Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **493,** 5195–5207 (2020)
Advance Access publication 2020 February 3

doi:10.1093/mnras/stz3537

Ages and kinematics of chemically selected, accreted Milky Way halo stars

Payel Das [1]★ Keith Hawkins [2] and Paula Jofré [3]

[1] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK*
[2] *Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, Texas 78712-1205, USA*
[3] *Núcleo de Astronomía, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Ejército 441, Santiago de Chile*

Accepted 2019 December 11. Received 2019 November 25; in original form 2019 March 21

ABSTRACT

We exploit the [Mg/Mn]-[Al/Fe] chemical abundance plane to help identify nearby halo stars in the 14th data release from the APOGEE survey that have been accreted on to the Milky Way. Applying a Gaussian Mixture Model, we find a 'blob' of 856 likely accreted stars, with a low disc contamination rate of ∼7 per cent. Cross-matching the sample with the second data release from Gaia gives us access to parallaxes and apparent magnitudes, which place constraints on distances and intrinsic luminosities. Using a Bayesian isochrone pipeline, this enables us to estimate new ages for the accreted stars, with typical uncertainties of ∼20 per cent. This does not account for systematic uncertainties. Our new catalogue is further supplemented with estimates of orbital parameters. The blob stars span [Fe/H] between −2.5 to −0.5, and [Mg/Fe] between −0.1 to 0.5. They constitute ∼30 per cent of the metal-poor ([Fe/H] < −0.8) halo at [Fe/H] ∼ −1.4. Our new ages mainly range between 8 to 13 Gyr, with the oldest stars the metal-poorest, and with the highest [Mg/Fe] abundance. If the blob stars are assumed to belong to a single progenitor, the ages imply that star formation lasted 5 Gyr after which the system merged with our Milky Way around 8 Gyr ago. Dynamical arguments suggest that such a single progenitor would have had a total mass of ∼ 10^{11} $M_{\odot}$, similar to that found by other authors using chemical evolution models and simulations.

Key words: Galaxy: abundances – Galaxy: evolution – Galaxy: halo – Galaxy: kinematics and dynamics.

1 INTRODUCTION

The stellar halo contributes about 1 per cent of the stellar mass budget of the Milky Way, believed to be primarily in the form of old and metal-poor stars. Some of these stars may have formed *in situ*, but others arrived from disrupted satellite galaxies and globular clusters (Ibata, Gilmore & Irwin 1995; Majewski et al. 2003; Abadi, Navarro & Steinmetz 2006; Belokurov et al. 2007; Bell et al. 2008; Cooper et al. 2011). The engulfed systems disperse in positions and velocities but their ages and metallicities hardly alter (Bell et al. 2010). Therefore, an essential step towards distinguishing between potential origins of stars in the stellar halo is to analyse its chemodynamical and age structure.

The density profile of the stars in the halo, for example, reflects the accumulation of stellar mass there. Star counts of main-sequence turn-off stars (Bell et al. 2008; Sesar, Jurić & Ivezić 2011), RR Lyrae stars (Watkins et al. 2009; Sesar et al. 2013), BHB stars (Deason, Belokurov & Evans 2011), and K giants (Kafle et al. 2013) suggest a break in the number density around $R = 15$–25 kpc with a power-law index $\alpha \sim 2$–3 in the inner halo and $\alpha \sim 3.8$–5 further out. An examination of BHBs and blue stragglers from SDSS by Deason et al. (2014) found evidence that $\alpha \sim 6$ beyond 50 kpc and $\alpha \sim 6$–10 at still larger radii. Deason et al. (2013a), proposed that the existence of the 'break radius' in the Milky Way halo is associated with a 'pile up' of stellar apocenters at a comparable Galactocentric distance. The observed existence of a break radius in the Milky Way halo and the absence of such a break in the Andromeda galaxy (M31) suggests that the latter had a more prolonged accretion history than the former. More recently however, Xue et al. (2015) and Das & Binney (2016) showed with a sample of SEGUE K giants that if the flattening of the stellar-halo component is allowed to vary, a break is not required in the density profile.

The anisotropy profile of the stars preserves dynamical processes that had an impact on stellar orbits. If we consider the anisotropy parameter, $\beta = 1 - (\sigma_\theta^2 + \sigma_\phi^2)/2\sigma_r^2$, $\beta = -\infty$ for purely circular orbits and $\beta = 1$ for purely radial orbits. Chiba & Yoshii (1998) obtained $\beta = 0.52$ for a small sample of local halo red giants and RR Lyrae observed by the *Hipparcos* space mission. Using the SDSS Stripe 82 proper motions, and thus going deeper, Smith et al. (2009) measured $\beta = 0.69$ using ∼2000 nearby subdwarfs. Combining the SDSS observations with the digitized photographic plate measurements, Bond et al. (2010) increased the stellar halo sample further and derived $\beta = 0.67$. For slightly larger volumes

★ E-mail: payeldas.astro@gmail.com

probed with more luminous tracers, similar values of $\beta \sim 0.5$ were obtained (see e.g. Deason et al. 2012; Kafle et al. 2012). Beyond 15–20 kpc from the Sun, there have been several attempts to measure β from just line-of-sight velocity measurements (see e.g. Sirko et al. 2004; Williams & Evans 2015), although Hattori et al. (2017) showed that this may depend heavily on the functional form assumed for the β parameter. Only a few studies rely on proper motion measurements. These include a very small number of stars with proper motions measured with the *Hubble Space Telescope* finding a drop to $\beta \sim 0$ at ∼20 kpc (see Deason et al. 2013b; Cunningham et al. 2016). Das & Binney (2016) and Das, Williams & Binney (2016) fit equilibrium dynamical models to K giants and blue horizontal branch (BHB) stars using proper motions from UCAC5. They find orbits becoming more isotropic going outwards in the K giant population, and possibly also in the BHB population. Both populations are found to be only mildly radially anisotropic in the inner halo.

The assembly history of the stellar halo is also reflected in its abundance structure. Carollo et al. (2007) claimed a negative metallicity gradient, with the outer halo significantly more metal poor than the local halo. This was later confirmed by several studies (e.g. de Jong et al. 2010; Kafle et al. 2013; Allende Prieto et al. 2014; Chen et al. 2014), with some also finding metal-poorer stars in retrograde motion and metal-richer stars in prograde motion (Carollo et al. 2007; Kafle et al. 2013). Schönrich, Asplund & Casagrande (2014) argued however that metal-poorer stars can be seen at greater distances than metal-richer stars, and if this effect is not correctly included in the adopted selection function, a metallicity gradient can be erroneously inferred. The latest studies on the topic by Xue et al. (2015) and Das & Binney (2016), who take the selection function into account in a sample of SEGUE K giant stars, only find a modest metallicity gradient.

The claims of a break in the density profile, a changing anisotropy profile, and a gradient in the metallicity have been interpreted as evidence for the existence of a dual halo, comprised of both *in-situ* and *accreted* components in the stellar halo. The metal-rich, in-situ stars dominate the inner halo, and metal-poor, accreted stars dominate the outer halo. A number of cosmological simulations support a similar composition of the stellar halo (e.g. Zolotov et al. 2009; Font et al. 2011; Tissera, White & Scannapieco 2012; Pillepich, Madau & Mayer 2015). The stars in the halo have generally been found to be old, 10–12 Gyr in age, born early in the Universe (e.g. Jofré & Weiss 2011; Kalirai 2012), Nissen & Schuster (2010) and Schuster et al. (2012) identified the presence of a high-α population of metal-poor stars in the Solar Neighbourhood with ages 2–3 Gyr larger than a low-α population at similar metallicities. The high-α halo stars were also found at smaller radii and heights and a range of eccentricities. The low-α halo stars, on the other hand, are clumped at eccentricities greater than 0.85. Hawkins et al. (2014) further showed that a bifurcation appears in the age-metallicity relation such that in the low-metallicity regime the α-rich and α-poor populations are coeval, while in the high-metallicity regime the α-rich population is older than the α-poor population. This suggests that the α-rich halo population, which has a shallow age-metallicity relation, could have formed in a rapid event with a high star formation rate (SFR) such as the thick disc of our Milky Way, while the α-poor stars were formed in an environment with a slower chemical evolution timescale such as in dwarf spheroidal galaxies. Similar conclusions have been reached by Fernández-Alvar et al. (2018) and Haywood et al. (2018) in their analysis of the chemical compositions and kinematics of the metal-rich nearby halo using data from the 14$^{\text{th}}$ data release (DR14) from the APOGEE survey combined with the second data release (DR2) from Gaia (Gaia Collaboration et al. 2018). Hayes et al. (2018) further showed that the metal-rich low-[Mg/Fe] halo stars were distinct in multiple abundance planes constructed with APOGEE DR14.

A number of works published in the last year have associated the metal-rich, α-poor inner halo specifically with the debris of a single, relatively massive stellar system that was accreted on to the Milky Way (Helmi et al. 2018; Belokurov et al. 2018; Mackereth et al. 2019). Helmi et al. (2018) examine stars from the cross-match between Gaia DR2 and APOGEE DR14 that are primarily counter-rotating and loosely bound. They suggest that the selected stars belong to a single massive progenitor, 'Gaia-Enceladus', due to the tightness of their sequence in the [α/Fe]-[Fe/H] plane. The system has a large spread in metallicity, implying it had an extended star formation history (SFH). Using chemical evolution models, they estimate star formation produced $0.3\,\mathrm{M_\odot\,yr^{-1}}$ and lasted about 2 Gyr, implying that the progenitor has a stellar mass of $\sim 6 \times 10^8\,\mathrm{M_\odot}$, comparable to the Large and Small Magellanic Clouds.

Belokurov et al. (2018) used a large sample of Main Sequence stars extracted from the first data release (DR1) of Gaia together with SDSS, to study the kinematic properties of the local stellar population within ∼10 kpc from the Sun. They estimate distances from the photometry using the relations in Ivezić et al. (2008) and proper motions from combining SDSS data and Gaia DR1. They examine the tangential (v_θ) and radial (v_r) components of the sample sliced by metallicity and vertical distance from the plane (z). They find two components in the metal-richer stars ($-1.33 <$ [Fe/H] < -1.00), one co-rotating with the disk and with a small spread in v_r, and the other showing minimal rotation and a large spread in v_r. By examining a set of cosmological simulations, they claim that the very high radial anisotropy of the 'Gaia Sausage' stars ($\beta \sim 0.9$) is inconsistent with a series of minor accretion events but rather is the result of a massive satellite that sinks deeper into the potential well of the Milky Way as a result of dynamical friction. The radialization is then further enhanced in the presence of a growing disc as a result of loss in angular momentum. Mackereth et al. (2019) came to similar conclusions based on their analysis of EAGLE simulations. They show that the median orbital eccentricities of debris are largely unchanged since merger time, implying that this accretion event likely happened at $z \sim 1.5$. Using Gaia DR2, Myeong et al. (2018) found 12 halo GCs to be on highly eccentric orbits, at $e \sim 0.85$, thus making them consistent with an origin in a single massive accretion event.

