The contribution of N-rich stars to the Galactic stellar halo using APOGEE red giants

Danny Horta [1]★ J. Ted Mackereth [2,3] Ricardo P. Schiavon,[1] Sten Hasselquist,[4,5] Jo Bovy [2,3]
Carlos Allende Prieto,[6,7] Timothy C. Beers,[8] Katia Cunha,[9,10] D. A. García-Hernández,[6,7]
Shobhit S. Kisku,[1] Richard R. Lane,[11,12] Steven R. Majewski,[13] Andrew C. Mason,[1] David M. Nataf [14]
Alexandre Roman-Lopes[15] and Mathias Schultheis[16]

[1]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*
[2]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada*
[3]*Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada*
[4]*New Mexico State University, Las Cruces, NM 88003, USA*
[5]*Department of Physics & Astronomy, University of Utah, Salt Lake City, UT 84112, USA*
[6]*Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain*
[7]*Departamento de Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain*
[8]*Department of Physics and JINA Center for the Evolution of the Elements, University of Notre Dame, Notre Dame, IN 46556, USA*
[9]*University of Arizona, Tucson, AZ 85719, USA*
[10]*Observatório Nacional, São Cristóvão, Rio de Janeiro, Brazil*
[11]*Instituto de Astronomía y Ciencias Planetarias, Universidad de Atacama, Copayapu 485, Copiapó, Chile*
[12]*Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuna Mackenna 4860, 782-0436 Macul, Santiago, Chile*
[13]*Department of Astronomy, University of Virginia, Charlottesville, VA 22904-4325, USA*
[14]*Center for Astrophysical Sciences and Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA*
[15]*Departamento de Astronomía, Universidad de La Serena, Cisternas 1200, La Serena, Chile*
[16]*Laboratoire Lagrange, CNRS, Observatoire de la Côte d'Azur, Université Côte d'Azur, Blvd de l'Observatoire, F-06304 Nice, France*

ABSTRACT

The contribution of dissolved globular clusters (GCs) to the stellar content of the Galactic halo is a key constraint on models for GC formation and destruction, and the mass assembly history of the Milky Way. Earlier results from APOGEE pointed to a large contribution of destroyed GCs to the stellar content of the inner halo, by as much as 25 per cent, which is an order of magnitude larger than previous estimates for more distant regions of the halo. We set out to measure the ratio between nitrogen-rich (N-rich) and normal halo field stars, as a function of distance, by performing density modelling of halo field populations in APOGEE DR16. Our results show that at 1.5 kpc from the Galactic Centre, N-rich stars contribute a much higher $16.8^{+10.0}_{-7.0}$ per cent fraction to the total stellar halo mass budget than the $2.7^{+1.0}_{-0.8}$ per cent ratio contributed at 10 kpc. Under the assumption that N-rich stars are former GC members that now reside in the stellar halo field, and assuming the ratio between first and second population GC stars being 1:2, we estimate a total contribution from disrupted GC stars of the order of $27.5^{+15.4}_{-11.5}$ per cent at r = 1.5 kpc and $4.2^{+1.5}_{-1.3}$ per cent at $r = 10$ kpc. Furthermore, since our methodology requires fitting a density model to the stellar halo, we integrate such density within a spherical shell from 1.5 to 15 kpc in radius, and find a total stellar mass arising from dissolved and/or evaporated GCs of $M_{GC,total} = 9.6^{+4.0}_{-2.6} \times 10^7$ M⊙.

Key words: Galaxy: evolution – Galaxy: formation – globular clusters: general – Galaxy: halo.

1 INTRODUCTION

It is common knowledge that the globular cluster (GC) system of the Milky Way holds crucial insights into the formation and assembly history of the Galaxy. A milestone in this field was the work by Searle & Zinn (1978), who used the Galactic GC population to infer that GCs found in the outer Galactic halo region formed over a longer period than those found in the inner halo region,

leading to the conclusion that the former population originated from accreted satellite systems. This concept has since been refined with improved measurements of ages and chemical abundances, leading to the discovery of a bifurcation in the age–metallicity distribution of Galactic GCs (Marín-Franch et al. 2009), where the young and metal-poor branch traces the population of GCs that are thought to result from satellite galaxy accretion (e.g. Forbes & Bridges 2010; Leaman, VandenBerg & Mendel 2013). Similarly, this division is also reflected in the α and Fe compositions of GCs with metallicities between $-1.5 <$ [Fe/H] < -1, whereby accreted GCs typically present lower [α/Fe] for a given [Fe/H] in comparison to their

★ E-mail: D.HortaDarrington@2018.ljmu.ac.uk

in situ counterparts (e.g. Horta et al. 2020a). The results obtained from these observational studies fit well in the present cosmological framework, the ΛCDM model (e.g. Schaeffer & Silk 1985; Frenk & White 2012). In this cosmological paradigm, larger galaxies engulf smaller satellite galaxies and grow through a process of hierarchical mass assembly. As a result, a substantial fraction of the GCs from those accreted galaxies may survive the merger, depending on the details of the accretion, as suggested by various observational results (e.g. Brodie & Strader 2006; Koppelman et al. 2019; Kruijssen et al. 2019b; Massari, Koppelman & Helmi 2019; Myeong et al. 2019; Forbes 2020).

To elucidate the role GCs play in the formation and mass assembly of their host galaxies, it is vital to first understand how GCs form and evolve in a cosmological context. The leading GC formation scenarios propose a framework in which GCs formed in the turbulent discs of their host galaxies at $z \sim 2$–3, where, due to tidal shocks (Gnedin 2001) and the so-called cruel cradle effect (Kruijssen et al. 2012), GCs formed *in situ* were largely destroyed (Elmegreen 2010; Kruijssen et al. 2011; Kruijssen 2014, 2015). As galaxies evolved, mergers redistributed GC systems of accreted satellites on to the host galaxies, typically depositing them in the outer regions of the stellar halo component, where mass-loss via evaporation takes place in a longer time-scale.

The question of how much stellar mass in the halo arises from GC dissolution and/or evaporation emerges as a natural implication from the above theoretical framework. In an attempt to answer such a question, over the last decade there has been substantial work focusing on the identification of GC stars that have been stripped from their parent GC, and now reside in the halo field of the Galaxy (e.g. Martell & Grebel 2010; Martell et al. 2016; Schiavon et al. 2017; Koch, Grebel & Martell 2019; Hanke et al. 2020, Kisku et al., in preparation). In these studies, stars arising from GC dissolution and/or evaporation were identified either by selecting field stars that present high nitrogen and low carbon abundances, or by identifying stars with strong CN bands, a feature that implies a higher N content and one that is typically found in the so-called second population (SP) GC stars (for a full review, see Bastian & Lardo 2018). Such studies have focused on both the outer regions of the stellar halo (Martell & Grebel 2010; Martell et al. 2016; Koch et al. 2019), and within the inner few kiloparsecs from the Galactic Centre (Schiavon et al. 2017). Specifically, Martell et al. (2016) focused on metal poor ([Fe/H] $\leq$ −1.3) sample from the 12th APOGEE data release and identified halo field stars that present high [N/Fe] and [Al/Fe] values, thus selecting former SP GC escapee candidates. Using the number ratio of nitrogen-rich (N-rich) stars to halo field stars, Martell et al. (2016) obtained an estimate for the contribution from dissolved SP GC stars to the outer regions of the stellar halo to be of the order of 2–3 per cent. A similar estimate was also obtained in a more recent study using the SEGUE data (Yanny et al. 2009) by Koch et al. (2019), who found a value of~ 2.6 per cent. Separately, Schiavon et al. (2017) analysed APOGEE data for stellar populations within ~3 kpc of the Galactic Centre and found a much larger fraction of N-rich stars than their outer halo counterpart. In their study, Schiavon et al. (2017) identified field stars that presented high [N/Fe] and low [C/Fe] compared to the main population at a given metallicity, and determined a minimum contribution to the stellar halo from dissolved and/or evaporated SP GC stars of approximately 13–17 per cent. This results in a much higher contribution rate than in the outer halo, by a factor of ~4–5.

Furthermore, recent theoretical studies making use of the suite of E-MOSAICS hydrodynamical cosmological simulations (Pfeffer et al. 2018; Kruijssen et al. 2019a) have assessed the contribution

of stars dissolved from GCs to the stellar halo field for ~25 Milky Way analogues. The results from these works predict a smaller contribution from SP GC stars than observational estimates, predicting a total contribution of the order of ~ 1 per cent (Reina-Campos et al. 2019) for the outer halo region and ~0.2–9 per cent for the inner Galaxy (Hughes et al. 2020). Although the predicted fractions at face value do not match that of the observed fractions, an inner-to-outer halo ratio discrepancy is still predicted.

In connection with the mass contribution of dissolved GC stars to the Galactic stellar halo, the question of the spatial distribution of GC escapees is also an important one, and similarly, its stellar density as a function of Galactocentric distance. Answering these questions is non-trivial due to the difficulty in correcting the observed number of a tracer population to the total underlying sample. Along the same lines, estimates of the stellar mass are rigidly connected to the derived mass normalization per tracer star, which generally relies on intricate calibration and stellar models. Moreover, GCs are typically associated with the stellar halo component of the Galaxy, and with no previous knowledge of the density profile of GC escapees, it is important to first understand how the stellar halo is spatially distributed. This information can be used as a guide to understand the density profile of GC-escapee stars, and will serve later for assessing the contribution of N-rich stars to the stellar halo field.

Recently, Iorio et al. (2018) modelled the density of the stellar halo using a cross-matched sample of *Gaia* 2MASS RR Lyrae stars. In that work, several density profiles ranging in parameter complexity were tested, and it was found that the stellar halo was best fit by a triaxial ellipsoid, rotated with respect to the Galactic reference frame (in which the *X*-axis connects the Sun to the Galactic Centre) by $\sim$ 70°; such rotation was thought to be introduced by a single massive merger event. Using red giant branch (RGB) stars from the APOGEE 14th data release, Mackereth & Bovy (2020) also found this to be the case. The same study estimated a total halo stellar mass of the order of $\sim 1.3 \times 10^9$ M⊙ when considering the halo as an assembly of many individual components in the [Mg/Fe]–[Fe/H] plane. This result is in agreement with the recent work of Deason, Belokurov & Sanders (2019), based on RGB number counts and extrapolation of the Einasto (1965) profile, which also suggests a higher mass for the stellar halo of $\sim 1.4 \times 10^9$ M⊙. These works utilized existing density models to understand the shape and density profile of the stellar halo component of the Galaxy. Using such density modelling procedures, they were able to estimate the total mass of the stellar halo within a given Galactocentric radius. Such methods can be applied in theory to any tracer stellar population, enabling the estimation of mass as a function of Galactocentric radius while accounting for selection effects induced by observational data.

In this paper, we build on the previous work by Mackereth & Bovy (2020), and present a density profile, mass estimate, and a percentage contribution to the stellar halo from likely SP GC (i.e. N-rich) stars arising from GC dissolution and/or evaporation as a function of Galactocentric distance. This is accomplished via density modelling of APOGEE red giant stars, selected on the basis of their chemical compositions. Our methodology allows for the assessment of the stellar mass contained in the complete halo and the N-rich samples, as well as their corresponding mass as a function of Galactocentric distance. As explained in Mackereth & Bovy (2020), the APOGEE red giant star counts are corrected for their respective normalization, via the reconstruction of the APOGEE DR16 selection function and using stellar evolution models, which enables the estimation of the stellar density profiles to good accuracy.

The paper is organized as follows: Section 2 presents our selection of likely halo APOGEE red giant stars, and our criteria to determine N-rich stars residing in the stellar halo field. In Section 3, we briefly describe the density modelling procedure, based on the work by Bovy et al. (2016) and Mackereth et al. (2017), including allowances for the APOGEE-2 selection function. In Section 4, we present the main results of the paper. Section 5 includes the discussion of our results in the context of previous studies, for which we then summarize our results in Section 6, and provide our conclusions.

2 DATA AND SAMPLE

We use data from the 16th data release of Sloan Digital Sky Survey-IV (SDSS-IV; Ahumada et al. 2019), which contains refined elemental abundances (Jönsson et al. 2020) for stars observed both by the Apache Point Observatory Galactic Evolution Environment (APOGEE; Majewski et al. 2017) and its successor, APOGEE-2, one of four SDSS-IV (Blanton et al. 2017) experiments. With twin spectrographs (Wilson et al. 2019) mounted to the 2.5-m Sloan Telescope (Gunn et al. 2006) at Apache Point Observatory in New Mexico and the 2.5-m du Pont Telescope at Las Campanas Observatory, APOGEE-2, includes high-resolution ($R \sim 22\,500$), high signal-to-noise ratio (SNR > 100 pixel^{-1}), near-infrared (1.5–1.7 µm) spectra of over 450 000 stars sampling all parts of the Milky Way. These spectra are used to derive accurate stellar atmospheric parameters, radial velocities, and the abundances for up to 25 atomic elements. Targets were selected from the 2MASS point source catalogue with a dereddened ($J - Ks$) $\geq$ 0.3 colour cut in up to three apparent H-band magnitude bins. Reddening corrections were determined using the Rayleigh–Jeans colour excess method (RJCE; Majewski, Zasowski & Nidever 2011). Corrections are obtained by applying the method to the combined 2MASS (Skrutskie et al. 2006) and Spitzer–IRAC surveys GLIMPSE-I, -II (Churchwell et al. 2009), and -3D when available, or 2MASS combined with *WISE* photometry (Wright et al. 2010). A more in-depth description on the APOGEE survey, data reduction pipeline, and the target selection can be found in Majewski et al. (2017), Holtzman et al. (2015), Nidever et al. (2015), Jönsson et al. (2020), and Zasowski et al. (2013, 2017). All the APOGEE data products used in this paper are those output by the standard data reduction and analysis pipeline. The data are first reduced with a custom pipeline (Nidever et al. 2015; Jönsson et al. 2020). The data are then thoroughly checked, before being fed into the APOGEE Stellar Parameters and Chemical Abundances Pipeline (ASPCAP; García Pérez et al. 2016; Jönsson et al. 2020). ASPCAP makes use of a specifically computed spectral library (Zamora et al. 2015; Holtzman et al. 2018; Jönsson et al. 2020), calculated using a customized H-band line list (Shetrone et al. 2015; Smith et el., in preparation), from which then the outputs are then analysed, calibrated, and tabulated (Holtzman et al. 2018).

