The Pristine Survey – VIII. The metallicity distribution function of the Milky Way halo down to the extremely metal-poor regime

K Youakim, E Starkenburg, Nicolas F. Martin, G Matijevič, D Aguado, C Allende prieto, A Arentsen, Piercarlo Bonifacio, R. Carlberg, J González hernández, et al.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **492,** 4986–5002 (2020) doi:10.1093/mnras/stz3619
Advance Access publication 2020 January 16

The *Pristine* Survey – VIII. The metallicity distribution function of the Milky Way halo down to the extremely metal-poor regime

K. Youakim [1]★ E. Starkenburg,[1] N. F. Martin,[2] G. Matijevič,[1] D. S. Aguado,[3] C. Allende Prieto,[4,5] A. Arentsen [1], P. Bonifacio [6], R. G. Carlberg,[7] J. I. González Hernández,[4,5] V. Hill,[8] G. Kordopatis,[8] C. Lardo [9], J. F. Navarro,[10] P. Jablonka,[6,9] R. Sánchez Janssen [11], F. Sestito,[1,2] G. F. Thomas[12] and K. Venn[10]

[1]*Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, Potsdam D-14482, Germany*
[2]*Observatoire astronomique de Strasbourg, Université de Strasbourg, 11 rue de l'Université, Strasbourg F-67000, France*
[3]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[4]*Instituto de Astrofísica de Canarias, Vía Láctea, La Laguna E-38205, Tenerife, Spain*
[5]*Departamento de Astrofísica, Universidad de La Laguna, La Laguna E-38206, Tenerife, Spain*
[6]*GEPI, Observatoire de Paris, Université PSL, CNRS, Place Jules Janssen, Meudon F-92190, France*
[7]*Department of Astronomy & Astrophysics, University of Toronto, Toronto, ON M5S 3H4, Canada*
[8]*Observatoire de la Côte d'Azur, CNRS, Université Côte d'Azur, Lagrange, Bd de l'Observatoire, CS34229, Nice F-06304, France*
[9]*Laboratoire d'astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire, Institute of Physics, Versoix CH-1290, Switzerland*
[10]*Department of Physics and Astronomy, University of Victoria, P.O. Box 3055, STN CSC, Victoria BC V8W 3P6, Canada*
[11]*Royal Observatory Edinburgh, UK Astronomy Technology Centre, Blackford Hill, Edinburgh EH9 3HJ, UK*
[12]*NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Road, Victoria, BC, V9E 2E7, Canada*

Accepted 2019 December 10. in original form 2019 October 1

ABSTRACT

The *Pristine* survey uses narrow-band photometry to derive precise metallicities down to the extremely metal-poor regime ([Fe/H] < -3), and currently consists of over 4 million FGK-type stars over a sky area of $\sim 2500\,\mathrm{deg}^2$. We focus our analysis on a subsample of ∼80 000 main-sequence turn-off stars with heliocentric distances between 6 and 20 kpc, which we take to be a representative sample of the inner halo. The resulting metallicity distribution function (MDF) has a peak at [Fe/H] $= -1.6$, and a slope of $\Delta(\mathrm{LogN})/\Delta[\mathrm{Fe/H}] = 1.0 \pm 0.1$ in the metallicity range of $-3.4 <$ [Fe/H] < -2.5. This agrees well with a simple closed-box chemical enrichment model in this range, but is shallower than previous spectroscopic MDFs presented in the literature, suggesting that there may be a larger proportion of metal-poor stars in the inner halo than previously reported. We identify the Monoceros/TriAnd/ACS/EBS/A13 structure in metallicity space in a low-latitude field in the anticentre direction, and also discuss the possibility that the inner halo is dominated by a single, large merger event, but cannot strongly support or refute this idea with the current data. Finally, based on the MDF of field stars, we estimate the number of expected metal-poor globular clusters in the Milky Way halo to be 5.4 for [Fe/H] < -2.5 and 1.5 for [Fe/H] < -3, suggesting that the lack of low-metallicity globular clusters in the Milky Way is not due simply to statistical undersampling.

Key words: stars: abundances – Galaxy: abundances – Galaxy: evolution – Galaxy: formation – Galaxy: halo – galaxies: dwarf.

1 INTRODUCTION

The extended stellar halo of our Galaxy is diffuse and full of stellar substructures, continually being discovered as overdensities and kinematical and/or chemical substructures by ongoing large surveys, especially now in the era of the *Gaia* mission (e.g. Belokurov et al. 2018; Haywood et al. 2018; Helmi et al. 2018; Koppelman, Helmi & Veljanoski 2018; Malhan, Ibata & Martin 2018; Myeong et al. 2018). Together, they relate a complex history in which the halo is built up through many merger events of smaller systems. While the majority and largest events are thought to have happened in the distant past, some are still continuing in the present day (with the most obvious example being the Sagittarius dwarf

★ E-mail: kyouakim@aip.de

galaxy; Ibata, Gilmore & Irwin 1994). It is a complicated task to characterize such a complex structure, especially taking into account our vantage point of observing the halo from within the stellar disc. We know that the halo is highly structured, that its average metallicity is much lower than that of the disc(s) or inner Galaxy, and that its stellar population has a very low net rotation (see for reviews on the stellar halo Helmi 2008; Belokurov 2013; Bland-Hawthorn & Gerhard 2016). It has further been suggested that the character of the stellar halo is dual in nature, consisting of two broadly overlapping components in density, orbits, and chemistry (Carollo et al. 2007; Beers et al. 2012), but this view has also been challenged (Schönrich, Asplund & Casagrande 2011, 2014) and is complicated further by the discovery of large substructures in the *Gaia* DR2 data, standing out in their (partly retrograde) orbits and chemistry in different ways (e.g. Belokurov et al. 2018; Haywood et al. 2018; Helmi et al. 2018; Myeong et al. 2019). Furthermore, simulations of merging galaxy systems have shown that in some cases a significant part of the inner halo can consist of *in situ* material from the original progenitor (as suggested by Abadi, Navarro & Steinmetz 2006), rather than accreted material from merging satellites.