With the abundant spectroscopic and astrometric data for a large number of halo stars, we are now in a good position to estimate individual ages, and as such, obtain a complete picture of the nature of accreted halo stars. In this paper, we revisit the cross-match between Gaia DR2 and APOGEE DR14 and develop a new method for identifying a clean accreted sample of stars that is kinematically unbiased (subject to the selection function of the survey). We calculate new ages for these stars, and re-examine their distribution with respect to a number of chemical abundance planes, age, and kinematic variables.

2 THE DATA

Here we introduce the spectroscopic data that will be used to select accreted stars, and the kinematic data that will be used to further analyse the sample.

2.1 Chemical abundances, spectral parameters, and line-of-sight velocities

We use abundances and line-of-sight velocities measured by the near-infrared APOGEE survey. Data Release 14 (DR14, Abolfathi et al. 2018; Holtzman et al. 2018) contains high S/N, moderate resolution ($R = \lambda/\Delta\lambda \sim 22\,500$) spectra, line-of-sight velocities, stellar photospheric parameters, and up to 19 element abundances for over 270 000 stars in the H-band (1.5–1.7 μm). Most stars are red giants, with a significant contribution from red dwarf stars.

Spectra are reduced and analysed through the APOGEE data reduction pipeline (Nidever et al. 2015), and the APOGEE Stellar Parameters and Chemical Abundances Pipeline (ASPCAP, García Pérez et al. 2016). ASPCAP uses a library of synthetic stellar spectra (Zamora et al. 2015) that is precomputed using a customised linelist (Shetrone et al. 2015) to measure stellar parameters, 19 element abundances, and heliocentric line-of-sight velocities (Holtzman et al. 2018). There is a two-step process. First, the stellar parameters, $T_{\rm eff}$, $\log g$, v_μ (micro-turbulent velocity), [M/H], [α/M], [C/Fe], and [N/Fe] are determined via a global fit. The individual element abundances are then calculated by adjusting the [M/H] ([C/Fe] and [N/Fe] for C and N, and [α/M] for α elements) of the best-fit spectrum. The abundances for species other than C and N can then be calibrated internally relative to open cluster observations, to account for systematic abundance variations with $T_{\rm eff}$. Surface gravities are calibrated using stars with independent asteroseismology determinations from the Kepler mission (Haas et al. 2010). We apply the following quality control cuts, requiring:

(i) `ASPCAPFLAG = 0`,
(ii) `STARFLAG = 0`,
(iii) `M_H, ALPHA_M, MG_FE, C_FE, N_H, MN_FE, and AL_FE must be known`,
(iv) `M_H_ERR, ALPHA_M_ERR, MG_FE_ERR, C_FE_ERR, N_H_ERR, MN_FE_ERR, and AL_FE_ERR must be less than 0.15`,
(v) `TEFF > 4000 K`,
(vi) `LOGG > 0.5 dex`,

where `ASPCAPFLAG` and `STARFLAG` are flags in the APOGEE data that report potential issues with the star and/or with the specific parameter determination process for that star. `M_H`, `ALPHA_M`, `MG_FE`, `C_FE`, `N_H`, `MN_FE`, and `AL_FE` are the metallicity and abundance ratios, respectively, as measured in the APOGEE pipeline. We require these particular abundance ratios to be known as they form the basis of our chemical selection criteria for studying the halo stellar populations. We also impose a maximum uncertainty of 0.15 on these abundance ratios. Otherwise tails towards more extreme chemical abundances can appear, which can be difficult to capture in the Gaussian Mixture Models (GMM) discussed in Section 4. `TEFF` and `LOGG` are the APOGEE effective temperature and surface gravity, respectively. Spectral parameters and abundances are generally more reliable for the imposed ranges on effective temperature and $\log g$. Imposing these quality cuts reduces the sample to 136 212 stars.

2.2 Parallaxes and proper motions

The second data release of the ESA-Gaia mission, Gaia DR2 (Gaia Collaboration et al. 2018), provides five-parameter astrometry (proper motions, positions and parallaxes) for over 1.3 billion objects in the Milky Way. Many improvements were made to the data-processing pipeline between Gaia DR1 (Lindegren et al. 2016) and DR2 that have reduced the uncertainty on astrometric parameters to ∼0.2 to 0.3 mas for stars in the middle of the covered range of magnitudes (going up to ∼2 mas for the faintest sources). We use the catalogue provided by Sanders & Das (2018), who perform a 5″ radius cross-match between Gaia DR2 and six ground-based spectroscopic surveys (APOGEE, GALAH, Gaia-ESO, RAVE, LAMOST, SEGUE; approximately three million stars) by utilising the Gaia proper motions and accounting for the respective epochs of the surveys. They extract the parallax, proper motions, and the uncertainty covariance matrix for the astrometry from the Gaia DR2 source catalogue.

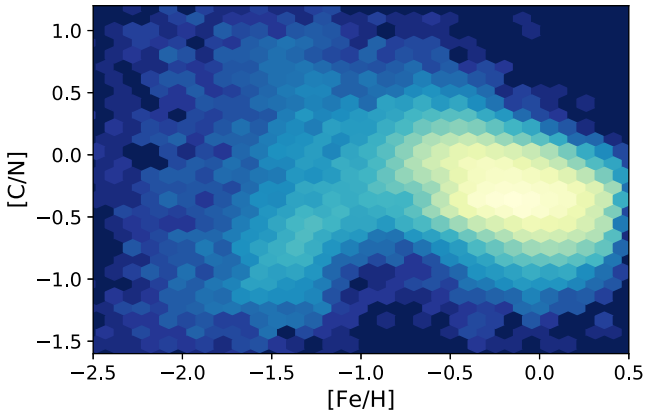

Figure 1. Two-dimensional distribution of APOGEE DR14 stars in the plane of [Fe/H] and [C/N].

3 MASSES, AGES, DISTANCES, AND ORBITAL PARAMETERS

Here we discuss the method for obtaining the age and distance of each star using spectroscopic mass estimates and isochrones. We also discuss how we derive orbital parameters for each star.

3.1 Spectroscopic mass estimates

Sanders & Das (2018) supplement their catalogue with spectroscopic mass estimates for stars metal-richer than −1.5. The estimates are based on the relation between the mass of a giant star and [C/N], which has been discussed several times in the literature (e.g. Masseron & Gilmore 2015; Martig et al. 2016; Ness et al. 2016; Das & Sanders 2019; Lagarde et al. 2019). Adopting the procedure presented in Das & Sanders (2019), they build a Bayesian artificial neural network (ANN) that predicts mass estimates and associated uncertainties informed by asteroseismology data (Vrard, Mosser & Samadi 2016), from spectroscopic parameters $T_{\rm eff}$, $\log g$, [M/H], [α/M], [C/Fe], and [N/Fe], and their uncertainties.

Fig. 1 shows the distribution of the APOGEE DR14 stars in the [Fe/H]-[C/N] plane. Stellar populations become richer in metals with time, and the mass of red giants reflects the time taken to complete the Main Sequence. Therefore, in red giants, [Fe/H] is roughly correlated with age and mass. In Fig. 1, [C/N] decreases with metallicity along two sequences for stars metal-richer than approximately −0.8. The two sequences are likely the thick and thin disks of the Milky Way. Metal-poorer stars are likely to mainly belong to the halo and therefore originate in different stellar systems. We therefore do not expect a regular sequence with metallicity as seen in the thick and thin disk stars. However, the scatter in

their [C/N] is far greater than seen in the thick and thin disk stars, which is unexpected for stars that should generally be old. This consolidates recent work (Shetrone et al. 2019) that suggests efficient thermohaline mixing in metal-poor stars may lead to a large spread in [C/N] at low metallicities. Inhomogeneities in the ISM (e.g. Revaz & Jablonka 2012) also become increasingly important at these low metallicities, e.g. stars formed close to SNII will have very different abundances from those further away. This implies that [C/N] is a good age and mass indicator only in red giants metal-richer than -0.8. We therefore only use the spectroscopic mass estimates from Sanders & Das (2018) for stars metal-richer than -0.8.

3.2 Ages and distances

Sanders & Das (2018) employed the Bayesian isochrone pipeline of Burnett & Binney (2010) to derive the distance, metallicity, mass, age, and line-of-sight extinction probability distribution functions (PDFs) of each star from the combined catalogues and the spectroscopic mass estimates for stars metal-richer than -1.5. Here, we use this catalogue, but rederive the distance, metallicity, mass, and age PDFs for stars metal-poorer than -0.8 and metal-richer than -1.5, omitting the unreliable spectroscopic mass estimates. We use a similar Bayesian isochrone pipeline presented in MADE (Das & Sanders 2019). A key difference in the MADE pipeline is that it assumes a point estimate for the line-of-sight extinction from previous studies rather than make a new estimate. We also account for the systematic parallax offset of 0.03 mas (Lindegren et al. 2018), because this may have a minor effect on the ages of halo stars for which parallaxes are smaller. The effect of the offset is small and we therefore do not worry about stars metal-poorer than -1.5, for which the offset has not been accounted for.

The parameters, ϕ^i, of the star i estimated from the Bayesian isochrone pipeline are given by

$$\phi^i = (\mathcal{M}^i, \tau^i, [\mathrm{M/H}]^i, s^i), \tag{1}$$

where $\mathcal{M}$ is the *initial* mass, τ is the age, [M/H] is the metallicity, and s is the distance. These parameters predict the following observed properties of the star i

$$v^i = (m^i, H^i, (J-K_s)^i, \varpi^i), \tag{2}$$

where m is the current mass, H is the H-band magnitude, $J-K_s$ is a colour, and ϖ is the parallax. The accompanying measured observed stellar properties are denoted by $\tilde{v}^i$. These are associated with measurement uncertainties, $\tilde{\rho}^i$.

