The origin of accreted stellar halo populations in the Milky Way using APOGEE, Gaia , and the EAGLE simulations

J Ted Mackereth, Ricardo Schiavon, Joel Pfeffer, Christian Hayes, Jo Bovy, Borja Anguiano, Carlos Allende Prieto, Sten Hasselquist, Jo Holtzman, Jo Johnson, et al.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **482,** 3426–3442 (2019)
Advance Access publication 2018 November 1

doi:10.1093/mnras/sty2955

The origin of accreted stellar halo populations in the Milky Way using APOGEE, *Gaia*, and the EAGLE simulations

J. Ted Mackereth [1]★ Ricardo P. Schiavon,[1] Joel Pfeffer [1] Christian R. Hayes,[2] Jo Bovy [3] Borja Anguiano [2] Carlos Allende Prieto,[4,5] Sten Hasselquist,[6] Jon Holtzman,[7] Jennifer A. Johnson,[8] Steven R. Majewski,[2] Robert O'Connell,[2] Matthew Shetrone,[9] Patricia B. Tissera [10] and J. G. Fernández-Trincado[11,12,13]

[1]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[2]*Department of Astronomy, University of Virginia, Charlottesville, VA 22904-4325, USA*
[3]*Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada*
[4]*Instituto de Astrofisica de Canarias, Via Lactea, 38205 La Laguna, Tenerife, Spain*
[5]*Universidad de La Laguna, Departamento de Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[6]*Department of Physics & Astronomy, University of Utah, Salt Lake City, UT 84112, USA*
[7]*New Mexico State University, Las Cruces, NM 88003, USA*
[8]*Department of Astronomy and Center for Cosmology and AstroParticle Physics, Ohio State University, Columbus, OH, USA*
[9]*University of Texas at Austin, McDonald Observatory, Fort Davis, TX 79734, USA*
[10]*Department of Physics, Universidad Andres Bello, 700 Fernandez Concha, Chile*
[11]*Departamento de Astronomía, Casilla 160-C, Universidad de Concepción, Concepción, Chile*
[12]*Instituto de Astronomía y Ciencias Planetarias, Universidad de Atacama, Copayapu 485, Copiapó, Chile*
[13]*Observatoire de Besançon, Institut Utinam, CNRS UMR 6213, Université Bourgogne-Franche-Comté, OSU THETA Franche-Comté, BP 1615, 25010 Besançon Cedex, France*

Accepted 2018 October 30. Received 2018 October 26; in original form 2018 August 2

ABSTRACT

Recent work indicates that the nearby Galactic halo is dominated by the debris from a major accretion event. We confirm that result from an analysis of APOGEE-DR14 element abundances and *Gaia*-DR2 kinematics of halo stars. We show that ∼2/3 of nearby halo stars have high orbital eccentricities ($e \gtrsim 0.8$), and abundance patterns typical of massive Milky Way dwarf galaxy satellites today, characterized by relatively low [Fe/H], [Mg/Fe], [Al/Fe], and [Ni/Fe]. The trend followed by high-e stars in the [Mg/Fe]–[Fe/H] plane shows a change of slope at [Fe/H] ~ -1.3, which is also typical of stellar populations from relatively massive dwarf galaxies. Low-e stars exhibit no such change of slope within the observed [Fe/H] range and show slightly higher abundances of Mg, Al, and Ni. Unlike their low-e counterparts, high-e stars show slightly retrograde motion, make higher vertical excursions, and reach larger apocentre radii. By comparing the position in [Mg/Fe]–[Fe/H] space of high-e stars with those of accreted galaxies from the EAGLE suite of cosmological simulations, we constrain the mass of the accreted satellite to be in the range $10^{8.5} \lesssim M_* \lesssim 10^9\,\mathrm{M}_\odot$. We show that the median orbital eccentricities of debris are largely unchanged since merger time, implying that this accretion event likely happened at $z \lesssim 1.5$. The exact nature of the low-e population is unclear, but we hypothesize that it is a combination of *in situ* star formation, high-$|z|$ disc stars, lower mass accretion events, and contamination by the low-e tail of the high-e population. Finally, our results imply that the accretion history of the Milky Way was quite unusual.

Key words: Galaxy: abundances – Galaxy: formation – Galaxy: halo – Galaxy: kinematics and dynamics – Galaxy: stellar content.

1 INTRODUCTION

It is now well established that accretion of lower mass systems is a fundamental component of the evolution and mass build up of galaxies (White & Frenk 1991). Due to its very long dynamical

★ E-mail: J.E.Mackereth@2011.ljmu.ac.uk

time-scale, the stellar halo of the Milky Way keeps a record of the Galaxy's past accretion activity. That record can be accessed through the collection of precision 6D phase space and multi-element abundance information for very large samples of halo stars, which together enable fundamental tests of galaxy formation models. While this field has a long history (e.g., Eggen, Lynden-Bell & Sandage 1962; Searle & Zinn 1978), we highlight only a few of the main contributions from the past decade, for brevity. Nissen & Schuster (2010) and Schuster et al. (2012) were the first to identify the presence of an older, high α and a younger, low-α halo population at metallicity lower than that of the Galactic disc in the solar neighbourhood. The kinematics of those stellar populations suggested an *in situ* [1] or accreted origin, respectively. More recently, Hawkins et al. (2015) proposed abundance ratio diagnostics to distinguish accreted from *in situ* halo stars, arguing that the accreted population dominates the nearby halo. Using APOGEE data, Fernández-Alvar et al. (2018b) and Hayes et al. (2018) studied the chemical compositions and kinematics of the metal-rich nearby halo, suggesting that much of the low [Mg/Fe] halo population is associated with the debris of accreted satellites, likely with a similar star formation history to the Large Magellanic Cloud (LMC).

Studies of the Galactic halo are being revolutionized by the advent of large astrometric, photometric, and spectroscopic surveys of the stellar populations of the Galaxy. The *Gaia* astrometric satellite has opened new avenues for exploration of substructure in phase space, with the potential for new discoveries further amplified by the addition of chemical information from spectroscopic surveys. Indeed, combining *Gaia* parallaxes and proper motions (Lindegren et al. 2016; Gaia Collaboration et al. 2018) with spectroscopic data from SDSS (York et al. 2000) and APOGEE (Majewski et al. 2017), two groups have identified what seems to be the accretion of a relatively massive stellar system that dominates the stellar populations of the nearby halo. Analysing a sample of SDSS–*Gaia* DR1 main sequence stars, Belokurov et al. (2018) showed that the velocity ellipsoid of halo stars becomes strongly anisotropic for stars with [Fe/H] > -1.7. Comparing their data to a suite of N-body only cosmological numerical simulations, they concluded that such orbital configurations are likely to result from the accretion of a massive satellite at about the time of the formation of the Galactic disc, roughly between $z = 1$ and 3.

Based on *Gaia* DR2 data (Gaia Collaboration et al. 2018), Myeong et al. (2018a) determined the configuration of MW globular clusters (GCs) in action space. They find that 12 GCs in the halo are consistent with an origin in a single massive accretion event, consistent with the conclusions reached by Belokurov et al. (2018). Myeong et al. (2018a) find that these clusters have highly eccentric orbits, at $e \gtrsim 0.85$, and suggest that the fact that all the clusters occupy a similar region in action space supports the idea that this highly anisotropic stellar population in the halo is mainly formed from the debris of a single accretion event.

In a follow-up study, Deason et al. (2018) estimated the orbital parameters for a sample of nearby main-sequence and distant horizontal-branch stars by combining *Gaia* DR2 data with spectroscopic outputs from SDSS-DR9 (Ahn et al. 2012). They found that the apocentre radii of a significant population of stars in the halo appear to 'pile up' at an $r_{\rm ap} \sim 20$ kpc. The authors link this population with that found by Belokurov et al. (2018). This result has special significance in light of the analysis of numerical simulations by Deason et al. (2013), who proposed that the existence of a 'break radius' in the Milky Way halo, beyond which the stellar density drops precipitously, is associated with the 'pile up' of stellar apocentres at a comparable Galactocentric distance. Deason et al. (2013) argue that 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 much more prolonged accretion history than the former. Follow-up work using the same sample suggests that, inside this break radius, roughly 50 per cent of the halo is made up of stars from this accretion event (Lancaster et al. 2018).

An independent analysis of *Gaia* DR2 data conducted by Koppelman, Helmi & Veljanoski (2018) identified the presence of a large, old, and metal-poor slightly counter-rotating structure in phase space. They concluded that this population is associated with a relatively massive object which, they hypothesized, may have been responsible for the heating of the thick disc. Following up on that result, Helmi et al. (2018) used kinematic, chemical, and age information for a large sample of stars in *Gaia* and APOGEE to identify a population of metal-poor stars with the same phase-space characteristics as those reported by Koppelman et al. (2018). The distribution of this stellar population in the [α/Fe]–[Fe/H] plane, with relatively low [α/Fe] and a large spread in [Fe/H], suggests the chemical evolution trend of a relatively massive system.[2] Moreover, the positions of the stars in the HR diagram are consistent with old ages (10–13 Gyr). According to Helmi et al. (2018), the accretion of a dwarf galaxy with a mass similar to that of the Small Magellanic Cloud (see also Hayes et al. 2018) $\sim$10 Gyr ago may have been responsible for the heating of the thick disc. The notion that the thick disc was formed from the vertical heating of a thinner progenitor disc competes with the so-called 'upside-down' formation scenario (see, e.g. Bird et al. 2013; Navarro et al. 2017) according to which the early gaseous disc was thick as a result of strong stellar feedback (and/or clumpy gas accretion, e.g. Brook et al. 2004), and slowly settled as the star formation waned to form the thinner components of the disc. Despite their differences, both scenarios are consistent with at least some heating of the stellar disc by satellite mergers, which, in turn, are also likely necessary to explain the flaring of high [α/Fe] mono-age disc populations (e.g. Minchev et al. 2015; Mackereth et al. 2017). In general, recent observational results do not seem to point towards a scenario where the thick disc formed thin and was heated entirely by mergers, which would produce a plateau in the age or α abundance against scale–height relationship that is not currently borne out by the data (e.g. Bovy, Rix & Hogg 2012a; Bovy et al. 2012b,c, 2016; Casagrande et al. 2016; Mackereth et al. 2017)

It is also important to note that the same population was identified by Haywood et al. (2018), who studied Gaia DR2 and APOGEE-based colour–magnitude diagrams, kinematics, and chemistry to identify a population of metal-poor stars with high transversal velocities and typically low or retrogade rotation, which they associate with the last significant merger undergone by the Milky Way.

In addition to, and in support of these findings, Kruijssen et al. (2018) recently inferred that the Milky Way has had a rather atypical assembly history given its mass, based on analysis of the age–metallicity relation of Galactic GCs. They found that the assembly rate of the Milky Way was among the uppermost quartile of galaxies

[1] By stars formed *in situ* we mean those that were formed within the Galaxy, either from gas originally associated with its dark matter halo or that which was accreted onto it.

[2] The recent paper by Fernández-Alvar et al. (2018a) hints at a similar conclusion, also on the basis of APOGEE data.

in their simulation, and identified three recent massive accretion events. Those authors proposed that two of these accretion events correspond to the Sagittarius dwarf and Canis Major.[3] The most massive of those accretion events is suggested to have no known debris, and have a stellar mass $>10^9\,M_\odot$, and correspond to GCs that reside close to the Galactic centre. It indeed may be possible, on the basis of the analysis of the Galactic GC population by Kruijssen et al. (2018), that the accreted satellite identified by Belokurov et al. (2018) and Koppelman et al. (2018) is associated with Canis Major, given that the GCs identified by Myeong et al. (2018a) are further out in the halo, and those potentially associated with Canis Major are located at Galactocentric distances >10 kpc. The finding that the assembly history of the Milky Way is atypical is also consistent with the work of Mackereth et al. (2018), who found that Milky Way stellar mass galaxies in the EAGLE simulation with [α/Fe] abundance patterns similar to the Milky Way had atypical accretion histories, characterized by early, rapid accretion, which slowed at late times.

In summary, the local stellar halo has been shown to be dominated by a population of moderately metal-poor, low [α/Fe], old stars on highly eccentric orbits. This population is the likely remnant of a major accretion event that took place at about the same time that the Galactic disc was itself forming. These results have important implications, which prompted us to examine the chemical and kinematic properties of the newly discovered stellar population in detail. In this paper, we present an analysis of the abundance pattern of stars in common between the *Gaia*-DR2 and APOGEE-DR14 catalogues, and discuss the implications of their kinematic properties in light of the EAGLE suite of numerical cosmological simulations. We extend the studies of the element abundances in these populations to include odd-*Z* and Iron peak elements, and examine the detailed kinematics of stars in sub-populations defined by abundances and orbit eccentricity. We also extend previous theoretical work on this population by examining the kinematics of accreted debris from a fully self-consistent cosmological simulation that provides a cosmologicallly motivated sample of accreted satellite debris onto Milky Way mass haloes. In Section 2, we describe our sample selection and orbital parameter determination, as well as the details of the EAGLE simulations. In Section 3, we discuss the chemical and kinematic properties of this population. In Section 4, we contrast the kinematic properties of the newly discovered stellar population with the expectations from cosmological numerical simulations. Our conclusions are summarized in Section 5.