Robust observational results on the metallicity structure of the Galactic halo can help us understand not only its structure – as stellar substructures often stand out in metallicity space – but also the processes involved in its formation. The extremely low-metallicity tail of the metallicity distribution function (MDF), for instance, constrains the pace of chemical evolution at its earliest epochs (for a review on this topic see Frebel & Norris 2015). Metal-rich halo stars, on the other hand, are more likely to be deposited by larger rather than smaller galaxies (or building blocks). Simulations in a Λ-CDM framework demonstrate that the median metallicity of a stellar halo is primarily determined by the mass of the most massive building block (Robertson et al. 2005; Deason, Mao & Wechsler 2016; Oman, Starkenburg & Navarro 2017; D'Souza & Bell 2018b). An interesting pair of galaxies in this respect is the Milky Way and its close neighbour Andromeda. While the Milky Way has many features in common with our neighbouring galaxy Andromeda, one of the more striking differences is the metallicity of their stellar haloes, namely that the Andromeda stellar halo is significantly more metal-rich out to large radii (e.g. Ibata et al. 2014). This could be linked to the presence of stellar substructure on vast scales throughout its large stellar halo and to M32 and its Giant Southern Stream (e.g. McConnachie et al. 2018; D'Souza & Bell 2018a). Additionally, a stellar halo mass–metallicity relation is also seen in stellar haloes of galaxies outside the Local Group (Harmsen et al. 2017).

To understand the merging history of the Galaxy in detail, it is thus of clear importance to measure the MDF of the Milky Way stellar halo at different radii and through different lines of sight, minimizing selection biases as much as possible. Such measurements will also help to understand the globular cluster (GC) population with respect to the field stars, and constrain chemical evolution in the early Milky Way. Unfortunately, the metal-poor end of the MDF in the Galactic halo remains especially poorly constrained. Several previous studies have used large spectroscopic surveys for this purpose, selecting a halo sample either kinematically to build up a local halo sample, by metallicity to study the metal-poor tail, or by stellar type and magnitude to study the halo at different distances (e.g. Ryan & Norris 1991; Schörck et al. 2009; Allende Prieto et al. 2014). These different approaches result in very different distributions, especially in the very metal-poor regime. One difficulty is to understand all of the selection biases that affect the spectroscopic survey selection and correct for them appropriately. Especially further out and at lower metallicities, spectroscopic samples quickly become sparser. Additionally, distances are difficult to constrain – even in the era of the *Gaia* mission – as the Galactic halo stretches out far beyond the regime where parallax information is reliable.

Until recently, observational studies constraining the metal-poor tail of the halo MDF have mostly been done using spectroscopic samples since photometric metallicity calibrations have been limited to [Fe/H] > -2 (e.g. Ivezić et al. 2008). However, several recent studies have shown that it is possible to provide reliable photometric metallicities using SDSS broad-band photometry down to [Fe/H] $= -2.5$, if deeper *u*-band photometry is used (An et al. 2013, 2015; Ibata et al. 2017; Zuo et al. 2017). Subsequently, this metallicity limit has been pushed even further with the implementation of narrow-band filters targeting the Ca II H&K absorption lines in surveys such as *Pristine* and *SkyMapper* (Keller et al. 2007; Casagrande et al. 2019), which have shown that with this technique it is possible to derive metallicities down to the regime of [Fe/H] < -3 (Starkenburg et al. 2017; Da Costa et al. 2019).

In this paper, we present MDFs derived using data from the *Pristine* survey (Starkenburg et al. 2017). *Pristine* uses a narrow-band filter to obtain photometric metallicities for millions of stars over large regions of the northern Galactic halo. We convolve its results with spectroscopic follow-up information from thousands of stars to remove spurious features that may compromise the photometric metallicity determinations. Because of its photometric nature, the *Pristine* survey is (almost) completely unbiased in its selection of stars, although its metallicity determination is most reliable in the metal-poor regime ([Fe/H] < -1). We therefore focus our work on the derivation of the MDF at low metallicities for a sample of stars that has been selected to be consistent with halo ages and distances.

We describe the photometric survey, as well as the spectroscopic samples used to rescale the MDF in Section 2. In Section 3, we describe the selection of main-sequence turn-off (MSTO) stars and a halo sample, and show the resulting MDFs for different distance ranges. In Section 4, we apply corrections to the photometric MDF, to account for uncertainties in the photometric metallicities and metallicity biases induced by the colour cut used to select the halo sample. We investigate substructures seen in the MDF as a function of observational direction in Section 5. We discuss the potential remaining biases from the halo sample selection in Section 6, and compare the metal-poor halo MDF to those derived in other works. In this section, we also discuss the possibility of the inner halo being dominated by a single, large accretion event, the duality of the stellar halo, and the implications that the halo MDF has for the lack of metal-poor GCs in the Milky Way. In Section 7, we summarize and conclude the paper.