We make use of the distances for the APOGEE DR16 catalogue generated by Leung & Bovy (2019a), using the ASTRONN PYTHON package (for a full description, see Leung & Bovy 2019b). These distances are determined using a previously trained ASTRONN neural network, which predicts stellar luminosity from spectra using a training set comprised of stars with both APOGEE spectra and *Gaia* DR2 parallax measurements (Gaia Collaboration 2018). The model is able to predict simultaneously distances and account for the parallax offset present in *Gaia* DR2, producing high precision, accurate distance estimates for APOGEE stars, which match well with external catalogues and standard candles.

The sample adopted in this study is restricted to stars contained in the statistical sample[1] of APOGEE DR16 located within fields with extinction information and that satisfy the following criteria: $0.5 < \log g < 3.5$, 3500 K $< T_{\rm eff} <$ 4750 K, $-2.5 <$ [Fe/H] < -1, $d_{\rm err}/d < 0.2$ (where d and $d_{\rm err}$ are the distance and distance error, respectively). Here, the $T_{\rm eff}$ and $\log g$ cuts were made in order to target RGB stars whose C, N, and Al abundances are unaffected by dredge-up processes. We note here that we use the calibrated ASPCAP abundances, rather than those from other abundance pipelines (e.g. BACCHUS, ASTRONN).

We also remove stars that are in fields in close proximity to the Galactic Centre by excluding fields within 10° in/from the Galactic Centre and $|b| < 10°$ for which dust extinction is most difficult to correct, as well as those which contained GCs used for APOGEE calibration. Furthermore, since our aim is to study the Galaxy's stellar halo field population, we remove 1781 stars in the APOGEE DR16 statistical sample that are known to reside in GCs, adopting the GC memberships established by Horta et al. (2020a). Lastly, since we are focused on selecting dissolved and/or evaporated SP GC stars in [N/Fe] space, we applied further cuts to ensure we were only considering stars where [N/Fe] is robustly determined, and removed any stars that had a N_FE_FLAG warning set by ASPCAP. We also ensured stars in our sample had reliable Fe, C, and Al abundances by checking the C_FE_FLAG and AL_FE_FLAG warnings were set to zero, and by removing any stars with [Fe/H], [C/Fe], and/or [Al/Fe] set to -9999. Our final working halo sample is comprised of 1455 stars.

Next, we define SP GC stars by identifying N-rich stars present in the stellar halo field population of the Milky Way. To identify N-rich stars, we make use of the publicly available code XDGMM[2] (Holoien, Marshall & Wechsler 2017), and fit a two-component Gaussian mixture model (GMM) to our halo sample [N/Fe]–[Fe/H] distribution, with the expectation that N-rich stars will stand out in this plane. This software uses the extreme deconvolution (XD) algorithms (Bovy, Hogg & Roweis 2011) to identify components in an n-dimensional space, and allows us to determine the N-rich star field sample statistically by accounting for the uncertainties and correlations in the measurement errors. We then refine our N-rich star selection by only considering stars with carbon abundances below [C/Fe] $< +0.15$, to ensure that these stars present the low [C/Fe] abundances typical of SP GC stars, and remove four stars from the sample. Our final sample of N-rich stars is comprised of 46 stars. Fig. 1 shows the distribution of the selected N-rich stars and the halo field population in the [N/Fe]–[Fe/H] plane, where the lowest N-rich star nitrogen abundance is [N/Fe]$\sim +0.5$, in agreement with values from other samples (e.g. Martell et al. 2016; Schiavon et al. 2017; Nataf et al. 2019). To check whether the metallicity distribution function (MDF; shown in the top panel of Fig. 1) of our likely halo and N-rich star samples is statistically equal, we perform a two-sided Kolmogorov–Smirnov (KS) test, for which we obtain a resulting *p-value* of approximately 0.75. Therefore, we can reject the null hypothesis that these two samples originate from different parent distributions and trust that they are statistically the same with a high degree of confidence; this indicates that the N-rich stars display a similar MDF as the halo field population. However, the N-rich star sample is clearly separated from the halo field sample in their nitrogen distribution function (right-hand panel of Fig. 1).

---

[1]Stars belonging to the statistical sample are those which were selected at random from 2MASS and are included in fields with completed cohorts.
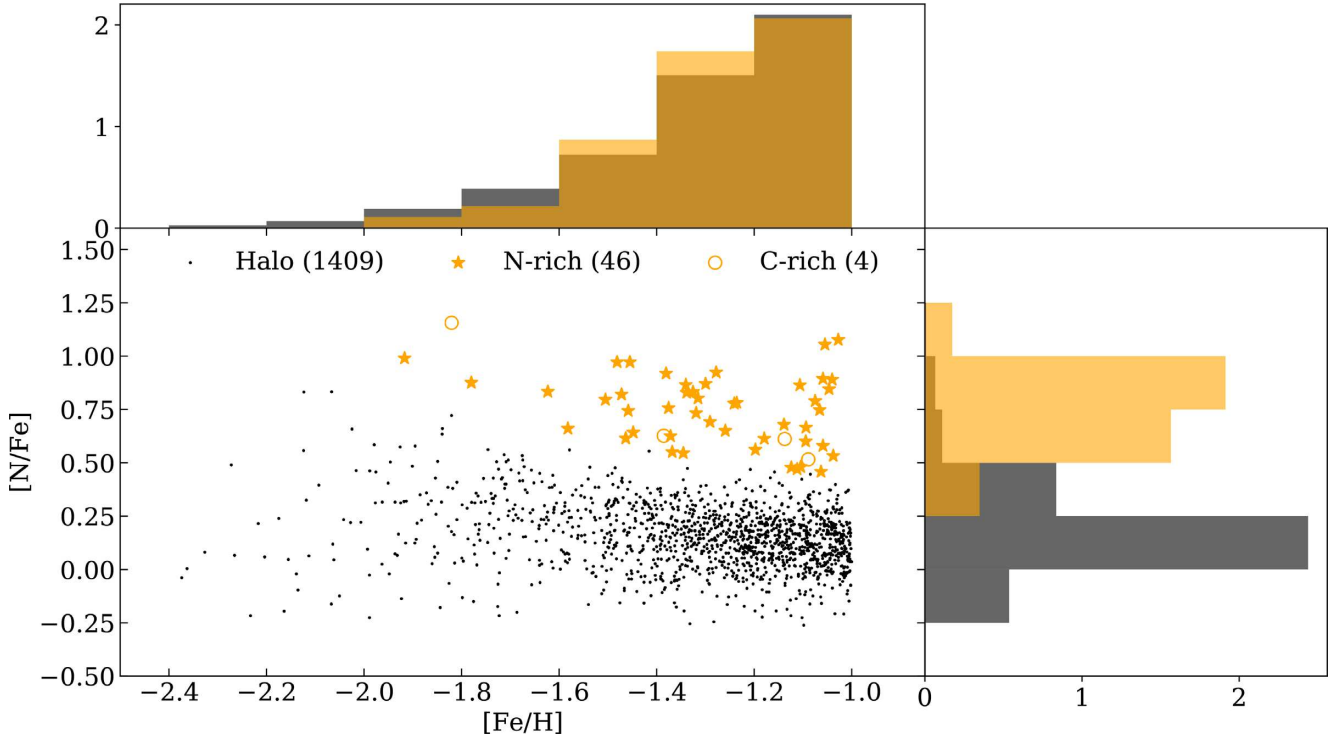[2]https://github.com/tholoien/XDGMM

Figure 1. [N/Fe] versus [Fe/H] distribution of the halo field star sample (black dots) and the selected N-rich stars (orange stars) used in this work. Histograms highlight the MDF (top) and nitrogen distribution function (right) of both samples, normalized by the total star number of stars in each sample. Both samples share the same MDF, however, can be clearly distinguished in the [N/Fe] distribution, with the mean N-rich star [N/Fe] value sitting ∼0.7 dex higher than the mean halo field [N/Fe] abundance. Open circles at high [N/Fe] values are the N-rich stars that did not satisfy the [C/Fe] < 0.15 criterion. The numbers stated in brackets quantify the number of stars in each sample.

We plot the nitrogen abundances of the normal and N-rich halo field stars as a function of their respective carbon and aluminium abundances in Figs 2 and 3, respectively. Those abundances are overlaid on data for stars residing in three GCs (NGC 6205, NGC 6904, NGC 7078) whose metallicities span the same range as our N-rich star sample (i.e. $-2.5 <$ [Fe/H] < -1), and for which we have substantial GC members, using the APOGEE GC sample from Horta et al. (2020a). By comparing our N-rich star sample to GC stars, we aim to discern if our N-rich stars present the typical light element variations of SP GC stars (Bastian & Lardo 2018). The data for the GC members display two clear sequences in Fig. 2, corresponding to the loci of the first population (FP) and SP stars, with the latter typically showing [N/Fe] $> +0.5$. Stellar evolution along the giant branch runs in parallel along each sequence, where more evolved, more luminous, lower $\log g$ stars have higher [N/Fe] and lower [C/Fe]. The well-known N–C anticorrelation in GCs connects stars of same evolutionary stage in each sequence, so that FP stars have higher [C/Fe] and lower [N/Fe] than their SP counterparts. With that in mind, the bottom panel suggests that the stars at the bottom of the SP sequence, with [N/Fe] $\sim +0.5$ and [C/Fe] $\sim +0.2$ have no counterparts in the FP sequence. Those would be stars with lower [N/Fe] and higher [C/Fe] than the bottom of the FP sequence, which itself is located at [C/Fe] $\sim +0.1$ and [N/Fe] ~ -0.1. We hypothesize that such stars are lacking in our sample due to a combination of effects. First, these stars have relatively high $T_{\rm eff}$ and $\log g$, which makes CN lines, upon which nitrogen abundances rely, too weak for reliable N abundances. Secondly, the ASPCAP grid has an [N/Fe] floor at -0.25.

Fig. 3 shows a correlation between [N/Fe] and [Al/Fe] for our N-rich star sample, with the majority of these presenting [Al/Fe]
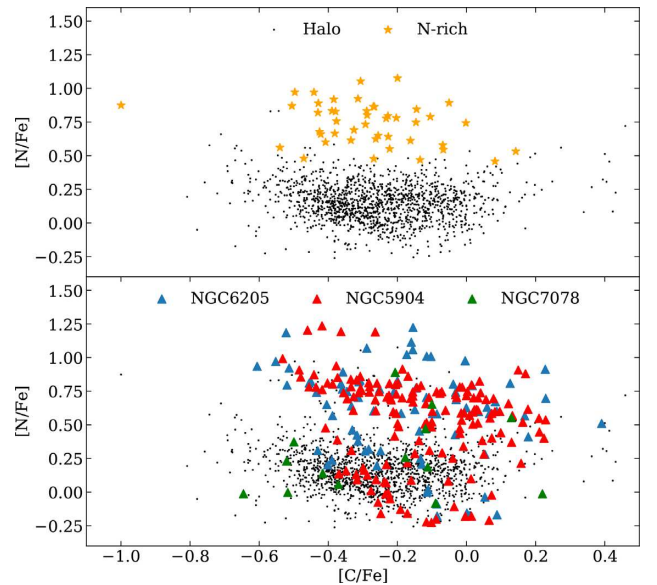

Figure 2. [N/Fe] versus [C/Fe] distribution for the N-rich stars (top) and for stars from the APOGEE GC sample for three representative clusters (bottom) derived using the same procedure as in Horta et al. (2020a). The black points are the same in both panels, and represent the halo field population. The N-rich star sample mimics the behaviour of SP GC stars of similar metallicity, occupying the same locus on this plane. Note that the N-rich stars with the highest [C/Fe] have no counterparts in the FP sequence. This is due to a combination of factors. Stars within that high [C/Fe] regime have higher $T_{\rm eff}$ and $\log g$ and low [N/Fe], which makes CN lines weaker. Moreover, there is an [N/Fe] floor in ASPCAP at -0.25.
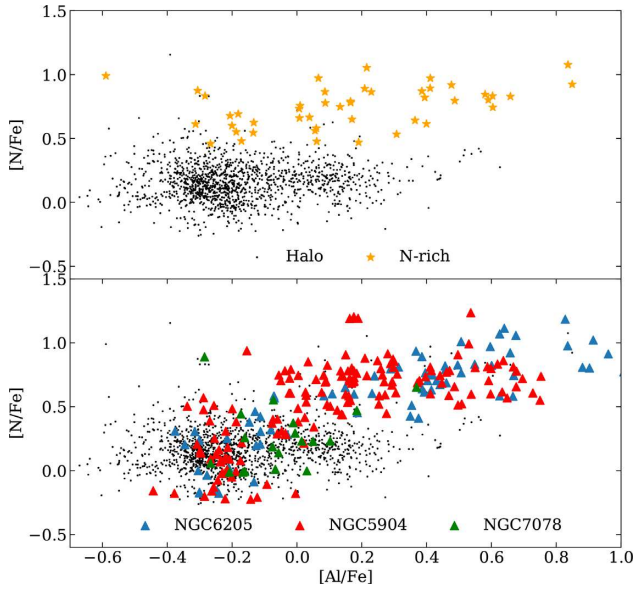
Figure 3. The same stars plotted from Fig. 2 in the [N/Fe] versus [Al/Fe] plane. Despite a small fraction of N-rich stars displaying [Al/Fe] < 0 dex, the majority occupy the same locus as the SP GC stars that display high [Al/Fe] abundances.

