & 0.026 \end{pmatrix}$

Note. The order of the data vector is $\{V_X, V_Y, V_Z, [\text{Fe}/\text{H}]\}$.

residual plot. The dashed lines show the approximate trace of the branches and they are at the same levels in both plots. They show that over[under]estimation of the density occurs at the locations of physical under[over]densities. The Sirius, Coma, Hyades, Pleiades, Hercules, and HR 1614 moving groups are clearly matched with our over/under estimated regions and this shows that our GMM model is classifying structures on larger scales than individual moving groups, and could potentially be used to remove the underlying smooth component for better study of these groups.

Figure 12 shows the augmented Toomre diagram for the best-fit five component Gaussian mixture model of Gaia–APOGEE. The components identified are strikingly similar to those picked out by the best-fit model for the mock catalog

(Figure 5). This is borne out by examining the distributions of various other stellar properties in Figure 13: though (sadly) we cannot show the d_{form} values and ages for the real catalog, APOGEE provides magnesium-to-iron ratios, Gaia provides heights above the disk plane at present day, and the astroNN Value-Added Catalog provides ages and estimated orbital properties (Leung & Bovy 2019; Mackereth et al. 2019). For this catalog, in addition to quality cuts that we apply on the observed and mock catalogs, we select the astroNN ages for stars with $[\text{Fe}/\text{H}] > -0.5$ because there is not any training set stars for the catalog with low metallicities. The age distributions panel in Figure 13 is from the astroNN Value-Added Catalog and it shows the same pattern as in the mock catalogs: an exclusively old halo component, a thin disk containing

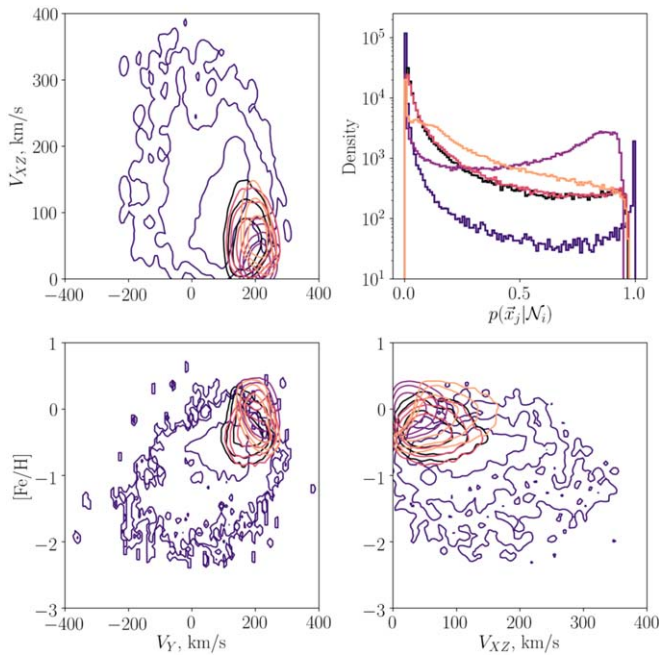


Figure 12. Augmented Toomre diagram for the five component mixture model of the real Gaia–APOGEE catalog (compare to Figure 5 for the mock catalog).

nearly all of the youngest stars, and three intermediate groups with distinct distributions. The rose and the black components are close to each other in age, but the orange component is much younger on average.

Figure 14 shows the orbital eccentricity distributions of the real data for subpopulations classified using our mixture model. From the orbital eccentricity distributions, we see that stars in the halo component have more elliptical orbits and thin disk stars are on circular orbits, also as expected. The three thick disk components in this space are also different from each other; the orange component contains orbits that are generally more circular than those in the black and rose components.

Examining the alpha abundance distributions bears out our intuition from the mock catalogs as well. The Milky Way famously has a bimodal distribution in $[\alpha/Fe]$ (e.g., Nidever et al. 2015; Hayden et al. 2015) and we see that the least alpha-enhanced stars are associated with the thin disk component, while the halo includes the most enhanced ones. The black and rose components include some stars with both alpha abundances, while the orange component is mainly composed of stars with *some* alpha enhancement, but less than the other two. However, the black and rose components are not completely identical: the metallicity distribution for the black component extends to solar metallicity and above, while the rose component truncates at lower $[Fe/H]$. The black component also has a larger tail at high V_{XZ} , extending as far as the orange component does, while the rose component is somewhat less extended in random energy. Interestingly, the orange component has a higher mean V_Y than all but the thin disk, yet also the widest spread in V_{XZ} of any but the halo component.