Applying Bayes' law to each star gives

$$p(\phi^i|\tilde{v}^i, l^i, b^i) = \frac{p(\tilde{v}^i|\phi^i, l^i, b^i)p(\phi^i|l^i, b^i)}{p(\tilde{v}^i)}, \tag{3}$$

where (l^i, b^i) are the predicted sky positions of star i in Galactic coordinates (which we assume to be the same as the observed sky positions), and $p(\tilde{v}^i)$ is an unimportant normalization. MADE does not estimate extinctions, but instead corrects the apparent magnitudes of star i for extinction using the state-of-the-art `Combined15` map compiled by Bovy et al. (2016) in the `mwdust` package. The likelihood of the star's observed properties, $p(\tilde{v}^i|\phi^i, l^i, b^i)$, is assumed to be the product of the separate likelihoods. Each likelihood is represented by a Gaussian distribution

$$G(\tilde{v}^i_j, v^i_j, \tilde{\rho}^i_j) = \frac{1}{\sqrt{2\pi}\tilde{\rho}^i_j}\exp\left(-\frac{(\tilde{v}^i_j - v^i_j)^2}{2(\tilde{\rho}^i_j)^2}\right). \tag{4}$$

Thus

$$p(\tilde{v}^i|\phi^i, l^i, b^i) = \prod_j G(\tilde{v}^i_j, v^i_j, \tilde{\rho}^i_j). \tag{5}$$

The pipeline employs PARSEC isochrones v1.1 (assuming a mass-loss efficiency, $\eta = 0.2$, Bressan et al. 2012) evaluated for 57 metallicities ranging between -2.192 and 0.696, and 353 ages ranging between $\log_{10}\tau = 6.60$ and 10.12 (i.e. a spacing of $\Delta\log_{10}\tau = 0.01$) for which we create a dictionary of interpolants in PYTHON that returns luminosity, $\log g$, $T_{\rm eff}$ and apparent magnitudes, given the metallicity, age, and mass of a star. We use a prior $(\phi^i|l^i, b^i)$ informed by the Milky Way model presented in Das & Sanders (2019) (the superscript i is omitted in the following)

$$p(\mathcal{M}, \tau, [\mathrm{M/H}], s|l, b) = s^2\epsilon(\mathcal{M})\sum_{k=1}^{4} p_k([\mathrm{M/H}])p_k(\tau)p_k(R, z), \tag{6}$$

where $k = 1, 2, 3, 4$ correspond to a bulge, thin disc, thick disc, and stellar halo respectively. The s^2 term accounts for the Jacobian of the transformation of spatial coordinates, and $\epsilon(\mathcal{M})$ is the Kroupa (Kroupa, Tout & Gilmore 1993) initial mass function (IMF)

$$\epsilon(\mathcal{M}) = \begin{cases} 0.035\mathcal{M}^{-1.5} & \text{if } 0.08 \leq \mathcal{M}/\mathcal{M}_\odot < 0.5, \\ 0.019\mathcal{M}^{-2.2} & \text{if } 0.5 \leq \mathcal{M}/\mathcal{M}_\odot < 1.0, \\ 0.019\mathcal{M}^{-2.7} & \text{if } \mathcal{M}/\mathcal{M}_\odot \geq 1.0. \end{cases} \tag{7}$$

The priors for the four components are

Bulge ($k = 1$):

$$p_1([\mathrm{M/H}]) = G([\mathrm{M/H}], \mu_{[\mathrm{M/H}],1}, \sigma_{[\mathrm{M/H}],1}),$$

$$p_1(\tau) = G(\tau, \mu_{\tau,1}, \sigma_{\tau,1}),$$

$$p_1(R, z) \propto \frac{(1+m)^{(\gamma-\delta)}}{m^\gamma}\exp[-(mr_0/r_t)^2],$$

$$\text{where } m(R, z) = \sqrt{(R/r_0)^2 + (z/qr_0)^2}. \tag{8}$$

Thin disc ($k = 2$):

$$p_2([\mathrm{M/H}]) = G([\mathrm{M/H}], \mu_{[\mathrm{M/H}],2}, \sigma_{[\mathrm{M/H}],2}),$$

$$p_2(\tau) \propto \begin{cases} \exp(\frac{\tau}{8.4}) & \text{if } \tau/\mathrm{Gyr} \leq 8 \\ 2.6\exp\left(-0.5\frac{(\tau-8)^2}{1.5^2}\right) & \text{if } \tau/\mathrm{Gyr} > 8 \end{cases}$$

$$p_2(R, z) \propto \exp\left(-\frac{R}{R_{\mathrm{d},2}} - \frac{|z|}{z_{\mathrm{d},2}}\right). \tag{9}$$

Thick disc ($k = 3$):

$$p_3([\mathrm{M/H}]) = G([\mathrm{M/H}], \mu_{[\mathrm{M/H}],3}, \sigma_{[\mathrm{M/H}],3}),$$

$$p_3(\tau) = G(\tau, \mu_{\tau,3}, \sigma_{\tau,3}),$$

$$p_3(R, z) \propto \exp\left(-\frac{R}{R_{\mathrm{d},3}} - \frac{|z|}{z_{\mathrm{d},3}}\right). \tag{10}$$

Stellar halo ($k = 4$):

$$p_4([\mathrm{M/H}]) = G([\mathrm{M/H}], \mu_{[\mathrm{M/H}],4}, \sigma_{[\mathrm{M/H}],4}),$$

$$p_4(\tau) \propto G(\tau, \mu_{\tau,4}, \sigma_{\tau,4})$$

$$p_4(R, z) \propto r^{-3.39}. \tag{11}$$

Table 1. Parameters of the four-component Milky Way prior.

Component	Parameter	Value
Bulge	$\mu_{[M/H],\,1}$/dex	−0.3
	$\sigma_{[M/H],\,1}$/dex	0.3
	$\mu_{\tau,\,1}$/Gyr	5.0
	$\sigma_{\tau,\,1}$/Gyr	5.0
	q	0.5
	γ	0.0
	δ	1.8
	r_0/kpc	0.075
	r_t/kpc	2.1
Thin disc	$\mu_{[M/H],\,2}$/dex	0.0
	$\sigma_{[M/H],\,2}$	0.2
	$R_{d,\,2}$/kpc	2.6
	$z_{d,\,2}$/kpc	0.3
Thick disc	$\mu_{[M/H],\,3}$/dex	−0.6
	$\sigma_{[M/H],\,3}$/dex	0.5
	$\mu_{\tau,\,3}$/Gyr	10.
	$\sigma_{\tau,\,3}$/Gyr	2.
	$R_{d,\,3}$/kpc	3.6
	$z_{d,\,3}$/kpc	0.9
Stellar halo	$\mu_{[M/H],\,4}$/dex	−1.6
	$\sigma_{[M/H],\,4}$/dex	0.5
	$\mu_{\tau,\,4}$/Gyr	11.0
	$\sigma_{\tau,\,4}$/Gyr	2.0

The parameters of the prior are given in Table 1. The thin disc is normalized to have a local density of 0.04 $M_\odot\,pc^{-3}$ (Bovy 2017). The thick disc and stellar halo are normalized so that their local densities have ratios of 0.04 and 0.005 with the thin disc, respectively (Bland-Hawthorn & Gerhard 2016; Bovy 2017). The bulge component is normalized to have a central density of 35.45 $M_\odot\,pc^{-3}$ (Robin et al. 2012). An overarching prior is imposed that constrains metallicities and ages to the range covered by the isochrones.

We calculate $p(\phi^i|u^i)$ on a grid of all 353 ages, τ^i, all metallicities, $[M/H]^i$, lying within 3σ of the measured metallicity, 2000 initial masses, $\mathcal{M}^i$, ranging between the minimum and maximum mass of the relevant isochrone, and 30 distances, s^i, based on a linear grid of parallaxes lying within 3σ of the the measured parallax. We calculate the first and second moments of the logarithm of age, $\log_{10}\tau$, metallicity, [M/H], logarithm of the current mass, $\log_{10}m$, and the distance modulus, μ as our outputs and output uncertainties, e.g.

$$\langle \log_{10}\tau \rangle = \left(\int \mathrm{d}\phi\ \log_{10}\tau\ p(u|\phi,l,b)p(\phi|l,b) \right) \Big/ \left(\int \mathrm{d}\phi\ p(u|\phi,l,b)p(\phi|l,b) \right). \quad (12)$$

The isochrone pipeline successfully determines outputs for 203,127 stars in the overlap between APOGEE DR14 and Gaia DR2. Fig. 2 compares the one-dimensional age distributions determined here and in Sanders & Das (2018) for these stars. We further compare the age distribution of the stars that are metal-poorer than −0.8. The distributions are very similar when we look at the whole sample because it is dominated by young, metal-rich disc stars. But when looking at just the metal-poor stars, which pick out mainly thick-disc and halo stars, we find that Sanders & Das (2018) predict a significantly higher number of young stars.

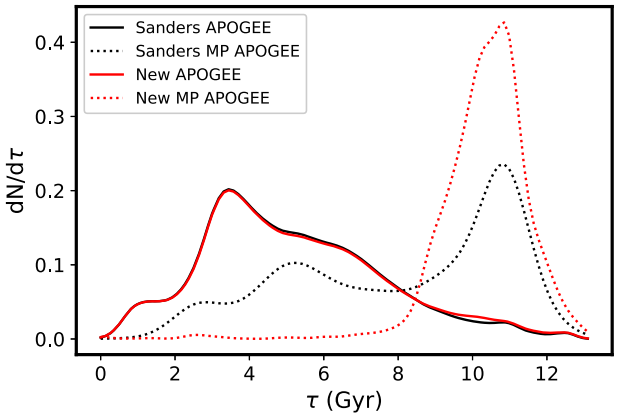

Figure 2. One-dimensional age distributions for all APOGEE stars (solid), and stars metal-poorer than −0.8 (dotted) in Sanders & Das (2018) (black) and the present work (red). The distributions have been normalized so that the area under each is one.

3.3 Estimating orbital parameters

We make estimates of orbital parameters, called actions, using the AGAMA software package of Vasiliev (2019) from six-dimensional phase-space coordinates in the APOGEE DR14-Gaia DR2 cross-match, and assuming the gravitational potential of Piffl et al. (2014). In a near-integrable potential, actions, which are constants of motion, efficiently package the orbital properties of each star into three labels. Assuming an axisymmetric potential, the three actions are the radial action J_r, vertical action J_z, and z component of angular momentum, L_z. The radial and vertical actions approximately describe excursions of an orbit in the radial and vertical directions. We apply a last quality control cut to create our final catalogue by requiring:

(i) `Age, `J_r`, `J_z`, and `L_z` are known,`

Our final sample contains 132 380 stars. Typical uncertainties are 0.03 dex in [M/H], 0.02 in [Mg/Fe], 0.06 in [Al/Fe], 0.04 in [C/Fe], and 0.05 in [N/Fe]. Typical uncertainties in mass are 5 per cent and in age are 18 per cent, although these do not consider systematic uncertainties.

4 SELECTING ACCRETED STARS IN THE MILKY WAY

Here we examine the Milky Way stars in a number of abundance planes, and use a Gaussian Mixture Model to identify stars that are likely to be accreted.

4.1 Abundance planes

Hawkins et al. (2015) present a chemical abundance distribution study in 14 α, odd-Z, even-Z, light, and Fe-peak elements of approximately 3200 intermediate-metallicity giant stars from APOGEE-DR12. They suggest a set of chemical abundance planes constructed from combinations of [α/Fe], [C + N/Fe], [Al/Fe], and [Mg/Mn] that may be able to chemically label the Galactic components in a clean way independent of kinematics.