2 SAMPLE AND DATA

2.1 APOGEE DR14

Our study is based on a cross-match between the SDSS-APOGEE DR14 and *Gaia* DR2 catalogues. APOGEE (Majewski et al. 2017) is a near-infrared spectroscopic survey of the stellar populations of the Galaxy and its Local group neighbours. The data employed in this paper come from the DR14 catalogue (Abolfathi et al. 2018), which comprises a re-reduction and analysis of APOGEE-1 data (from SDSS-III, Eisenstein et al. 2011), alongside a set of newly reduced and analysed observations from APOGEE-2 (taken as part of SDSS-IV, Blanton et al. 2017). APOGEE DR14 contains high S/N, R $\sim$ 22 500 spectra, radial velocities, stellar photospheric parameters, and element abundances for over 270 000 stars in the *H*-band (1.5–1.7μm). Observations are carried out using the 2.5 m SDSS Telescope at Apache Point Observatory (APO) (Gunn et al. 2006), and fibre-fed to the APOGEE spectrograph (Wilson et al. 2010). Targeting is performed so as to simplify as much as possible the survey selection function whilst preferentially selecting red giant stars, by employing selection bins in the apparent *H*-band magnitude, and a simple colour selection in dereddened $(J-K)_0$ (Zasowski et al. 2013, 2017). Spectra are reduced, combined, and then 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 relies on a pre-computed library of synthetic stellar spectra (Zamora et al. 2015) computed using a customized linelist (Shetrone et al. 2015) to measure stellar parameters, 19 element abundances, and heliocentric radial velocities of MW stars (Holtzman et al. 2015). Abundances are well tested against samples from the literature (Jönsson et al. 2018). We use distances for stars in APOGEE DR14 measured by the Brazilian Participation Group (BPG, Santiago et al. 2016), included in a publicly available Value-Added Catalogue (VAC). These distances are measured using an early version of the StarHorse code (Queiroz et al. 2018), and combine spectroscopic and photometric information to make a Bayesian distance estimation. The precision of these measurements is expected to be $\sim$15 per cent, which we determine to be similar to distance estimates derived from the current *Gaia* parallaxes at the range of distances spanned by our sample of interest. Given that the parallax measurements can be uncertain in some cases, and especially so for distant stars, the use of spectro-photometric distances is well motivated.

We examined the distribution of our sample stars in chemical composition space considering all elemental abundances available in the APOGEE DR14 catalogue. However, in this paper, we choose to focus on the abundances of Fe, Mg, Al, and Ni, which are the ones providing interesting insights into the nature of the accreted halo stellar population. The elemental abundances are measured as part of the ASPCAP pipeline, which uses a two-step process. First, the stellar parameters $T_{\rm eff}$, $\log(g)$, v_μ, [M/H], [α/M], [C/M], and [N/M] (where v_μ is the micro-turbulent velocity) are determined via a global fit to the aforementioned spectral library (Zamora et al. 2015). The individual element abundances are then calculated by adjusting the [M/H] ([C/M] and [N/M] for Carbon and Nitrogen, and [α/M] for α elements) of the best-fit spectrum, and finding the best match for the observed spectrum in windows around features in the spectrum which are dominated by each element. The abundances are then all estimated consistently, and can then be calibrated internally relative to open cluster observations. The internal calibrations are performed to account for systematic abundance variations with $T_{\rm eff}$. In DR14, an external calibration is applied that forces the abundance ratios of solar metallicity stars located near the solar circle to be equal to solar (Holtzman et al. 2018). This small zero-point correction should be taken into consideration when making comparisons between our results and other data.

2.2 *Gaia* DR2 and cross-matching

The ESA–*Gaia* mission is a space-based astrometric survey which is providing an unprecendented mapping of MW stars in phase space. The second data-release, *Gaia* DR2 (Gaia Collaboration et al. 2018),

[3]It was acknowledged by Kruijssen et al. (2018) that 'Canis Major' as it is known to the community is no longer considered as a genuine accreted stellar population. The term is used there historically, as many of the GCs were those originally associated with 'Canis Major' and so, for clarity, we adopt it here when discussing those results.

provides five-parameter astrometry (proper motions, positions, and parallaxes) for over 1.3 million objects in the Galaxy. Combined with accurate radial velocities and spectro-photometric distance estimates from APOGEE DR14, these data make possible the calculation of 6D phase-space coordinates for objects in common between the surveys. Many improvements were made to the data-processing between *Gaia* DR1 (Lindegren et al. 2016) and DR2, examples of which include: improvements to the source detection algorithm, better modelling of the spacecraft attitude, and the fact that DR2 uses its own reference frame based on quasars (whereas DR1 was tied to the *Tycho-2* and HIPPARCOS catalogues for proper motion measurements). As a result of these improvements, the typical uncertainty on astrometric parameters is expected to be ∼0.2–0.3 mas in the middle of the magnitude range (going up to ∼2 mas for the faintest sources). While the exact selection function of *Gaia* is as yet not well known, DR2 has improved completeness in bright stars, and the survey is expected to be complete between $G = 12$ and 17.

We perform a cross-match between APOGEE DR14 and *Gaia* DR2 using the CDS X-match service[4] and adopting a conservative position mismatch tolerance of 0.5″. We find that the full, uncut APOGEE DR14 catalogue has 254 789 matched objects in *Gaia* DR2 (∼99 per cent), 83 189 of which have full 6D phase-space coordinates (using APOGEE radial velocities), have no warning or bad flags from the APOGEE reduction and ASPCAP analysis, and were not observed during commissioning of the APOGEE instrument (the main factor that reduces the sample size are the APOGEE data quality flag cuts). Of these objects, 81 491 have reliable distance measurements in the APOGEE DR14 distance VAC. We transform the observed data into the Galactocentric coordinate frame, assuming the solar motion of Schönrich, Binney & Dehnen (2010), propagating the observational uncertainties while accounting for the correlation between errors in the *Gaia* data. The sample extends from Galactocentric cylindrical radii $R \sim 3$ kpc out to $R > 15$ kpc, reaching up to a maximum of 10 kpc away from the midplane. Throughout the paper, we assume the solar radius $R_0 = 8$ kpc, and its distance from the midplane $z_0 = 0.025$ kpc. The combined spectra of our final sample have a minimum SNR ∼ 40, and a median SNR ∼ 150, corresponding to median uncertainties on [Fe/H] ∼ 0.01 dex, [Mg/Fe] ∼ 0.02 dex, [Al/Fe] ∼ 0.05 dex, and [Ni/Fe] ∼ 0.02 dex.

Stars located within $r = 3\,r_{\rm tidal}$ of the centres of known globular clusters were excluded from the sample. Tidal radii, $r_{\rm tidal}$, and cluster centres were adopted from the 2010 edition of the Harris (1996) catalogue.[5] This conservative cut removes 19 stars, after removal of stars not belonging to the main APOGEE sample (e.g. stars from APOGEE ancillary science programs).

2.3 Orbital parameters

To study the kinematical structure of the halo population in the APOGEE–*Gaia* catalogue, it is necessary to estimate the orbital parameters of the stars and their associated uncertainties robustly. To make these estimates, we use the fast orbit parameter estimation method of Mackereth & Bovy (2018), which adapts the Stäckel fudge method for estimating action-angle coordinates in axisymmetric potentials (presented in Binney 2012) to directly estimate the orbital eccentricity, e, apo-, and pericentre radii, $r_{\rm ap}$ and $r_{\rm peri}$, and the maximum vertical excursion, $Z_{\rm max}$, to high precision and without recourse to orbit integration (which can make the proper propagation of uncertainties computationally costly at this scale). We also estimate the orbital actions, J_R, L_Z, and J_Z for each star via the same method. All estimates are performed using the implementation of the Stäckel approximation in the python package `galpy` (Bovy 2015), and assuming the `MWPotential2014` Milky-Way mass model included in `galpy`. For each star, we Monte-Carlo (MC) sample the errors by constructing the covariance matrix of the observed data. We sample 100 realizations of the observed coordinates of each star, and compute the orbit parameters and actions for each sampled point, cataloguing the median value of the samples for each parameter, their standard deviation, and the correlation between parameters. Performing the orbital parameter estimation with many more than 100 samples makes little difference to the median and standard deviation obtained, so this number was selected for computational efficiency.

2.4 The EAGLE simulations

In Section 4, we undertake a simple analysis of simulated galaxies from the EAGLE suite (Crain et al. 2015; Schaye et al. 2015). EAGLE models the formation and evolution of galaxies in the context of ΛCDM cosmology. The simulations are run using a version of the smoothed particle hydrodynamics (SPH) and TreePM gravity solver GADGET 3 (described most recently by Springel 2005), modified to include the Hopkins (2013) pressure–entropy formulation of SPH, a time-step limiter (Durier & Dalla Vecchia 2012), and switches for artificial viscosity and conduction (as proposed by Cullen & Dehnen 2010; Price 2010). EAGLE produces a realistic population of galaxies by virtue of the calibration of its feedback efficiency to the observed galaxy stellar mass function (GSMF), the stellar versus black hole mass relation and galaxy disc sizes, but has been shown to reproduce a broad range of observed galaxy properties and scaling relations, such as the Tully–Fisher relation (Ferrero et al. 2017), colour–magnitude relationships (Trayford et al. 2015, 2016, 2017), and galaxy size evolution (Furlong et al. 2017). All particle data and galaxy catalogues are now available publicly (see McAlpine et al. 2016) and are accessible online.[6]

The EAGLE suite offers various box sizes ranging from 12.5 to 100 cMpc on a side. Here, we examine the higher resolution simulation adopting the 'Recalibrated' model parameters (see Schaye et al. 2015), referred to here as L025N752-Recal (we use the lower resolution, 'Reference' model simulations to test the numerical convergence of our results in Appendix B). The motivation behind the use of the higher resolution simulation is the better resolution of small dwarf galaxies, and therefore better sampling of the accreted galaxy population within the EAGLE haloes. L025N752-Recal has dark matter particles of mass $1.21 \times 10^6\,{\rm M}_\odot$ and an initially equal number of SPH particles with mass $2.26 \times 10^5\,{\rm M}_\odot$, adopting a Plummer-equivalent gravitational softening length $\varepsilon_{\rm com} = 1.33$ ckpc, limited to a maximum proper length of $\varepsilon_{\rm prop} = 0.35$pkpc. We select a sample of galaxies which have virial masses at $z = 0$ roughly equal to that proposed for the Milky Way, between $M_{200} = 0.8$ and $2.0 \times 10^{12}\,{\rm M}_\odot$. In the L025N752-Recal simulation, this corresponds to $N = 22$ galaxies with a wide range of stellar masses, and a range of assembly and accretion histories. The mean halo mass of the galaxies is $0.9 \times 10^{12}\,{\rm M}_\odot$. The majority (81 per cent) of the galaxies are disc-dominated systems, as defined using the methodology of Correa et al. (2017), by calculating the fraction of

[4] http://cdsxmatch.u-strasbg.fr/xmatch
[5] http://physwww.mcmaster.ca/ harris/mwgc.dat
[6] http://galaxy-catalogue.dur.ac.uk

kinetic energy of a galaxies stars invested in ordered co-rotation in the plane perpendicular to the vertical component of the angular momentum, κ_{co}. Disc-dominated galaxies are defined as those with $\kappa_{co} > 0.4$. The remaining 19 per cent are a combination of galaxies undergoing mergers and those without significant disc components. We do not cut the sample based on this information, as we are interested in the debris accreted in the lifetime of the galaxies, and not their eventual morphology. Moreover, we show in Section 4 that our general conclusions are upheld regardless of the $z = 0$ galaxy morphology. This sample makes a good testbed for understanding the origin of satellite debris in the Milky Way, allowing us to study the $z = 0$ characteristics of a diverse sample of satellite galaxies accreted onto Milky-Way-like haloes.

To calculate the eccentricities for the simulated star particles, we solve (using the bisection method) the equation

$$L^2 + 2r^2[\Phi(r) - E] = 0, \quad (1)$$

where L is the angular momentum, $\Phi(r)$ the gravitational potential, and E the total energy of the particle. For a bound orbit, the equation has two solutions (the peri- and apocentre distances, r_{peri} and r_{ap}), which are equal for a perfectly circular orbit (Binney & Tremaine 2008, equation (3.14)). Unbound particles are disregarded. The eccentricity is then calculated as

$$e = \frac{r_{ap} - r_{peri}}{r_{ap} + r_{peri}}. \quad (2)$$

In practice, performing this estimation requires the assumption that the potential is spherically symmetric, calculated using the mass enclosed in the sphere with radius, r, which is equal to the distance of the particle to the galaxy centre. As most of the accreted particles reside in the haloes of the central galaxies, this provides a good approximation of the potential. Performing an equivalent analysis to that performed on the data would require a procedure for fitting an axisymmetric potential to EAGLE galaxies, which is beyond the scope of this study. As the assumed potential approaches spherical symmetry, these two approximations become equivalent, and so we are confident that the comparison is sound, given that most of the APOGEE–*Gaia* stars being considered here reside in the Galactic halo (where the potential can be assumed to be near-spherical).

3 THE HIGH ECCENTRICITY HALO POPULATION

As mentioned above, Helmi et al. (2018) used APOGEE-DR14 abundances to show that a stellar population they initially identified in phase space occupies a distinct locus in the [α/Fe]–[Fe/H] plane. Among all α elements available from APOGEE spectra, Mg is the one for which abundances are the most reliable in the metal-poor regime, due to the number of available lines and their strength at low metallicity. Therefore, we decide to first look at how stellar populations are distributed in the Mg–Fe plane as a function of orbital eccentricity. We show the [Mg/Fe]–[Fe/H] distribution of APOGEE DR14 in Fig. 1, where symbols are colour-coded according to each particle's orbital eccentricity.

Three main groups of stars are apparent in [Mg/Fe]–[Fe/H] space, which we demarcate using solid and dashed black lines (the dotted grey line indicates the upper limit imposed on [Fe/H] for the analysis in Section 3.1.2). The focus of this paper is on the group we call *halo stars* – a term we adopt merely as a label, which does not necessarily imply a definitive assignation of every member star to the Galactic halo. These stars occupy the same locus as the stellar population identified by other groups as associated to a major accretion event (see Section 1), as well the 'LMg' stars classed as accreted halo by Hayes et al. (2018) and Fernández-Alvar et al. (2018b) on the basis of APOGEE DR13 chemistry and kinematics. This group extends between ([Fe/H], [Mg/Fe]) $\sim(-2.0, 0.3)$ to $\sim(-0.5, -0.1)$ and is dominated by high-eccentricity stars. The other groups are the well-known, high-α. and low-α *disc* populations, characterized in detail in several previous studies (e.g. Nidever et al. 2014; Hayden et al. 2015; Masseron & Gilmore 2015; Mackereth et al. 2017), and commonly conflated with the thick and thin discs, respectively. As expected, the latter populations are dominated by stars in very circular orbits, although the high-α group contains a non-negligible population of stars in fairly eccentric orbits ($e > 0.7$) – a topic that will be explored in a future study.