2 DESCRIPTION OF THE DATA

2.1 Photometric sample

The *Pristine* survey and the derivation of photometric metallicities are described in detail in Starkenburg et al. (2017). In summary, *Pristine* utilizes a narrow-band filter that is nearly top-hat in shape and centred on the Ca II H&K absorption lines at 3900−4000 Å. These absorption lines are some of the strongest features in a stellar spectrum and their size is directly dependent on the metallicity of the star. The coverage of the survey is $\sim$2500 deg^2 in the Northern hemisphere and goes to a maximum depth of $g_0 \sim 21.0$ at signal-to-noise ratio $= 10$, although to avoid differential depths of the

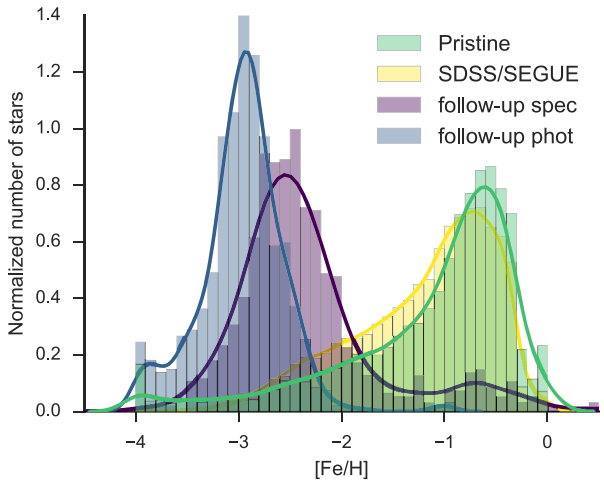

Figure 1. Normalized distributions of metallicities from *Pristine* and the spectroscopic samples. The distribution of the *Pristine* spectroscopic follow-up sample is shown with photometric metallicities (blue) and spectroscopic metallicities (purple), along with the distribution of *Pristine* photometric metallicities for the matched SDSS/SEGUE sample (yellow) and the distribution of the photometric metallicities for the whole *Pristine* sample (green).

observed fields and the resulting selection effects, we limit the sample to $g_0 < 20$ in this work. All detected sources from the *Pristine* survey are then matched to stars in the Sloan Digital Sky Survey (SDSS; York et al. 2000). The metallicity of each star is determined using a colour–colour space combining the narrow-band information and broad-band information to solve for the temperature of the star (in Starkenburg et al. 2017 and in this work we use SDSS broad-band colours, but a similar principle can be followed using *Gaia* G_{BP} and G_{RP} colours, as demonstrated in Bonifacio et al. 2019). Our method is calibrated through the use of overlapping SDSS/SEGUE spectra as well as some stars in the footprint that were additionally observed with high-resolution spectroscopy. Overall, we find that we can determine metallicities for F-, G-, and K-type stars with $4000\,\mathrm{K} < T_{\mathrm{eff}} < 7000\,\mathrm{K}$, $0.15 < (g_0 - i_0) < 1.5$) and $-3.0 <$ [Fe/H] < -0.5 with ~ 0.2 dex precision (Starkenburg et al. 2017). Spectroscopic follow-up has demonstrated that *Pristine* is also successful at identifying stars with [Fe/H] < -3.0 (Youakim et al. 2017; Aguado et al. 2019). The raw photometric MDF for all stars in the current *Pristine* footprint is shown as the green histogram in Fig. 1.

2.2 Spectroscopic sample

In addition to the photometric component, the *Pristine* survey includes a large spectroscopic follow-up campaign targeting metal-poor star candidates (Caffau et al. 2017; Youakim et al. 2017; Aguado et al. 2019; Bonifacio et al. 2019; Venn et al. 2020; Arentsen et al. 2020). Although the efficiency of the *Pristine* survey to discover extremely metal-poor stars is very good (for more details see Youakim et al. 2017; Aguado et al. 2019) and we take precautions to clean the sample of outliers such as variable stars (using Hernitschek et al. 2016) and white dwarfs (through additional colour cuts), the metal-poor tail of the MDF is still, to some degree, contaminated by more metal-rich stars. This is to be expected, as the metal-poor population is much less numerous than the more metal-rich stars so that stars of higher metallicity are more likely to scatter down to lower metallicities than the other way around, assuming similar uncertainties. In Aguado et al. (2019), we estimate this contamination to be ∼18 per cent for stars that are predicted to be extremely metal-poor ([Fe/H] < -3) with photometry, but that are found with spectroscopy to be metal-rich ([Fe/H] > -2). It might also be the case that some unidentified contamination remains, if stars misclassified by photometry are also misclassified by the medium-resolution follow-up spectroscopy, although these cases are likely to be rare. Fig. 1 shows the raw MDFs, where the blue and purple histograms show the spectroscopic follow-up sample of *Pristine* metal-poor candidates with their predicted photometric metallicities and derived spectroscopic metallicities, respectively. The purple histogram is shifted, showing that the spectroscopic metallicities are more metal-rich than the predicted photometry on average (this effect has been previously described in Starkenburg et al. 2017), and also has a tail that extends to high metallicities. For the purpose of building a representative MDF down to the lowest metallicities from photometry, it is important to take these misclassifications into account and correct for them.

3 SELECTION OF THE HALO SAMPLE

The *Pristine* survey is a photometric survey, and as such observes every star in the footprint that is in the right magnitude range (excluding holes caused by chip gaps and stars occluded by bright foreground stars, etc.). However, for each line of sight, the total sample will consist of a mixture of stars in different Galactic components, and of different evolutionary stages that probe a large variety of distances at the same magnitude. As it is our goal to produce an MDF of the Galactic halo, we need a way to separate out halo stars from the far more numerous nearby disc stars. This can in principle be done kinematically, chemically, or spatially. In this work, we choose to select a halo sample based primarily on distance from the Galactic plane. The selection of these distance samples and the Galactic environments that they probe are discussed in this section.