Fig. 3 presents some of the abundance planes considering these ratios. We overplot abundances for 714 F and G dwarf and subgiant stars in the Solar Neighbourhood from Bensby et al. (2014), who conducted a high-resolution ($R = 40\,000$–$110\,000$) spectroscopic

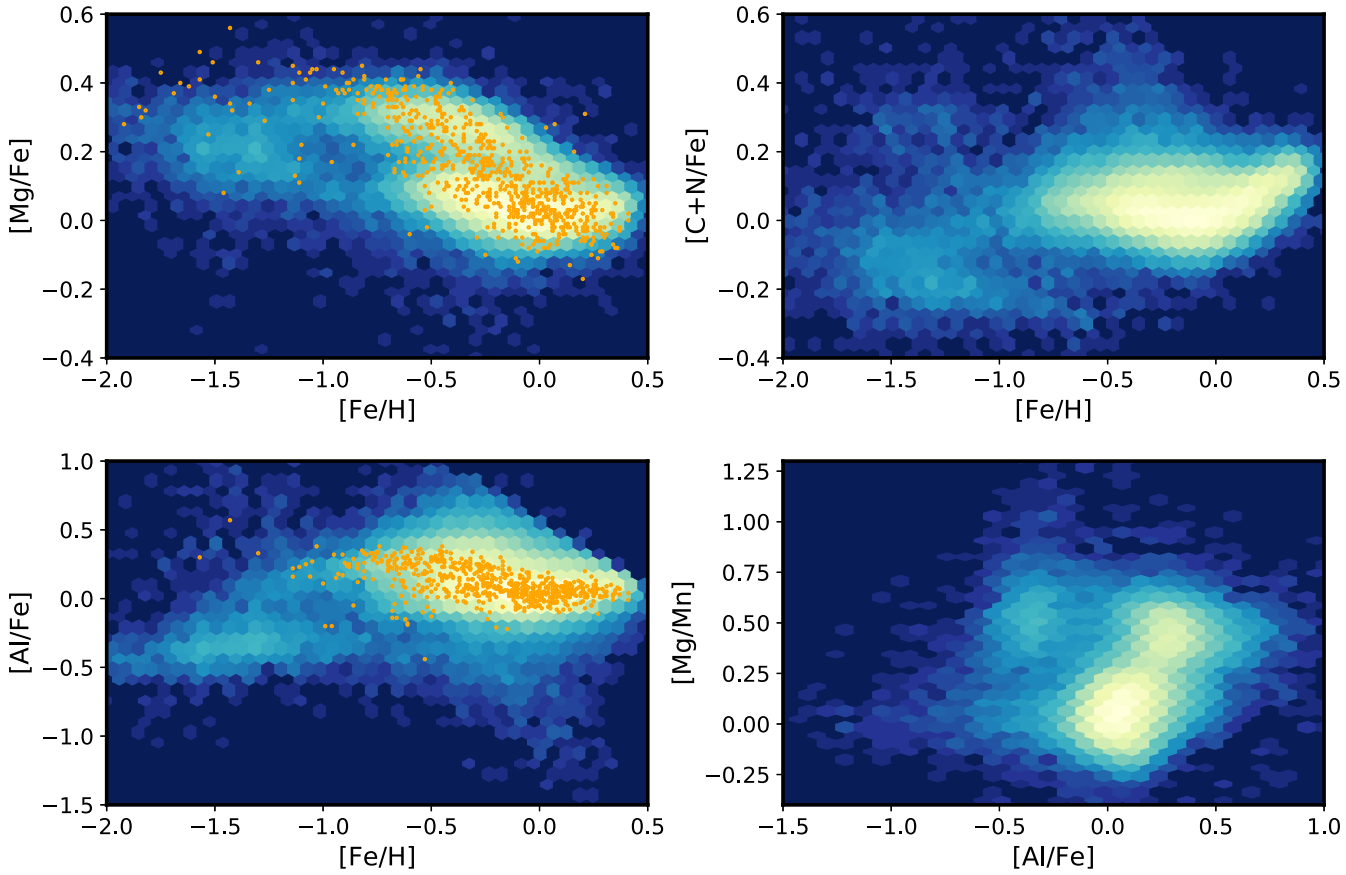

Figure 3. Two-dimensional abundance distribution of APOGEE DR14 stars in the plane of [Fe/H] and [Mg/Fe] (top left), [Fe/H] and [Al/Fe] (bottom left), [Fe/H] and [C + N/Fe] (top right), and [Mg/Mn] and [Al/Fe] (bottom right). Filled orange circles show abundances for 714 kinematically-selected disc F and G dwarfs and subgiants from Bensby, Feltzing & Oey (2014).

study using spectra from a number of telescopes. The sample was kinematically selected to trace the Galactic thin and thick disks to their extremes, and therefore should reflect the underlying range of abundances expected for the Milky Way thin and thick discs. Typical uncertainties are likely similar between the two samples, but are difficult to quantitatively compare in detail (see e.g. Jofré, Heiter & Soubiran 2019, for a recent review) because the systematic uncertainties are likely to differ. The abundances in the sample of Bensby et al. (2014) are extracted from lines in the optical wavelength range while APOGEE abundances are drawn from lines in the infrared wavelength range. We expect the resulting abundance trends in the discs from both samples to be similar however.

The [Fe/H]-[Mg/Fe] plane shows two overdensities at around solar metallicity. The locations of the disk stars selected by Bensby et al. (2014) imply the two overdensities are likely to be the thin (α-poor) and thick (α-rich) discs of the Milky Way. These overdensities are well known and have already been characterized in detail (for example in Hayden et al. 2014). There is also a faint overdensity at intermediate [Mg/Fe] and low [Fe/H]. These stars have a slightly lower α abundance than the thick-disc stars and are at significantly lower metallicities. Discussions about this metal-poor and relatively α-poor population has notably been motivated by Nissen & Schuster (2010), who conjectured an external origin for these stars. Some evidence for this lies in the slopes of the sequences traced by each of the overdensities. α elements are produced in the cores of short-lived massive stars during the post-Main-Sequence burning phase. They are dispersed in the interstellar medium (ISM) via Type II, core-collapse supernovae (SNII). Some iron is also produced. Type 1a supernovae (SN1a) of longer-lived stars are the main production site for iron-peak elements, but only contribute a minimal amount of α elements. Therefore as a stellar population evolves, the [α/Fe] decreases with [Fe/H], as can be seen in both the thick and thin discs. For massive systems, enrichment by iron happens on a shorter timescale compared to smaller systems due to a higher SFR and a higher ability to retain the expelled material in the potential well. Therefore the slope of the sequence of stars in [α/Fe]-[Fe/H] is shallower in smaller systems, such as those from which accreted stars have originated. Assuming that the IMF is universal, the oldest stars in each system will however all start as very metal-poor and high in α. This can be clearly seen in the sample from Bensby et al. (2014) where the most α-rich stars extend down to very low metallicities.

The [Fe/H]-[C + N/Fe] plane shows a single overdensity around zero metallicity, and a second overdensity at lower metallicity and lower [C + N/Fe]. APOGEE stars are giants, and therefore their surface C and N abundances have been partially affected during their evolution as a result of dredge-up processes, as discussed earlier. However, the initial [C + N/Fe] is approximately conserved throughout the evolution of those stars (see Masseron & Gilmore 2015), due to minimal change in ^{16}O. Therefore, [C + N/Fe] should depend on metallicity and environment. C is mostly made by He burning in the cores of stars and dispersed in the ISM via SNII at very low metallicity as well as by asymptotic giant branch (AGB) stars at around $\sim$−1.50 dex in metallicity. Hence, [C + N/Fe] is

expected to start decreasing as SN1a start dominating, which do not produce C. It may then start increasing as a result of production in AGB stars. There trends are not clearly evident in the overdensities, but the metal-poorer overdensity appears to have lower [C + N/Fe] at similar metallicities, which is expected for stars born in less massive systems.

The [Fe/H]-[Al/Fe] plane is very similar to the [Fe/H]-[C + N/Fe] plane. The locations of the disc stars selected by Bensby et al. (2014) in this plane show that the metal-richer overdensity is likely to be the discs of the Milky Way. [Al/Fe] decreases with metallicity in the sample from Bensby et al. (2014), but it is unclear what is happening at metallicities lower than −1.0. Na and Al are thought to be produced in the cores of stars and dispersed in the ISM via SNII. However, according to Kobayashi et al. (2006), the production quantities of those elements is strongly dependent on the initial C and N in the gas cloud that forms the stars. Therefore, it is expected that Na and Al are primarily correlated with C+N as observed in the case of Al. Na and Al are also expected to be partially produced by AGB stars at intermediate metallicities (Nomoto, Kobayashi & Tominaga 2013). Moreover, since SNIa do not produce Al and Na as efficiently as Fe, the [Al/Fe] and [Na/Fe] tend to decrease towards higher metallicities, which can be clearly seen in the sample from Bensby et al. (2014). Nissen & Schuster (2010) demonstrate the effective ability of Na and Al to characterize accreted stars born in less massive systems, where, as with C+N, their overall abundances are lower. At these metallicities, typical APOGEE [Na/Fe] abundance uncertainties in this metallicity range approach ∼±0.20 dex. In contrast, the precision of Al abundances in the APOGEE data is effectively very high, ∼±0.06, and thus offer an alternative to Na.

In summary, the abundance ratio of the α elements, generally decrease with metallicity, but at rates that depend on the SFR in the local environment. The overall abundances of C+N and Al of stars depend on the mass of the system in which they were born. Therefore at lower metallicities the [α/Fe]-[Fe/H] sequences of stars born in different environments may overlap, but they diverge towards higher metallicities. Also, low-[Al/Fe] stars are expected in both accreted systems, and in the metal-rich thin disc, but their α abundances should be different. We note that Hawkins et al. (2015) do not discuss [Al/Fe] ratios at high metallicities as their study was dedicated to stars with metallicities around −1.

We use APOGEE DR14 stars in the [Mg/Mn]-[Al/Fe] plane (Fig. 3) to select our accreted stars. We choose [Mg/Mn] to highlight the α-poor population rather than [Mg/Fe] as Mn is a pristine tracer of SN1a. Iron, on the other hand, has other formation channels in stellar evolution. Usually iron abundances are much easier to measure than manganese abundances. In APOGEE spectra however, manganese abundances can be measured precisely as there are many weak lines in the infrared that are not heavily blended and also do not suffer from strong hyperfine splitting effects (see discussions in Hawkins et al. 2016; Jönsson et al. 2018; Jofré et al. 2019). We choose [Al/Fe] as it is generally lower in accreted systems. This plane clearly separates high-α, thick-disc stars from high-α, accreted stars (see further discussion in Hawkins et al. 2015), due to the difference in [Al/Fe]. It also separates low-α, thin-disc stars from low-α, accreted stars due to their difference in [Mg/Mn].

4.2 Applying a GMM

We formalize the identification of the overdensities in the 2-D [Mg/Mn]-[Al/Fe] plane using the GMM-fitting routines available in PYTHON's `skikit-learn` package. A GMM is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussians with unknown parameters. We explore GMMs with up to 20 components, and find that the Bayesian Information Criterion favours a model with 14 components. Performing the GMM in the 2-D chemical abundance defined by [Mg/Mn] and [Al/Fe] was motivated by Hawkins et al. (2015). In principal, addition dimensions (e.g. [Fe/H]) could have also been added.