Careful inspection of the 'halo stars' in Fig. 1 suggests that there is some dependence of [Mg/Fe] on eccentricity, in that stars with higher e have slightly lower [Mg/Fe]. Further examination of the data suggests that the same is the case for [Al/Fe] and possibly [Ni/Fe] and to a lesser extent [(C + N)/Fe]. This finding motivates us to attempt an identification of sub-structure in chemo-kinematic space in an objective and data-driven fashion, so as to avoid 'cherry-picking' arbitrary selection limits in various parameters. In this way, we hope to characterize the populations within the halo star locus by sub-dividing its members into meaningful groups. We use the `scikit-learn` *k*-means clustering algorithm (Pedregosa et al. 2011) in the space of [Fe/H], [Mg/Fe], [Al/Fe], [Ni/Fe], and eccentricity e. The choice of these parameters is guided by the fact that, on one hand, these elements fall into the main groups: α, odd-z, and Fe peak elements, respectively, and on the other hand it appears, at least from Fig. 1, that eccentricity is a useful discriminator between disc and halo populations. We limit the maximum [Fe/H] to –0.7, to minimize contamination by disc stars, finding that high-e stars in the low [α/Fe] disc locus at higher [Fe/H] are not chemically similar to those in the 'halo stars' population and can be confidently disregarded. We set the assumed number of clusters at $k = 4$, to anticipate the expected separation of the high and low [α/Fe] disc, and then to allow for subdivision of the halo population into any potentially meaningful groups. We find that the algorithm groups the high and low [α/Fe] disc stars separately, and find two clear groups in the lower [Fe/H] space. Setting $k > 4$ subdivides the lower [Fe/H] groups in an unstable manner, while $k = 4$ provides very good stability over many iterations of the algorithm. We also test the algorithm at $k < 4$, finding that the stability is decreased also at lower k. If the 'halo stars' group from Fig. 1 is isolated, using $k = 2$ clustering can still recover the two groups found when clustering the whole data set. We retain the full data set and use the $k = 4$ clustering to avoid resorting to an arbitrary selection of the division between halo and disc stars. Re-scaling the data and re-running the clustering algorithm leads to negligible differences in the results. It is worth noticing that both halo populations have lower abundance ratios than thick disc stars (in the region where they overlap in [Fe/H]) for C + N, Si, K, Ca, and possibly also Mn.

We show the eccentricity distributions of the *k*-means groups in Fig. 2. The high and low [α/Fe] disc groups are combined into a single group in this plot, as these stars are not of interest to this work, except as a comparison sample. The halo stars naturally split up into two groups, one with intermediate eccentricities, peaking at $e \sim 0.5$, and another with very high eccentricities, peaking at $e \sim 0.9$. The lower e population distribution is fairly broad whereas the high-e population is very strongly peaked towards the highest e values (with some small skew to low e). The high- and low-e groups contain 679 and 318 stars, respectively. We also show the *k*-means groups in [Al/Fe]–[Mg/Fe] space in Fig. 3. This abundance plane

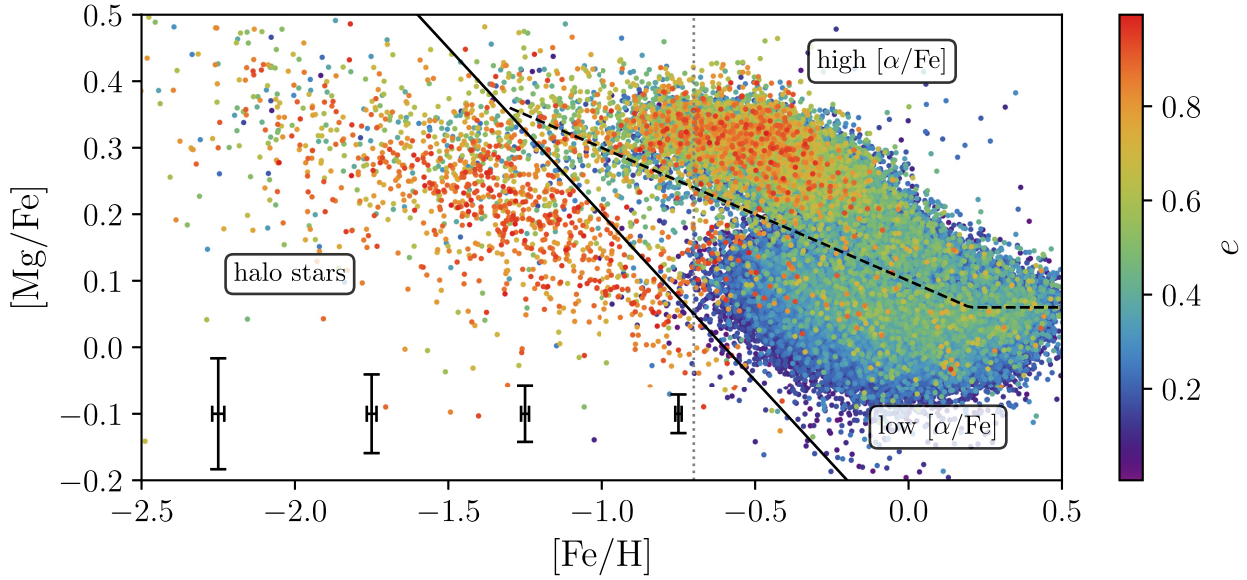

Figure 1. The [Mg/Fe]–[Fe/H] plane in APOGEE DR14, coloured by orbital eccentricity e, as estimated using the method of Mackereth & Bovy (2018). The points are plotted with the highest e stars overlaying the points at lower e, such that the highest e populations stand out. Plotted also are the mean error bars within [Fe/H] bins of width 0.5 dex between −2.5 and −0.5 dex. It is clear that a population extends from ([Fe/H],[Mg/Fe]) ∼ (−2.0, 0.3) to ∼ (−1.0, 0.1) that appears to consist mainly of stars on highly eccentric orbits, with a distinct element abundance pattern to that of the Galactic disc (at [Fe/H] > −0.7). The dotted grey line reflects the cut in [Fe/H] which is imposed to perform the k-means analysis.

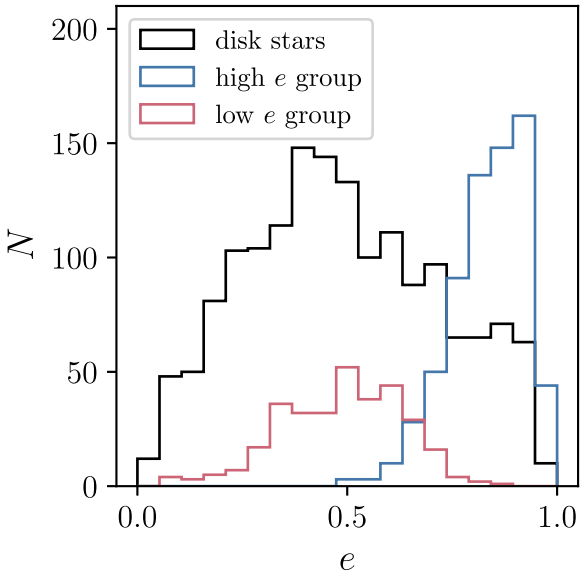

Figure 2. Eccentricity distributions of stars in three groups identified by performing a k-means detection of structure in the eccentricity-[Fe/H]-[Mg/Fe]-[Al/Fe]-[Ni/Fe] space. The k-means algorithm cleanly separates the accreted halo component into two groups, one characterized by low eccentricities (in red), and the other with a peak at very high e (in blue). We show the stars assigned to the disc (with [Fe/H] < −0.7) in black as a comparison sample.

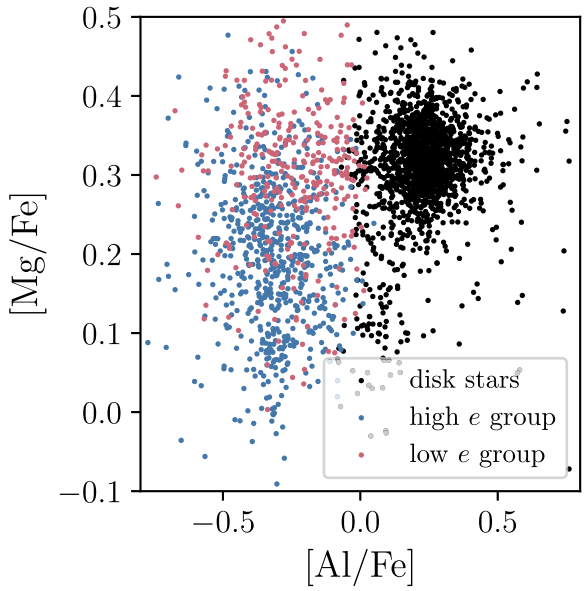

Figure 3. Distribution of the stars in the different k-means groups in the [Mg/Fe]–[Al/Fe] plane. High-e stars occupy preferentially the low [Al/Fe], low [Mg/Fe] region of the plot, a locus which was identified in Hawkins et al. (2015) as being common to accreted halo populations. Low-e group stars tend to be more distributed at higher [Mg/Fe] and slightly higher [Al/Fe]. Disc stars (in black) are enhanced in Al and, to a lesser extent, Mg, relative to the high- and low-e groups.

was shown by Hawkins et al. (2015) to discriminate well between accreted and *in situ* stellar populations in the Milky Way, with the former occupying the low [Al/Fe], low [Mg/Fe] region of the plot. Indeed, we find that the halo population separates out from the disc population very clearly in this plane, being characterized by lower Al and Mg abundances. Moreover, while there is considerable overlap between high- and low-e populations in this abundance plane, the high-e group occupies a lower [Mg/Fe] locus, on average, than the low-e group.

In conclusion, the results above show that once APOGEE chemical compositions are combined with kinematics inferred from Gaia DR2 proper motions, ∼2/3 of the stars in the accreted halo population identified in previous studies are characterized by highly

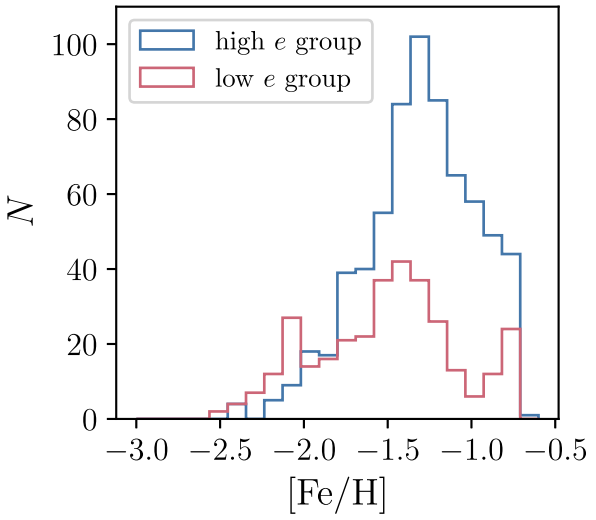

Figure 4. Raw metallicity distribution functions of the high- and low-*e* populations. Unlike their low-*e* counterparts, the high-*e* population shows a strong peak at about [Fe/H] ~ -1.3, resembling MDFs of Local Group dwarfs. The low-*e* group shows peaks at both [Fe/H] ~ -2.1 and ~ -1.4, and stars piled up at the imposed upper [Fe/H] limit at [Fe/H] $= -0.7$.

eccentric orbits. In the following sub-sections, we examine the properties of this population in both orbital and element abundance space compared to those of their low-*e* counterparts.

3.1 Chemical compositions

3.1.1 Metallicity distributions

In this section, we examine more closely the chemical compositions of the halo stellar populations, keeping track of the differences between the high- and low-*e* groups. We start by examining the metallicity distribution function (MDF) of these populations. The *raw* MDFs are displayed in Fig. 4, and are *not* corrected for the APOGEE selection function. Correction for selection function effects cannot be performed to the full sample of stars studied here, as many of the sample stars belong to magnitude bins on plates where observations were incomplete in DR14 and so are not part of the statistical sample (see Zasowski et al. 2013; Bovy et al. 2014, for details on the selection function of APOGEE and the targeting strategy, respectively). As a result, a selection function correction would leave us with a much smaller sample, resulting in more uncertain MDFs, particularly in the case of the smaller low-*e* population sample. We nevertheless compared the lower *N*, corrected, and raw MDFs for the larger high-*e* sample, which is statistically more robust, finding them to be in good overall agreement with similar peak position and width. Therefore, we conclude that selection function effects do not affect substantially our halo sample. Moreover, it is reasonable to assume that relative differences between the MDFs of the high- and low-*e* populations are not affected significantly by selection biases from the targeting strategy, and that any differences in the MDFs are likely to be real. Therefore, we can confidently compare the MDFs of the high- and low-*e* populations, at least to first order. It should be noted here that [Fe/H] is included in the *k*-means analysis, and so these MDFs are reflective of the metallicities of the groups as defined by that procedure.

With the above caveats in mind, examining Fig. 4 we learn that the MDF of the high-*e* population peaks at [Fe/H]$\sim$–1.3 and has a width of approximately 0.6 dex (FWHM), with a tail towards the metal-poor end, resembling the MDF of a classical closed-box model. Its MDF is in fact qualitatively similar to those of relatively massive dwarf galaxies in the Local Group (e.g. Ross et al. 2015), and in particular quite similar to those of the Small Magellanic Cloud (SMC, Dobbie et al. 2014) and Leo I (Kirby et al. 2013). This is an interesting result given that Hayes et al. (2018) showed that their LMg population (of which the high-*e* stars are a sub-group) had a star formation history similar to that of massive dwarfs such as the LMC prior to its accretion, which if accreted would have likely been around present-day SMC mass ($\sim 3 \times 10^8\,M_\odot$, e.g. Stanimirović, Staveley-Smith & Jones 2004; Rubele et al. 2018) at the time that the LMg population was accreted (van der Marel, Kallivayalil & Besla 2009; Helmi et al. 2018).

It is also instructive to compare the MDFs of the high- and low-*e* populations. The most striking difference is that the MDF of the low-*e* population has no strong peak, but rather two smaller (likely significant) peaks at lower metallicities. The sharp edge at [Fe/H] = –0.7 is due to the [Fe/H] limit adopted in our *k*-means grouping, and the smaller peak of the low-*e* population at [Fe/H] > -1.0 is probably due to disc contamination. Interestingly, the two peaks towards the lower metallicity end happen at similar metallicities to those previously assigned to the inner and outer halo populations (e.g. Carollo et al. 2007; Allende Prieto et al. 2014; Fernández-Alvar et al. 2015). Most importantly, the MDFs of the high- and low-*e* populations are strikingly different and, as argued above, this difference is almost certainly insensitive to target selection effects, which then strongly suggests that the two populations have different origins.