3.1 Distance slices into the Galaxy using MSTO stars

To obtain heliocentric distances for our sample, we make a colour selection to target MSTO stars. Fig. 2 shows isochrones of different metallicities and a fixed age of 12 Gyr, produced using the PAdova and TRieste Stellar Evolution Code (PARSEC; Bressan et al. 2012). We convert [M/H] to [Fe/H] for these isochrones assuming [α/Fe] $= +0.4$. The dotted purple line shows the isochrone with [Fe/H] $= -2.5$ and an age of 10 Gyr to show how the turn-off shifts as a function of age. We therefore make a colour selection such that we select the full MSTO from the tip of the turn-off up to the subgiant branch for the most metal-poor isochrone, giving a colour cut of $0.15 < (g_0 - i_0) < 0.6$, shown by the combined dark and light grey regions in Fig. 2. The absolute magnitude range spanned by this selection is shown by the black dotted lines, and corresponds to a heliocentric distance range of ∼4−25 kpc for a magnitude of $g_0 = 19.5$ (the middle of the faintest magnitude bin $19 < g_0 < 20$). By selecting different magnitude bins, we separate the Galaxy into different distance ranges, allowing us to target different Galactic environments.

3.2 The halo sample

Although the faintest MSTO sample selected with $0.15 < (g_0 - i_0) < 0.6$ and $19 < g_0 < 20$ is likely composed of mostly halo stars, it still probes a wide distance range of ∼4−25 kpc, and thus a wide range of Galactic environments. In order to narrow this range, we impose

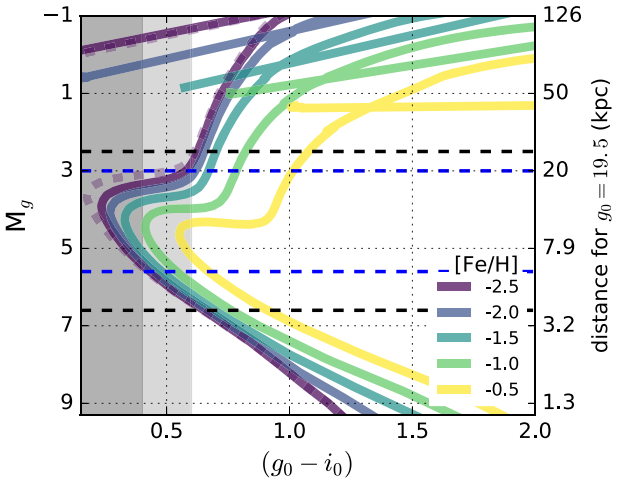

Figure 2. The MSTO colour selection and corresponding distances. PARSEC isochrones of decreasing metallicity and an age of 12 Gyr are plotted as the solid lines, and the [Fe/H] = −2.5 isochrone with an age of 10 Gyr is shown by the dotted line. The absolute SDSS g magnitude is shown on the left axis, and the corresponding distances at an apparent magnitude of $g_0 = 19.5$ on the right axis. The grey shaded region shows the colour cuts used to select the MSTO stars and the dark shaded grey region shows the narrower colour cut used to select the halo sample. The black and blue dashed lines show the corresponding magnitude ranges spanned by the MSTO and halo selections, respectively.

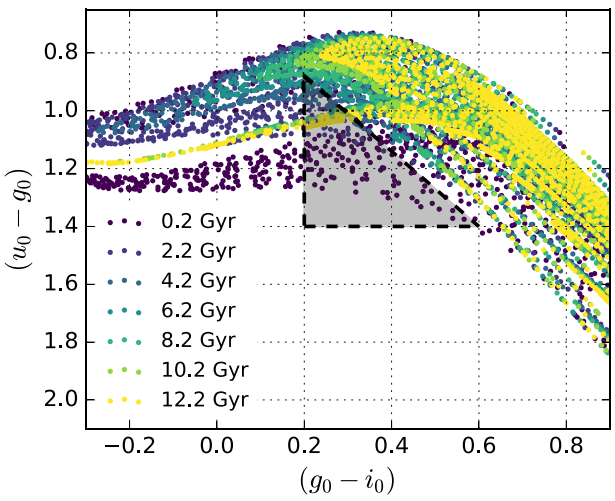

Figure 3. Colour–colour plot of PARSEC isochrones sampling ages of 0.2−12.2 Gyr and a metallicities of −2.5 < [Fe/H] < −0.5. The isochrones are coloured by age, but if instead they were coloured by metallicity, the metal-poor populations would be at the top of the plot and the metal-rich ones at the bottom. The grey-shaded triangle shows the region removed (see text for details), which targets young disc stars and old horizontal branch stars.

an additional colour cut at $0.15 < (g_0 - i_0) < 0.4$ (depicted by the dark grey region in Fig. 2) to more precisely target the tip of the MSTO. The blue dotted lines show the narrower magnitude range spanned by this sample, which corresponds to a narrower distance range of ∼6−20 kpc. Therefore, this sample does not probe as far out into the halo as the faintest MSTO sample, but has a larger minimum distance and thus contains less of a contribution from solar neighbourhood and disc stars.

In a further attempt to clean the sample of young disc stars, we identify a region in the (u_0-g_0) versus (g_0-i_0) diagram that is dominated by young stars. Fig. 3 shows this colour–colour space with PARSEC isochrones of metallicities ranging from $-2.5 <$ [Fe/H] < -0.5 and ages from 0.2 to 12.2 Gyr. The isochrones are colour coded by age, but colour coding them instead by metallicity would show that the more metal-poor isochrones would be at the top of the diagram and the more metal-rich ones would be at the bottom. The grey triangle shows the region that is almost solely populated by young, metal-rich stars. In the upper left of this triangle, there are also some older populations, but these are horizontal branch (HB) stars, which constitute a much smaller population compared to the more numerous MSTO stars. We therefore decide to remove all stars that fall within this triangular region from the analysis in both the spectroscopic and photometric data sets to clean the young disc dwarfs from the sample. The underlying assumption is that there are far more young stars in the disc than young metal-poor stars in the halo. This cut removes ∼ 10 per cent of the halo sample selected with *Pristine* stars and only ∼ 3 per cent of the spectroscopic sample.