A star is assigned to a component if the probability of belonging to it is greater than 0.7. There is a cluster of components centred on what is likely to be the thick disc (i.e. high-α, high-[Al/Fe] stars), and a second cluster of components centred on what is likely to be the thin disc (i.e. low-α, high-[Al/Fe] stars). Two components are centred on lower [Al/Fe], one at lower values of [Mg/Mn], and one at higher values. Fig. 4 replots the density maps of Fig. 3 with contours of stars belonging to the two low-[Al/Fe] Gaussian components overplotted. As explained above, stars belonging to the component with lower values of [Mg/Mn] are likely to be the low-[Al/Fe], thin-disc stars, also identified in Hawkins et al. (2015) through their kinematics. Like other thin-disc stars, these stars are low in [Mg/Mn] because they are born after the ISM has been enriched by Mn expelled by Type 1a SNe. They are low in [Al/Fe], and therefore likely to comprise a mixture of accreted stars and metal-richer, thin-disc stars.

The high-[Mg/Mn], low-[Al/Fe] component or 'blob' is likely to be a relatively pure sample of accreted stars. We examine it in more detail in Section 5. There are 856 stars allocated to the blob.

5 THE BLOB STARS

Here, we examine the properties of the stars belonging to the blob.[1]

5.1 Ages and kinematics in the abundance planes

Fig. 5 shows the location of the blob stars in the [Mg/Fe]-[Fe/H] (left) and [Mg/Mn]-[Al/Fe] (right) abundance planes, coloured by age (top) and the z component of angular momentum (bottom). It is obvious from the [Mg/Fe]-[Fe/H] plots that some stars that were allocated to the blob are disc contaminants, either because they are younger than ∼7 Gyr and/or because they are co-rotating with the disc. They are low in [Al/Fe] and [Mg/Mn] and are therefore allocated to the blob, but are comparatively metal-rich. They can be mainly removed by imposing an additional criterion in metallicity, i.e. [Fe/H] < −0.8. 61 of the 856 stars allocated to the blob are suspected contaminants (i.e. ∼ 7 per cent). We continue with this modified definition of the blob stars, and revisit the properties of the metal-rich contaminants in the blob in Section 5.3.

The remaining blob stars follow a sequence in the [Mg/Fe]-[Fe/H] plane, i.e. metal-richer stars are generally poorer in α elements, similar to that seen in the thick and thin discs, but over a larger range of metallicities and [Mg/Fe]. [Al/Fe] does not change as significantly over the same range in metallicities. The blob stars with the highest α and lowest metallicity are around 13 Gyr, and metal-richer stars are as young as 7 to 8 Gyr. This correlation is not in the Milky Way model prior that was used to calculate the new ages, and therefore must be driven by the data. The bottom panel of plots shows the z component of angular momentum, L_z, of

[1] A catalogue of the spectroscopic, age, and kinematic properties of the blob stars can be downloaded from https://drive.google.com/open?id = 1TYncwKeWlBx7F2b8Y8KN2LWyMKpSKKIk.

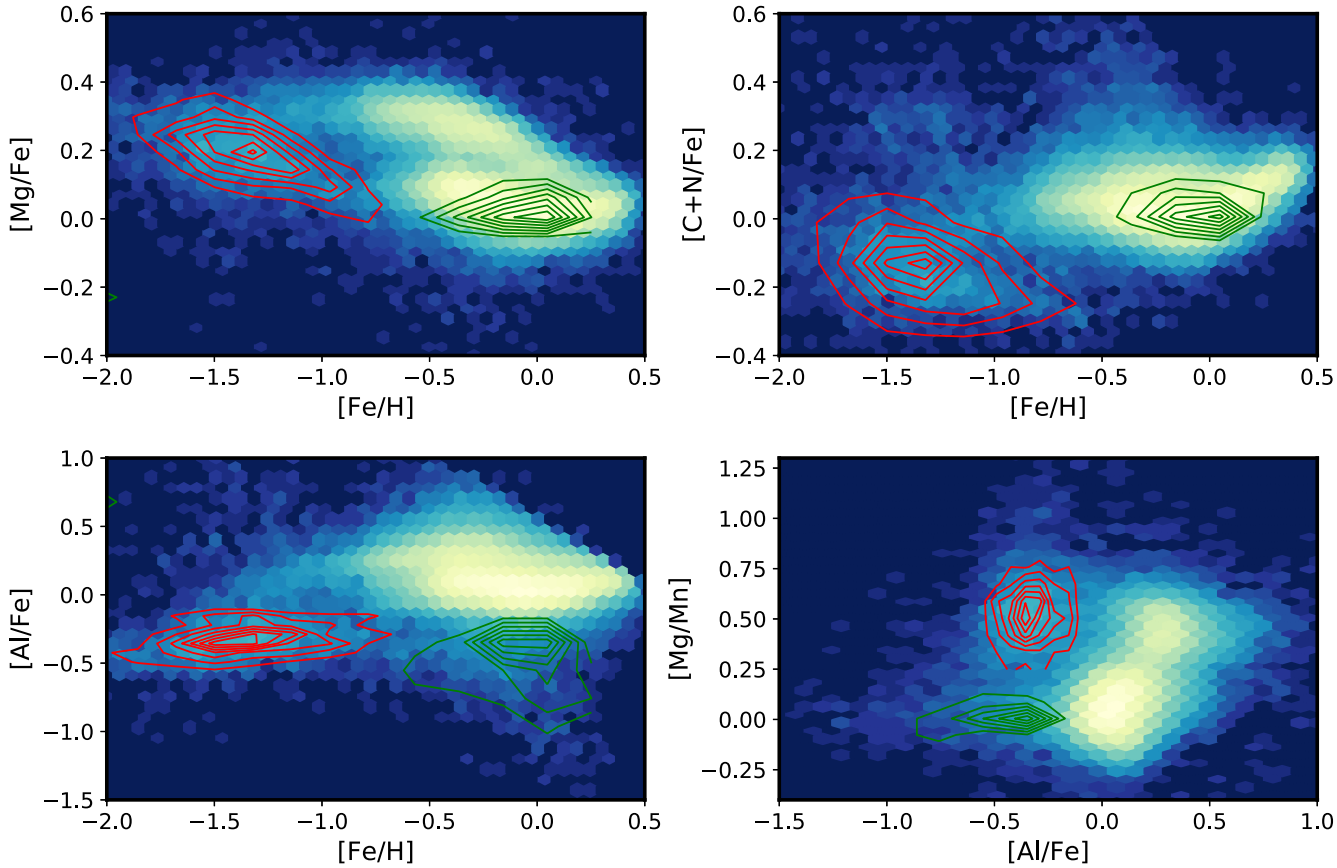

Figure 4. Two-dimensional abundance distribution of APOGEE DR14 stars in the plane of [Fe/H] and [Mg/Fe] (top left), [Fe/H] and [Al/Fe] (bottom left), [Fe/H] and [C + N/Fe] (top right), and [Mg/Mn] and [Al/Fe] (bottom right). The contours show the distributions of stars belonging to the two low-[Al/Fe] Gaussian components found by the GMM. The red contours represent the 'blob', and the green contours, likely low [Al/Fe] thin-disc stars.

the blob stars. Both co-rotating and counter-rotating blob stars are found throughout the sequence, however the metal-richer stars may have a higher proportion of counter-rotating stars.

5.2 Contribution to the metal-poor population of APOGEE

Fig. 6 shows how the proportion of APOGEE DR14 stars belonging to the blob depends on metallicity. The contribution increases from ∼24 per cent at a metallicity of around −2.5 to 32 per cent at a metallicity just below −1.5. It then decreases towards solar metallicity. Therefore a significant fraction of the metal-poor stars in APOGEE DR14 belong to the blob.

5.3 1-D abundance, age, and kinematic distributions

Figs 7 and 8 show kernel density estimates of the 1-D abundance, age, and kinematic distributions of all APOGEE stars, blob stars without the metal-rich contaminants, the metal-rich contaminants of the blob stars, and the metal-poor stars not allocated to the blob. The area under each distribution has been normalized to one. Interpreting these distributions should usually consider the selection function. For example, there is a bias towards seeing stars on more circular orbits around the Solar Neighbourhood, as stars with larger Galactocentric distance can only be seen at the Sun if they are on more eccentric and/or inclined orbits. However as all the distributions are subject to the same overall selection function, we make some tentative comparisons here.

The metallicity distribution (Fig. 7) of all APOGEE stars peaks around solar as the sample is dominated by thin-disc stars. The blob stars peak at a metallicity of ∼−1.4, but have a wide range of metallicities. The metal-poor stars in APOGEE not in the blob have a similarly extended tail of metal-poor stars, but peak at higher metallicities. The metal-rich blob contaminants peak at the minimum metallicity of −0.8, and have a secondary peak at metallicities just below solar.

[α/Fe] of all APOGEE stars peaks around solar, because of the dominance of the α-poor, thin-disc stars. There is a secondary peak around 0.2, which is likely to correspond to thick-disc stars. The blob stars have a peak just above 0.2, with a wide range of α abundances. The width of the [α/Fe] distribution of metal-poor stars in APOGEE is similar to the blob stars but it peaks at slightly higher [α/Fe] values. [α/Fe] of the metal-rich blob contaminants peaks just below 0.1, implying that they are dominated by thin-disc stars.

[Mg/Mn] behaves in a similar way to [α/Fe], but the ranges for blob stars and metal-poor APOGEE stars are narrower than in the case for [α/Fe]. This may be a result of a smaller dispersion in Mn abundances, which is typically a purer tracer of SN1a than iron. The metal-rich blob contaminants almost all have lower values of [Mg/Mn] compared to the rest of the blob stars, and are therefore likely to belong to the other low-[Al/Fe] component shown in Fig. 3.