6. Discussion

In this paper we derived a best-fit, five component Gaussian mixture model of the solar neighborhood in augmented Toomre space (velocities plus iron abundances). Despite extremely limited assumptions, the components picked out by the model in both the mock and real Gaia–APOGEE catalogs generally

reflect common interpretations of the origin of various kinematic subpopulations. Based on our parallel analysis of simulated data, the model appears to be flexible enough to accommodate both asymmetric drift and thick disk subpopulations with different histories.

6.1. Origins of the Families in Our Neighborhood

In Section 5 we discuss the properties of the five components identified by our mixture model in the Gaia–APOGEE cross-match. By analogy with similar distributions in our simulations, we propose the following scenario that is consistent with the properties of these different components and with other observations of the Galaxy.

- (a) The *halo component* (H in Table 1; shown in dark purple) is extremely old yet has a relatively high mean metallicity, resembling the early history of m12m. This supports the idea that the majority of the halo component is from *one or two early, massive mergers* (e.g., Belokurov et al. 2018; Helmi et al. 2018; Myeong et al. 2019; Horta et al. 2021; Santistevan et al. 2020).
- (b) The most *metal-rich thick-disk-like component* (I2 in Table 1; shown in orange) has very high random energy yet still orbits nearly at the solar circular velocity, and appears to be made up mostly of stars with solar values of $[\alpha/Fe]$ or slightly higher with younger estimated ages. This resembles intermediate components in our m12f simulation that contain stars disturbed by a late merger. We propose that this component contains stars that have been heated by interactions with satellite galaxies such as Sagittarius (e.g., Villalobos & Helmi 2008; Ma et al. 2017; Antoja et al. 2018; El-Badry et al. 2018; Laporte et al. 2018; Sheffield et al. 2018).
- (c) Of the two *metal-poor thick-disk-like components*, one component (I3, shown in black) has a random energy distribution as wide as I2 and extends to a similarly high metallicity, but has a lower orbital velocity and contains far more alpha-enhanced stars. It also seems to contain stars with a preferentially negative V_Z , while the other component (I1, shown in rose) has mostly stars with a positive V_Z (Figure 15). I1 also has a slightly narrower V_{XZ} distribution and a slightly higher mean V_Y , but is otherwise elementally and spatially similar to I3. These two components comprise the velocity-asymmetric, alpha-enhanced thick disk (Widrow et al. 2012), with one component used to represent each side of the “wave.” Consistent with observations of the velocity asymmetry (e.g., Figure 3 of Widrow et al. 2012), we find that the negative-velocity component (I3) has a wider velocity spread and includes stars both above and below the disk, while the positive-velocity component (I1) is more coherent in velocity, more spatially confined, and slightly prefers positive z (Figure 15). In terms of formation mechanism, these components resemble the radially mixed thick disks seen in all three simulations, but given the MW’s quiet recent accretion history, they seem to most closely resemble the black component of the three m12i simulations.
- (d) The *thin disk component* (D in Table 1; shown in magenta) contains young, non-alpha-enhanced stars consistent with recent star formation in a cold gas disk, after a late influx of gas that reset the local $[\alpha/Fe]$ ratio