3.3 Distance ranges of the samples

To further investigate the distance ranges probed by these samples, we generate a simulated data set using the GALAXIA code (Sharma et al. 2011), which implements the analytical Besançon model (Robin et al. 2003) of the Galaxy. We run GALAXIA using the default parameters, and selecting the regions of the *Pristine* footprint with a magnitude limit of $14 < g_0 < 22$, an oversampling factor of 3.0, and disc flare and warp turned on. We mimic the survey depth and observational selection function of the *Pristine* survey by selecting a subset of stars from the GALAXIA catalogue to match the stellar density of *Pristine*. This is done using SDSS as a reference sample, under the assumption that it is complete at the magnitude limit of the *Pristine* survey ($V < 21$). We then bin the stars spatially (by each 1 × 1 deg^2 *Pristine* field), by colour [bins of 0.15 in $(g_0 - i_0)$] and by magnitude (bins of 0.2 in g_0) and count the number of stars in each bin. A random subset of GALAXIA stars is selected from a given bin, such that the ratio of the subset to the total is equal to the ratio of *Pristine* stars to SDSS stars in that given bin (i.e. $N_{\rm select}/N_{\rm GALAXIA} = N_{\rm Pristine}/N_{\rm SDSS}$ for each bin in RA and Dec., colour, and magnitude). Since the *Pristine* sources are cross-matched to SDSS, there are no bins in which $N_{\rm Pristine} > N_{\rm SDSS}$. In addition, the oversampling of Galaxia with fsample = 3 minimizes the cases where $N_{\rm select} > N_{\rm Galaxia}$, and results in very few bins in which this is the case. When this does happen, we simply take all of the GALAXIA stars in that bin. The result of this is a GALAXIA generated mock sample that has the same coverage as the *Pristine* footprint, and mimics the depth, magnitude, and colour distributions of *Pristine*.

We then apply the same selections for the MSTO and halo samples and compute the distance distributions for each in the Galaxia mock catalogue. Fig. 4 shows the heliocentric, Galactocentric, and height above the disc distance distributions for each of the samples, and Table 1 summarizes the specific heliocentric distance ranges probed, computed with isochrones and with the Galaxia catalogue. Both methods yield similar distances, particularly at the lower limit of the distance range, but the isochrone technique has slightly higher upper limits due to the extension of the beginning of the subgiant branch, which in Galaxia only represents a small number of stars in the tail of the distributions. Using heliocentric distance ranges can sometimes make it difficult to interpret exactly which Galactic environments are being probed, however, as can be seen in the

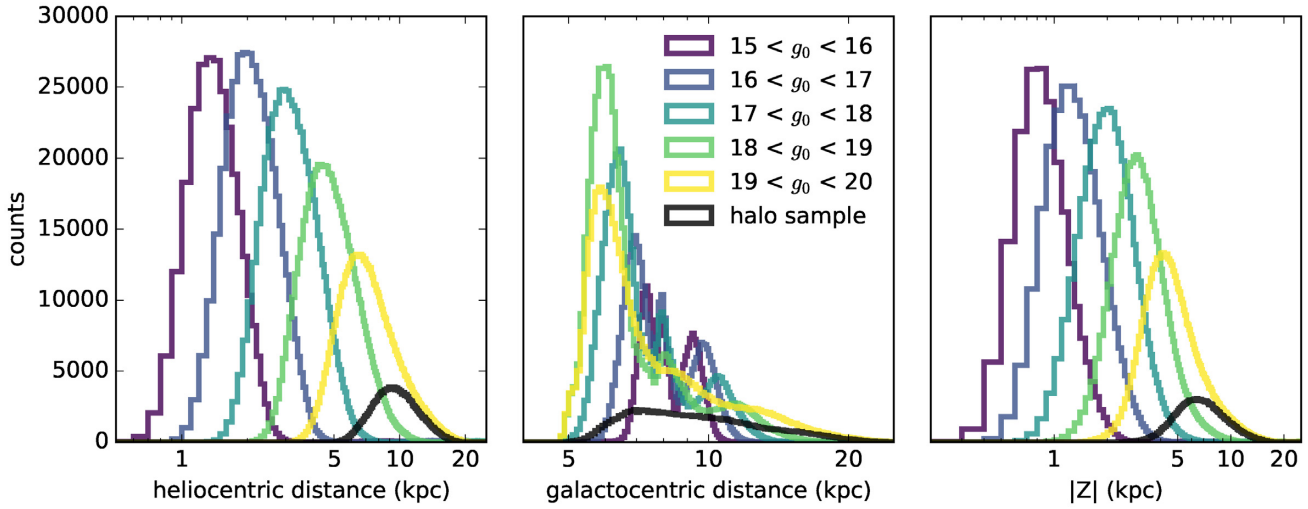

Figure 4. Distance distributions for the MSTO cuts with $0.15 < (g_0 - i_0) < 0.6$ applied to the Galaxia mock catalogues. The left-hand panel shows the heliocentric distances, the middle panel the Galactocentric distances, and the right-hand panel the height above/below the disc, |Z|. The histograms are coloured according to the magnitude range shown, and the black histogram shows the distance distribution for the halo sample with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$.

Table 1. Heliocentric distance ranges probed by the MSTO samples selected with $0.15 < (g_0 - i_0) < 0.6$ and the shown magnitude cuts. The bottom row shows the heliocentric distance range of the halo sample, selected with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$. Distance ranges are computed both using isochrones as well as the distances from the Galaxia mock data. The bold numbers indicate the peaks of the distributions in the left-hand panel of Fig. 4.