> 0, again occupying the same locus as the SP GC stars. We also report that N-rich stars display a much greater mean [Al/Fe] ($\simeq 0.3$) than the normal halo field population ([Al/Fe] $\simeq -0.2$). We also checked whether the N-rich population differed from its normal halo counterpart in terms of its neutron-capture element abundances, since those were found to be enhanced in SG GC stars by Cunha et al. (2017). The only neutron-capture element for which APOGEE DR16 provides abundances for our sample is Ce. We found no statistically significant difference in [Ce/Fe] between the two samples.

In a recent paper, Fernández-Trincado et al. (2019) suggest that SP GC stars must have [Al/Fe] > +0.5 in order to be considered true SP GC star candidates. In fact, adoption of such a restrictive criterion would lead to a severe underestimate of the N-rich population in both our field and GC sample. Stars belonging to NGC 6205, NGC 5904, and NGC 7078 from the Horta et al. (2020a) GC sample display a clear bimodality in the carbon–nitrogen plane, where the SP GC stars occupy a locus positioned at higher [N/Fe], congregating above their FP counterparts. This bimodality is also observed in the aluminium–nitrogen plane. In the bottom panel of Fig. 3, the SP GC stars (i.e. the filled triangles not overlapping with the halo population) clearly present [Al/Fe] abundances that fall below [Al/Fe] = +0.5. Thus, since our N-rich star sample occupies the same locus as the SP GC stars in both the carbon–nitrogen and aluminium–nitrogen planes, applying an [Al/Fe] > +0.5 cut to our sample would lead to a similarly severe underestimate in the number of N-rich stars in the field. In addition, it has been shown by several authors e.g. Carretta et al. (2009), Nataf et al. (2019), and Mészáros et al. (2020) that the extent of the Mg–Al anticorrelation varies from GC to GC, making a strict cutoff in Al difficult to implement.

Finally, we show the spatial distribution of the halo field population and the N-rich star sample in Fig. 4, represented with the same symbols as in Fig. 1. Before correcting for the APOGEE selection function or modelling the stellar density, it is immediately clear that the halo field population (black) occupies a vast range of heights above the Galactic mid-plane, reaching values of $Z \sim 20$ kpc. The same applies to the N-rich stars, however, these do not display as

large a Galactocentric distance range. Interestingly, we find that a significant fraction of the N-rich stars occupy a position close to the Galactic Centre, within an approximate radius of $R_{GC} \sim 3$ kpc. The remaining N-rich stars are scattered at higher Z and Y values. It has already been shown that N-rich stars make up a significant fraction of the inner Galaxy (Schiavon et al. 2017) and to a lesser extent of the outer halo (Martell et al. 2016; Koch et al. 2019), therefore our initial findings are in agreement with results from previous studies.

3 DENSITY MODELLING

In this section, we briefly describe the method employed for fitting the underlying number density of stars in the Milky Way's halo from APOGEE observations, which we represent here as $\nu_*(X, Y, Z|\tau)$, in units of stars kpc^{-3}. The computation of this quantity requires allowances to be made for the survey selection function, which is non-trivial due to the pencil beam nature of APOGEE, the presence of inhomogeneous dust extinction along the lines of sight, the target selection criteria imposing different H-magnitude limits, and the use of tracers, RGB stars, that are not standard candles. Our main goal is to determine density laws that describe the spatial distributions of the stellar halo and the N-rich samples separately. In this way, the number of observed normal field and N-rich stars can be converted into actual densities, in units of stars pc^{-3} at the solar radius (N_{R_0}). This density law can then be used to estimate the number of halo or N-rich stars APOGEE would have observed if it covered the full sky. Moreover, using stellar evolution models we can estimate the mass density. This value can then be estimated within a chosen volume by integrating an accurate density law for the halo and N-rich star samples, respectively. It is also straightforward to use the separate density laws to determine the ratio of mass density between N-rich stars and halo field stars, resulting in the fractional number of N-rich to total halo field stars, given as a percentage, as a function of position in the Galaxy. By following this procedure, we aim to determine if the ratio of N-rich stars is constant throughout the Milky Way, and if N-rich stars follow a similar density profile to that of the stellar halo field population.

Our methodology is an adaptation of that used by Mackereth et al. (2017) and Bovy et al. (2012b, 2014, 2016). We employ a modified version of their publicly available code.[3] Despite our method following that of these previous studies, we summarize the key steps here for clarity and completeness. For a full description of the fitting method, we refer the interested reader to section 3.1 from Mackereth & Bovy (2020). The main steps are summarized as follows:

(i) We fit different density functional forms for the density profile to the APOGEE N-rich star sample using a maximum likelihood fitting procedure, based on the assumption that star counts are well modelled by an inhomogeneous Poisson point process,[4] which takes into account the APOGEE selection function. For the halo sample, we adopt the functional form from Mackereth & Bovy (2020; see equation 1 in Section 3.1) for two reasons: First, the sample employed in their study is extremely similar to that used here; and secondly, their model describes our data well (see Section 4).

(ii) We determine the best-fitting density model for our N-rich star sample by calculating the Bayesian Information Criterion (BIC;

---

[3]Available at https://github.com/jmackereth/halo-mass and https://github.com/jobovy/apogee-maps, respectively.

[4]See Daley & Vere-Jones (2003) for a more detailed definition of an inhomogeneous Poisson point process.
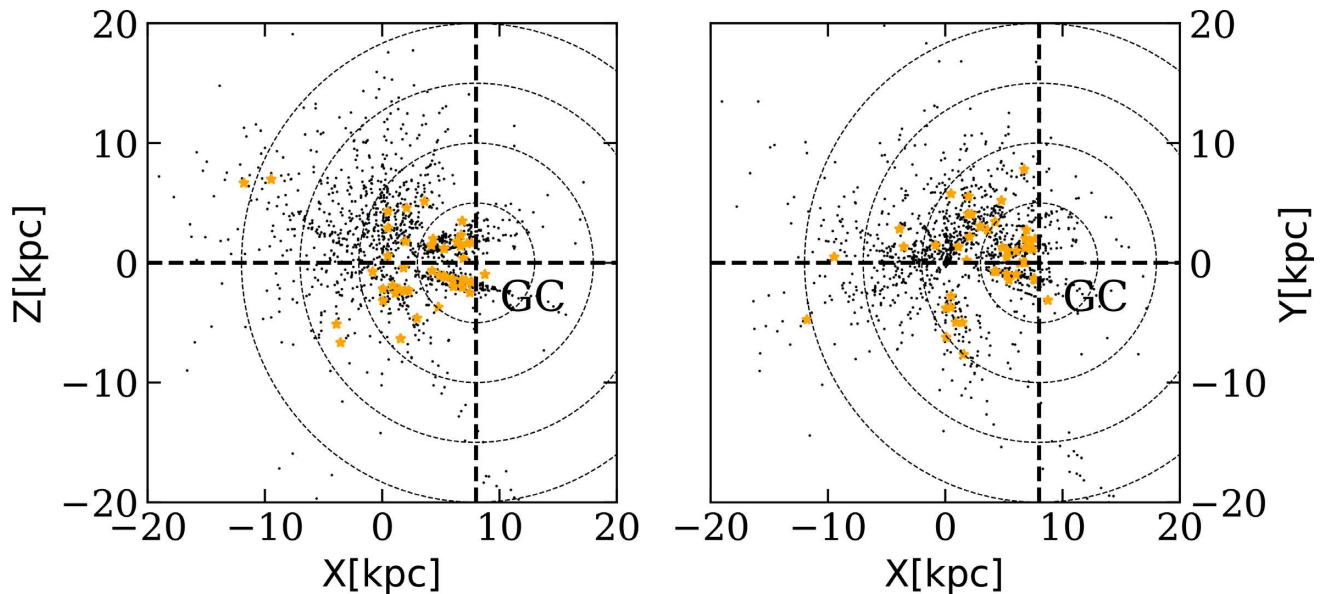
Figure 4. Spatial distribution in heliocentric *X–Z* and *X–Y* of our APOGEE DR16 field (black dots) and N-rich star (orange stars) samples between $-2.5 <$ [Fe/H] < -1. The crosshair signifies the position of the Galactic Centre. The dashed circled lines denote spherical radius bins of 5 kpc in size, signifying a 5, 10, 15, and 20 kpc distance from the Galactic Centre.

Schwarz 1978) and the logarithmic maximum likelihood value, $\ln(\mathcal{L}_{\max})$, for each model, and choosing that with the lowest BIC value (see Sections 3.2 and 3.3 for a description of the models tested and the selection of our best-fitting model, respectively). Our selected halo density profile includes a single exponential disc profile to account for the contribution of the (thick) disc component to our sample. We find that by including this disc fraction parameter to our model, we are able to quantify the fraction of the sample that belongs to the (thick) disc and that would affect the fitting procedure. Moreover, given our metallicity cut, the contribution by the thin disc to our sample is negligible, so we do not include that component in our model. Using this density law for the halo, and the resulting best-fitting model for the N-rich stars, we initiate a Markov chain Monte Carlo (MCMC) sampling of the posterior probability distribution function (PDF) of the parameters fitted in each density law. We adopt the median and standard deviation of one-dimensional projections of the MCMC chain as our best-fitting parameter values and uncertainties.

(iii) As the fitting procedure does not fit for the normalization of the density N_{R_0} (the number surface density of halo or N-rich stars at the solar neighbourhood in stars pc^{-3}), we compute this quantity by comparing the observed number of halo or N-rich stars in the metallicity bin adopted to that which would be observed in APOGEE for the fitted density model if $N_{R_0} = 1$ star pc^{-3}. We then employ this number density quantity (N_{R_0}) to obtain the true number of stars given by our best-fitting halo or N-rich density model within a given volume. This number density can then be converted into a mass density using stellar evolution models. Since our aim is to determine the contribution of N-rich stars as a function of Galactocentric distance, we take the ratio of the mass densities yielded for the N-rich star sample and the halo sample. This allows us to compare the ratio of GC dissolution in different spatial regions of the Galaxy, and potentially place constraints on the origin of the contribution of dissolved GC stars to the Galaxy's stellar halo mass budget, and the mass assembly history of the Milky Way.

Our method corrects for the effects induced by interstellar extinction using combined 3D maps of the Milky Way derived by

Marshall et al. (2006) for the inner disc plane and those derived for the majority of the APOGEE footprint by Green et al. (2019), adopting conversions $A_H/A_{K_s} = 1.48$ and $A_H/E(B - V) = 0.46$ (Schlafly & Finkbeiner 2011; Yuan, Liu & Xiang 2013). Any fields with no dust data (10 fields) were removed from the analysis. We adopt the combination of Marshall et al. (2006) and Green et al. (2019) dust maps because Bovy et al. (2016), who assessed the relative merits and limitations of a number of available maps in the literature, determined that this combination of dust maps provides the best fits when performing density modelling on APOGEE data for a trace population.

3.1 Halo density model

We aim to discern the profile that best describes our halo field sample. Recent work modelling the Galactic stellar halo using RR Lyrae stars found that the stellar halo is well modelled by a triaxial density ellipsoid, angled at roughly $70°$ with respect to the axis connecting the Sun and the Galactic Centre (Iorio et al. 2018; Iorio & Belokurov 2019), but can also be well modelled by a single power law (SPL). Along the same lines, Mackereth & Bovy (2020) used the APOGEE DR14 data and the corresponding APOGEE selection function, and found that the mono-abundance populations in the [Fe/H] range selected in that work are well modelled by a triaxial SPL density ellipsoid with a 'cut-off' term. The cut-off term accounts for our ignorance regarding the extent of the sample outside of the observed range, and includes a scale parameter β corresponding to the scale length (such that $h_{r_e} = 1/\beta$). This density profile also includes a disc term, modelled by an exponential disc profile with scale height $h_z = 0.8$ kpc and scale length $h_R = 2.3$ kpc (Mackereth et al. 2017), to account for any contamination from the thicker components of the high [α/Fe] disc, parametrized by the factor f_{disc}. Hereafter, we refer to this model as TRI-CUT-DISC.