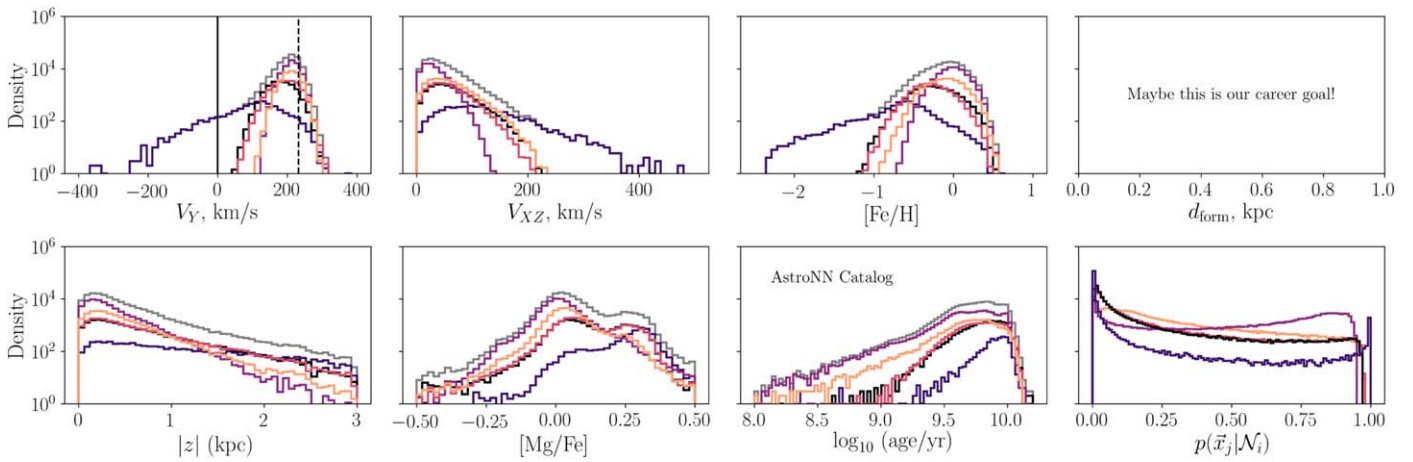


Figure 13. Distribution of various properties of stars in each component of the best-fit mixture model, for the real Gaia–APOGEE catalog. We do not have the d_{form} values and ages for the real catalog; APOGEE provides magnesium-to-iron ratios, Gaia provides heights above the disk plane at present day, and the astroNN Value-Added Catalog provides ages.

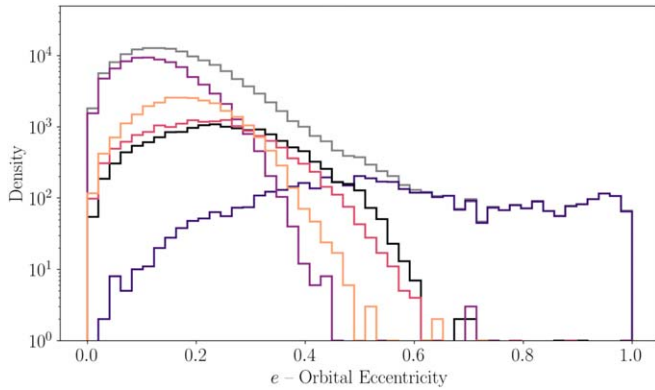


Figure 14. The distribution of orbital eccentricity from the astroNN catalog. Stars of the halo component have more elliptical orbits and thin disk stars are on circular orbits. The three thick disk components in this space are also different from each other.

(e.g., Mackereth et al. 2017, 2018, A. Wetzel et al. 2021, in preparation).

Finally from the weights τ in Table 1 we can estimate the proportion of stars with each of these origins in the APOGEE (i.e., evolved-star) view of our solar neighborhood. According to our mixture, the APOGEE sample is 4% halo (a), 22% interaction-heated (b), 27% radially mixed, asymmetric thick disk (c, d), and 46% thin disk (e). We caution that this breakdown is modulated by the APOGEE selection function; we plan to correct for this in future work using our new mock catalogs as a testbed.

6.2. Further Applications of the Mixture Model

Our strategy offers a probabilistic approach for selecting stars that are likely to belong to a particular population, as an alternative to making hard cuts on the data, through analogy with state-of-the-art simulations of galaxies. The resulting model can be used either to classify and study the stars within the modeled data set (in this case the Gaia–APOGEE catalog), or to predict the composition of other data sets that measure the same parameters.

The model in Table 1 offers several options for use as a classifier, of which the optimal method will depend on the application. One option is simply to assign each star to the component that produces the highest probability for that star, as we have done here to produce Figures 5 and 12. Another is to use the upper right-hand panel of Figure 12 to set a limit on the acceptable probability that a star belongs in a particular component (particularly useful for strongly bimodal probability distributions like the one seen for the halo). Yet another option is to include *all* stars in the analysis of a component of interest, weighted by their probability of belonging to that component. This last approach showcases the real power of the GMM strategy, since it can handle the ambiguity of weakly assigned stars on the same footing with those that more clearly belong to a particular population. A short Python script implementing the model will be provided as part of the public data release (with Sloan DR17).