It is not entirely clear how the difference in shape between the MDFs of the high- and low-*e* populations can be understood in terms in of the halo accretion history. While the MDF of the high-*e* population lends support to the notion that those stars were injected into the halo as part of one major accretion event, it is hard to draw any strong conclusion regarding the low-*e* population.

3.1.2 Abundance ratios

In this section, we inspect abundance ratios of the high- and low-*e* groups, to gain further insights into the origins of these two populations. Fig. 5 displays the run of [Mg/Fe], [Al/Fe], and [Ni/Fe] as a function of [Fe/H]. For each element, we take the running median and interquartile range of [X/Fe] as a function of [Fe/H]. We calculate the median abundance in bins of 50 stars sorted by increasing [Fe/H], which is shown by the red and blue coloured bands for the high- and low-*e* populations, respectively.

The top panel shows that in general, [Mg/Fe] decreases with [Fe/H], and both the high- and low-*e* populations have [Mg/Fe] lower than that of the high-α disc stars at same [Fe/H], in the metallicity interval where those populations overlap ($-1.5 \lesssim$ [Fe/H] $\lesssim -0.5$), as previously discussed by Hayes et al. (2018) and Fernández-Alvar et al. (2018b). Moreover, the high- and low-*e* populations occupy slightly separated sequences, such that the former has lower [Mg/Fe] at fixed [Fe/H]. It appears from the simple running median that the high-*e* population exhibits a change in slope of [Mg/Fe] against [Fe/H], at [Fe/H] ~ -1.3, whilst there is only slight evidence change of slope in the lower edge of the low-*e* relation. In order to test this further, and to determine more rigorously the location of the change in slope, we use a Bayesian inference to determine a generative model for the [Mg/Fe]–[Fe/H] distribution of both populations. The full procedure is outlined in Appendix A, the basic principle being that we fit a piecewise-linear

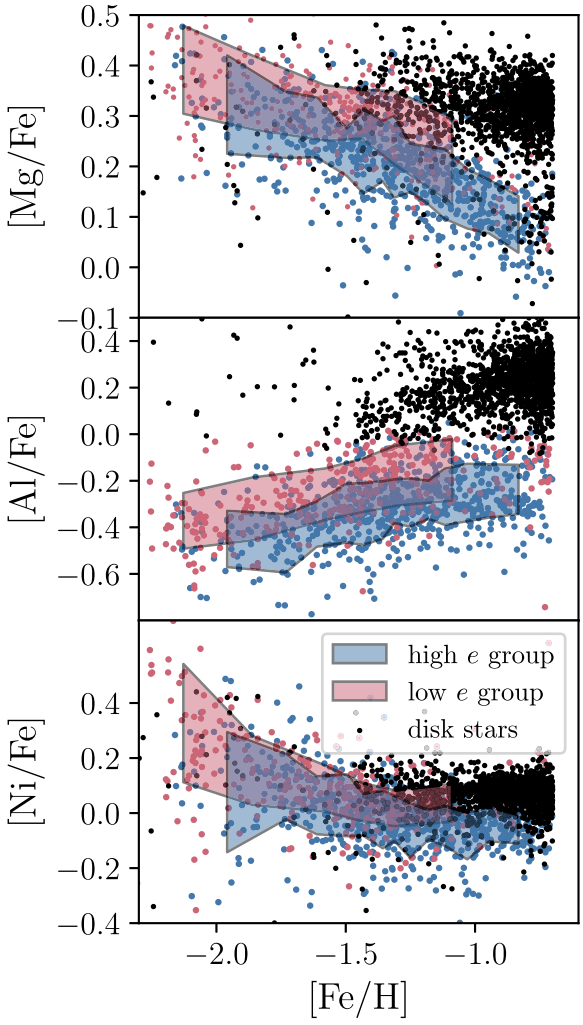

Figure 5. Abundance ratios for the k-means groups as a function of [Fe/H]. These elements are good representations of α, odd-Z, and Fe peak elements that are measured by APOGEE. The coloured bands show the interquartile range in bins of 50 stars sorted in increasing [Fe/H]. The high- and low-e populations are characterized by lower Mg and Al than the high-α stars. The bottom panel shows that the high-e group has slightly lower [Ni/Fe] than the high-α population, whereas the low-e group is essentially consistent with the high-α population in [Ni/Fe]. In all cases, the high-e group median is slightly lower than the low-e group at fixed [Fe/H]. The slightly higher [Fe/H] of the high-e population is also evident here. Again, both groups are depleted in all these elements relative to the disc stars at that [Fe/H].

model to the data of the form:

$$[\mathrm{Mg/Fe}]([\mathrm{Fe/H}]) = \begin{cases} m_1[\mathrm{Fe/H}] + b_1 & [\mathrm{Fe/H}] < [\mathrm{Fe/H}]_0 \\ m_2[\mathrm{Fe/H}] + b_2 & [\mathrm{Fe/H}] > [\mathrm{Fe/H}]_0, \end{cases} \quad (3)$$

where $m_{[1,2]}$ and $b_{[1,2]}$ are the slope and y-intercept on either side of the break, which is positioned at $[\mathrm{Fe/H}] = [\mathrm{Fe/H}]_0$. The fitting procedure accounts for the error on both [Mg/Fe] and [Fe/H], allowing a proper assessment of the significance of any fitted breaks. If there is no break in the range of [Mg/Fe] and [Fe/H] considered, the data are fit by a simple linear relation, as the break is pushed to the edge of the [Fe/H] range.

We show the models fitted to the high- and low-e groups in Fig. 6. The data for each group are plotted in the black points, and the best fit and 95 per cent confidence interval are shown by the

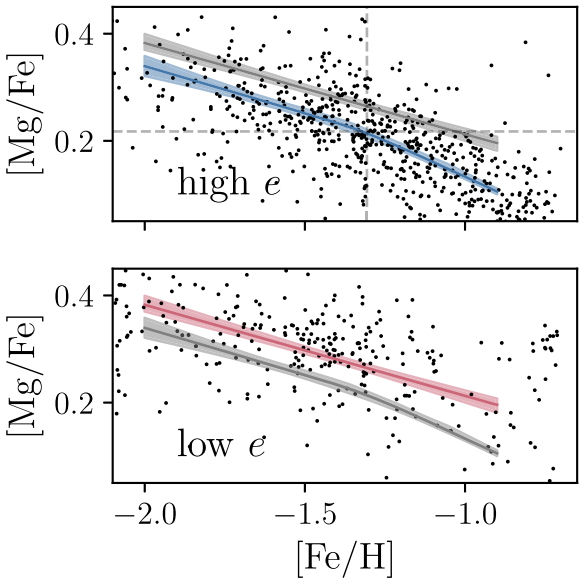

Figure 6. Best-fit models for [Mg/Fe] as a function of [Fe/H] for the high- and low-e k-means groups (shown top and bottom, respectively). The best-fit model is shown by the blue and red lines, with the 95 per cent confidence interval marked by the banded blue and red regions either side of this line (the fit for the other group is shown in each panel for reference). The raw data are shown by the black scatter points. The high-e group is well-fit by a model with a break at $[\mathrm{Fe/H}] = -1.31^{+0.03}_{-0.06}$ and $[\mathrm{Mg/Fe}] = 0.22^{+0.01}_{-0.08}$, as indicated by the dashed crosshair. The low-e group is better fit by a single linear relation in the same range of [Fe/H]. The position of the models also demonstrates the slightly higher [Mg/Fe] of the low-e group.

line and coloured bands in each panel. The high-e group model (shown in the top panel) has a change of slope, indicated by the dashed crosshair, positioned at $[\mathrm{Fe/H}] = -1.31^{+0.03}_{-0.06}$ and $[\mathrm{Mg/Fe}] = 0.22^{+0.01}_{-0.08}$. The slope to the metal-poor side of the break is found to be -0.15 ± 0.01, and is roughly consistent with the slope before the break in the high [Mg/Fe] disc star sequence (e.g. that seen in Hayden et al. 2015). A similar metal-poor end slope is also present in the work of Fernández-Alvar et al. (2018b), who used APOGEE DR13 data to determine also that a change of slope is present in low [Mg/Fe] stars, located at [Fe/H] ~ -1.0 dex, in rough agreement with that found here. At $[\mathrm{Fe/H}] > [\mathrm{Fe/H}]_0$, the slope steepens to become -0.26 ± 0.01. The low-e population is well fit in this range of [Fe/H] by a simple linear relation with a slope equal to -0.14 ± 0.01, with no significant break found. As it is likely that at least some of the stars at the highest [Fe/H], and at high [Mg/Fe] in this population are high [α/Fe] disc contaminants, we also perform the same fit using stars at [Fe/H] $< -1.$ dex, and still find no evidence of a break at the same position as that found in the high-e group. Fitting only low-e stars with [Fe/H] > -1.1 dex, we find that the best fit is a single linear relation with a slope equal to that found when using the whole low-e group.

This change of slope in the [Mg/Fe]–[Fe/H] relation of the high-e population offers important clues on the nature of this stellar population. This feature is most commonly interpreted as being due to the onset of Fe enrichment by SNe Ia. The [Fe/H] at which the change of slope occurs is primarily related to the star formation efficiency (e.g. Weinberg, Andrews & Freudenburg 2017), and additionally depends on the gas inflow and outflow rates of its parent galaxy. It follows that it should also be a function of the mass of the parent galaxy, which regulates gas density and gas inflow/outflow. Data for stellar samples from Local Group dwarf galaxies indicate

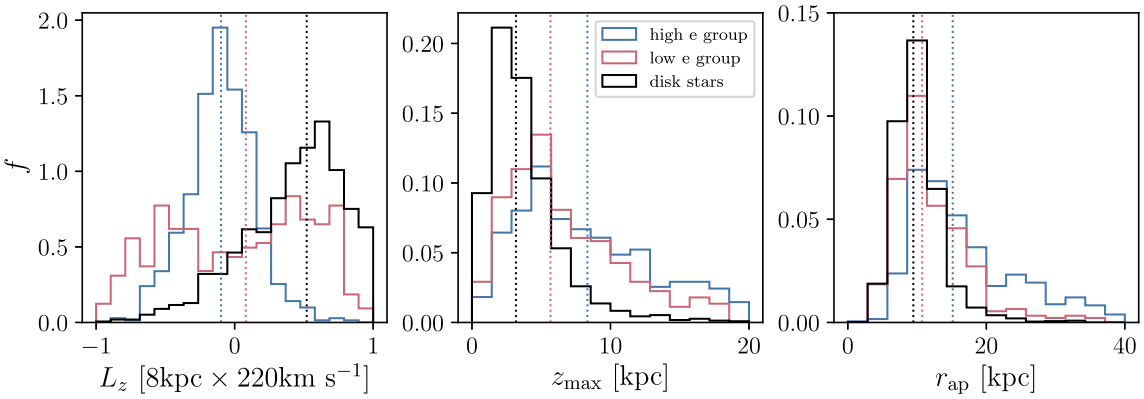

Figure 7. Kinematics of the k-means-selected stars from Fig. 1. The distribution of azimuthal angular momentum L_z (*left-hand*), maximum vertical excursion from the disc plane $z_{\rm max}$ (*centre*), and the spherical apocentre radius of orbits $r_{\rm ap}$ (*right-hand*), are shown for the high- and low-[α/Fe]-disc (red and yellow, respectively) and the accreted halo population (blue). The halo stars clearly occupy a very different orbital distribution, having low-L_z, and distributions of $z_{\rm max}$ and $r_{\rm ap}$ that extend to very large distances. The median L_z is slightly negative, while the median $z_{\rm max}$ and $r_{\rm ap}$ are ∼10 and ∼20 kpc, respectively. All histograms are normalized such that the summed probability under *each* group is equal to unity.

the presence of this same change of slope on the Mg–Fe plane, and it is suggested that the [Fe/H] at which it occurs is higher in galaxies with larger luminosity, and presumably higher stellar mass (see, e.g. Tolstoy, Hill & Tosi 2009). Given the [Fe/H] at which data for Local Group dwarf galaxies show a change of slope (see, e.g. Tolstoy et al. 2009), we estimate that the mass of the progenitor of this high-*e* population was somewhere between 10^8 and 10^9 M$_\odot$. This rough mass estimate is in good agreement with that inferred from arguments based on the MDF of the high-*e* population in Section 3.1.1. It moreover places the mass involved in this accretion event within the range of the total stellar mass of the Galactic halo ($\sim 4 - 7 \times 10^8$ M$_\odot$ Bland-Hawthorn & Gerhard 2016) which argues against the high-*e* population belonging to more than one major accretion event, as multiple accretions at this mass would exceed this mass limit.

In contrast, the lack of a clear change of slope in the distribution of the low-*e* stars on the Mg–Fe plane, argues strongly for this group being in fact some mix of stellar populations with various origins, including stars accreted in different smaller events, stars formed *in situ*, ejected members of the disc, and contamination by members of the low-*e* tail of the high-*e* group.

The middle panel of Fig. 5 shows the run of [Al/Fe] with [Fe/H]. As pointed out by Hayes et al. (2018), their LMg (and thus also our high *e*) population is characterized by much lower [Al/Fe] than the disc populations at same [Fe/H]. In fact, [Al/Fe] in those populations is comparable to that of stars in the Sagittarius dwarf spheroidal (Hasselquist et al. 2017). In addition, the high- and low-*e* populations occupy separate loci in [Al/Fe]–[Fe/H] space, with the high-*e* population showing lower [Al/Fe]. For both populations, [Al/Fe] increases with [Fe/H]. As an odd-*Z* element, Al is an important component in the explosive ejecta of massive stars, synthesized in their carbon and nitrogen hydrostatic burning phases. It is also a product of Asymptotic Giant Branch (AGB) stellar nucleosynthesis, particularly at [Fe/H] > -0.2 dex (see, e.g. Andrews et al. 2017). Importantly, it is known that odd-*Z* element yields should be metallicity dependent to some extent due to their dependence on a high neutron surplus (e.g. Arnett 1996). This fact may explain the slight correlation observed between [Al/Fe] and [Fe/H] in each group and also in the disc populations. The depletion in [Al/Fe] relative to the MW disc in both halo groups has no straightforward explanation, and we speculate that it may indicate that AGB stars contributed more importantly to the enrichment of MW disc ISM.