Since we are modelling a similar metal-poor halo sample as that one modelled in Mackereth & Bovy (2020), with a very similar MDF, we adopt the functional form from their best-fitting profile to model

the stellar halo field population. That density profile is given by

$$v_*(r_e) \propto (1 - f_{disc})r_e^{-\alpha} \exp(-\beta r_e) + f_{disc}v_{*,disc}. \tag{1}$$

The density is normalized such that the density at the Sun (i.e. $v_*(r_{e,0})$) is equal to unity, and r_e defines the ellipsoidal surfaces on which the density is constant, given by

$$r_e^2 = X_g^2 - \frac{Y_g^2}{p^2} - \frac{Z_g^2}{q^2}, \tag{2}$$

where p and q describe the Y_g-to-X_g and Z_g-to-X_g axis ratios, respectively, and (X_g, Y_g, Z_g) are the Cartesian coordinates relative to the Galactic Centre. The best-fitting model found in Iorio et al. (2018) included an allowance for the variation in q with r_e, with a scale length of ~13 kpc. In this work, we assume q does not vary with elliptical radius, and constrain both shape parameters q and p to be <1, forcing the longest axis to be that in the X_g direction. As in Mackereth & Bovy (2020), we allow the orientation of the density ellipsoid to vary, applying a transformation defined by the unit vector $\hat{z}$ along the transformed Z-axis, and the angle of rotation (from the original X-axis) of the ellipsoid about this transformed axis, ϕ. In practice, $\hat{z}$ is sampled uniformly (with the MCMC algorithm) by deprojecting samples from an equal-area rectangular projection of the unit sphere. Employing this method, $\hat{z}$ is represented by two parameters, η and θ, such that $\hat{z} = (\sqrt{1 - \eta^2}\cos\theta, \sqrt{1 - \eta^2}\sin\theta, \eta)$. Here, η is sampled uniformly between -1 and 1, and θ between 0 and 2π. Such a transformation generally has little impact on the measurement of the total mass, as the parameters tend to define an ellipsoid which is roughly aligned with the Sun's position. For all parameters, we adopt uninformative priors, and set the allowed exponent α and parameter β range to be positive. For the remaining parameters, only a range of [0, 1] is permitted.

3.2 N-rich star density models

With no preconception of the density distribution of N-rich stars in the Galaxy, it is important to test several different density models for the N-rich sample, and statistically evaluate which model best fits the data. In this subsection, we describe the range of density models fitted to the N-rich star data, and provide the analytical function for clarity and completeness.

We begin by fitting the simplest model to the N-rich star sample, a single power law (SPL). This model is described analytically as

$$v_*(r_e) \propto r^{-\alpha}, \tag{3}$$

where r is given by equation (2) when equating p and q to unity. This profile assumes the density is prescribed by spherical shells, with no flattening along any axis. A flattening parameter along the Z-axis (q) can be introduced to obtain an axisymmetric profile (AXI), which is described by equation (3), assuming r is given by equation (2), and equating p to one. Similarly, a separate flattening parameter along the Y-axis (p) can be introduced to obtain a triaxial profile (TRI), governed by the shape of a triaxial ellipsoid. Alongside these three SPL density models, we choose also to fit the model described in Section 3.1 (TRI-CUT-DISC), employed to fit the stellar halo field data.

In addition to the profiles mentioned above, we fit two further profiles that are not part of the SPL family. The first is an exponential disc profile from Mackereth et al. (2017; DISC), given by

$$v_*(r_e) \propto \exp(-h_R(R - R_0) - h_z|z|), \tag{4}$$

where R and z are the radial and vertical axes in a cylindrical coordinate system, respectively, h_R and h_z are the radial and vertical

Table 1. Summary of the results for a sample of density profiles used to fit the N-rich star sample. The models tested are: a single power law (SPL), an axisymmetric power law (AXI), a triaxial power law (TRI), a triaxial power law with a disc and cut-off term (TRI-CUT-DISC), a broken power law (BPL), and an exponential disc (DISC). For each model, we report the Bayesian Information Criterion (BIC) and maximum logarithmic likelihood differences between the best-fitting model (bold) and the remaining models, calculated by taking 10 000 MCMC samples and using the median posterior parameter values.

Density profile	Δ BIC	$\Delta \ln(\mathcal{L}_{max})$
SPL	34.3	19.3
AXI	0.0	0.0
TRI	2.2	0.5
TRI-CUT-DISC	10.0	2.6
BPL	60.9	17.6
DISC	9.1	12.0

scale lengths, and R_0 is the Galactocentric radius of the Sun. We believe that by testing a disc profile, we will be able to discern if the N-rich star sample is better described by a disc density model, and thus if it has a high contribution to the Galactic disc component. The second additional profile fitted is a broken power law (BPL). This density model resembles that of an SPL (see equation 3), however, has a break radius denoting a change in the slope of the model, given by a change in the value of α. The BPL is given by

$$v_*(r_e) \propto \begin{cases} r^{-\alpha_{in}} & \text{if } R < R_{break} \\ r^{-\alpha_{out}} & \text{if } R > R_{break}, \end{cases} \tag{5}$$

where α_{in} is the power-law exponent inside the break radius (i.e. R_{break}), and α_{out} is the exponent outside the break radius. By fitting a broken profile, we will be able to test the hypothesis in which the N-rich star sample is governed by a break between an inner and outer component, potentially linked to the different N-rich star ratios found in previous studies for different spatial regions.

For all the models presented, we adopted uninformative priors, as performed in Section 3.1. For the parameters of the exponential disc profile, we set a permitted scale length range of [0, 10] (kpc). For the permitted range of the break radius of the BPL, we permit values between [0, 20] (kpc), a range of Galactocentric distances that is relatively well covered by our data.

3.3 Identification of best-fitting density model

In this subsection, we describe the methodology employed to select the best-fitting N-rich star profile from the density models described in Section 3.2. We choose the Bayesian evidence ratio as our figure of merit to identify the best-fitting model. We assume that the posterior distributions are nearly Gaussian, and that therefore the Bayesian evidence ratio can be approximated by the BIC (Schwarz 1978), defined as

$$\text{BIC} = d(\tau)ln(N_*) - 2ln(\mathcal{L}_{max}), \tag{6}$$

where $d(\tau)$ is the number of free parameters in the density model, N_* is the number of stars in our sample, and $\ln(\mathcal{L}_{max})$ is the logarithm of the maximum value of the likelihood.

The best-fitting model is that for which the BIC value is the lowest. The BIC statistic penalizes models with a large number of free parameters, such that for two models with the same $\ln(\mathcal{L}_{max})$ value, the one fewer free parameters is preferred. The results from this comparison are listed in Section 4.2, displayed in Table 1. For

the fit of each profile to the N-rich star data, we refer the reader to Appendix A.

3.4 Mass estimation

We follow the work by Bovy, Rix & Hogg (2012a) and Mackereth & Bovy (2020), and estimate the mass for our halo and N-rich samples. Although the methodology is fully described in Mackereth & Bovy (2020), we briefly explain the procedure for estimating the mass within an APOGEE star sample for clarity. Upon finding a best-fitting model and its associated uncertainty for a distribution of stars, the measurement of the mass can be computed by employing the normalization of the rate function:

$$\lambda(O|\tau) = \nu_*(X, Y, Z|\tau) \times |J(X, Y, Z; l, b, D)|$$
$$\times \rho(H, [J - K_s]_0, [Fe/H]|X, Y, Z) \times S(l, b, H), \quad (7)$$

where $\nu_*(X, Y, Z|\tau)$ is the stellar number density in rectangular coordinates, in units of stars kpc^{-3}. $|J(X,Y,Z;l,b,D)|$ is the Jacobian of the transformation from rectangular (X, Y, Z) to Galactic (l, b, D) coordinates, and $\rho(H, [J - K_s]_0, [Fe/H]|X, Y, Z)$ represents the density of stars in magnitude, colour, and metallicity space given a spatial position (X, Y, Z), in units of stars per arbitrary volume in magnitude, colour, and metallicity space. $S(l, b, H)$ is the survey selection function (see Bovy et al. 2016; Mackereth et al. 2017, for details), which denotes the fraction of stars successfully observed in the survey's colour and magnitude range, and includes dust extinction effects.

We perform this mass estimation procedure on both our total halo and N-rich star samples by calculating the number of stars seen by APOGEE for a given density model normalized to unity at the Solar position, $N(\nu_{*,0} = 1)$. That number is obtained by integrating the rate function over the observable volume of the survey. This integral is given by

$$N(\nu_{*,0} = 1) = \int_{fields} \text{dfield} \, dD \, \lambda(\text{field}, D)$$
$$= \int \text{dfield} \, d\mu \, \frac{D^3(\mu)\log(10)}{5} \nu_*([R, \phi, z](\text{field}, \mu|\tau)) \times \mathfrak{S}(\text{field}, \mu), \quad (8)$$

where the density and effective selection function (namely, $\mathfrak{S}(\text{field}, \mu)$) are calculated along APOGEE sightlines on a grid linearly spaced in distance modulus μ. Since the true number of observed stars is given by

$$N_{obs} = A \, N(\mu_{*,0} = 1), \quad (9)$$

comparison of the expected number count for a normalized density model with the true observed number of stars in the sample provides the proper amplitude, A, which is then equivalent to the true number density of RGB stars at the Sun, $\mu_{*,0}$, when considering either our halo or N-rich star sample.

The number density in concentric triaxial ellipsoids can also be calculated as a function of Galactocentric distance. Once a best-fitting model for a specific sample of APOGEE stars is estimated, the number density can be computed on a grid of $\nu_*([R, \phi, \theta])$ over a chosen volume. We can then compute this value as a function of Galactocentric radius by integrating along the ϕ- and θ-axis, which can later be converted into stellar mass density using stellar evolution models. In either case, the number counts in RGB stars can be converted into the mass of the entire underlying population. To do so, we use the PARSEC isochrones (Bressan et al. 2012; Marigo et al. 2017), weighted with a Kroupa (2001) initial mass function (IMF).

The average mass of RGB stars $\langle M_{RGB} \rangle$ observed in APOGEE is then calculated by applying the same cuts in $(J - K_s)_0$ and log g to the isochrones. The fraction of the stellar mass in giants, ω, is given by the ratio between the IMF weighted sum of isochrone points within the RGB cuts and those of the remaining population. The conversion factor between RGB number counts and total stellar mass can then simply be calculated using

$$\chi([Fe/H]) = \frac{\langle M_{RGB}\rangle([Fe/H])}{\omega([Fe/H])}. \quad (10)$$

As explained in Mackereth & Bovy (2020), the factor for each field and each selection in [Fe/H] is computed by adjusting the limit in $(J - K_s)_0$ to reflect the minimum $(J - K_s)_0$ of the bluest bin adopted in that field, and only considering isochrones that fall within $-2.5 <$ [Fe/H] < -1. The edges in colour binning for each field are accounted for by the integration over $\rho[(J - K_s)_0, H]$ for the effective selection function. Mackereth & Bovy (2020) used *Hubble Space Telescope* photometry to show that the factors determined using their method are reliable against systematic uncertainty arising from the choice of stellar evolution models.

Once the normalization for our sample is obtained, we integrate the normalized density models described by 400 samples from the posterior distributions of their parameters to attain the total mass within a population, and the total mass as a function of Galactocentric radius for that same population. To avoid overextrapolation from our fits to the halo density, we only integrate for the mass within a 1.5 kpc $< r <$ 15 kpc range, for which our data are well contained.