[Al/Fe] is lower for the blob stars than the APOGEE stars as a whole, with a very small region of overlap. The metal-poor stars not

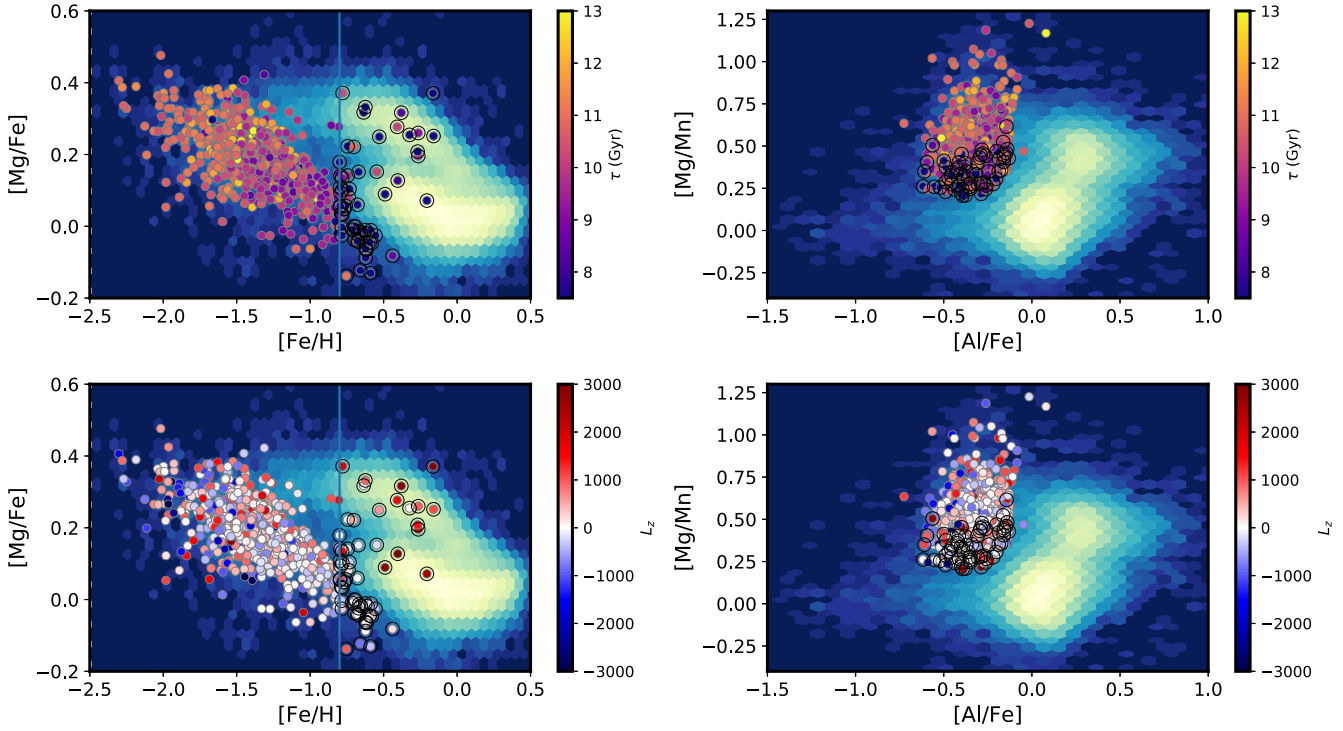

Figure 5. Density maps of APOGEE DR14 stars in the [Mg/Fe]-[Fe/H] (left) and [Al/Fe]-[Fe/H] (right), planes with the ages (top) and z component of angular momentum (bottom) of the blob stars superimposed. The vertical blue line shows the additional criterion imposed to remove metal-rich contaminants in the blob. The black circles highlight the blob stars that are removed as a result.

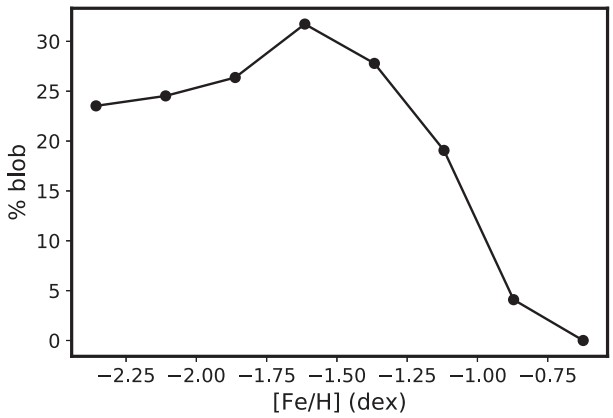

Figure 6. The variation in the fraction of APOGEE DR14 stars allocated to the blob with metallicity.

belonging to the blob have two peaks that coincide with the other two, suggesting that it is a mixture of metal-poor, disc stars, and possibly blob-like stars.

The age distribution for all APOGEE stars peaks towards younger stars due to the dominance of thin-disc stars. The ages for blob and metal-poor stars peak around 11 Gyr, with most stars between 8 to 13 Gyr. The metal-poor stars not belonging to the blob have a very similar distribution in ages, but the metal-rich blob contaminants have an age distribution more similar to the whole APOGEE sample, which is dominated by thin-disc stars. In the age distribution of blob stars, four stars appear to be young (i.e. < 7 Gyr). We will examine them in more detail in Section 6.

Fig. 8 shows that APOGEE stars are found primarily within about 10 kpc of the Sun, due to the selection function in distance. This is similar to the distance range examined in Belokurov et al. (2018). The distribution of blob stars and other metal-poor stars are very similar in R, z, and s. They are more extended in Galactocentric radius and tend to be found at higher absolute z. The bias towards positive values of z reflects the location of the spectroscopic fields in APOGEE DR14. As blob stars are typically found at larger vertical heights, there is a bias towards seeing them with larger J_z. Their distances from the Sun extend out to about 20 kpc, but most are within 10 kpc.

The z-component of angular momentum for all APOGEE stars peaks at around 2000 km s^{-1} kpc, i.e. they are dominated by disc stars. The blob stars peak at slightly negative values of angular momentum, but cover the full extent of angular momenta of APOGEE stars. The metal-poor stars in APOGEE cover a similar range to the blob stars, but with a primary peak around zero and a secondary peak closer to 2000. This implies that it is a mixture of stars with blob-like kinematics, and disc-like kinematics. The metal-rich blob contaminants have a tail towards more positive values, further verifying the disc contamination.

The radial actions are significantly higher for the blob stars and metal-poor stars not in the blob, compared to all APOGEE stars, particularly in blob stars without the metal-rich contaminants. Therefore these stars make greater radial excursions. The blob stars without the metal-rich contaminants also appear to have a narrower distribution in J_r. The vertical actions are significantly higher for the blob stars compared to the other metal-poor stars and all APOGEE stars. Therefore these stars make greater vertical excursions. This again supports that the metal-rich blob contaminants are likely to be primarily disc stars that are on orbits closer to the Galactic plane.

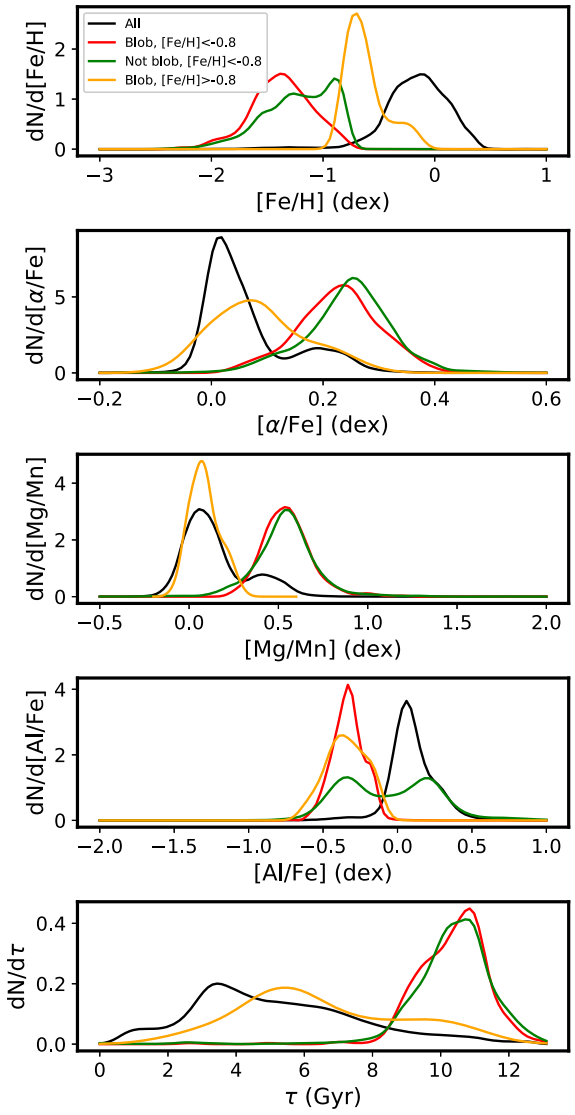

Figure 7. Going down from the top: kernel density estimates of the 1-D distributions of metallicity, α-abundance, [Mg/Mn], [Al/Fe], and age for all APOGEE DR14 stars (black, solid line), the blob stars without the metal-rich contaminants (red, solid line), metal-poor stars not belonging to the blob (green), and the metal-rich contaminants allocated to the blob (orange). The distributions are normalized so that the area under each is one.

6 DISCUSSION

Here, we discuss the ability to select purely accreted stars in the proposed abundance plane, estimate accurate ages, and possible origins of the blob stars.

6.1 Ability to identify purely accreted stars

The distribution of Milky Way stars in the [Mg/Fe]-[Fe/H] plane in Fig. 5 suggests stars in the APOGEE sample follow one of three primary chemical evolution paths. Although the selection function can play a role here (e.g. Schönrich, Asplund & Casagrande 2011), its role should differ between various abundance planes. The stars in the metal-poor overdensity in the [Mg/Fe]-[Fe/H] plane are also present in the distinct metal-poor overdensity visible in the [Al/Fe]-[Fe/H] plane. The [Mg/Mn]-[Al/Fe] plane was found to be a powerful way to separate metal-poor, high-α thick-disc stars from metal-poor, high-α accreted stars, and metal-rich, low-α thin-disc stars from the metal-poor, low-α accreted stars. However, there is some contamination from the higher [Mg/Mn] end of this second contaminant in the selection of blob stars. About half of the contaminants are co-rotating with the disc. However, some (about half) are young, not co-rotating with the disk, metal-rich, and low in [Mg/Fe]. Overall, however, the contamination rate is likely to only be around ~ 7 per cent, and this may be significantly reduced if [Fe/H] is used as a third dimension in the GMM.

There are a few stars in the blob for which we obtain ages that are smaller than 6 Gyr (Table 2). Their kinematics do not support a disc origin. Two of these stars can be associated with the young and α enriched population identified by Martig et al. (2015), though they only probe stars metal-richer than -1.0. A possible explanation for the origin of these stars is the blue straggler scenario, i.e. the stars are not young but their higher masses are a result of mass transfer (e.g. Jofré et al. 2016; Izzard et al. 2018). Martig et al. (2015), Chiappini et al. (2015), and the work presented here use red giants, and blue stragglers are notoriously blue. However, in an old population in which binaries coexist with isolated stars, it is inevitable that some of the evolving binaries will interact, and transfer mass at some point, creating overmassive stars. Izzard et al. (2018) illustrated the creation of such stars with a population synthesis model that includes binaries. If we use isochrones of single stars to date these 'overmassive' stars, they will appear young. From the spectroscopic point of view, such stars are indistinguishable from the rest of the parent population in terms of chemical composition (e.g. Yong et al. 2016; Matsuno et al. 2018). Jofré et al. (2016) show that some of these types of stars have variations in their line-of-sight velocities. However those that are the result of mergers between stars, or have periods that are larger than the time frame used to monitor the line-of-sight velocities, would not show these variations.