A trained Gaussian mixture model can be used to identify members of its components in any data set where the same features are available. To identify structural components of the Galaxy, we built a mixture model using 3D kinematics and metallicities of stars observed with Gaia–APOGEE, but this model can be used to probabilistically classify any star with a measured 3D velocity and $[\text{Fe}/\text{H}]$. Care should be taken to use the same metallicity scale when combining data from multiple sources. The “validity volume,” where the GMM has been constructed, should be considered as well. An already trained mixture model is especially useful for identifying members of Galactic components in smaller surveys, or ones that have a more complicated selection function and are thus unlikely to independently constrain a mixture model. This is likely to be especially useful given that while Gaia provides all-sky coverage for proper motions and distances, radial velocities and elemental abundances for most stars are determined by an ensemble of ground-based spectroscopic surveys, each with a different selection function, sky coverage, and target depth. Mixture modeling will thus supply a crucial tool to relate stars in the same population that have been observed by different instruments, unifying our chemodynamical view of the Galaxy.

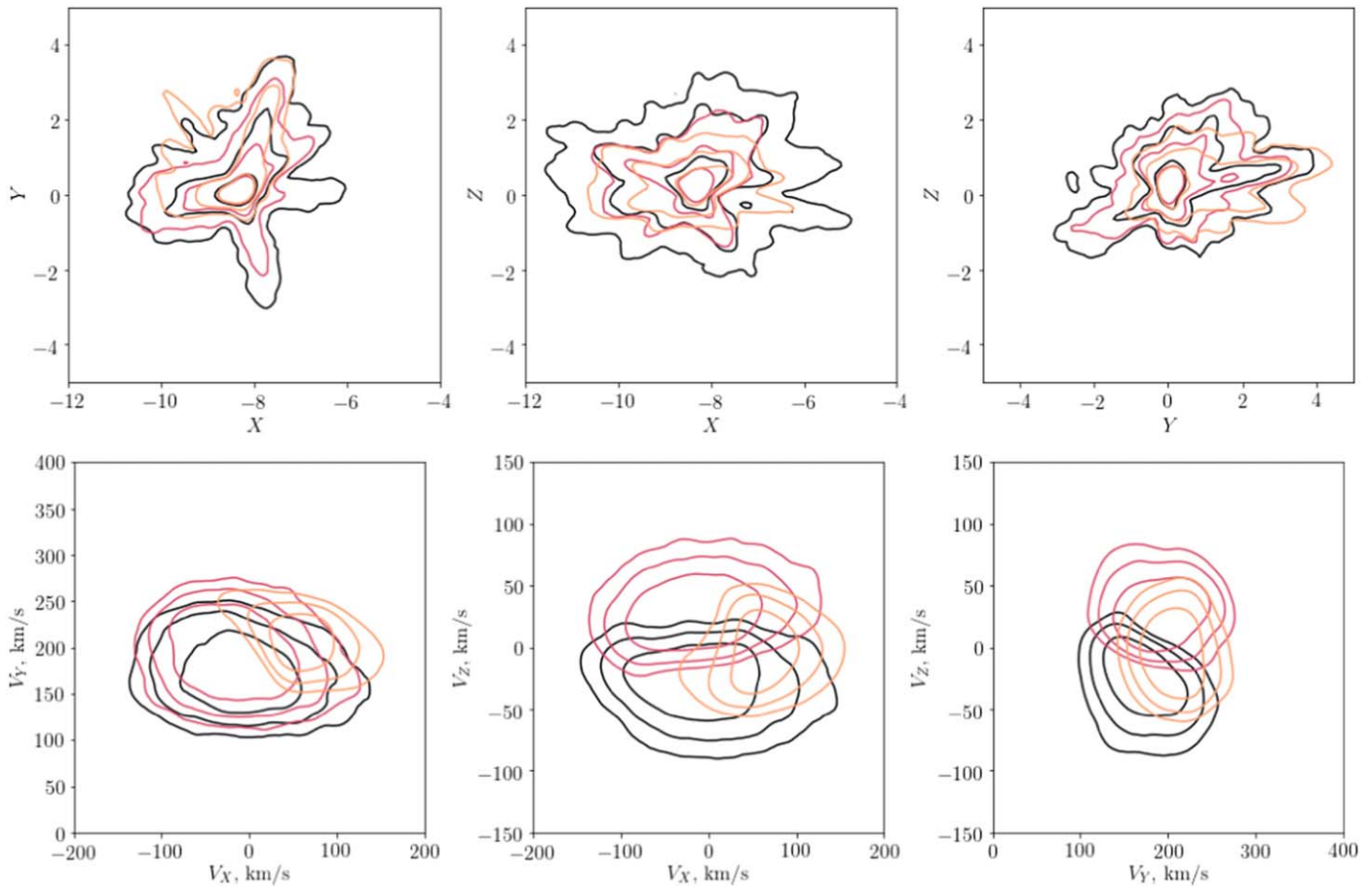


Figure 15. Spatial (top row) and velocity (bottom row) distributions of components I1, I2, and I3 in the mixture model of the real data. The black component (I3) contains stars with a preferentially negative V_Z , while the rose component (I1) has mostly stars with a positive V_Z . These two components comprise the velocity-asymmetric, alpha-enhanced thick disk, with one component used to represent each side of the “wave.” The most metal-rich thick-disk-like component (I2; orange) has very high random energy orbits nearly at the solar circular velocity, and contains stars that have been heated by interactions with satellite galaxies such as Sagittarius.

The authors thank Leandro Beraldo e Silva, Tobias Buck, José G. Fernández-Trincado, Kathryn V. Johnston, Ivan Minchev, Melissa Ness, Jeff Newman, the Milky Way as A Galaxy (MWAG) working group of SDSS-IV, and the Dynamics group at the Center for Computational Astrophysics, Flatiron Institute for valuable insights in preparing this paper. We also would like to thank the referee for the thoughtful and insightful comments.

F.N. acknowledges support from the National Science Foundation Graduate Research Fellowship under grant No. DGE-1845298.