While the trend in Ni abundances as a function of [Fe/H] of the high-*e* group appears to be roughly similar to that of the low-*e* group, it is noteworthy that the trend is negative in both groups, such that [Ni/Fe] decreases with increasing [Fe/H]. Similarly to the other elements, Ni is slightly depleted in the high-*e* group relative to the low-*e* group, but this depletion is very small. At fixed [Fe/H], the high-*e* group has a lower [Ni/Fe] than the disc stars, whereas the low-*e* group is nearly consistent with them.

We also studied the other α, odd-*Z*, and Fe peak elements available in APOGEE. These trends appear to be consistent between most of the elements. It is noteworthy that in almost all elements, the low-*e* group appears to show slightly higher abundance ratios than the high-*e* group, at fixed [Fe/H], which may be difficult to explain by invoking lower mass accretion events, unless the star formation in these stellar populations was very short and intense. Our discussion of the EAGLE simulations in Section 4 sheds some light on this matter.

3.2 Kinematics

Fig. 7 shows the distribution of the *k*-means groups in (from left to right) azimuthal angular-momentum L_z (equivalent to the azimuthal action J_ϕ), the maximum vertical excursion of orbits above the midplane $z_{\rm max}$, and the apocentric radius $r_{\rm ap}$, which is the maximum orbital distance from the Galactic centre. The median value of each distribution is shown by the coloured vertical dashed line.

It is clear from these plots that the halo groups (again, shown in red and blue) have a very different orbit distribution than the disc stars (in black). The latter are characterized by highly prograde orbits with relatively low vertical excursions. In turn, the two accreted halo populations differ substantially, particularly in terms of their L_z distributions. The high-*e* stars all have extremely low angular momentum, such that many of the stars in this population have negative L_z. The lower *e* population has a large spread in L_z, such that many stars are on prograde orbits while others have negative L_z, thus moving on retrograde orbits. The lack of a distinct single peak in L_z space for the low-*e* group may indicate again that this group is in fact a superposition of populations with different origins, including the debris of smaller mass-accreted satellites, whereas the

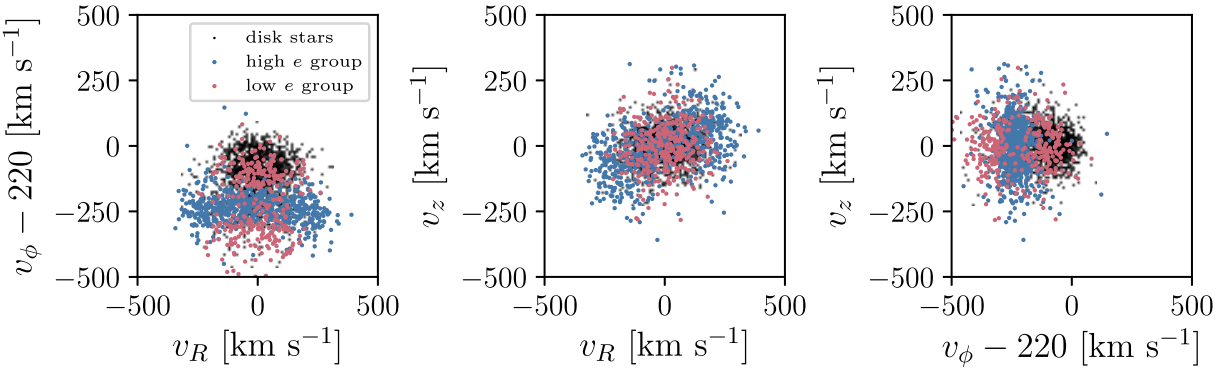

Figure 8. Distribution of disc and accreted halo populations in spherical polar coordinate planes. On the left-hand panel, the high-e population occupies the same locus in $v_R - v_\phi$ space (left-hand panel) as the population identified by Belokurov et al. (2018). The low-e population splits into two populations according to v_ϕ, with one prograde and one retrograde component, suggesting that this population may in fact be from a mix of disc contaminants and debris from smaller satellites.

single, clear peak in L_z for the high-e group supports the notion that this population is mainly the debris of a single satellite. The median value of L_z for the high-e stars is $-176\,\mathrm{km\,s^{-1}}$ kpc. A slightly retrograde motion for stars in the halo at roughly these eccentricities was noted by Helmi et al. (2018) and, as they suggest, could be either an effect of the assumption of the solar motion (which is likely subject to some systematic uncertainties), or could be a true feature of this population, which further supports its origin in a single accreted satellite. We find that a slight retrograde motion of the high-e stars is present when assuming either the Schönrich et al. (2010) or the Hogg et al. (2005) solar motion measurements, which differ markedly.

The $z_{\rm max}$ distribution shows that the stars belonging to the high- and low-e populations are on orbits that take them very far above the midplane of the Galaxy, extending to $z_{\rm max} > 10$ kpc, as expected for halo stars. The median $z_{\rm max}$ of the high-e population is higher than the other two populations, at 8.8 kpc. Similarly, the apocentre radii of these stars extend to very large distances from the Galactic centre with a median $r_{\rm ap} = 15.1$ kpc. There appears to be a secondary peak in the distribution at $r_{\rm ap} > \sim$25 kpc, which is roughly consistent with the 'apocentre pile up' suggested by Deason et al. (2018). Considering that roughly 2/3 of the stars in the halo population are on highly eccentric orbits ($e > 0.8$) which are strongly out of the disc plane ($z_{\rm max} > 10$ kpc), and have a median apocentre radius which is comparable to that of the accreted halo population analysed by Deason et al. (2018), it seems reasonable to conclude that the high-e stars are part of that population.

Further insights into the nature of the high- and low-e stellar populations can be gained by inspection of Fig. 8, where they are displayed together with disc populations on three cylindric coordinate velocity planes. The left-hand panel shows the $v_R - v_\phi$ plane, where the high-e population has a wide spread in v_R and a much narrower distribution in v_ϕ. This is the same locus as that of the population identified by Belokurov et al. (2018, their fig. 2). The low-e group, on the other hand, has a clearly bimodal distribution, with two concentrations above and below the high-e population. The concentration at higher v_ϕ shows *prograde* rotation, overlapping well with the disc population. The grouping at lower v_ϕ is markedly *retrograde*. These two sub-populations are also clearly seen on the right-hand panel, where one can notice that their distributions in v_z are very similar. The presence of such obvious structure in velocity space calls for a closer scrutiny of those two sub-populations in chemical composition space.

We compared the low-e prograde and retrograde populations in MDF, Mg–Fe, Al–Fe, and Ni–Fe spaces and found them to be, on one hand, indistinguishable from each other but, on the other hand, clearly distinct from the high-e population in all diagnostic plots. It is likely that there is some degree of inter-contamination between all these groups, so that the left-hand panel of Fig. 8 suggests strongly that the prograde population can be contributed entirely or in part by the disc. This is also suggested by the presence of a secondary peak at the high [Fe/H] end of the low-e MDF (see Fig. 4) which is dominated by prograde stars. However, quite puzzlingly, the prograde population is chemically indistinguishable from its retrograde counterpart, which is unlikely to be contributed by the disc. We test this further by re-fitting the relation between [Mg/Fe] and [Fe/H] for the separated prograde and retrograde components, using the procedure described in Appendix A. We find that in both cases, the best-fit relation matches that found for the combined population within the uncertainties, and furthermore, find no evidence of a change in slope in either population. Without further diagnostics with which to make a call on the nature of the distinct low-e populations, we summarize the situation by concluding that it is likely to be a mix of stars contributed by contaminants from the high-e population, smaller accretion events, puffed up disc, and stars formed *in situ*.

4 ACCRETED STELLAR POPULATIONS IN EAGLE

We have shown that the halo stellar population in the combined APOGEE and Gaia DR2 samples have MDFs and abundance patterns that are similar to those of dwarf galaxies in the Local Group. Furthermore, we have found that this population contains two distinct sub-groups, with different chemical and kinematic characteristics. One of these groups is characterized by highly eccentric and slightly retrograde orbits, thus sharing similar properties with stellar populations ascribed to a major accretion event in the distant past. In this section, we examine predictions by numerical simulations, drawing parallels with the observations to gain insights into the nature of this accreted halo population. The questions we intend to address with this exercise are the following: is the high-e population the result of a single relatively massive accretion event, as claimed by other groups? Do the simulations provide insight into the nature of the low-e population? Is it composed of a collection of small accretion events, or is it predominantly formed *in situ*?

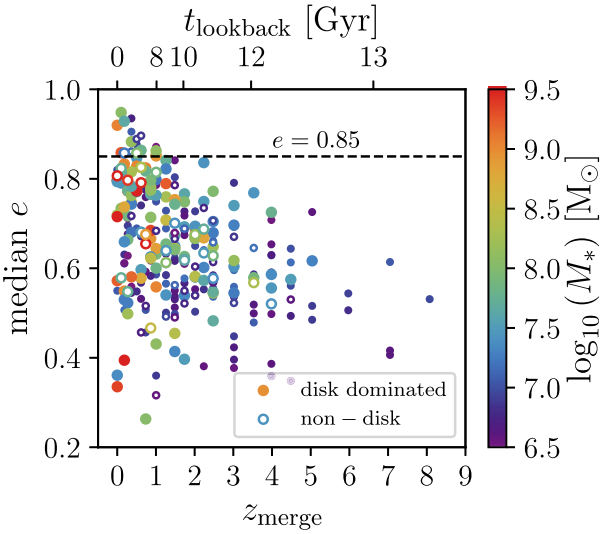

Figure 9. Merger redshift and lookback time ($z_{\rm merge}$ and $t_{\rm lookback}$) of 296 satellite galaxies accreted onto 22 Milky Way mass haloes in EAGLE against the median eccentricity e of their stellar debris at $z = 0$. The points are coloured by the stellar mass of the accreted galaxies at the time of the merger, M_*. The larger points denote those galaxies which have $N > 100$ star particles in EAGLE, whereas the smaller points are those galaxies with $20 < N < 100$ particles. Open points indicate satellites accreted onto galaxies which do not have a clear disc component at $z = 0$ (defined as described in Section 2.4). The dashed horizontal line indicates $e = 0.85$, which is the median eccentricity of the high-e group characterized in Section 3. Only the latest merged haloes have debris at the highest and lowest e, whereas early mergers tend to occupy intermediate e. The most massive mergers also occur later, with early mergers dominated by low-mass haloes.

The EAGLE simulations are a useful tool to address these questions, as they provide a cosmologically motivated history of satellite accretion for Milky Way-like galaxies, simulated self-consistently in a cosmological context. Therefore, we expect the $z = 0$ dynamical state of accreted systems around Milky Way analogues in the simulations to be a good approximation to the observations in the Milky Way halo. In order to attempt to shed light on the problem using EAGLE, we track star particles formed in galaxies which merged onto disc galaxies that eventually reach a virial mass roughly equal to that of the Milky Way ($\sim 10^{12}$ M$_\odot$). At $z = 0$, we measure the orbital eccentricity e of the star particles resulting from all well-resolved accretion events (i.e., those having >20 star particles) of systems with stellar mass M_*, occuring at redshift $z_{\rm merge}$.

In Fig. 9, we show the median e of star particle orbits against $z_{\rm merge}$ for accreted satellites onto the 22 simulated galaxies with Milky Way-like virial mass ($\sim 10^{12}$ M$_\odot$) in the EAGLE L025N752-Recal simulation. We colour each point according to the stellar mass reached by the progenitor population prior to merging onto the main progenitor. The median eccentricity of the high-e group in the Milky Way halo is indicated as a dashed horizontal line. To ensure that we properly resolve the orbital properties of the stellar component of the accreted satellites, we only show galaxies with 20 or more star particles, amounting to a total of 296 accreted satellites. The satellites with $20 < N < 100$ particles are shown by small symbols, whereas large symbols are adopted for satellites with $N > 100$. Satellites accreted onto galaxies that are not dominated by a disc component (as discussed in Section 2.4), are plotted as open symbols.

There is a striking trend of median e with $z_{\rm merge}$ in Fig. 9. The maximum median e decreases with lookback time, creating an upper envelope to the distribution of the data. As a result, $z = 0$ stellar particles resulting from the earliest accretions typically have median $e \sim 0.5$, whereas the most recently accreted satellites span a wide range of median e, from ~0.3 up to >0.9. As we discuss below, this means that early accreted satellites were accreted onto intermediate- to low-e orbits. Most importantly, this result implies that an accreted satellite whose stars at $z = 0$ are on very radial orbits is unlikely to have been accreted before $z \sim 1.5$. We verify that this trend is not due to numerical noise in Appendix B, by demonstrating that the same trend is realized in a larger volume simulation (at lower resolution), with a far larger sample of accreted satellites (see Fig. B3). We discuss the implications of this finding for the formation of the Milky Way thick disc, as suggested by Helmi et al. (2018), in Section 5.

As expected from hierarchica galaxy mass growth in a standard Λ-CDM universe, it can be seen in Fig. 9 that, at the highest redshifts, merged satellite galaxies are relatively low mass. Conversely, the most massive galaxies were accreted recently. Almost all accretion of stellar systems with $M_* \gtrsim 10^9 M_\odot$ happened at $z \lesssim 1$. This is mainly due to the time taken for satellites to build up their stellar mass. The fact that the satellite debris with the highest median e must have been accreted late also means that these satellites are likely to be high mass. We find that the median mass of accreted satellites with median $e > 0.85$ is 4.1×10^8 M$_\odot$, and a minimum mass of 0.77×10^8 M$_\odot$. The median accretion time of high-e debris is $z \sim 0.6$. This suggests that the high-e Milky Way halo population we identify is likely to have had $M_* \gtrsim 10^8$ M$_\odot$ and have been accreted after $z \sim 1.5$. Accretion events such as this are relatively rare in EAGLE, with only 5 of the 22 (~23 per cent) central galaxies having accreted any system with $10^8 < M_* < 10^9$ M$_\odot$ after $z = 1.5$ *and* retaining a median $e > 0.8$ at $z = 0$. For two of those five simulated galaxies, the aggregate of all accreted stellar mass far exceeds that of the Galactic halo (e.g. by a factor of 2 or more), leaving only 3 galaxies (~14 per cent) with accretion profiles resembling that suggested by the observations reported in this paper. It is also worth noticing that the high-e populations of the latter three galaxies, although dominated by one massive accretion event, had not had insignificant contributions by accretions of smaller mass systems. The latter suggests that it is possible that the high-e population of the Galactic halo may have contributions by low mass systems as well.