6.2 The accuracy of the ages

There is significant uncertainty associated with estimating ages with isochrones for metal-poor red giants. An ideal alternative would be to obtain mass and hence age estimates with asteroseismology (Miglio et al. 2017). There are however only a very limited number of metal-poor stars with asteroseismology data (Valentini et al. 2019, see also Worley et al (in preparation)), although this is starting to change (Rendle et al. 2019; Sharma et al. 2019). The situation will further improve in the near future thanks to the ongoing and future asteroseismology campaigns (K2, TESS and Plato). In this work, we improve age estimates for metal-poor red giants by incorporating accurate estimates of the parallax with Gaia DR2. This places strong constraints on the stellar luminosity, which leads to improved age estimates compared to when only atmospheric parameters are used with the isochrones.

6.3 What is the origin of the blob stars?

Helmi et al. (2018) suggests that the metal-rich halo stars could belong to a single progenitor that merged with the Milky Way. Mackereth et al. (2019), however, noted two populations, which overlap slightly in [Mg/Fe]-[Fe/H], and separate in eccentricity. Myeong et al. (2019) suggest a retrograde accretion event linked to the 'Sequoia' galaxy discovered by Barbá et al. (2019).

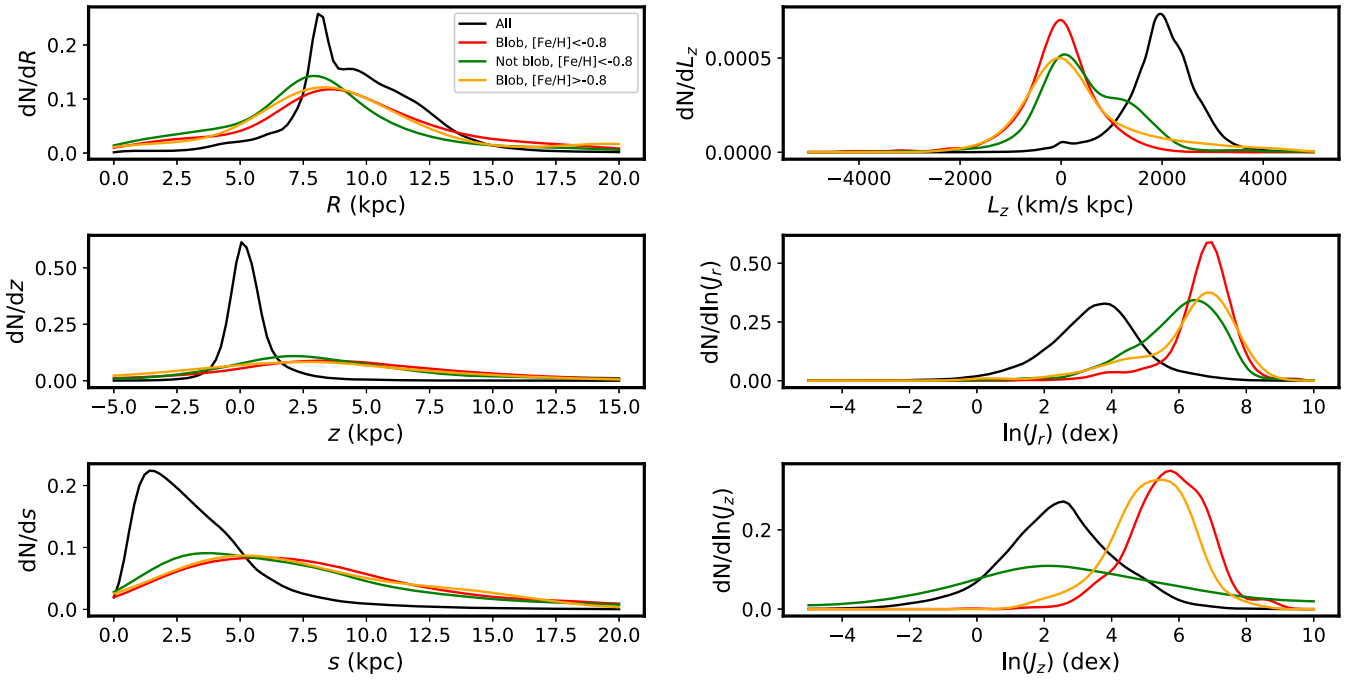

Figure 8. Going anti-clockwise from top left: kernel density estimates of the 1-D distributions of R, z, s, $\ln J_z$, $\ln J_R$, and the z component of angular momentum for all APOGEE DR14 stars (black, solid line), the blob stars without the metal-rich contaminants (red, solid line), metal-poor stars not belonging to the blob (green), and the metal-rich contaminants allocated to the blob (orange). The distributions are normalized so that the area under each is one.

Table 2. Properties of the young blob stars. Italic rows show the young, high-α candidates.

	APOGEE ID	$V_{\rm los}$ error (km s^{-1})	Age (Gyr)	[Fe/H]	[α/Fe]	[Al/Fe]	L_z (km s^{-1} kpc)
1	2M02091357+1446138	0.020	2.8 ± 1.2	-0.95 ± 0.06	0.12 ± 0.02	-0.36 ± 0.08	-288
2	*2M11515780–0147375*	*0.021*	*2.6 ± 0.4*	*-1.40 ± 0.05*	*0.21 ± 0.03*	*-0.30 ± 0.06*	*-185*
3	*2M13423140+2822543*	*0.026*	*4.9 ± 2.7*	*-1.41 ± 0.05*	*0.23 ± 0.03*	*-0.21 ± 0.06*	*748*
4	2M16404349+4255040	0.012	5.1 ± 1.8	-0.83 ± 0.06	0.05 ± 0.02	-0.32 ± 0.08	-231

One argument against a scenario of multiple massive progenitors, is that not many massive systems could fit into the halo of the Milky Way. Simulations only predict two satellites as large as the Large Magellanic Cloud (LMC) for 6.6 per cent of Milky-Way sized hosts (Rodríguez-Puebla, Avila-Reese & Drory 2013).

Here, we examine the abundance, age, and kinematic distributions of the blob stars to discover what can be learned about the merger history.

6.3.1 *The abundance and age distributions*

The distributions of stars in the [Mg/Fe]-[Fe/H]-age planes in Fig. 5 show that the oldest stars are metal poor and α rich, and younger stars become progressively metal richer and poorer in α. This could be consistent with a single progenitor origin.

Multiple progenitors could create a similar distribution in the [Mg/Fe]-[Fe/H]-age plane, but they would have to have similar SFH. Multiple progenitors are also likely to lead to an increase in the scatter in these relations. The 1D distributions in Fig. 7, suggest that there may be two progenitors. However as a result of uncertainties in the abundances and ages, and the unknown selection function, we cannot make a strong statement on the possibility of more than one progenitor. A forward-fitting approach where a composite model consisting of multiple progenitors is fit to the data would be the desired way to proceed.

Assuming the simplest scenario in which the blob stars originate from a single progenitor, the SFH cannot be directly inferred from the age distribution due to the age-metallicity selection function. However the range implies that star formation occurred over approximately 5 Gyr, and that a single progenitor would have merged with the Milky Way about 8 Gyr ago. The SFH must be rather different from the LMC, which is similar in mass. The LMC is likely to have had a recent burst of star formation (Sanders & Das 2018; Nidever et al. 2019), which leads to an increase in [α/Fe] between a metallicity of -1.5 and -1.0.

6.3.2 *Kinematic distributions*

The blob stars are mainly found within 10 kpc, though many are as far as 20 kpc. Like other metal-poor stars, they are found at higher values of radial action, which may be a result of the selection function, i.e. metal-poor stars are generally found further away. Therefore they are only observed close to the Solar Neighbourhood if they are on more radially anisotropic orbits. This can only be properly understood in the context of models. Their distribution in J_r is noticeably narrower however. As noted in Mackereth et al. (2019) there are potentially more retrograde objects at higher [Mg/Fe]. As stated above, a forward-fitting composite model of an assembly of progenitors that takes into account uncertainties is required to distinguish between scenarios.

If we again assume a single progenitor, the width in J_r can be used to make an estimate of the mass of a single progenitor (Eyre & Binney 2011). The radial action is given by

$$J_r = \frac{1}{\pi} \oint v_r \mathrm{d}r, \tag{13}$$

where the integration path is one complete integration. If we consider a particle whose radial velocity differs from the system's average by σ, i.e. $\delta v_r \sim \sigma_r$, then we can take a finite difference over equation (13). This gives us an expression for the difference between the radial action of a particle from the system's average

$$\delta J_r \sim \frac{1}{\pi} \sigma_r \Delta r, \tag{14}$$

where $\Delta r = R_p - R_a$ is the amplitude of the radial oscillation, R_p is the pericentre of the orbit, and R_a is the apocentre of the orbit. Then from the virial theorem, $\mathrm{KE_{avg}} = -\frac{1}{2}\mathrm{GPE_{avg}}$, which gives:

$$\frac{1}{2} m (\sigma_r)^2 = \frac{1}{2} \frac{G M_b m}{\frac{1}{2} R_b}, \tag{15}$$

where M_b is the mass of the blob, and R_b is the radius of the single system from which the blob originates. Approximating the action width, δJ_r, by the dispersion in J_r, and the radial oscillation, Δr, by the 95 per cent range in R, and assuming the accreted system to be a large dwarf spheroidal with a radius of $\sim$5 kpc, equation (15) gives a total progenitor mass of $\sim 3.4 \times 10^{11}\,\mathrm{M}_\odot$. This is similar to that recently estimated for the LMC (Erkal et al. 2019). Using an abundance-matching based mapping from total mass to stellar mass (e.g. Read et al. 2017), this implies a stellar mass $\sim 10^{9.5}\,\mathrm{M}_\odot$. Chemical evolution models applied to the distribution of accreted stars in the [Mg/Fe]-[Fe/H] plane find a stellar mass of $\sim 6 \times 10^8\,\mathrm{M}_\odot$ (Fernández-Alvar et al. 2018; Vincenzo et al. 2019). Mackereth et al. (2019) compare the distribution of the accreted stars in the [Mg/Fe]-[Fe/H] plane with those of accreted galaxies from the EAGLE suite of cosmological simulations, and constrain the stellar mass of a single progenitor to be between $10^{8.5}\,\mathrm{M}_\odot$ and $10^9\,\mathrm{M}_\odot$.

7 CONCLUSIONS

We use a GMM to select a 'blob' of accreted stars in the [Mg/Mn]-[Al/Fe] abundance plane, using abundances published in APOGEE DR14. The blob stars are found to span a range of metallicities from -0.5 to -2.5 and [Mg/Fe] from -0.1 to 0.5. They constitute $\sim$30 per cent of the metal-poor ([Fe/H] < -0.8) halo at metallicities of ~ -1.4. The ages from Sanders & Das (2018) imply a significant fraction of young ages, potentially as a result of the effects of thermohaline mixing and stochastic chemical evolution at lower metallicities. Our new ages are found to mainly range from 8 to 13 Gyr, with the oldest stars the metal-poorest, and with the highest [Mg/Fe] abundance. The blob stars are mainly found within 10 kpc, though many are as far as 20 kpc. Like other metal-poor stars, they are found at higher values of radial action, but their distribution in radial action is noticeably narrower. Their metallicities and distances from the Sun imply that the sample is similar to that examined from the SDSS-Gaia sample in Belokurov et al. (2018).