R.E.S. acknowledges support from NASA grant 19-ATP19-0068, NSF grant AST-2009828, and HST-AR-15809 from the Space Telescope Science Institute (STScI), which is operated by AURA, Inc., under NASA contract NAS5-26555.

We performed this work in part at the Aspen Center for Physics, supported by NSF grant PHY-1607611, and at KITP, supported by NSF grant PHY-1748958.

A.W. received support from NASA through ATP grants 80NSSC18K1097 and 80NSSC20K0513; HST grants GO-14734, AR-15057, AR-15809, and GO-15902 from STScI; a Scialog Award from the Heising-Simons Foundation; and a Hellman Fellowship.

J.T.M. acknowledges support from the Banting Postdoctoral Fellowship program administered by the Government of Canada, and a CITA/Dunlap Institute fellowship. The Dunlap

Institute is funded through an endowment established by the David Dunlap family and the University of Toronto.

A.B. acknowledges support from NASA through HST grant HST-GO-15930.

S.L.M. acknowledges support from the UNSW Scientia Fellowship program and the Australian Research Council through grant DP180101791.

C.A.F.G. was supported by NSF through grants AST-1715216 and CAREER award AST-1652522; by NASA through grant 17-ATP17-0067; by STScI through grant HST-AR-16124.001-A; and by the Research Corporation for Science Advancement through a Cottrell Scholar Award and a Scialog Award.

The authors thank the Flatiron Institute Scientific Computing Core for providing computing resources that made this research possible, and especially for their hard work facilitating remote access during the pandemic. Analysis for this paper was carried out on the Flatiron Institute’s computing cluster *rusty*, which is supported by the Simons Foundation.

Simulations used in this work were run using XSEDE supported by NSF grant ACI-1548562, Blue Waters via allocation PRAC NSF.1713353 supported by the NSF, and NASA HEC Program through the NAS Division at Ames Research Center.

This project was developed in part at the 2017 Heidelberg Gaia Sprint, hosted by the Max-Planck-Institut für Astronomie, Heidelberg.












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 the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

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

SDSS 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, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/ University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical 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.