4 RESULTS

4.1 Fit to the halo sample

We perform the fitting procedure described in Section 3 to our APOGEE halo sample, to ensure that the prescribed density provides a good fit to the data. Since it has recently been shown that the APOGEE DR14 star sample between $-2.5 <$ [Fe/H] < -1 is well defined by a triaxial SPL density ellipsoid, which included a cut-off term and freedom to rotate around its axis (Mackereth & Bovy 2020), we choose to fit this profile to our sample. This profile also included a parameter to account for the fraction of the disc in our sample, described by an exponential disc profile. The resulting best-fitting profile, obtained running 10 000 realizations of the model, follows a moderately steep power law, yielding a value of $\alpha \simeq 3.48 \, ^{+0.05}_{-0.07}$, with a slight flattening along the Y-axis ($p = 0.76 \, ^{+0.03}_{-0.04}$), and slightly more flattened along the Z-axis, with $q = 0.46 \pm 0.01$, which is in good agreement with previous results (e.g. Deason, Belokurov & Evans 2011; Xue et al. 2015; Iorio et al. 2018; Mackereth & Bovy 2020). We find the β cut-off parameter to be ~0.01, indicating a scale length which is well outside the range of our data, of the order of $h_{r_e} \sim 100$ kpc, and that the triaxial ellipsoid is consistent with a minor rotation around the Z-axis, $\theta = 3.6 \pm 0.5$ deg, and a slightly larger rotation along the X-axis, $\phi = 14 \, ^{+11}_{-9}$ deg. Our results also show that our halo sample has low contamination from the disc, given by the low $f_{disc} = 0.08 \, ^{+0.16}_{-0.06}$ value. The samples from the posterior distribution of parameters given by the data are illustrated in a corner plot in Fig. 5, which shows that the posterior distributions are well behaved and converge towards a single value. Alongside the samples, we show in the top right panel the observed distance modulus μ distribution as predicted by a triaxial disc cut-off profile (black) and an SPL (blue) with an exponent that best fits the data ($\alpha \simeq$ 3.2), compared to the real APOGEE data. Given that our triaxial disc cut-off power law fits our data well, and yields parameter estimates
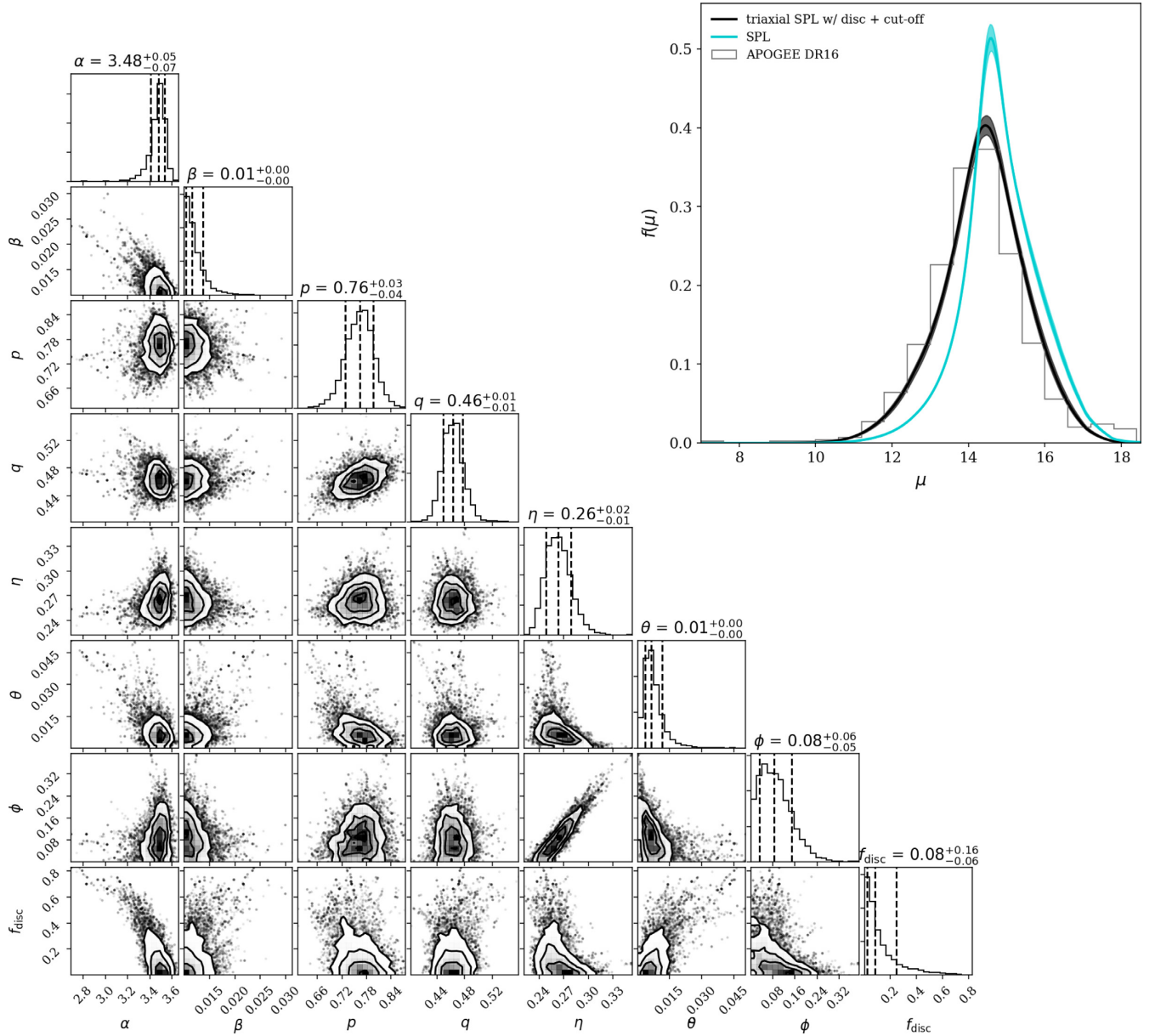
Figure 5. Corner plot showing the posterior 10 000 samples of the parameters for the adopted triaxial disc cut-off SPL (TRI-CUT-DISC) model when fit to the full statistical likely halo sample (1409 stars) between $-2.5 < $ [Fe/H] < -1. The posterior distributions are well behaved, and converge to a median value. The best-fitting profile has a mildly steep power law of $\alpha \simeq 3.5$, and is slightly flattened along its *Y*- and *Z*-axes. The inset panel (top right) shows the distance modulus μ distribution predicted by the best-fitting model (black), and the best-fitting SPL (blue) of similar $\alpha \simeq 3.2$, where the thickness of each line signifies the 1σ spread. The grey histogram represents the real APOGEE data, and shows that the TRI-CUT-DISC profile provides a better quality fit.

within agreement with values from previous studies, we choose to use this model as our best-fitting density profile.

It is interesting to compare our best estimate of the halo enclosed mass and normalization with those by Mackereth & Bovy (2020). Using the triaxial disc cut-off power-law profile and the APOGEE star counts, we obtain a halo normalization $\rho_{*,0} = 1.3 \, ^{+0.1}_{-0.2} \times 10^{-4}$ $M_\odot$ pc^{-3}, and a total stellar halo mass $M_{*,\mathrm{halo}} = 8.3 \, ^{+1.5}_{-1.3} \times 10^8$ $M_\odot$, by using equation (8) and integrating over the full observable volume (i.e. $r \sim 1.5$–15 kpc). We find our total stellar halo mass estimate to be larger by a factor of 1.5 than that obtained by Mackereth & Bovy (2020) for the same density profile using APOGEE DR14 data, and suggest it is likely due to the different volumes employed to integrate the density (Mackereth & Bovy 2020 integrated the halo volume within a Galactocentric distance ranging from 2 kpc $< r$

< 70 kpc). Furthermore, within the uncertainties, we find our total halo mass estimate to be in borderline agreement with the estimated 4–7×10^8 $M_\odot$ from the review by Bland-Hawthorn & Gerhard (2016), and to be smaller than the $\sim 10^9$ $M_\odot$ estimate presented in the work of Deason et al. (2019), or the sum of individual MAPS in Mackereth & Bovy (2020).

4.2 Fitting the N-rich star sample

In this subsection, we describe the density modelling of the chemically defined N-rich stars. We test six different density profiles ranging in the number of free parameters, and calculate the $\ln(\mathcal{L}_{\mathrm{max}})$ and BIC values in order to estimate the best-fitting density profile.
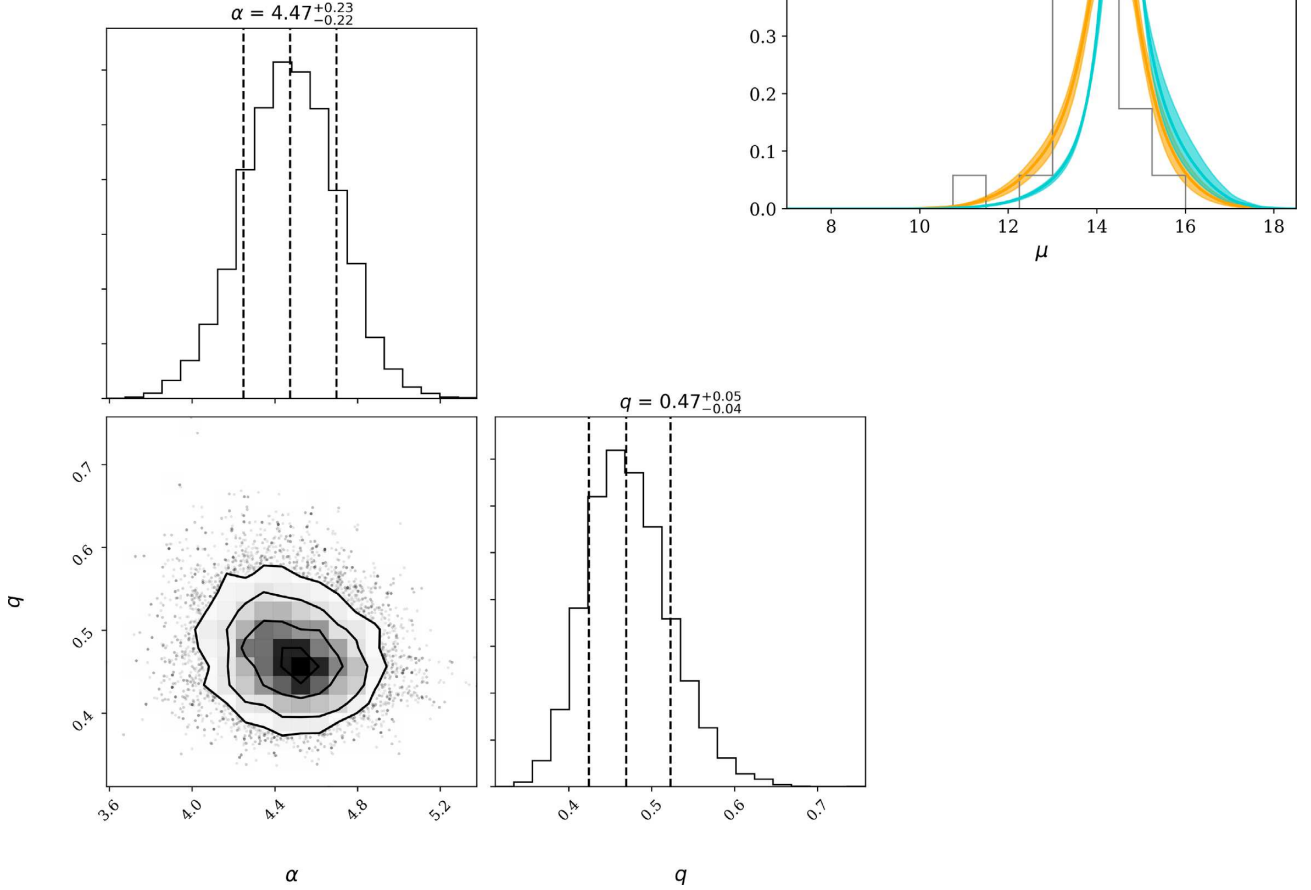
Figure 6. The same as in Fig. 5, however, for the 46 N-rich stars displayed in Fig. 1 fitted by an AXI. As in Fig. 5, the posterior distributions are well behaved. The best-fitting model displays a steeper exponent $\alpha \simeq 4.5$ in comparison to the halo sample, but shows that it is similarly flattened along the Z-axis. As in Fig. 5, the inset panel displays the predicted distance modulus μ distribution predicted by the best-fitting profile (orange) and a best-fitting SPL with $\alpha \simeq 3.9$ (blue), compared to the N-rich star data.

The models of choice are described in detail in Section 3.2, and are summarized as follows: an SPL, an AXI, a TRI, the triaxial disc cut-off profile used to fit the halo sample (TRI-CUT-DISC), a BPL, and an exponential disc profile (DISC). Employing the method described in Section 3.3, we have obtained the BIC and $\ln(\mathcal{L}_{max})$ for the six models tested, which are listed in Table 1. The fit of each profile to the N-rich star data is shown in Fig. A1 in Appendix A. We draw 10 000 samples from the posterior of each model, for which we take the median value as the best-fitting parameters for each profile. We then compute the BIC and $\ln(\mathcal{L}_{max})$. As is evident from the resulting BIC values from Table 1 (and from Fig. A1), the AXI profile is the best-fitting profile. However, we find the N-rich star sample can also be well fitted by the TRI-CUT-DISC and TRI profiles.

Once we obtain our best-fitting profile, we perform the density fitting procedure described in Section 3 on the N-rich star sample,

and obtain a density profile that provides an exquisite fit to the data. After running 10 000 MCMC iterations, we obtain a density profile with a slope of $\alpha = 4.47^{+0.23}_{-0.22}$, which is much steeper than that of the halo profile. Furthermore, we find the N-rich profile to have roughly the same flattening along the Z-axis as the halo profile (namely, $q = 0.47^{+0.05}_{-0.04}$).

The resulting samples from the posterior distribution of parameters given by the data are shown in the corner plot illustrated in Fig. 6, which shows that the posterior distributions are well behaved and converge to a median value. In addition to the resulting MCMC samples, we show the distance modulus μ distribution predicted by the best-fitting model (orange) and an SPL (blue) with an exponent that best fits the data ($\alpha \simeq 3.9$), compared to the real N-rich star APOGEE data. Using the best profile and APOGEE N-rich star counts, we obtain an N-rich normalization $\rho_{0,\text{N-rich}} = 3.1^{+0.6}_{-0.5} \times 10^{-6}$
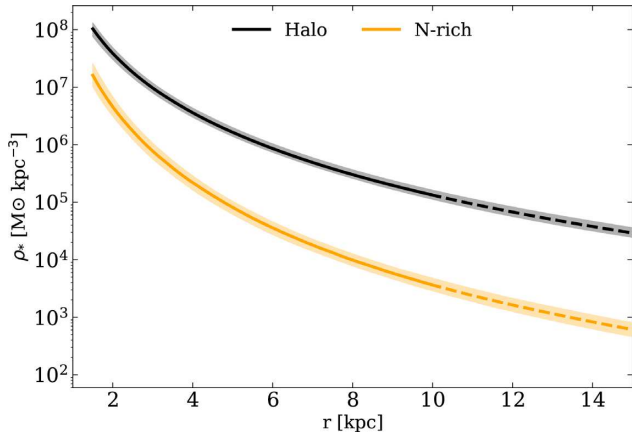
Figure 7. Integrated density as a function of spherical radius for the best-fitting density profiles of the halo (black) and the N-rich (orange) star samples. The shaded regions mark the 16th and 84th percentile uncertainties. The dashed line indicates the region where, due to low sample numbers, the density is not strongly constrained. Both the halo and N-rich density profile follow a similar pattern within the outer $r \gtrsim 3$ kpc region, however, the N-rich density decreases more steeply.