Perhaps, most importantly, the above discussion indicates that the accretion event which deposited the high-e Milky Way debris was quite unusual for a galaxy of its $z = 0$ virial mass. If one further considers that the Milky Way is currently accreting a similarly massive galaxy (the Sagittarius Dwarf), and may have accreted yet another massive system in the distant past (the *Kraken*), as proposed by Kruijssen et al. (2018), our results suggest that the overall accretion history of the Milky Way has been quite atypical when compared to the rest of the galaxies of same halo mass. This is an interesting result in light of the findings reported by Mackereth et al. (2018), which suggest that the distribution of Milky Way disc stars on the [α/Fe]–[Fe/H] plane indicate that it has undergone an unusual accretion history, for it's stellar mass. We discuss the implications of that work further in Section 5.

In Fig. 10, we show how the eccentricity of satellite debris changes between the time of the merger and $z = 0$. We measure the initial eccentricity by taking the median eccentricity of the star particles in the snapshot immediately following that in which the satellite is identified as being merged onto the main progenitor. It is clear from this figure that the orbit eccentricities are not greatly changed following accretion. The median change in eccentricity

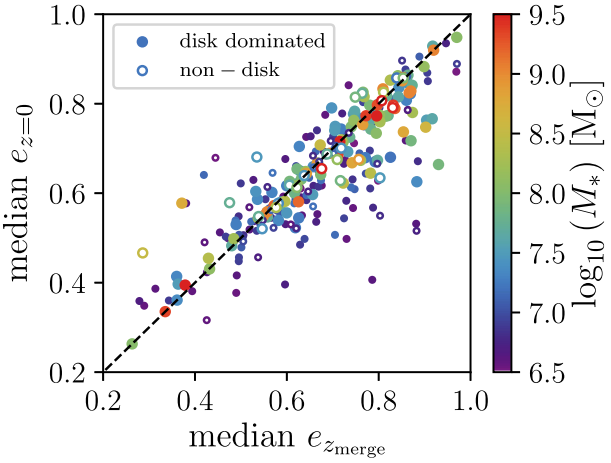

Figure 10. Median eccentricities at merger time $e_{z_{\rm merge}}$ against median eccentricity at $z = 0$ $e_{z=0}$ for satellites accreted onto Milky Way mass haloes in EAGLE. The points are coloured by the stellar mass of the accreted galaxies at the time of the merger, M_*. As in Fig. 9, point sizes denote the number of stellar particles in the satellite, and the open points show satellites accreted onto galaxies with no clear disc component. Debris which falls below the dashed unity line has been circularized following accretion, whereas the debris above this line has been radialized after accretion. The majority of debris has a similar e at $z = 0$ to the time at which it was accreted, but a few accreted satellites are circularized/radialized following accretion. High mass satellite debris has undergone the least change in median e.

is small, at ∼4 per cent, although a few accreted satellites have changed their median eccentricity by as much as ∼20 per cent. It is noticeable that the debris whose eccentricity has changed most is that from lower mass satellites. As noted previously, these are the satellites commonly accreted earlier, although we find no significant trend between $z_{\rm merge}$ and the change in eccentricity, meaning that changes in the median e are not strongly time-dependent. By the same token, the highest mass satellite debris are those which have undergone the least change in median e, likely because they were accreted late, when the gravitational potential of the host dark matter halo is no longer varying significantly. Furthermore, no satellites accreted at lower e than ∼0.7 have debris that is radialized to higher than $e \sim 0.8$ by $z = 0$. The finding that median e changes very little is important, as it means that the trend between $z_{\rm merge}$ and median e at $z = 0$ (Fig. 9) is cosmological in nature, and not due to dynamical effects, substantiating the idea that high orbit eccentricity is a good indicator of late-time accretion. The trend between $z_{\rm merge}$ and median e is likely a product of the changing merger cross-section with cosmic time, whereby the gravitational potential of the central halo dictates the maximum impact parameter for a successful accretion (as opposed to a 'fly-by'). Galaxies in the early universe are smaller and less massive, and so approaching satellites must have a smaller impact parameter to be accreted – in fact so small at very high redshift that very eccentric orbits are essentially impossible to achieve. For example, if we assume that at high redshifts (z ∼3–4), galaxies have small virial radii, $r_{\rm vir} \sim 50$ kpc, a high-e orbit ($\gtrsim 0.8$) can only be attained by a satellite impacting at $r_{\rm peri} \lesssim 5.5$ kpc (assuming that the maximum $r_{\rm apo}$ is set by the virial radius). Conversely, for Milky-Way-like centrals with larger virial radii at $z = 0$, closer to $r_{\rm vir} \sim 250$ kpc, an impacting galaxy can merge anywhere with $r_{\rm peri} \lesssim 28$ kpc for an orbit with similar eccentricity, meaning that high eccentricity mergers are more likely at lower redshifts. This logic supports our contention that the high-e group in the Milky Way must have been a relatively late accretion event.

It is noteworthy that the morphology of the central galaxy does not seem to impact the trends seen in Figs. 9 or 10, since the galaxies without dominant disc components (open symbols) trace roughly the same locus as their disc-dominated counterparts (filled symbols). This is of interest because it suggests that the galaxy morphology is generally not necessarily reflective of the galaxy assembly history in terms of merger time, mass, and orbital properties following the merger, at least on the relatively long time-scales shown in these figures. Moreover, disc and non-disc galaxies have essentially the same mean and standard deviations around the identity relation in Fig. 10 (to within a fraction of a per cent). This result suggests that the evolving galaxy morphology has little effect on the orbital evolution of accreted debris once the satellite is fully unbound.

Regarding the low-e group, the numerical simulations reinforce the notion that it may be contributed by a combination of smaller and/or earlier accretion events. Given that intermediate-e populations can be accreted at almost any $z_{\rm merge}$ in EAGLE, and that the element abundances of this group show no clear [Mg/Fe] change of slope in the [Fe/H] range observed and no clear MDF peak at moderately high [Fe/H], it is then most likely that these stars were accreted at earlier times than the high-e stars, and in small satellites. In general, earlier and lower mass-accreted satellites in the simulation show α-element enhancements, due to the fact that these must have formed quickly (with high SFE) to be accreted at early times.

In Fig. 11, we plot the *mean* [Mg/Fe] and [Fe/H] abundances of the accreted debris shown in Figs. 9 and 10. For reference, the dashed lines display the locus occupied by disc stellar populations of central galaxies in the MW mass range (For example, galaxies like that studied in Fig. 6 of Mackereth et al. 2018). In the right-hand panel, we colour the points by the stellar mass of the accreted satellite, as in the previous figures, as well as adopting the same size and filling of the points. The left-hand panel shows the same points, but coloured by the median e of the debris at $z = 0$. The black error bar represents the median and standard deviation of the stars which are members of the high-e group from the APOGEE–*Gaia* data.

We first focus on the behaviour of simulated galaxies. It is immediately clear that the chemical compositions of the EAGLE debris are broadly consistent with those seen in the APOGEE–*Gaia* halo stars (Fig. 1). The right-hand panel clearly shows that there is a significant trend between the mass and mean [Fe/H] of the accreted satellites, such that more massive satellites have higher [Fe/H]. Comparing the left- and right-hand panels, one concludes that the higher mass satellites, which are characterized by lower [Mg/Fe] and higher [Fe/H], are those which are more likely to have higher median e. Very few of the low-mass, high [Mg/Fe] satellites have median $e \gtrsim 0.7$. These trends indicate that we can use both the kinematics and chemistry of the high-e population to constrain the mass of the accreted satellite. Looking at the right-hand panel of Fig. 11 one would conclude that the mass would be around 10^8 M$_\odot$. However, the EAGLE simulations overpredict the metallicity at fixed mass below 10^9 M$_\odot$ by about 0.5 dex (Schaye et al. 2015, Fig. 13). Correcting for that discrepancy, the mass inferred for the satellite associated with the high-e population would be somewhere between $10^{8.5}$ M$_\odot$ and 10^9 M$_\odot$.

As a further interesting point, we note that the lowest mass satellite debris in Fig. 11 achieve the highest [Mg/Fe] abundances. We interpret this result as being due to the fact that these systems form quickly with star formation characterized by short bursts and an early cessation, so that SN Ia could not contribute substantially to the chemical enrichment of the gas. As a caveat, however, we point

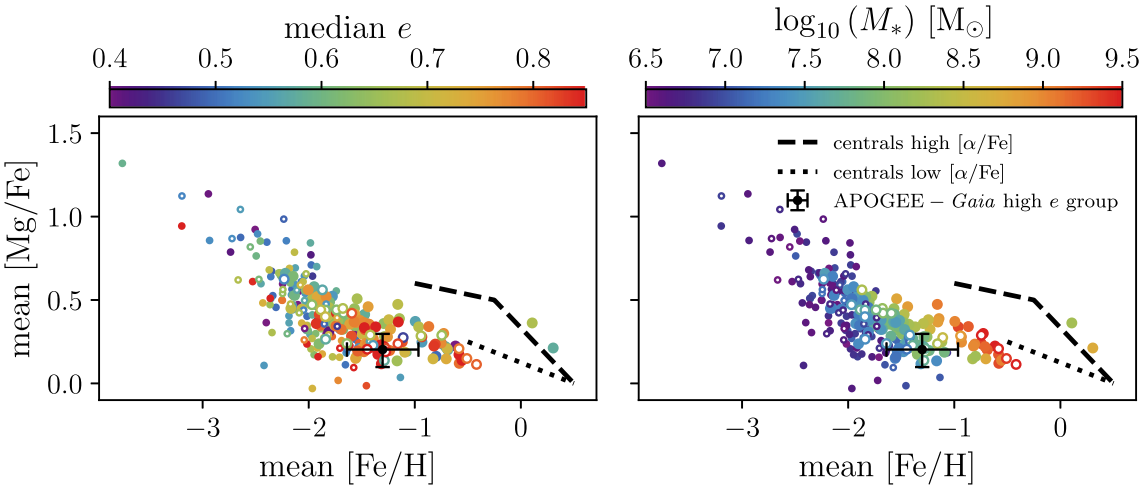

Figure 11. Mean [Mg/Fe] and [Fe/H] of the accreted satellites shown in Figs. 9 and 10. In both panels, point size and filling indicate the number of particles of accreted systems, and morphology of central galaxy, respectively, as in previous figures. On the right-hand panel, colour denotes the masses of accreted satellites. The same data are shown in the left-hand panel, colour-coded by the median eccentricity at $z = 0$ of the debris. The dashed and dotted lines indicate roughly the locus of the disc stars of MW-like central galaxies in the L025N752–Recal simulation (see, e.g. Mackereth et al. 2018). It is clear that the abundances of the stars in the accreted satellites are roughly consistent with those seen in halo stars in the APOGEE data (the median and 1σ scatter of the high-e group is shown by the black error bars). Less massive satellites have a large spread in [Mg/Fe], but all have lower [Fe/H] than the more massive satellites. The increasing [Fe/H] with satellite mass is very clear. The left-hand panel shows that there is a distinct lack of high eccentricity satellite debris at high [Mg/Fe] and low [Fe/H], such that, in general, high-e debris occupies a similar locus to that of the high-e group in the observed data.

out that the scatter in [Mg/Fe] also increases as the satellites become less massive (and [Fe/H] decreases). We emphasize that these low-mass satellites are sampled by the fewest star particles ($N \sim 20$), and so the sampling of the enrichment from supernovae in EAGLE becomes somewhat stochastic.

We test the numerical convergence of these results in Appendix B, showing that the result in Fig. 9 is robust against changes to the simulation resolution, sub-grid model, and volume. Importantly, we show that degrading the resolution of the simulation does not significantly change the result shown in Fig. 10, which indicates that the simulations are adequately resolving the interaction of the central galaxy and the accreted satellite debris. It is important to emphasize that the time at which the initial e is measured is, of course, an important factor which affects the conclusions drawn from this result and its comparison with previous work, which we discuss in Section 5.

5 SUMMARY AND CONCLUSIONS

This paper presents an analysis of chemical compositions and orbital information of halo stellar populations, based on APOGEE and Gaia data. We applied k-means clustering to identify subgroups in chemical composition and kinematics space, considering the abundances of α, odd-Z, and Fe peak elements, and orbital eccentricity. We have shown that $\sim 2/3$ of the accreted halo stars exhibit very high orbital eccentricity and display chemical compositions that are characteristic of those seen in massive dwarf galaxy satellites of the Milky Way today, suggesting that this population is likely the progeny of a single, massive accretion event which occurred early in the history of the Milky Way Galaxy, as suggested by other groups (e.g. Belokurov et al. 2018; Deason et al. 2018; Helmi et al. 2018; Myeong et al. 2018b). The remaining 1/3 of the sample consists of stars with low orbital eccentricities and slightly higher abundance ratios than the high-e population. The latter stars likely result from a mixture of different origins, including the remnants of less massive accretion events, *in situ* star formation, disc heating, and likely some contamination from the high-e population.

We further examine this scenario by studying a numerical simulation from the EAGLE suite. We demonstrate that satellite galaxies accreted into MW mass haloes show clear trends between the time of accretion and the median eccentricity of the accreted stars at $z = 0$. According to the simulation, only satellites accreted at $z \lesssim 1.5$ result in median debris as high as those of the high-e population identified in our data (>0.8). This constraint also means that such satellites are likely to be accreted at relatively high stellar mass ($M_* \gtrsim 10^8\,M_\odot$). A stricter constraint is obtained by comparing the median position of high-e stars with those of EAGLE-accreted systems on the Mg–Fe plane, whereby we infer a mass in the range $10^{8.5}-10^9\,M_\odot$. We also showed that, according to the simulation, the median eccentricity of accreted debris generally does not appear to evolve significantly over time, further suggesting that a high orbital eccentricity is a good indicator of a relatively recent merger. Analysis of the numerical simulation further suggests that a massive accretion event such as that identified in the APOGEE/Gaia data is not very common for a Milky Way-like galaxy, indicating an unusual accretion history in same vein as suggested in our previous work (Mackereth et al. 2018).