Software: Numpy (van der Walt et al. 2011; Harris et al. 2020), matplotlib (Hunter 2007), scikit-learn (Pedregosa et al. 2011), astropy (Astropy Collaboration et al. 2013, 2018), gizmo-analysis (Wetzel & Garrison-Kimmel 2020a), halo-analysis (Wetzel & Garrison-Kimmel 2020b).

ORCID iDs

Farnik Nikakhtar  <https://orcid.org/0000-0002-3641-4366>
 Robyn E. Sanderson  <https://orcid.org/0000-0003-3939-3297>
 Andrew Wetzel  <https://orcid.org/0000-0003-0603-8942>
 Sanjib Sharma  <https://orcid.org/0000-0002-0920-809X>
 Rachael Beaton  <https://orcid.org/0000-0002-1691-8217>
 Vijith Jacob Poovelil  <https://orcid.org/0000-0002-9831-3501>
 Gail Zasowski  <https://orcid.org/0000-0001-6761-9359>
 Ana Bonaca  <https://orcid.org/0000-0002-7846-9787>
 Sarah Martell  <https://orcid.org/0000-0002-3430-4163>
 Henrik Jönsson  <https://orcid.org/0000-0002-4912-8609>
 Claude-André Faucher-Giguère  <https://orcid.org/0000-0002-4900-6628>

References

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, *ApJ*, **597**, 21
 Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2020, *ApJS*, **249**, 3