$M_\odot$ pc^{-3}. Given the reasonable fit to the N-rich star data by the AXI, we choose to use this profile as our best-fitting N-rich density profile. However, we point out that the TRI and TRI-CUT-DISC profiles also provide a good-quality fit to the N-rich star data, and yield a normalization value that is consistent, within the uncertainties, with the value obtained for the AXI profile. Moreover, we find that when assessing the contribution of N-rich stars to the stellar halo field using either the TRI or TRI-CUT-DISC profiles as the best-fitting N-rich star model, we obtain the same results as for the AXI profile, within the uncertainties.

The resulting density profiles for the halo (black) and N-rich star (orange) samples are displayed as a function of spherical radius in Fig. 7. It is clear that, while the halo density appears to decrease at a slower rate with increasing radius, the N-rich star sample exhibits a much greater decrease at high radii, and a much higher density within $r \lesssim 3$ kpc, in comparison to the halo field population.

4.3 Contribution of dissolved/evaporated globular clusters to the stellar halo of the Milky Way

Under the assumption that N-rich stars result from GC dissolution, it is interesting to estimate the total stellar mass contained in dissolved GC stars, in order to assess their total contribution to the total stellar halo mass budget. Besides an estimate of the contribution to the stellar halo budget by stars that once belonged to GCs, another output from our method is the spatial distribution of those stars. In this section, we estimate the mass density contributed from the halo and N-rich samples, respectively, as given by our best-fitting density models as a function of distance from the Galactic Centre. Then we estimate the ratio between the mass in N-rich stars and the total stellar halo mass. To determine the total mass contributed by field stars from GC origin to the stellar halo, we assume the ratio of FP to SP stars in GC to be 2 SP stars for every 1 FP star, adopting the minimal scenario proposed by Schiavon et al. (2017). Although other FP-to-SP GC star ratios have been proposed (e.g. Bastian & Lardo 2015; Cabrera-Ziri et al. 2015), we chose to focus solely on the 'minimal scenario' as it has been shown by Schiavon et al. (2017) that other scenarios can be excluded due to the predicted total GC star number
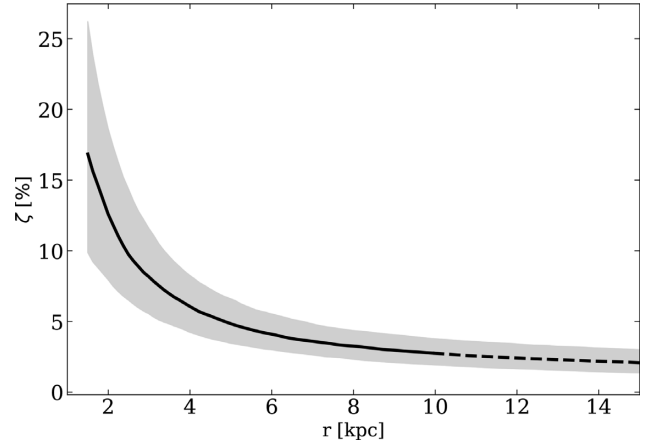

Figure 8. Mass density percentage ratio of N-rich stars and halo field stars as a function of spherical radius. The black solid line signifies the median value, while the shaded regions show the 16th and 84th percentile uncertainties. The dashed line indicates the Galactocentric distance range where the density is not well constrained due to low numbers of N-rich stars. The mass density percentage ratio drops from $\zeta = 16.8^{+10.0}_{-7.0}$ per cent at $r = 1.5$ kpc to a value of $\zeta = 2.7^{+1.0}_{-0.8}$ per cent at $r = 10$ kpc. Under the minimal scenario assumption, one can multiply ζ by 1.5, and subtract the FP stars from the halo field, to obtain the total contribution from disrupted GC stars to the stellar halo.

counts exceeding the expected number of total halo stars. We then take the ratio of the disrupted GC mass with halo mass, to assess the mass contribution of the former to the halo field as a function of Galactocentric distance. Performing this exercise will not only lead to a deeper understanding of the origin of disrupted GC stars (and in consequence N-rich stars), but it will also provide a more clear understanding of the mass contribution from GC destruction to the total stellar halo mass budget.

We determine the mass in spherical annuli for both N-rich and halo stars, which in turn enables us to determine the mass from each sample as a function of spherical radius, as well as the ratio between the samples. As mentioned above, to estimate the mass contribution by stars that once belonged to GCs, we need to account for the contribution by former GC stars whose abundance patterns do not differ from that of the field population at same metallicity, which can be accounted for by assuming the FP–SP GC star ratio predicted from the minimal scenario.

Now that we have determined the density of both N-rich and halo stars as a function of Galactocentric distance, we focus our attention on estimating the ratio between N-rich stars and the total stellar halo. The results from this ratio estimation are shown as the black solid line in Fig. 8, where the N-rich to halo ratio, hereafter denoted as ζ, is given as a percentage. The shaded grey regions signify the 16th and 84th percentiles. In a similar fashion to Fig. 7, we choose to compute ζ between a 1.5 kpc $< r <$ 15 kpc range, for which is covered relatively well by the data. Our results show that ζ increases rapidly with decreasing radius, growing from a value of $\zeta = 2.7^{+1.0}_{-0.8}$ per cent at $r = 10$ kpc to a value of $\zeta = 16.8^{+10.0}_{-7.0}$ per cent at $r = 1.5$ kpc. Moreover, around the solar neighbourhood (i.e. $r = 8$ kpc), we find the ratio to be $\zeta = 3.3^{+1.1}_{-1.0}$ per cent. Therefore, our estimates reveal approximately an eight-fold increase of the N-rich star to halo contribution in the inner Galactocentric regions, in agreement with previous findings (Schiavon et al. 2017).

Under the assumption of the minimal scenario presented in Schiavon et al. (2017), whereby the ratio of FP to SP is 1-to-2, we estimate the ratio of dissolved GC stars to halo field stars to be

$\zeta_{\text{tot}} = 27.5^{+15.4}_{-11.5}$ per cent at $r = 1.5$ kpc. Along the same lines, we obtain an estimate of $\zeta_{\text{tot}} = 4.2^{+1.5}_{-1.3}$ per cent at $r = 10$ kpc. Our results show that when accounting for selection effects in the observational data, the contribution of stars arising from GC dissolution and/or evaporation to the total stellar halo field is greater by a factor of $\sim$7–9 in the inner few kiloparsecs when compared to the outer regions of the Galaxy. Furthermore, we report that when repeating the methodology employing another well-fitting density profile for the N-rich stars (in this case the TRI and TRI-CUT-DISC models), we find that, within the uncertainties, our results remain unchanged, thus validating both our estimates and procedure.

Now that we have shown the mass fraction contribution of stars arising from GC disruption to the halo field of the Galaxy as a function Galactocentric distance, we can compute an estimation of the mass contributed from GC escapees to any given volume or spherical shell. We chose to estimate the mass contributed by stars arising from GC destruction, under the minimal scenario assumption, within a volume spanning a radius between 1.5 and 3 kpc from the Galactic Centre. For this shell volume, we obtain a total mass from stars arising from GC dissolution of $M_{\text{GC,inner}} = 5.6^{+2.8}_{-1.8} \times 10^7$ M$\odot$. Similarly, we compute the total mass of stars arising from GC disruption within a shell of $\sim$13.5 kpc in radius (from 1.5 to 15 kpc), and find a mass estimate of $M_{\text{GC,total}} = 9.6^{+4.0}_{-2.6} \times 10^7$ M$\odot$. Thus, our results show that disrupted GC stars contribute a significant amount of mass to the stellar halo of the Galaxy. This notable mass contribution is observed at all Galactocentric scales, however, is more dominant at smaller radii, as shown in Fig. 8.

5 DISCUSSION

The results obtained in Fig. 8 suggest that the fraction of N-rich star mass as a function of halo field mass is much greater in the inner regions of the Galaxy when compared to the outer regions. In agreement with observational estimates (Martell & Grebel 2010; Carollo et al. 2013; Martell et al. 2016; Schiavon et al. 2017; Fernández-Trincado et al. 2019; Koch et al. 2019; Lee et al. 2019), and theoretical predictions (Tremaine, Ostriker & Spitzer 1975; Gnedin, Ostriker & Tremaine 2014; Reina-Campos et al. 2019; Hughes et al. 2020), we hypothesize that N-rich stars result from the destruction of pre-existing GCs. Assuming this hypothesis is correct, our findings may have important repercussions for the current understanding of the formation and evolution of the Galactic GC system, the presence of multiple stellar populations in GCs, possibly the formation and evolution of the Milky Way bulge, and even the mass assembly history of the Galaxy. In the following subsections, we discuss and compare the results from our stellar halo density modelling and mass estimation to previous work.

5.1 Comparison with previous halo density work

In this section, we perform a detailed comparison of our findings with those from previous works. The contrasting method employed in this work allows for an interesting comparison, one that will help shed light into the nature of the density of the stellar halo of the Galaxy.

The majority of early work focused on modelling the density of the Galactic stellar halo employed different stellar types as tracers of the halo population, including either main-sequence turn-off (MSTO) stars (Morrison et al. 2000; Jurić et al. 2008) or blue horizontal branch (BHB) stars (Yanny et al. 2000), for which the photometry is

relatively easily calibrated to provide accurate distances.[5] By fitting SPL models of the form:

$$\rho(r_e) = r_e^{-\alpha}, \qquad (11)$$

these groups found an exponent of $\alpha \sim 3$ and find a flattening requirement (given by equation 2 when equating $p = 1$) to fit star counts well. A pivotal work by Deason et al. (2011) showed that BHB stars in SDSS could be best fit by introducing a break radius at $r_b \sim 27$ kpc in their exponential profile, thus fitting a broken exponential, for which the inner slope is shallow $\alpha_{\text{in}} \sim 2.3$, and the outer slope is steep, $\alpha_{\text{out}} \sim 4.6$. Along those lines, Whitten et al. (2019) used a BHB sample from a cross-match of Pan-STARRS and GALEX data, and studied their relative age distribution across the Galactic halo. The results from this work showed that the BHB relative age distribution of the Milky Way is also best modelled by a broken profile, finding an older population within the ($R_{\text{GC}} \sim 14$ kpc) break radius. Another crucial study by Xue et al. (2015) used SDSS-SEGUE giants to show that such a profile can be accommodated with an SPL of steep exponent $\alpha \sim 4.2$, with flattening parameter q varying as a function of Galactocentric distance between $q \sim 0.5$ and 0.8. Later, Iorio et al. (2018) found that a similar profile provided a good fit to a sample of RR Lyrae stars within a Galactocentric distance of $\sim$30 kpc, such that $\alpha \sim 2.96$ with $0.57 < q < 0.75$. A recent study found that the APOGEE DR14 giant data were best fit by a similar SPL profile, however, with the inclusion of a disc and cut-off term (Mackereth & Bovy 2020), such as the model employed to fit the halo sample in this study. The power-law slope obtained for this model was found to be steep $\alpha \sim 3.5$, presenting a flattened triaxial ellipsoid, $p \sim 0.73$ and $q \sim 0.56$, slightly rotated in the Z- and Y-axes.

The behaviour resulting from the fit to the APOGEE data is similar to work from the literature. Fitting a likely halo sample of stars with metallicities between $-2.5 < $ [Fe/H] < -1, we find the data are well matched by a triaxial profile with a moderately steep exponent $\alpha \sim 3.5$, flattened with $p \sim 0.8$ and $q \sim 0.5$. We also find the triaxial ellipsoid to be minorly rotated to the Z- and Y- axes, by $\theta \sim 3$–4 deg and $\phi \sim 14$ deg, respectively. Finally, since our data are all contained within a $\sim$20 kpc radius, the issue of flattening beyond the break radius found in other studies is irrelevant to this work. Thus, we can conclude that the resulting profile and its associated parameters values are consistent with the results found in previous studies.

5.2 The contribution of GCs to the stellar halo of the Galaxy

5.2.1 Mass ratio of stars arising from GC dissolution and/or evaporation

Employing the method described in Section 3, we have assessed the fraction of the halo stellar mass contributed by N-rich stars as a function of Galactocentric distance between 1.5 kpc $< r <$ 15 kpc. We have shown that N-rich stars contribute a $\zeta = 16.8^{+10.0}_{-7.0}$ per cent to the halo field at $\sim$1.5 kpc from the Galactic Centre (see Fig. 8), resulting in a much larger contribution when compared to the outer regions of the Galaxy (i.e. $r \sim 10$ kpc) for which we find a ratio of $\zeta = 2.7^{+1.0}_{-0.8}$ per cent. Under the assumption of the minimal scenario proposed by Schiavon et al. (2017), whereby the ratio of FP-to-SP GC star is assumed to be 2 SP stars for every 1 FP star, we estimate the total contribution from disrupted star clusters $\sim$1.5 kpc

---

[5]Listed in this subsection are only a few of the more recent examples of an extensive list of studies focused on estimating the density of the stellar halo. For a more comprehensive account of previous work, see Helmi (2008).

from the Galactic Centre to be $\zeta_{\text{tot}} = 27.5^{+15.4}_{-11.5}$ per cent. Along those lines, we estimate the contribution from GC dissolved stars to the stellar halo at $r = 10$ kpc to be $\zeta_{\text{tot}} = 4.2^{+1.5}_{-1.3}$ per cent. Within the uncertainties, our estimate for the contribution by stars that once belonged to GCs to the stellar halo mass budget at Galactocentric distances of $r = 10$ kpc is in agreement with the low bound of the 2–5 per cent estimate from the literature for the halo of the Galaxy (Martell et al. 2016; Koch et al. 2019; Reina-Campos et al. 2019). Similarly, our estimate for the contribution at $r = 1.5$ kpc from the Galactic Centre is in agreement with the 19–24 per cent estimate from Schiavon et al. (2017), but is only partly in agreement with the theoretical predictions from Hughes et al. (2020). Specifically, within the uncertainties, our estimate falls in the upper range of the predictions provided by the E-MOSAICS suite of simulations, which predict a contribution ranging between 0.3 and 14 per cent for Milky Way analogues, depending on a galaxy's accretion history.