The high-e population identified in the combined APOGEE/Gaia DR2 sample appears to be the same population as those discovered by Belokurov et al. (2018), Deason et al. (2018), Helmi et al. (2018), and Haywood et al. (2018). We have shown that the kinematics of the high-e group is consistent with that found in these studies, having very low, slightly negative mean L_z consistent with the Helmi et al. (2018) population, and $r_{\rm ap}$ as high as 40 kpc, with a median of 15.1 kpc, and a suggestive secondary peak at $r_{\rm ap} > 20$, in rough consistency with the Deason et al. (2018) population. Our high-e population has a median eccentricity in good agreement with those of the populations discussed by Belokurov et al. (2018) and the clusters measured by Myeong et al. (2018a).

Using cosmological zoom-in simulations, Belokurov et al. (2018) found that the growing discs of central galaxies act to radialize the orbits or accreted satellite debris as they accrete. This result is in seeming contradiction with our findings that debris eccentricity is

relatively unchanged after satellite accretion (Fig. 10). However, the $e_{z_{\rm merge}}$ we measure in the simulation is that of the debris once the satellite is fully unbound and merged to the central galaxy, and therefore likely already having undergone any radialization from its initial orbit before accretion. Our result simply shows that the orbits do not evolve greatly following accretion onto the galaxy, and is therefore not in direct contradiction with the work of Belokurov et al. (2018) (or Amorisco 2017, who showed a similar effect in an earlier study). However, the finding that the galaxies in EAGLE which do not have a significant disc component at $z = 0$ appear to follow the same trends in $e(z = 0)$-$z_{\rm merge}$ space (Fig. 9) suggests that disc growth may not be the fundamental factor in driving these trends.

Kruijssen et al. (2018) suggest that the age–metallicity distribution of the Galactic globular cluster population can be used to infer the formation and assembly history of the Milky Way. They identify one very massive ($M_* > 10^9\,{\rm M}_\odot$) accretion event (*Kraken*), which is associated with the overabundance of metal-rich GCs in the accreted cluster branch of the age–metallicity relation and has no known debris. They argue that the *Kraken* is associated with GCs that reside within 5 kpc of the Galactic centre. Kruijssen et al. (2018) further point out that the GCs identified by Myeong et al. (2018a) (many of which are located at $r_{GC} > 10$ kpc) closely match the sample they ascribe to the *Canis Major* accretion event. That would suggest that the stellar population reported in this paper would by association also be a part of their proposed *Canis Major* accretion event (as would also be the case of the system identified by, e.g., Nissen & Schuster 2010; Belokurov et al. 2018; Hayes et al. 2018; Helmi et al. 2018).

In addition to these results based on simulations, it is worth pointing out that Leaman, VandenBerg & Mendel (2013) associated metal-poor GCs in the Milky Way with an accretion event with mass between $10^8 < M_* < 10^9\,{\rm M}_\odot$ (consistent with our estimation for the high-e population progenitor), showing that these clusters were on very low angular momentum orbits. In summary, all of these results point towards confirmation that the assembly history of the Milky Way has been very active, and quite atypical compared to other galaxies of similar mass.

As mentioned previously, the notion of a massive accretion event onto the Galaxy is not an entirely new one. The earlier work of Nissen & Schuster (2010) and Schuster et al. (2012) demonstrated that the halo divides into groups in its α-element abundances, with the kinematics of the lower [α/Fe] group resembling that of an accreted population. Even earlier work by Brook et al. (2003) showed that a low angular momentum stellar population (resembling the one described in this paper) is present in the data used by Chiba & Beers (2000), and possibly also in those upon which Eggen et al. (1962) based their scenario for the formation of the halo. Brook et al. (2003) suggested that this population may indeed be the debris of an accreted dwarf galaxy.

Finally, Helmi et al. (2018) suggest that the merger event reported in this paper and by other groups has occurred approximately 10 Gyr ago and was responsible for the dynamical heating responsible for the formation of the thick disc. The results from analysis of EAGLE suggest that if the debris at $e \sim 0.85$ was accreted in a single satellite, then this is likely to have occurred around 8-9 Gyr ago (Fig. 9). The Helmi et al. (2018) merger time is based on the minimum isochronal age of the stars in their sample so it is possibly indicative of the final time of star formation rather than the actual accretion time. On the other hand, EAGLE gives the time at which the satellite became bound to the central halo, so these timescales are potentially consistent. Ages of stars in the Milky Way high-[α/Fe] disc population are generally found to be older than or similar to ∼10 Gyr (e.g. Haywood et al. 2013; Martig et al. 2016b; Mackereth et al. 2017), so this does suggest that this population was in place before the merger occurred. It is worth mentioning, however, that analysis of the origin of α-enhanced populations in EAGLE suggest that these stars must form in an early collapse in a period of rapid gas accretion, in order to foster the high-density ISM necessary to generate a short-enough gas consumption time to consume the high-[α/Fe] gas into stars before it is polluted by SN Ia (Mackereth et al. 2018), rather than forming in an initially thin disc that was later heated. This scenario is consistent with the thick disc being the result of the geometric combination of the thick, centrally concentrated high-[α/Fe] disc and the extended, flared low [α/Fe] disc (e.g. Minchev et al. 2015; Martig et al. 2016a; Mackereth et al. 2017). Further work on this newly found halo component, and the high-[α/Fe] disc, will surely shed more light on this discussion.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewer for a careful and helpful report. We thank Marie Martig for insightful discussions during the preparation of this paper, and Robert Crain for crucial comments on a late draft of the manuscript. We also thank Vasily Belokurov, Lachlan Lancaster, and Ryan Leaman for useful comments on the originally submitted mansucript. JTM acknowledges an STFC doctoral studentship. JP gratefully acknowledges funding from a European Research Council consolidator grant (ERC-CoG-646928-Multi-Pop). CRH acknowledges the NSF Graduate Research Fellowship through grant DGE-1315231. JB received support from the Natural Sciences and Engineering Research Council of Canada (NSERC; funding reference number RGPIN-2015-05235). JB also received partial support from an Alfred P. Sloan Fellowship. CAP is thankful to the Spanish Government for funding for his research through program AYA2017-86389-P. PBT acknowledges Fondecyt-Conicyt Regular 1150334. This project was developed in part at the 2018 Gaia Sprint, hosted by the Center for Computational Astrophysics of the Flatiron Institute in New York City. This research made use of the cross-match service provided by CDS, Strasbourg. Analyses and plots presented in this article used `iPython`, and packages in the `SciPy` ecosystem (Jones et al. 2001; Hunter 2007; Perez & Granger 2007; van der Walt, Colbert & Varoquaux 2011). The study made use of high performance computing facilities at Liverpool John Moores University, partly funded by the Royal Society and LJMU's Faculty of Engineering and Technology..We acknowledge the Virgo Consortium for making their simulation data available. The EAGLE simulations were performed using the DiRAC-2 facility at Durham, managed by the ICC, and the PRACE facility Curie based in France at TGCC, CEA, Bruyères- le-Chatel.

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-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.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration

including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, 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, Observatário 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.

REFERENCES

Abolfathi B. et al., 2018, ApJS, 235, 42
Ahn C. P. et al., 2012, ApJS, 203, 21
Allende Prieto C. et al., 2014, A&A, 568, A7
Amorisco N. C., 2017, MNRAS, 464, 2882
Andrews B. H., Weinberg D. H., Schönrich R., Johnson J. A., 2017, ApJ, 835, 224
Arnett D., 1996, Supernovae and Nucleosynthesis: An Investigation of the History of Matter from the Big Bang to the Present. Princeton Press, Princeton, NJ, USA
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Binney J., 2012, MNRAS, 426, 1324
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton University Press, Princeton, NJ, USA
Bird J. C., Kazantzidis S., Weinberg D. H., Guedes J., Callegari S., Mayer L., Madau P., 2013, ApJ, 773, 43
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Blanton M. R. et al., 2017, AJ, 154, 28
Bovy J., 2015, ApJS, 216, 29
Bovy J., Rix H.-W., Hogg D. W., 2012a, ApJ, 751, 131
Bovy J., Rix H.-W., Liu C., Hogg D. W., Beers T. C., Lee Y. S., 2012b, ApJ, 753, 148
Bovy J., Rix H.-W., Hogg D. W., Beers T. C., Lee Y. S., Zhang L., 2012c, ApJ, 755, 115
Bovy J. et al., 2014, ApJ, 790, 127
Bovy J., Rix H.-W., Schlafly E. F., Nidever D. L., Holtzman J. A., Shetrone M., Beers T. C., 2016, ApJ, 823, 30
Brook C. B., Kawata D., Gibson B. K., Flynn C., 2003, ApJ, 585, L125
Brook C. B., Kawata D., Gibson B. K., Freeman K. C., 2004, ApJ, 612, 894
Carollo D. et al., 2007, Nature, 450, 1020
Casagrande L. et al., 2016, MNRAS, 455, 987
Chiba M., Beers T. C., 2000, AJ, 119, 2843
Correa C. A., Schaye J., Clauwens B., Bower R. G., Crain R. A., Schaller M., Theuns T., Thob A. C. R., 2017, MNRAS, 472, L45
Crain R. A. et al., 2015, MNRAS, 450, 1937
Cullen L., Dehnen W., 2010, MNRAS, 408, 669
Deason A. J., Belokurov V., Evans N. W., Johnston K. V., 2013, ApJ, 763, 113
Deason A. J., Belokurov V., Koposov S. E., Lancaster L., 2018, preprint (arXiv:1805.10288)
Dobbie P. D., Cole A. A., Subramaniam A., Keller S., 2014, MNRAS, 442, 1680
Durier F., Dalla Vecchia C., 2012, MNRAS, 419, 465
Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748
Eisenstein D. J. et al., 2011, AJ, 142, 72
Fernández-Alvar E. et al., 2015, A&A, 577, A81
Fernández-Alvar E. et al., 2018a, preprint (arXiv:1807.07269)
Fernández-Alvar E. et al., 2018b, ApJ, 852, 50
Ferrero I. et al., 2017, MNRAS, 464, 4736
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Furlong M. et al., 2017, MNRAS, 465, 722
Gaia Collaboration et al., 2018, A&A, 616, A1
García Pérez A. E. et al., 2016, AJ, 151, 144
Goodman J., Weare J., 2010, Commun. Appl. Math. and Comp. Sci., 65
Gunn J. E. et al., 2006, AJ, 131, 2332
Harris W. E., 1996, AJ, 112, 1487
Hasselquist S. et al., 2017, ApJ, 845, 162
Hawkins K., Jofré P., Masseron T., Gilmore G., 2015, MNRAS, 453, 758
Hayden M. R. et al., 2015, ApJ, 808, 132
Hayes C. R. et al., 2018, ApJ, 852, 49
Haywood M., Di Matteo P., Lehnert M. D., Katz D., Gómez A., 2013, A&A, 560, A109
Haywood M., Di Matteo P., 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, Natur, 563, 85
Hogg D. W., Blanton M. R., Roweis S. T., Johnston K. V., 2005, ApJ, 629, 268
Hogg D. W., Bovy J., Lang D., 2010, preprint (arXiv:1008.4686)
Holtzman J. et al., 2018, AJ, 156, 125
Holtzman J. A. et al., 2015, AJ, 150, 148
Hopkins P. F., 2013, MNRAS, 428, 2840
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jones E., Oliphant T., Peterson P. et al., 2001, SciPy: Open source scientific tools for Python., Available at: http://www.scipy.org/
Jönsson H. et al., 2018, AJ, 156, 126
Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, ApJ, 779, 102
Koppelman H., Helmi A., Veljanoski J., 2018, ApJ, 860, L11
Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2018, MNRAS, 1537
Lancaster L., Koposov S. E., Belokurov V., Evans N. W., Deason A. J., 2018, preprint (arXiv:1807.04290)
Leaman R., VandenBerg D. A., Mendel J. T., 2013, MNRAS, 436, 122
Lindegren L. et al., 2016, A&A, 595, A4
Mackereth J. T., Bovy J., 2018, PASP, 130, 114501
Mackereth J. T. et al., 2017, MNRAS, 471, 3057
Mackereth J. T., Crain R. A., Schiavon R. P., Schaye J., Theuns T., Schaller M., 2018, MNRAS, 477, 5072
Majewski S. R. et al., 2017, AJ, 154, 94
Martig M., Minchev I., Ness M., Fouesneau M., Rix H.-W., 2016a, ApJ, 831, 139
Martig M. et al., 2016b, MNRAS, 456, 3655
Masseron T., Gilmore G., 2015, MNRAS, 453, 1855
McAlpine S. et al., 2016, Astron. Comput., 15, 72
Minchev I., Martig M., Streich D., Scannapieco C., de Jong R. S., Steinmetz M., 2015, ApJ, 804, L9
Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018a, MNRAS, 475, 1537
Myeong G. C., Evans N. W., Belokurov V., Amorisco N. C., Koposov S. E., 2018b, MNRAS, 475, 1537
Navarro J. F. et al., 2017, MNRAS, 476, 3648
Nelder J. A., Mead R., 1965, Comput. J., 7, 308
Nidever D. L. et al., 2014, ApJ, 796, 38
Nidever D. L. et al., 2015, AJ, 150, 173
Nissen P. E., Schuster W. J., 2010, A&A, 511, L10
Pedregosa F. et al., 2011, Journal of Machine Learning Research, 12, 2825
Perez F., Granger B. E., 2007, Computing in Science Engineering, 9, 21
Price D. J., 2010, MNRAS, 401, 1475