Akaike, H. 1974, *ITAC*, **19**, 716
 An, D., & Beers, T. C. 2020, *ApJ*, **897**, 39
 Antoja, T., Figueras, F., Fernández, D., & Torra, J. 2008, *A&A*, **490**, 135
 Antoja, T., Figueras, F., Torra, J., Valenzuela, O., & Pichardo, B. 2010, in *LNEA 4 ed. A. Ulla & M. Manteiga (Vigo, Spain: Tórculo Press)*, 13
 Antoja, T., Helmi, A., Romero-Gómez, M., et al. 2018, *Natur*, **561**, 360
 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123
 Astropy Collaboration, Robitaille, T. P., & Tollerud, E. J. 2013, *A&A*, **558**, A33
 Beane, A., Sanderson, R. E., Ness, M. K., et al. 2019, *ApJ*, **883**, 103
 Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, **478**, 611
 Bensby, T., Feltzing, S., & Lundström, I. 2003, *A&A*, **410**, 527
 Beraldo e Silva, L., Debattista, V. P., Nidever, D., Amarante, J. A. S., & Garver, B. 2021, *MNRAS*, **502**, 260
 Bird, J. C., Kazantzidis, S., Weinberg, D. H., et al. 2013, *ApJ*, **773**, 43
 Bird, J. C., Loebman, S. R., Weinberg, D. H., et al. 2021, *MNRAS*, **503**, 1815
 Bonaca, A., Conroy, C., Wetzel, A., Hopkins, P. F., & Kereš, D. 2017, *ApJ*, **845**, 101
 Bournaud, F., Elmegreen, B. G., & Martig, M. 2009, *ApJL*, **707**, L1
 Bovy, J., Rix, H.-W., & Hogg, D. W. 2012a, *ApJ*, **751**, 131
 Bovy, J., Rix, H.-W., Liu, C., et al. 2012b, *ApJ*, **753**, 148
 Bovy, J., Rix, H.-W., Schlafly, E. F., et al. 2016, *ApJ*, **823**, 30
 Bowen, I. S., & Vaughan, A. H. 1973, *ApOpt*, **12**, 1430
 Brook, C. B., Kawata, D., Gibson, B. K., & Freeman, K. C. 2004, *ApJ*, **612**, 894
 Brook, C. B., Stinson, G. S., Gibson, B. K., et al. 2012, *MNRAS*, **426**, 690
 Buck, T., Obreja, A., Macciò, A. V., et al. 2020, *MNRAS*, **491**, 3461
 Deason, A. J., Belokurov, V., Koposov, S. E., et al. 2017, *MNRAS*, **470**, 1259
 Dempster, A. P., Laird, N. M., & Rubin, D. B. 1977, *J. Roy. Statist. Soc. Ser. B*, **39**, 1, <https://www.jstor.org/stable/2984875>
 El-Badry, K., Bland-Hawthorn, J., Wetzel, A., et al. 2018, *MNRAS*, **480**, 652
 El-Badry, K., Wetzel, A., Geha, M., et al. 2016, *ApJ*, **820**, 131
 Elmegreen, B. G., Bournaud, F., & Elmegreen, D. M. 2008, *ApJ*, **688**, 67
 Escala, I., Wetzel, A., Kirby, E. N., et al. 2018, *MNRAS*, **474**, 2194
 Faucher-Giguère, C.-A. 2018, *MNRAS*, **473**, 3717
 Fernández-Trincado, J. G., Beers, T. C., Placco, V. M., et al. 2019, *ApJL*, **886**, L8
 Forbes, J., Krumholz, M., & Burkert, A. 2012, *ApJ*, **754**, 48
 Freeman, K., & Bland-Hawthorn, J. 2002, *ARA&A*, **40**, 487
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, **616**, A1
 Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, **595**, A1
 García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, *AJ*, **151**, 144
 Garrison-Kimmel, S., Hopkins, P. F., Wetzel, A., et al. 2018, *MNRAS*, **481**, A133
 Gerhard, O., & Martínez-Valpuesta, I. 2012, *ApJL*, **744**, L8
 Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, *AJ*, **131**, 2332
 Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
 Hayden, M. R., Bland-Hawthorn, J., Sharma, S., et al. 2020, *MNRAS*, **493**, 2952
 Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, *ApJ*, **808**, 132
 Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Natur*, **563**, 85
 Helmi, A., Veljanoski, J., Breddels, M. A., Tian, H., & Sales, L. V. 2017, *A&A*, **598**, A58
 Herzog-Arbeitman, J., Lisanti, M., & Necib, L. 2018, *JCAP*, **4**, 52
 Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, *MNRAS*, **480**, 800
 Horta, D., Schiavon, R. P., Mackereth, J. T., et al. 2021, *MNRAS*, **500**, 1385
 Hunter, J. D. 2007, *CSE*, **9**, 90
 Ishigaki, M. N., Aoki, W., & Chiba, M. 2013, *ApJ*, **771**, 67
 Iwamoto, K., Brachwitz, F., Nomoto, K., et al. 1999, *ApJS*, **125**, 439
 Izzard, R. G., Tout, C. A., Karakas, A. I., & Pols, O. R. 2004, *MNRAS*, **350**, 407
 Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, *AJ*, **160**, 120
 Kafle, P. R., Sharma, S., Robotham, A. S. G., et al. 2017, *MNRAS*, **470**, 2959
 Laporte, C. F. P., Johnston, K. V., Gómez, F. A., Garavito-Camargo, N., & Besla, G. 2018, *MNRAS*, **481**, 286
 Leung, H. W., & Bovy, J. 2019, *MNRAS*, **489**, 2079
 Loebman, S. R., Roškar, R., Debattista, V. P., et al. 2011, *ApJ*, **737**, 8
 Ma, X., Hopkins, P. F., Wetzel, A. R., et al. 2017, *MNRAS*, **467**, 2430
 Mackereth, J. T., & Bovy, J. 2020, *MNRAS*, **492**, 3631
 Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019, *MNRAS*, **489**, 176
 Mackereth, J. T., Bovy, J., Schiavon, R. P., et al. 2017, *MNRAS*, **471**, 3057
 Mackereth, J. T., Crain, R. A., Schiavon, R. P., et al. 2018, *MNRAS*, **477**, 5072
 Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, **154**, 94