Interestingly, the largest fractional contribution to the stellar mass budget by disrupted GCs is attained in the simulations by Hughes et al. (2020) by galaxies that underwent a phase of intense accretion in their early lives. Mackereth et al. (2018), analysing EAGLE simulations (Crain et al. 2015; Schaye et al. 2015) established a link between those types of accretion histories with the (rare) presence of an α bimodality in the simulated galaxies' disc populations. According to Hughes et al. (2020), the physical basis for this connection is the high gas pressure brought about by intense accretion which, on one hand lowers the gas consumption time-scale leading up to formation of an α-enhanced population, and on the other hand creates an environment that is hostile to GC survival.

In this context, it is important to notice that Kisku et al. (in preparation) studied the abundance patterns of N-rich stars in the inner Galaxy, suggesting that roughly 1/2 of the N-rich stars with [Fe/H] < -1 were actually accreted, and were likely associated with the progenitor of the inner Galaxy structure (i.e. IGS), identified by Horta et al. (2020b).

5.2.2 Mass in dissolved GCs

As mentioned in Section 4.3, our methodology enables us to determine the mass for both the stellar halo and for N-rich stars within a sphere of any given radius. Under the minimal scenario assumption (see Section 4.3), we estimate a total stellar mass arising from GC remnants, within a shell ranging from 1.5 to 15 kpc in radius around the Galactic Centre, to be $M_{\text{GC,total}} = 9.6^{+4.0}_{-2.6} \times 10^7$ M$_\odot$. Our estimates confirm theoretical predictions by different groups ($\sim 10^7$–10^8 M$_\odot$; Tremaine et al. 1975; Gnedin et al. 2014). Furthermore, we find that the estimated mass from stars arising from GC disruption and/or evaporation to be $\sim$3–4 times greater than the total mass in all existing Galactic GCs ($\sim 2.8 \times 10^7$ M$_\odot$; Kruijssen & Portegies Zwart 2009). This result would imply that the Galactic GC system was initially 4–5 times larger, where approximately only one fourth/fifth survived, resulting in the $\sim$150 GCs observed today. However, due to the metallicity cuts employed in this study, our estimated mass is biased low, since it does not include the mass from stars arising from disruption and/or evaporation of GCs with [Fe/H] > -1. Moreover, our estimate only accounts for GCs that were massive enough to develop multiple populations (MP). It is likely that GCs were formed with masses below that threshold that nevertheless were destroyed, contributing only FP stars to the field population. Those of course are not accounted for by our chemical tagging, so that our estimated contribution of GCs to the stellar mass budget of halo field populations should be taken as a lower limit.

By the same token, since our results show that there is a much higher contribution of dissolved GC stars in the inner galaxy, we estimated the total stellar mass arising from GC disruption, within a hollow sphere of 1.5–3 kpc in radius around the Galactic Centre, to be $M_{\text{GC,inner}} = 5.6^{+2.8}_{-1.8} \times 10^7$ M$_\odot$. We find our estimate to be greater than the predicted $\sim 10^7$ M$_\odot$ from Gnedin et al. (2014). Similarly, within the uncertainties, we find our result to be slightly smaller than the estimate given in Schiavon et al. (2017), who found that the mass in destroyed GCs within 2 kpc to be a factor of 6–8 higher than the mass of existing GCs. The difference is due to three factors: (1) the different volumes included in the calculation; (2) the different models for the inner halo adopted; and (3) the different metallicity ranges considered.

5.3 Supporting evidence for the GC origin of N-rich stars

Despite the growing evidence for the GC origin of N-rich stars, there have been many alternate scenarios proposed for the formation of such population. Such alternative formation channels range from the notion that N-rich stars are the oldest stellar population in the Galaxy which formed in high-density environments (Bekki 2019), to the idea that N-rich stars were formed in the same molecular clouds that GCs were formed in, however, were never gravitationally bound to them. In this subsection, we discuss the results from previous studies which have shown, using measurements other than chemical compositions, that N-rich stars are likely formed from the dissolution and/or evaporation of GCs.

The availability of the 6D phase space information provided by the *Gaia* survey has made possible the estimation of the integrals of motion (hereafter IoM) of stars in the Milky Way. Such properties are adiabatic invariants, and thus retain some information about the origin of their parent population. A recent study by Savino & Posti (2019) determined the IoM for a sample of 65 CN-strong stars, from the work by Martell & Grebel (2010), as well as for the Galactic GC system. In their study, a direct comparison of the IoM values, as well as the metallicity values, was performed for every N-rich-GC pair, associating a likelihood of these being from the same distribution. The results from that study showed that a considerable fraction of CN-strong stars display the same IoM values as existing GCs, thus supporting the notion that N-rich stars arise from GC dissolution and/or evaporation.

A similar chemodynamical analysis of halo stars and Galactic GCs was performed by Hanke et al. (2020). In that study, three separate methodologies were employed to attempt to link halo stars to existing GCs. These ranged from statistically linking halo stars positioned nearby existing Galactic GCs, to linking CN-strong stars (i.e. N-rich stars) kinematically to existing GCs, to attempting to find halo stars that displayed similar kinematics to those CN-strong stars. The authors showed that over 60 per cent of their N-rich star population presented IoM which could be statistically linked to known GCs, and that around $\sim$150 halo stars could be associated with a GC origin. Separately, a study by Tang et al. (2020) performed a comparison between the IoM of $\sim$100 CN-strong (metal poor) stars from the LAMOST DR5 data set with metal poor halo stars. The authors from this study concluded that the CN-strong stars do not present similar kinematics to the halo field population, but resembled that of the inner halo where there is a high density of GCs with similar chemical compositions. Based on their kinematic results, the authors supported a GC origin for the CN-strong stars. In a final remark, we find the findings of Tang et al. (2020) to be in agreement with the results from Carollo et al. (2013). Using an independent sample CN-strong stars, these authors found that these follow the velocity

distribution of the 'inner halo population' (IHP, as defined in their work).

The kinematic results from the aforementioned work, when coupled with the results from chemical composition studies (e.g. Martell & Grebel 2010; Martell et al. 2016; Schiavon et al. 2017; Koch et al. 2019; Lee et al. 2019), support the hypothesis that N-rich stars arise from the disruption and/or evaporation of GCs.

6 SUMMARY AND CONCLUSIONS

In this paper, we report a study of the spatial variation of the density of N-rich and normal halo field stars. In this way, we assessed the contribution of N-rich stars, and by inference that of dissolved GCs, to the stellar mass budget of the Galactic halo, as well as its variation as a function of Galactocentric distance. A summary of our results is listed as follows:

(i) We identified in a parent sample of 1455 halo stars with $-2.5 <$ [Fe/H] < -1, using a GMM, 46 N-rich stars that are distributed throughout the Galaxy (see Figs 1 and 4), and are not bound to existing GCs. The N-rich stars present an N–C abundance anticorrelation and an N–Al abundance correlation (see Figs 2 and 3, respectively), and are likely SP stars that once were bound to a GC (which may or may not still exist) and now reside in the halo field of the Galaxy.

(ii) We show that once the survey selection effects are accounted for (see appendix A in Mackereth et al. 2019 for details), the halo APOGEE data between $-2.5 <$ [Fe/H] < -1 are well fit by a triaxial single power law with exponent $\alpha \sim 3.5$, and flattening parameters $p \sim 0.8$, $q \sim 0.5$, with major axis only slightly misaligned with the axis connecting the Sun and the Galactic Centre.

(iii) Similarly, we show that the APOGEE N-rich star data between $-2.5 <$ [Fe/H] < -1 are well fit by an axisymmetric profile, defined by an SPL of slope $\alpha \sim 4.5$. The best-fitting model is flattened along the Z-axis, with a flattening value similar to that of the halo sample $q \sim 0.5$ (see Fig. 6).

(iv) We find a contribution of N-rich stars to the stellar halo of $\zeta = 16.8^{+10.0}_{-7.0}$ per cent at $r = 1.5$ kpc. However, this ratio drops by a factor of ~ 6 at large Galactocentric distances ($r = 10$ kpc), where we estimate a contribution of $\zeta = 2.7^{+1.0}_{-0.8}$ per cent.

(v) Assuming that the ratio of first population-to-second population stars in GCs is 1-to-2, we find that stars arising from GC disruption contribute $\zeta_{\mathrm{tot}} = 27.5^{+15.4}_{-11.5}$ per cent to the mass of the stellar halo at ~ 1.5 kpc from the Galactic Centre. Conversely, the contribution of GC escapees at larger Galactocentric distances (i.e. $r = 10$ kpc) is lower, namely $\zeta_{\mathrm{tot}} = 4.2^{+1.5}_{-1.3}$ per cent. Such estimates are in agreement (within the uncertainties) with previous estimates from the literature for the inner Galaxy (e.g. Schiavon et al. 2017) and the outer Galactic halo (e.g. Martell et al. 2016; Koch et al. 2019).

(vi) We integrate the total stellar mass between $-2.5 <$ [Fe/H] < -1 and 1.5 kpc $< r <$ 15 kpc using equation (8) and our best-fitting halo profile, and estimate the mass of the stellar halo to be $M_{*,\mathrm{halo}} = 8.3^{+1.5}_{-1.3} \times 10^8$ M⊙.

(vii) We integrate the total stellar mass between $-2.5 <$ [Fe/H] < -1 and 1.5 kpc $< r <$ 15 kpc using equation (8) and our best-fitting N-rich model, and estimate the mass from N-rich stars to be $M_{*,\mathrm{N\text{-}rich}} = 6.4^{+2.6}_{-1.7} \times 10^7$ M⊙.

(viii) Using the same technique, and under the assumption of a 1:2 first:second population GC ratio, we estimate the total stellar mass between $-2.5 <$ [Fe/H] < -1 and 1.5 kpc $< r <$ 15 kpc arising from GC dissolution and/or evaporation to be $M_{\mathrm{GC,total}} = 9.6^{+4.0}_{-2.6} \times$

10^7 M⊙. For a spherical volume ranging from 1.5 to 3 kpc, we obtain an estimated mass of $M_{\mathrm{GC,inner}} = 5.6^{+2.8}_{-1.8} \times 10^7$ M⊙. This total dissolved/evaporated GC mass is approximately 3–4 times greater than the total mass in all existing Galactic GCs (namely, $\sim 2.8 \times 10^7$ M⊙; Kruijssen & Portegies Zwart 2009).

(ix) We speculate that the increased contribution of GC dissolution towards the inner regions of the Galaxy may be associated with enhanced merger activity in the early life of the Milky Way. Some of these merging systems (e.g. IGS Horta et al. 2020b) may have brought with them a population of extragalactic N-rich stars (Kisku et al., in preparation). In addition, the enhanced merging activity in the early life of the Milky Way may also have given rise to conditions that led to efficient *in situ* formation and destruction of GCs, leaving behind a large population of N-rich stars in the inner Galaxy field.

In this paper, we have mapped the spatial distribution of N-rich stars in the Galactic halo, determining their contribution to the stellar mass budget as a function of Galactocentric distance.

Our study of the contribution of N-rich stars, commonly recognized as SP GC stars residing in the field, to the halo field constrains the mass contribution from GC disruption and/or evaporation to the total mass of the stellar halo, but also provides insights into the mass assembly history of the Milky Way. The order of magnitude increase in halo samples afforded by upcoming surveys such as WEAVE (Dalton et al. 2014) and 4-MOST (de Jong et al. 2019) will place key constraints on our understanding of the mass assembly of the Milky Way halo, and the role played by GCs in this process.

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

Software: NUMPY (Oliphant 06), MATPLOTLIB (Hunter 2007).