Magnitude range	Isochrone distance (kpc)	Galaxia distance (kpc)
$15 < g_0 < 16$	[0.6–4]	[0.8–**1.5**–3]
$16 < g_0 < 17$	[1–6.3]	[1–**2**–4.5]
$17 < g_0 < 18$	[1.5–10]	[1.6–**3**–7]
$18 < g_0 < 19$	[2.4–15.9]	[2.5–**4.5**–11.2]
$19 < g_0 < 20$	[3.8–25]	[3.5–**6.5**–18]
halo sample	[6–20]	[6–**9.5**–16]

middle and right-hand panels of Fig. 4, the high Galactic latitudes ($|b| > 20°$) and limited depth of the *Pristine* sample result in a distance distribution that still nicely samples the inner Galactic halo when converted to the reference frame of the Galactic Centre. Most importantly, the halo sample probes a distance range of ∼5−20 kpc from the Galactic Centre, and a minimum height above the disc of $|Z| > 4$ kpc, avoiding in large part the stellar populations of the bulge and disc.

3.4 The MDF at different distances

Fig. 5 shows the resulting MDFs for the MSTO samples at different magnitude ranges. These are raw photometric MDFs from *Pristine* data without any corrections applied to qualitatively show the decreasing metallicity gradient with increasing heliocentric distance. The coloured histograms are produced with the $0.15 < (g_0 - i_0) < 0.6$ colour cut and limiting the samples to the shown magnitude ranges. The black histogram shows the halo sample with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$. There is a clear trend that the MDFs shift to lower metallicities with increasing distance. There is also a clear transition from the strong peak at [Fe/H] ~ -0.7 in the

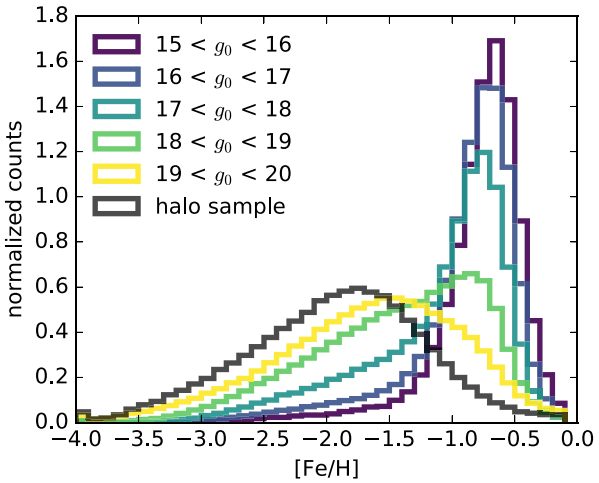

Figure 5. MDFs of MSTO samples selected with a colour cut of $0.15 < (g_0 - i_0) < 0.6$, separated into bins of magnitude. The black histogram shows the halo sample with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$.

brighter samples where the disc dominates, and the metal-poor peak at [Fe/H] ~ -1.6 in the fainter samples where the halo population dominates. The green and yellow histograms in the magnitude range of $18 < g_0 < 20$ show the transition, where both populations are visible.

4 CORRECTING THE MDF

For the current data set, there are two main corrections that we take into account in order to account for the metallicity biases present in the MDF. The first comes from determining metallicities photometrically and the inherent uncertainties associated with this procedure, particularly in the metal-poor regime, and the second are the biases introduced by making the colour cut of $0.15 < (g_0 - i_0) < 0.4$ to select the distance limited halo sample. The effect of these biases, as well as the methods used to correct them, are discussed in detail in this section.

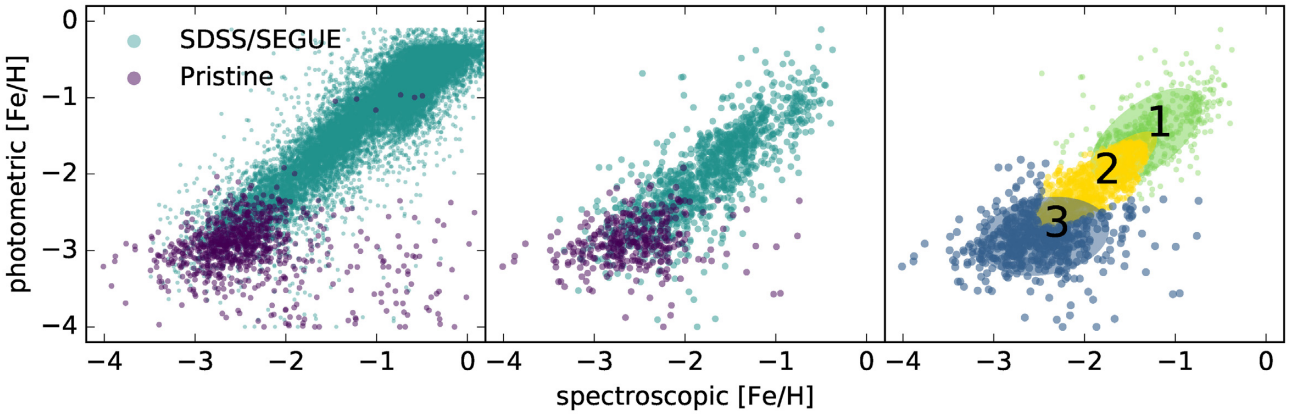

Figure 6. The spectroscopic sample used to rescale the *Pristine* photometric MDF for the halo sample. The left-hand panel shows the full spectroscopic samples from *Pristine* follow-up (purple) and SDSS/SEGUE (teal) and the middle panels shows the samples after applying the colour cut of $0.15 < g_0 < 0.4$ and a magnitude cut of $17 > g_0$ to select the halo sample. The right-hand panel shows the corresponding 3-component GMM used to fit the data. The points are coloured according to the Gaussian component to which they belong, and the 1σ contour of each is shown by the overplotted oval of the same colour. The varied point sizes are simply to enhance the clarity in crowded regions.