- Mannucci, F., Della Valle, M., & Panagia, N. 2006, *MNRAS*, **370**, 773
- Marigo, P. 2001, *A&A*, **370**, 194
- Minchev, I., Martig, M., Streich, D., et al. 2015, *ApJL*, **804**, L9
- Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2015, *MNRAS*, **454**, 2691
- Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018, *ApJL*, **863**, L28
- Myeong, G. C., Vasiliev, E., Iorio, G., Evans, N. W., & Belokurov, V. 2019, *MNRAS*, **488**, 1235
- Nidever, D. L., Holtzman, J. A., Allende Prieto, C., et al. 2015, *AJ*, **150**, 173
- Nissen, P. E., & Schuster, W. J. 2010, *A&A*, **511**, L10
- Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, *NuPhA*, **777**, 424
- Nordström, B., Mayor, M., Andersen, J., et al. 2004, *A&A*, **418**, 989
- Obreja, A., Dutton, A. A., Macciò, A. V., et al. 2019, *MNRAS*, **487**, 4424
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *J. Mach. Learn. Res.*, **12**, 2825, <http://jmlr.org/papers/v12/pedregosa11a.html>
- Posti, L., Helmi, A., Veljanoski, J., & Breddels, M. 2018, *A&A*, **615**, A70
- Sanderson, R. E., Garrison-Kimmel, S., Wetzel, A., et al. 2018, *ApJ*, **869**, 12
- Sanderson, R. E., Wetzel, A., Loebman, S., et al. 2020, *ApJS*, **246**, 6
- Santistevan, I. B., Wetzel, A., El-Badry, K., et al. 2020, *MNRAS*, **497**, 747
- Santistevan, I. B., Wetzel, A., Sanderson, R. E., et al. 2021, *MNRAS*, **505**, 921
- Schwarz, G. 1978, *AnSta*, **6**, 461
- Schwarzschild, K. 1907, *NWGot*, **5**, 614
- Searle, L., & Zinn, R. 1978, *ApJ*, **225**, 357
- Sestito, F., Buck, T., Starkenburg, E., et al. 2021, *MNRAS*, **500**, 3750
- Sestito, F., Longeard, N., Martin, N. F., et al. 2019, *MNRAS*, **484**, 2166
- Sestito, F., Martin, N. F., Starkenburg, E., et al. 2020, *MNRAS*, **497**, L7
- Sharma, S., Hayden, M. R., & Bland-Hawthorn, J. 2021, *MNRAS*, **507**, 5882
- Sheffield, A. A., Price-Whelan, A. M., Tzanidakis, A., et al. 2018, *ApJ*, **854**, 47
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, **131**, 1163
- Sparre, M., Hayward, C. C., Feldmann, R., et al. 2017, *MNRAS*, **466**, 88
- Starkenburg, E., Oman, K. A., Navarro, J. F., et al. 2017, *MNRAS*, **465**, 2212
- Steinhaus, H. 1957, *Bull. Acad. Pol. Sci.*, Cl. III, **4**, 801, http://www.laurent-duval.eu/Documents/Steinhaus_H_1956_j-bull-acad-polon-sci_division_cmp-k-means.pdf
- Stern, J., Faucher-Giguère, C.-A., Fielding, D., et al. 2021, *ApJ*, **911**, 88
- Su, K.-Y., Hopkins, P. F., Hayward, C. C., et al. 2017, *MNRAS*, **471**, 144
- Thorndike, R. 1953, *Psychometrika*, **18**, 267
- van den Hoek, L. B., & Groenewegen, M. A. T. 1997, *A&AS*, **123**, 305
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22
- Venn, K. A., Irwin, M., Shetrone, M. D., et al. 2004, *AJ*, **128**, 1177
- Villalobos, Á., & Helmi, A. 2008, *MNRAS*, **391**, 1806
- Wetzel, A., & Garrison-Kimmel, S. 2020a, *GizmoAnalysis: Read and analyze Gizmo simulations*, Astrophysics Source Code Library, [ascl:2002.015](https://ui.adsabs.org/abs/2020ASCl..2002..015W)
- Wetzel, A., & Garrison-Kimmel, S. 2020b, *HaloAnalysis: Read and analyze halo catalogs and merger trees*, Astrophysics Source Code Library, [ascl:2002.014](https://ui.adsabs.org/abs/2020ASCl..2002..014W)
- Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, *ApJL*, **827**, L23
- Widrow, L. M., Gardner, S., Yanny, B., Dodelson, S., & Chen, H.-Y. 2012, *ApJL*, **750**, L41
- Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, *PASP*, **131**, 055001
- Yu, S., Bullock, J. S., Wetzel, A., et al. 2020, *MNRAS*, **494**, 1539
- Zasowski, G., Cohen, R. E., Chojnowski, S. D., et al. 2017, *AJ*, **154**, 198
- Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, *AJ*, **146**, 81