DATA AVAILABILITY

All APOGEE DR16 data used in this study are publicly available and can be found at: https://www.sdss.org/dr16/

REFERENCES

Ahumada R. et al., 2019, preprint (arXiv:1912.02905)
Bastian N., Lardo C., 2015, MNRAS, 453, 357
Bastian N., Lardo C., 2018, ARA&A, 56, 83
Bekki K., 2019, MNRAS, 490, 4007
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Blanton M. R. et al., 2017, AJ, 154, 28
Bovy J., Hogg D. W., Roweis S. T., 2011, Ann. Appl. Stat., 5, 1657
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. 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
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Brodie J. P., Strader J., 2006, ARA&A, 44, 193
Cabrera-Ziri I. et al., 2015, MNRAS, 448, 2224
Carollo D., Martell S. L., Beers T. C., Freeman K. C., 2013, ApJ, 769, 87
Carretta E. et al., 2009, A&A, 505, 117
Churchwell E. et al., 2009, PASP, 121, 213
Crain R. A. et al., 2015, MNRAS, 450, 1937
Cunha K. et al., 2017, ApJ, 844, 145
Daley D., Vere-Jones D., 2003, An Introduction to the Theory of Point Processes Vol. I: Elementary Theory and Methods, Springer, New York
Dalton G. et al., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9147, Ground-Based and Airborne Instrumentation for Astronomy V. SPIE, Bellingham, p. 91470L
de Jong R. S. et al., 2019, Messenger, 175, 3
Deason A. J., Belokurov V., Evans N. W., 2011, MNRAS, 416, 2903
Deason A. J., Belokurov V., Sanders J. L., 2019, MNRAS, 490, 3426
Einasto J., 1965, Tr. Astrofiz. Inst. Alma-Ata, 5, 87
Elmegreen B. G., 2010, ApJ, 712, L184
Fernández-Trincado J. G., Beers T. C., Tang B., Moreno E., Pérez-Villegas A., Ortigoza-Urdaneta M., 2019, MNRAS, 488, 2864
Forbes D. A., 2020, MNRAS, 493, 847
Forbes D. A., Bridges T., 2010, MNRAS, 404, 1203
Frenk C. S., White S. D. M., 2012, Ann. Phys., Lpz., 524, 507
Gaia Collaboration, 2018, A&A, 616, A1
García Pérez A. E. et al., 2016, AJ, 151, 144
Gnedin O. Y., 2001, Astron. Astrophys. Trans., 20, 39
Gnedin O. Y., Ostriker J. P., Tremaine S., 2014, ApJ, 785, 71
Green G. M., Schlafly E. F., Zucker C., Speagle J. S., Finkbeiner D. P., 2019, ApJ, 887, 93
Gunn J. E. et al., 2006, AJ, 131, 2332
Hanke M., Koch A., Prudil Z., Grebel E. K., Bastian U., 2020, A&A, 637, A98
Helmi A., 2008, A&AR, 15, 145
Holoien T. W. S., Marshall P. J., Wechsler R. H., 2017, AJ, 153, 249
Holtzman J. A. et al., 2015, AJ, 150, 148
Holtzman J. A. et al., 2018, AJ, 156, 125
Horta D. et al., 2020a, MNRAS, 493, 3363
Horta D. et al., 2020b, MNRAS, 500, 1385
Hughes M. E., Pfeffer J. L., Martig M., Reina-Campos M., Bastian N., Crain R. A., Kruijssen J. M. D., 2020, MNRAS, 491, 4012
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Iorio G., Belokurov V., 2019, MNRAS, 482, 3868
Iorio G., Belokurov V., Erkal D., Koposov S. E., Nipoti C., Fraternali F., 2018, MNRAS, 474, 2142
Jönsson H. et al., 2020, AJ, 160, 120
Jurić M. et al., 2008, ApJ, 673, 864
Koch A., Grebel E. K., Martell S. L., 2019, A&A, 625, A75
Koppelman H. H., Helmi A., Massari D., Roelenga S., Bastian U., 2019, A&A, 625, A5
Kroupa P., 2001, MNRAS, 322, 231
Kruijssen J. M. D., 2014, Class. Quantum Gravity, 31, 244006
Kruijssen J. M. D., 2015, MNRAS, 454, 1658
Kruijssen J. M. D., Portegies Zwart S. F., 2009, ApJ, 698, L158
Kruijssen J. M. D., Pelupessy F. I., Lamers H. J. G. L. M., Portegies Zwart S. F., Icke V., 2011, MNRAS, 414, 1339
Kruijssen J. M. D., Maschberger T., Moeckel N., Clarke C. J., Bastian N., Bonnell I. A., 2012, MNRAS, 419, 841
Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., 2019a, MNRAS, 486, 3134
Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2019b, MNRAS, 486, 3180
Leaman R., VandenBerg D. A., Mendel J. T., 2013, MNRAS, 436, 122
Lee Y.-W., Kim J. J., Johnson C. I., Chung C., Jang S., Lim D., Kang Y., 2019, ApJ, 878, L2
Leung H. W., Bovy J., 2019a, MNRAS, 489, 2079
Leung H. W., Bovy J., 2019b, MNRAS, 483, 3255
Mackereth J. T., Bovy J., 2020, MNRAS, 492, 3631
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
Mackereth J. T. et al., 2019, MNRAS, 482, 3426
Majewski S. R., Zasowski G., Nidever D. L., 2011, ApJ, 739, 25
Majewski S. R. et al., 2017, AJ, 154, 94
Marigo P. et al., 2017, ApJ, 835, 77
Marín-Franch A. et al., 2009, ApJ, 694, 1498
Marshall D. J., Robin A. C., Reylé C., Schultheis M., Picaud S., 2006, A&A, 453, 635
Martell S. L., Grebel E. K., 2010, A&A, 519, A14
Martell S. L. et al., 2016, ApJ, 825, 146
Massari D., Koppelman H. H., Helmi A., 2019, A&A, 630, L4
Mészáros S. et al., 2020, MNRAS, 492, 1641

Morrison H. L., Mateo M., Olszewski E. W., Harding P., Dohm-Palmer R. C., Freeman K. C., Norris J. E., Morita M., 2000, AJ, 119, 2254
Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235
Nataf D. M. et al., 2019, AJ, 158, 14
Nidever D. L. et al., 2015, AJ, 150, 173
Oliphant T., 2006, NumPy: A Guide to NumPy. Trelgol Publishing, USA
Pfeffer J., Kruijssen J. M. D., Crain R. A., Bastian N., 2018, MNRAS, 475, 4309
Reina-Campos M., Hughes M. E., Kruijssen J. M. D., Pfeffer J. L., Bastian N., Crain R. A., Koch A., Grebel E. K., 2019, MNRAS, 493, 3422
Savino A., Posti L., 2019, A&A, 624, L9
Schaeffer R., Silk J., 1985, ApJ, 292, 319
Schaye J. et al., 2015, MNRAS, 446, 521
Schiavon R. P. et al., 2017, MNRAS, 465, 501
Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
Schwarz G., 1978, Ann. Stat., 6, 461
Searle L., Zinn R., 1978, ApJ, 225, 357
Shetrone M. et al., 2015, ApJS, 221, 24
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Tang B., Fernández-Trincado J. G., Liu C., Yu J., Yan H., Gao Q., Shi J., Geisler D., 2020, ApJ, 891, 28
Tremaine S. D., Ostriker J. P., Spitzer L., 1975, ApJ, 196, 407
Whitten D. D. et al., 2019, ApJ, 884, 67
Wilson J. C. et al., 2019, PASP, 131, 055001
Wright E. L. et al., 2010, AJ, 140, 1868
Xue X.-X., Rix H.-W., Ma Z., Morrison H., Bovy J., Sesar B., Janesh W., 2015, ApJ, 809, 144
Yanny B. et al., 2000, ApJ, 540, 825
Yanny B. et al., 2009, AJ, 137, 4377
Yuan H. B., Liu X. W., Xiang M. S., 2013, MNRAS, 430, 2188
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: FITTING THE N-RICH STAR SAMPLE

We set out to determine the best-fitting model to the sample of N-rich stars. Our methodology consists on fitting several different stellar halo density models and determining their logarithmic maximum likelihood ($\ln(\mathcal{L}_{max})$) and BIC (Schwarz 1978) values (see Section 3.3). We then compare the values obtained for each density model, and take the profile with the lowest BIC value to be our best-fitting model. If two models obtain the same $\ln(\mathcal{L}_{max})$ value, the BIC value gives preference to the model with the lowest number of free ranging parameters, therefore choosing the simplest model.

As described in Section 3.2, our N-rich sample is fit by an SPL, an AXI, a TRI, a rotated triaxial power law with the inclusion of a cut-off term and a disc contamination parameter (TRI-CUT-DISC), a BPL, and an exponential disc profile (DISC). The resulting BIC and $\ln(\mathcal{L}_{max})$ values obtained using the median posterior parameter of the 10 000 MCMC realizations for each model are listed in Table 1, whilst the resulting median and $1-\sigma$ uncertainties for the best fit parameters to each model are displayed in Table A1. From our results, we find that the best-fitting model to our N-rich sample is the AXI model, and therefore choose to use this model for the remainder of the analysis. However, as is apparent from Fig. A1, the TRI and TRI-CUT-DISC density profiles also provide good fits to the N-rich star data. Moreover, we repeat the methodology for determining the percentage ratio contribution of N-rich stars to the halo field employed in the main body of the paper using the TRI and TRI-CUT-DISC profile as the best-fitting N-rich models, in order to check if our results vary when adopting a different model for the N-rich star sample. The results from this comparison show that our initial findings remain unchanged, and suggest that the N-rich star sample can either be well modelled by the AXI, TRI, and TRI-CUT-DISC profiles.
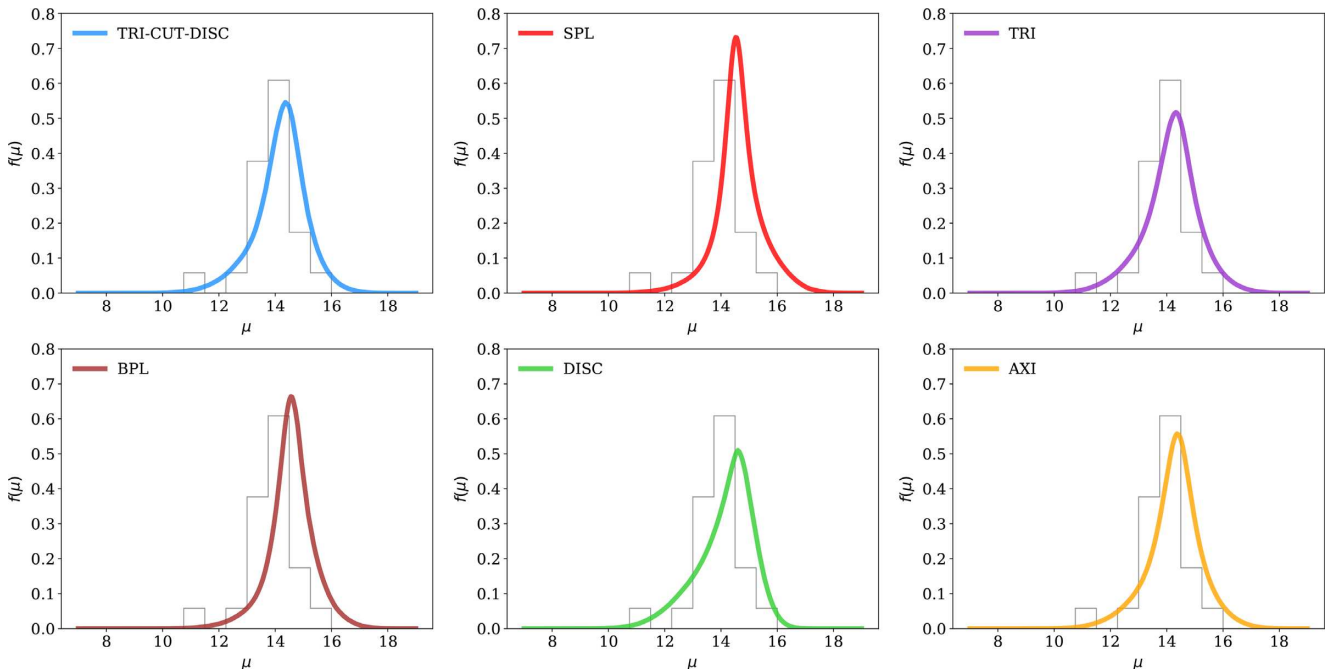

Figure A1. Density profile fits to the N-rich star data from Section 2. Each profile is obtained by taking the median posterior parameter value obtained from the 10 000 MCMC samples. The AXI profile is the best-fitting profile, closely followed by the TRI and TRI-CUT-DISC profile.

Table A1. Resulting best-fitting parameters for the different functional density forms fitted to the N-rich star sample. The best-fitting model (AXI) is highlighted in bold.

Density profile	Best-fitting parameters
SPL	$\alpha = 3.86^{+0.21}_{-0.18}$
AXI	$\boldsymbol{\alpha = 4.47^{+0.23}_{-0.22}, q = 0.47^{+0.05}_{-0.04}}$
TRI	$\alpha = 4.54^{+0.26}_{-0.24}, p = 0.85^{+0.09}_{-0.10}, q = 0.44^{+0.05}_{-0.05}$
TRI-CUT-DISC	$\alpha = 4.30^{+0.12}_{-0.31}, \beta = 0.03^{+0.03}_{-0.02}, p = 0.60^{+0.21}_{-0.09}, q = 0.41^{+0.11}_{-0.07}, \eta = 0.20^{+0.03}_{-0.04}, \theta = 0.06^{+0.2}_{-0.06}\ (°), \phi = 0.25^{+0.07}_{-2.38}\ (°),$ $f_{\text{disc}} = 0.04^{+0.11}_{-0.03}$
BPL	$\alpha_{\text{in}} = 3.19^{+0.52}_{-0.46}, \alpha_{\text{out}} = 6.05^{+1.25}_{-1.37}, R_{\text{break}} = 8.13^{+2.03}_{-1.82}\ (\text{kpc})$
DISC	$1/h_{\text{R}} = 0.41^{+0.04}_{-0.04}\ (\text{kpc}^{-1}), 1/h_{\text{Z}} = 0.64^{+0.09}_{-0.08}\ (\text{kpc}^{-1})$

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