The blob stars exhibit the expected age, metallicity, and α sequence for stars belonging to a single system. The blob stars tend to make larger radial excursions compared to the average APOGEE star. Belokurov et al. (2018) argue that the high level of radial anisotropy for 'blob' stars points towards a massive progenitor whose stellar orbits have been heavily radialized through a combination of dynamical friction and disc formation. We note, that all metal-poor stars tend to be more radially anisotropic as a result of a selection effect. The J_r distribution for blob stars appear to be more narrow however. Dynamical arguments analyzing this distribution suggest a single system with a total mass of $\sim 10^{11}\,\mathrm{M}_\odot$, similar to that found by other authors in the literature. Finally, we find four young (i.e. < 6 Gyr) stars in the blob. Two of these appear to extend the young α-enriched stars of Martig et al. (2015) to lower metallicities.

APOGEE includes elements from a range of nucleosynthesis processes that drive chemical evolution in the Galaxy, but is absent in heavy, neutron-rich elements ($Z \geq 30$), which are mainly produced by neutron capture via both the s- and r-processes. There are abundances for one heavy element (Rb; $Z = 37$) but these results are based upon very weak lines of Rb I. The s-process element cerium has a high density of absorption lines expected in the spectrum of red giant stars and consists primarily of two stable isotopes. The s-process accounts for $\sim$90 per cent of it, with the r-process contributing the other $\sim$10 per cent. This element may become available in future data releases (Cunha et al. 2017), and spectra of these stars in the optical part of the spectrum will also help reveal more s-process elements. Using these in the GMM procedure will better characterize the blob stars. Over the next few months we will be observing several of these stars in the optical part of the spectrum, giving access to a number of s-process and r-process elements.

In the future, we plan to formally disentangle a single from a multiple-progenitor scenario using chemical evolution models of the type ran recently by Vincenzo et al. (2019), taking the age-metallicity selection function into account.

ACKNOWLEDGEMENTS

PD would like to acknowledge support from the Science and Technology Facilities Council (ST/N000919/1). PJ acknowledges support of El Fondo Nacional de Desarrollo Científico y Tecnológico Grant Number 11170174. This work has made use of data from the European Space Agency mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

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

REFERENCES

Abadi M. G., Navarro J. F., Steinmetz M., 2006, MNRAS, 365, 747
Abolfathi B. et al., 2018, ApJS, 235, 42
Allende Prieto C. et al., 2014, A&A, 568, A7
Barbá R. H., Minniti D., Geisler D., Alonso-García J., Hempel M., Monachesi A., Arias J. I., Gómez F. A., 2019, ApJ, 870, L24
Bell E. F. et al., 2008, ApJ, 680, 295
Bell E. F., Xue X. X., Rix H.-W., Ruhland C., Hogg D. W., 2010, AJ, 140, 1850
Belokurov V. et al., 2007, ApJ, 654, 897
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611

Bensby T., Feltzing S., Oey M. S., 2014, A&A, 562, A71
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Bond N. A. et al., 2010, ApJ, 716, 1
Bovy J., 2017, MNRAS, 470, 1360
Bovy J., Rix H.-W., Green G. M., Schlafly E. F., Finkbeiner D. P., 2016, ApJ, 818, 130
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Burnett B., Binney J., 2010, MNRAS, 407, 339
Carollo D. et al., 2007, Nature, 450, 1020
Chen Y. Q., Zhao G., Carrell K., Zhao J. K., Tan K. F., Nissen P. E., Wei P., 2014, ApJ, 795, 52
Chiappini C. et al., 2015, A&A, 576, L12
Chiba M., Yoshii Y., 1998, AJ, 115, 168
Cooper A. P., Cole S., Frenk C. S., Helmi A., 2011, MNRAS, 417, 2206
Cunha K. et al., 2017, ApJ, 844, 145
Cunningham E. C. et al., 2016, ApJ, 820, 18
Das P., Binney J., 2016, MNRAS, 460, 1725
Das P., Sanders J. L., 2019, MNRAS, 484, 294
Das P., Williams A., Binney J., 2016, MNRAS, 463, 3169
de Jong J. T. A., Yanny B., Rix H.-W., Dolphin A. E., Martin N. F., Beers T. C., 2010, ApJ, 714, 663
Deason A. J., Belokurov V., Evans N. W., 2011, MNRAS, 416, 2903
Deason A. J., Belokurov V., Evans N. W., An J., 2012, MNRAS, 424, L44
Deason A. J., Belokurov V., Evans N. W., Johnston K. V., 2013a, ApJ, 763, 113
Deason A. J., Van der Marel R. P., Guhathakurta P., Sohn S. T., Brown T. M., 2013b, ApJ, 766, 24
Deason A. J., Belokurov V., Koposov S. E., Rockosi C. M., 2014, ApJ, 787, 30
Erkal D. et al., 2019, MNRAS, 487, 2685
Eyre A., Binney J., 2011, MNRAS, 413, 1852
Fernández-Alvar E. et al., 2018, ApJ, 852, 50
Font A. S., McCarthy I. G., Crain R. A., Theuns T., Schaye J., Wiersma R. P. C., Dalla Vecchia C., 2011, MNRAS, 416, 2802
Gaia Collaboration et al., 2018, A&A, 616, 22
García Pérez A. E. et al., 2016, AJ, 151, 144
Haas M. R. et al., 2010, ApJ, 713, L115
Hattori K., Valluri M., Loebman S. R., Bell E. F., 2017, ApJ, 841, 91
Hawkins K., Jofré P., Gilmore G., Masseron T., 2014, MNRAS, 445, 2575
Hawkins K., Jofré P., Masseron T., Gilmore G., 2015, MNRAS, 453, 758
Hawkins K., Masseron T., Jofré P., Gilmore G., Elsworth Y., Hekker S., 2016, A&A, 594, A43
Hayden M. R. et al., 2014, AJ, 147, 116
Hayes C. R. et al., 2018, ApJ, 852, 49
Haywood M., Matteo P. D., Lehnert M. D., Snaith O., Khoperskov S., Gómez A., 2018, ApJ, 863, 113
Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
Holtzman J. A. et al., 2018, AJ, 156, 125
Ibata R. A., Gilmore G., Irwin M. J., 1995, MNRAS, 277, 781
Ivezić Ž. et al., 2008, ApJ, 684, 287
Izzard R. G., Preece H., Jofre P., Halabi G. M., Masseron T., Tout C. A., 2018, MNRAS, 473, 2984
Jofré P., Weiss A., 2011, A&A, 533, A59
Jofré P. et al., 2016, A&A, 595, A60
Jofré P., Heiter U., Soubiran C., 2019, ARA&A, 57, 571
Jönsson H. et al., 2018, AJ, 156, 126
Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2012, ApJ, 761, 98
Kafle P. R., Sharma S., Lewis G. F., Bland-Hawthorn J., 2013, MNRAS, 430, 2973
Kalirai J. S., 2012, Nature, 486, 90
Kobayashi C., Umeda H., Nomoto K., Tominaga N., Ohkubo T., 2006, ApJ, 653, 1145
Kroupa P., Tout C. A., Gilmore G., 1993, MNRAS, 262, 545
Lagarde N. et al., 2019, A&A, 621, A24
Lindegren L. et al., 2016, A&A, 595, A4
Lindegren L. et al., 2018, A&A, 616, A2
Mackereth J. T. et al., 2019, MNRAS, 482, 3426
Majewski S. R., Skrutskie M. F., Weinberg M. D., Ostheimer J. C., 2003, ApJ, 599, 1082
Martig M. et al., 2015, MNRAS, 451, 2230
Martig M. et al., 2016, MNRAS, 456, 3655
Masseron T., Gilmore G., 2015, MNRAS, 453, 1855
Matsuno T., Yong D., Aoki W., Ishigaki M. N., 2018, ApJ, 860, 49
Miglio A. et al., 2017, Astron Nachr, 338, 644
Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018, ApJ, 863, L28
Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235
Ness M., Hogg D. W., Rix H.-W., Martig M., Pinsonneault M. H., Ho A. Y. Q., 2016, ApJ, 823, 114
Nidever D. L. et al., 2015, AJ, 150, 173
Nidever D. L. et al., 2019, ApJ, 874, 118
Nissen P. E., Schuster W. J., 2010, A&A, 511, L10
Nomoto K., Kobayashi C., Tominaga N., 2013, ARA&A, 51, 457
Piffl T. et al., 2014, MNRAS, 445, 3133
Pillepich A., Madau P., Mayer L., 2015, ApJ, 799, 184
Read J. I., Iorio G., Agertz O., Fraternali F., 2017, MNRAS, 467, 2019
Rendle B. M. et al., 2019, MNRAS, 490, 4465
Revaz Y., Jablonka P., 2012, A&A, 538, A82
Robin A. C., Marshall D. J., Schultheis M., Reylé C., 2012, A&A, 538, A106
Rodríguez-Puebla A., Avila-Reese V., Drory N., 2013, ApJ, 773, 172
Sanders J. L., Das P., 2018, MNRAS, 481, 4093
Schönrich R., Asplund M., Casagrande L., 2011, MNRAS, 415, 3807
Schönrich R., Asplund M., Casagrande L., 2014, ApJ, 786, 7
Schuster W. J., Moreno E., Nissen P. E., Pichardo B., 2012, A&A, 538, A21
Sesar B., Jurić M., Ivezić Ž., 2011, ApJ, 731, 4
Sesar B. et al., 2013, AJ, 146, 21
Sharma S. et al., 2019, MNRAS, 490, 5335
Shetrone M. et al., 2015, ApJS, 221, 24
Shetrone M. et al., 2019, ApJ, 872, 137
Sirko E. et al., 2004, AJ, 127, 914
Smith M. C. et al., 2009, MNRAS, 399, 1223
Tissera P. B., White S. D. M., Scannapieco C., 2012, MNRAS, 420, 255
Valentini M. et al., 2019, A&A, 627, 173
Vasiliev E., 2019, MNRAS, 482, 1525
Vincenzo F., Spitoni E., Calura F., Matteucci F., Aguirre V. S., Miglio A., Cescutti G., 2019, MNRAS, 487, L47
Vrard M., Mosser B., Samadi R., 2016, A&A, 588, A87
Watkins L. L. et al., 2009, MNRAS, 398, 1757
Williams A. A., Evans N. W., 2015, MNRAS, 454, 698
Xue X.-X., Rix H.-W., Ma Z., Morrison H., Bovy J., Sesar B., Janesh W., 2015, ApJ, 809, 144
Yong D. et al., 2016, MNRAS, 459, 487
Zamora O. et al., 2015, AJ, 149, 181
Zolotov A., Willman B., Brooks A. M., Governato F., Brook C. B., Hogg D. W., Quinn T., Stinson G., 2009, ApJ, 702, 1058

This paper has been typeset from a TEX/LATEX file prepared by the author.