Queiroz A. B. A. et al., 2018, MNRAS, 476, 2556
Ross T. L., Holtzman J., Saha A., Anthony-Twarog B. J., 2015, AJ, 149, 198
Rubele S. et al., 2018, MNRAS, 478, 5017
Santiago B. X. et al., 2016, A&A, 585, A42
Schaye J. et al., 2015, MNRAS, 446, 521
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Schuster W. J., Moreno E., Nissen P. E., Pichardo B., 2012, A&A, 538, A21
Searle L., Zinn R., 1978, ApJ, 225, 357
Shetrone M. et al., 2015, ApJS, 221, 24
Springel V., 2005, MNRAS, 364, 1105
Stanimirović S., Staveley-Smith L., Jones P. A., 2004, ApJ, 604, 176
Tolstoy E., Hill V., Tosi M., 2009, ARA&A, 47, 371
Trayford J. W. et al., 2015, MNRAS, 452, 2879
Trayford J. W., Theuns T., Bower R. G., Crain R. A., Lagos C. d. P., Schaller M., Schaye J., 2016, MNRAS, 460, 3925
Trayford J. W. et al., 2017, MNRAS, 470, 771
van der Marel R. P., Kallivayalil N., Besla G., 2009, in Van Loon J. T., Oliveira J. M., eds, IAU Symposium Vol. 256, The Magellanic System: Stars, Gas, and Galaxies. p. 81preprint (arXiv:0809.4268)
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Weinberg D. H., Andrews B. H., Freudenburg J., 2017, ApJ, 837, 183
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Wilson J. C. et al., 2010, in Ground-based and Airborne Instrumentation for Astronomy IIIs, 77351C
York D. G. et al., 2000, AJ, 120, 1579
Zamora O. et al., 2015, AJ, 149, 181
Zasowski G. et al., 2013, AJ, 146, 81
Zasowski G. et al., 2017, AJ, 154, 198

APPENDIX A: MODELLING [MG/FE] AS A FUNCTION OF [FE/H]

In order to test whether a change in slope is found in the relationship between [Mg/Fe] and [Fe/H] in the identified accreted halo groups, we use a Bayesian inference to fit a piecewise-linear model to the data. The form of the piecewise-linear function is given in equation (3). We follow the general procedure outlined in section 7 of Hogg, Bovy & Lang (2010) for fitting models to data with two-dimensional uncertainties. For completeness, we re-iterate here the mathematics. The best-fitting model is found by maximizing the likelihood function for the parameters $O = [[\mathrm{Fe/H}]_0, [\mathrm{Mg/Fe}]_0, \theta_1, \theta_2]$ given the data, which we assume here to be of the form

$$\ln \mathcal{L}(O|[\mathrm{Fe/H}], [\mathrm{Mg/Fe}]) = K - \sum_{i=1}^{N} \left(\frac{\Delta_i^2}{2\Sigma_i^2} + \ln |\Sigma_i^2| \right), \quad \text{(A1)}$$

where Δ_i^2 defines the distance between the data-point i and the model, and Σ_i^2 is the variance orthogonal to the model, determined by the covariance matric of the data points. K is a normalization constant, which is not necessary to consider in the optimization. We assume uninformative flat priors on $\theta_{[1,2]}$, and allow $[\mathrm{Fe/H}]_0$ and $[\mathrm{Mg/Fe}]_0$ to be free. In this case, Δ_i^2 is defined by

$$\Delta_i^2 = \hat{\mathbf{v}}^T \mathbf{Z}_i - b \cos\theta_{[1,2]}, \quad \text{(A2)}$$

where $\mathbf{Z}_i$ is the column vector made by $([\mathrm{Mg/Fe}],[\mathrm{Fe/H}])_i$, and $\hat{\mathbf{v}}$ is the unit vector orthogonal to the model:

$$\hat{\mathbf{v}} = \frac{1}{\sqrt{1 + m_{[1,2]}^2}} \begin{bmatrix} -m_{[1,2]} \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin\theta_{[1,2]} \\ \cos\theta_{[1,2]} \end{bmatrix}. \quad \text{(A3)}$$

where $\theta_{[1,2]}$, here and in equation (A2), is the angle between the linear model at that [Fe/H] and the x-axis, which is equal to $\arctan m_{[1,2]}$. Σ_i^2 is then simply defined as the projection of the data-point's covariance matrix $\mathbf{S}_i$ orthogonal to the model at that [Fe/H]

$$\Sigma_i^2 = \hat{\mathbf{v}}^T \mathbf{S}_i \hat{\mathbf{v}}. \quad \text{(A4)}$$

In this case, we assume that the uncertainties on [Mg/Fe] and [Fe/H] are uncorrelated, such that

$$\mathbf{S}_i = \begin{bmatrix} \delta[\mathrm{Fe/H}]_i & 0 \\ 0 & \delta[\mathrm{Mg/Fe}]_i \end{bmatrix}, \quad \text{(A5)}$$

where we use the catalogue values for the uncertainties on [Fe/H] and [Mg/Fe]. We minimize the negative log-likelihood using a downhill simplex algorithm (Nelder & Mead 1965), and use this optimal solution to initiate an Markov Chain Monte Carlo (MCMC) sampling of the posterior PDF of the parameters O using an affine-invariant ensemble MCMC sampler (Goodman & Weare 2010) as implemented in the python package `emcee` (Foreman-Mackey et al. 2013). We report the median and standard deviation of this posterior PDF as our best-fit parameters.

APPENDIX B: NUMERICAL CONVERGENCE TESTS

We examine in this appendix the effect on the results presented in Figs. 9 and 10 of varying the simulation resolution, box-size, and subgrid model parameters, in order to demonstrate that these results are well-converged numerically. We perform the equivalent analysis to that presented for L025N752–Recal (in the main body of the paper) on lower resolution volumes of the EAGLE simulations.

Schaye et al. (2015) define 'weakly' converged predictions as those which are unaffected by variation in the simulation resolution after re-calibrating the sub-grid physics. We test this by examining the equivalently sized, $L = 25$ cMpc, lower resolution run (at a factor of 10 lower in mass than L025N752–Recal) which adopts the 'Reference' sub-grid model, which we refer to as L025N376-Ref. We show the resulting equivalent of 9 for L025N356-Ref in Fig. B1. It is clear from this figure that the general prediction of 9 holds, in spite of the fact that the lower resolution simulation clearly does not resolve galaxies at masses as low as L025N752-Recal. The upper envelope of the distribution, which defines the maximum $z = 0$ eccentricity of a satellite merged at any given z, is still clear out to $z \sim 5$. Given that the sub-grid feedback model is adjusted between the 'Reference' and 'Recalibrated' models to maintain the predictions of the former, from this we gather that the result is 'weakly' converged.

In order to check that the prediction of EAGLE for the $z = 0$ eccentricity as a function of merger time is 'strongly' converged, we must also ascertain that the trends in Fig. B1 are conserved when the resolution is increased, but the sub-grid model held fixed. We test this by performing an equivalent analysis on the L025N752-Ref model, which has a resolution equal to the L025N752-Recal model, but assumes the same sub-grid physics as L025N376-Ref. The resulting $z = 0$ e against merger time is shown in Fig. B2. The trends which are seen in both the low-resolution 'Reference' model (Fig. B1) and high-resolution 'Recalibrated' model (Fig. 9 in the main text) are clearly conserved here also,[7] demonstrating the 'strong' numerical convergence of these results.

We also perform the analysis on a larger volume, $L = 50$ cMpc, EAGLE simulation, which adopts the 'Reference' sub-grid model at the same resolution as the smaller volume: L050N752-Ref. In

[7]There is a very slight change in the exact position of the maximum of the distribution as a function of z_{merge}, however, we contend that this is due to the stochastic sampling of the true underlying distribution.

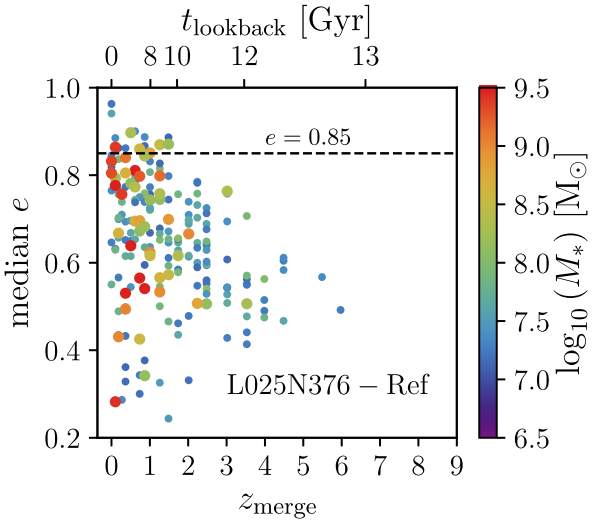

Figure B1. The merger time $z_{\rm merge}$ of satellites accreted onto Milky Way mass haloes in the L025N356-Ref simulation against the median eccentricity e of their stellar debris at $z = 0$, produced via an equivalent analysis to that which produced Fig. 9. The trends in Figs B1 and 9 are clearly conserved under an increase in resolution with no recalibration of the sub-grid feedback model, demonstrating the strong numerical convergence of these results.

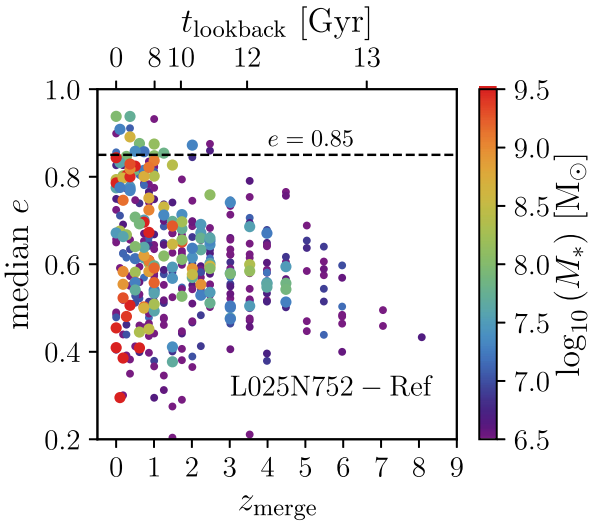

Figure B2. The merger time $z_{\rm merge}$ of satellites accreted onto Milky Way mass haloes in the L025N752-Ref simulation against the median eccentricity e of their stellar debris at $z = 0$, produced via an equivalent analysis to that which produced Fig. 9. The upper envelope of $z = 0$ eccentricity for a given merger time is reproduced in this lower resolution simulation, albeit with a different calibration of the subgrid model, demonstrating the 'weak' convergence of this result.

this much larger volume, we track the accretion of 1154 satellites onto 126 central haloes, again with halo masses which are roughly equivalent to that of the Milky Way. The resulting distribution of satellite debris in median $e(z = 0)$ against merger time is shown in Fig. B3. Again, the trend between the maximum $z = 0$ eccentricity and $z_{\rm merge}$ is seen, now more clearly, as a result of the better statistics offered by the increased sample size. This demonstrates that the results presented from EAGLE are strongly robust to variations in the simulation resolution, box-size and sub-grid physics.

Finally, we show in Fig. B4 that the findings of Fig. 10 are robust to degradation in the simulation resolution. It is possible that the dynamical interaction between the central galaxy and the accreting

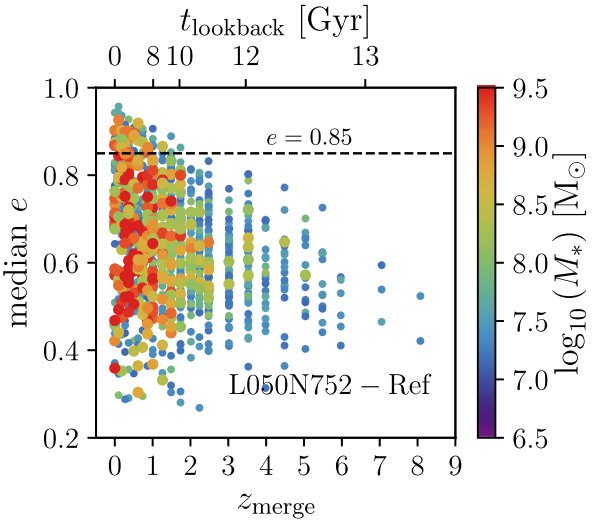

Figure B3. The merger time $z_{\rm merge}$ against median eccentricity e of the stellar debris at $z = 0$, of the 1154 accreted satellites (with > 20 particles) of 126 central galaxies from the L050N752-Ref simulation, produced via an equivalent analysis to that which produced Fig. 9. The trends seen in Figs 9, B1 and B2 are seen again, even though a much larger sample of accretion events is analysed.

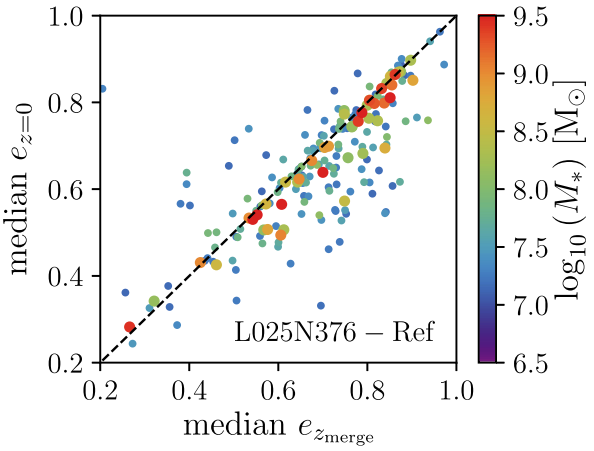

Figure B4. The change in eccentricity between the snapshot immediately after the satellites become unbound and $z = 0$ in the L025N376-Ref simulation (equivalent to Fig. 10 in the main text). Degrading the simulation resolution has little effect on this result, which suggests that the lack of any significant radialization or circularization of orbits after satellite infall is not due to poorly resolved dynamics, which would act to reduce any scatter upon degradation of the simulation resolution.

satellite may be poorly modelled by the simulations, which are of a relatively low resolution (as opposed to idealised simulations such as those of e.g. Amorisco 2017), reducing the action of the central galaxy on the accreting satellites. By this reasoning, a degradation in the simulation resolution should decrease any changes to the orbital properties of the satellites, as these effects would be worsened. In Fig. 10, showing the higher resolution simulation, we find the scatter around the dashed unity line to be 0.13. In the lower resolution, we find the scatter to be comparable, if not slightly increased, at 0.15. As a result, we contend that these interactions are likely well modelled by the simulation, even at this relatively low resolution.

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