4.1 Corrections for the photometric metallicities

Despite cleaning the photometric metallicity sample of contaminating objects such as white dwarfs and variable stars, some degree of contamination still persists, particularly at the metal-poor end. This is mostly due to metal-rich stars preferentially scattering into the metal-poor regime, but can also be caused by stars with unusual colours (e.g. blue stragglers, or stars with long-period variability not picked up in Hernitschek et al. 2016) or peculiar behaviour in their Ca II H&K lines (e.g. binaries, or stars with Ca II H&K in emission). Fortunately, most of these stars that are misclassified photometrically can be identified with the spectroscopic follow-up sample, and can subsequently be corrected for using a statistical rescaling of the photometric MDFs. The left-hand panel of Fig. 6 shows the relation between the *Pristine* photometric metallicities and the derived spectroscopic metallicities for the spectroscopic sample, which we assume to represent the true metallicity of the star. This sample consists of ∼1000 stars from *Pristine* spectroscopic follow-up, which predominantly cover the metallicity range of $-4 <$ [Fe/H] < -2.5 and are shown by the purple points (see Aguado et al. 2019, for more details on this sample). To supplement the sample at [Fe/H] > -2.5, we add ∼17 000 cross-matched SDSS/SEGUE stars over the *Pristine* footprint that have reliable metallicity determinations from the SEGUE Stellar Parameter Pipeline (SSPP; our selection of this sample is described in detail in Starkenburg et al. 2017), shown by the teal points in the Figure. For the 59 stars in common between the two samples, we find no systematic shift between the metallicities derived in the *Pristine* follow-up analysis and the SSPP, and a dispersion of ∼0.3 dex. The middle panel shows the same samples after the application of the halo selection, with $0.15 < (g_0 - i_0) < 0.4$. We additionally apply a cut of $g_0 > 17$ to the SDSS/SEGUE sample to remove the strong disc feature at [Fe/H] > -1, and to allow for the distribution to have a larger spread to compensate for the larger uncertainties at faint magnitudes. We do not apply a magnitude cut of $19 < g_0 < 20$ since there are too few stars in the spectroscopic sample at these faint magnitudes. Similarly, we do not make a magnitude cut on the *Pristine* spectroscopic sample since this sample is restricted to an even brighter magnitude range. Therefore, the implicit assumption that we make is that the relation between photometric and spectroscopic metallicity at [Fe/H] < -2.5 does not vary significantly between stars with $14 < g_0 < 18$ and $18 < g_0 < 20$. We also apply the $(u_0 - g_0)$ versus $(g_0 - i_0)$ cut to remove young disc stars as described in Section 3.2, but this only removes ∼3 per cent of the stars from the spectroscopic samples. The sizes of the samples after the cuts are applied are 1163 and 340 for the SDSS/SEGUE and *Pristine* samples, respectively.

To perform the rescaling, we build a probability density function (PDF) to describe the relation between the photometric and spectroscopic metallicities. We choose to use a Gaussian mixture model (GMM) for this purpose, because Gaussians are simple functions that are well suited to fitting smooth, continuous data.

We make the GMM using the PYTHON package *sklearn.mixture.GaussianMixture* (Pedregosa et al. 2011), which implements the expectation–maximization algorithm for fitting distributions composed of several Gaussian components, the number of which are provided as an input parameter. We use the default parameters, with the exception of increasing the maximum number of iterations to 500 to ensure that the solution converges.

To determine the number of components to use to best fit the data, we use the Bayesian information criterion (BIC). We compute GMMs for each of the samples using between 1 and 10 components, and then compute the BIC for each using the built-in attribute of the sklearn mixture model package. We repeat this 1000 times, taking the mean of the resulting distribution of test values, and then choose the number of components with the lowest computed BIC. For the halo sample, this gives a solution with 3-components, and the resulting GMM is shown in the right-hand panel of Fig. 6. All three of the components have similar weights, such that the GMM is not dominated by any 1 component. Table A1 in the Appendix shows the best-fitting parameters for each of the Gaussian components, including the means, weights, and covariance matrices.

To check that the GMM fits the data well, we use a 2D Kolmogorov–Smirnov (KS) test, which begins with the null hypothesis, H_0, that two samples are drawn from the same parent distribution. First, we draw a sample from the GMM and compute the KS statistic and p-value between this sample and the original

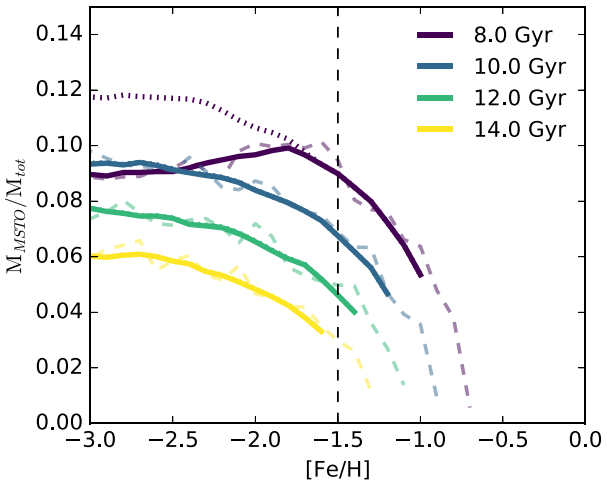

Figure 7. Fractional contribution of the mass from the colour cut of $0.15 < (g_0 - i_0) < 0.4$ to the total mass of the stellar simulated stellar population. The darker lines show the smoothed contours and the lighter dashed lines show the true values in each 0.1 dex metallicity bin. The dotted lines show the mass ratio for a colour cut of $0 < (g_0 - i_0) < 0.4$, and shows the degree of truncation for the MSTO as it extends below $(g_0 - i_0) = 0.15$.

spectroscopic one using a two-sample KS test.[1] We repeat this procedure 1000 times and take the mean p-value for the set of iterations. We find $p > 0.1$ for a GMM with 3 components, meaning that we cannot reject H_0 and the two samples are consistent with being drawn from the same distribution. If we use 2 components, then the resulting p-value is marginal and inconclusive, with $0.1 > p > 0.01$. Using only 1 component gives $0.01 > > p$, and H_0 is rejected. Taken together, these tests confirm that the GMM is an appropriate model for fitting a PDF to the spectroscopic sample, as long as the right number of components are used.

To apply the GMM correction, we take each photometric metallicity from the *Pristine* sample, and simply draw from the corresponding 1D PDF made by sampling along a horizontal line at that given metallicity in Fig. 6.

4.2 Corrections for the halo sample selection

For the colour cut used to select the halo sample, there is a metallicity bias that preferentially selects metal-poor stars over metal-rich stars since the turn-off point of the MSTO is significantly redder for metal-rich populations than for metal-poor ones. This can be clearly seen by the shifting isochrones in the grey region of Fig. 2, as a larger fraction of the isochrones fall outside this region with increasing metallicity. In addition, younger populations will have a bluer turn-off than older ones. In order to investigate this effect and its influence on the MDF, we use PARSEC simulated stellar populations with a total mass of 10 000 $M_\odot$ across a range of metallicities to compute the mass fraction selected by these colour cuts to the total mass of the stellar population. Fig. 7 shows the metallicity dependence for the $0.15 < (g_0 - i_0) < 0.4$ colour cut, for a series of ages. The lighter dashed lines show the true values for each bin, and the darker lines are smoothed with a mean kernel and a width of seven bins. The dotted lines (only visible for the 8 Gyr population) represent a colour cut of $0 < (g_0 - i_0) < 0.4$, and show the result when the entire MSTO is included at all ages.

For the 8 Gyr simulated population, we see the effect of a the tip of the MSTO falling outside our selected colour range ($g_0 - i_0 < 0.15$), shown by the deviation of the solid purple line from the dotted line at [Fe/H] < -1.7 in Fig. 7. For older ages, the MSTO is not truncated and the dotted lines lie underneath the solid lines. Therefore, even though we are not able to reliably assign metallicities to stars bluer than $(g_0 - i_0) = 0.15$ with *Pristine*, this will only result in a (small) bias at low metallicities even if there is a significant population of the halo that is younger than 10 Gyr. Furthermore, at all ages the shape of the curves in Fig. 7 shows that the relative mass selected decreases with increasing metallicity, as this colour selection preferentially selects the (blue) MSTO stars of metal-poor populations. At the most metal-poor end the trend flattens, as the isochrones begin to pile up on each other, i.e. the colour selection is largely unbiased at [Fe/H] < -3. On the right-hand side of the plot, the metal-rich MSTO populations are increasingly suppressed as they fall outside the colour selection, i.e. $0.4 < (g_0 - i_0)$. Table A2 in the Appendix gives correction factors for this effect at each metallicity bin. The factors are computed for each age separately by scaling each bin up to the value at [Fe/H] $= -3$. The correction we apply to the halo sample is the mean of the 10 and 12 Gyr populations, and is given in the rightmost column of Table A2.

When making the MSTO cuts, we also investigate whether using redder broad-band colours to make the MSTO selections could reduce the correction factors since isochrones of different metallicity do not show as strong of a colour difference as they do with bluer colours. Recomputing Fig. 7 for $(i_0 - z_0)$, we find that the curves are significantly flatter in the metal-rich regime, indicating a decrease in the metallicity bias when using these colours instead of $(g_0 - i_0)$ colours. However, at the limiting magnitude of $g_0 = 20$, uncertainties in the g magnitudes are not larger than 0.03, while uncertainties in the z magnitudes in SDSS can be as large as 0.15. Given that in $(i_0 - z_0)$ space the MSTO is more compressed, these uncertainties are equal in size to the entire width of the MSTO, and are therefore too large to make a reliable MSTO selection. For this reason, we choose to use the $(g_0 - i_0)$ colours to select the MSTO.

Finally, this colour range also includes HB stars (shown in the top part of Fig. 2), which are significantly brighter and more distant than the turn-off stars. Using the PARSEC simulated populations, we calculate the mass contribution of HB stars compared to the mass at the rest of the turn-off, and find that in the colour range of $0.15 < g_0 < 0.4$, the relative contribution of HB stars is at most ~ 0.2 per cent for the 8 Gyr population, and less than this for the older populations. Therefore, we conclude that the HB stars constitute only a very small amount of the sample and subsequently that distance cross-contamination between bins is small, resulting in no noticeable effect on the computed MDFs.

5 SUBSTRUCTURE IN THE HALO MDF

5.1 The MDF in different directions

We would like to further assess the impact of observational direction on the resulting MDF of the sample. Fig. 8 shows the footprint of the *Pristine* survey on the sky, coloured by distance to the Galactic Centre. On the left-hand y-axis is the Galactic latitude, and on the right-hand axis the corresponding distance above or below the Galactic plane [$Z = d \cdot \sin(b)$]. The distances shown are computed to the approximate peak of the halo sample distance distribution

[1]The code used for the 2D KS test was modified from https://github.com/syrte/ndtest.

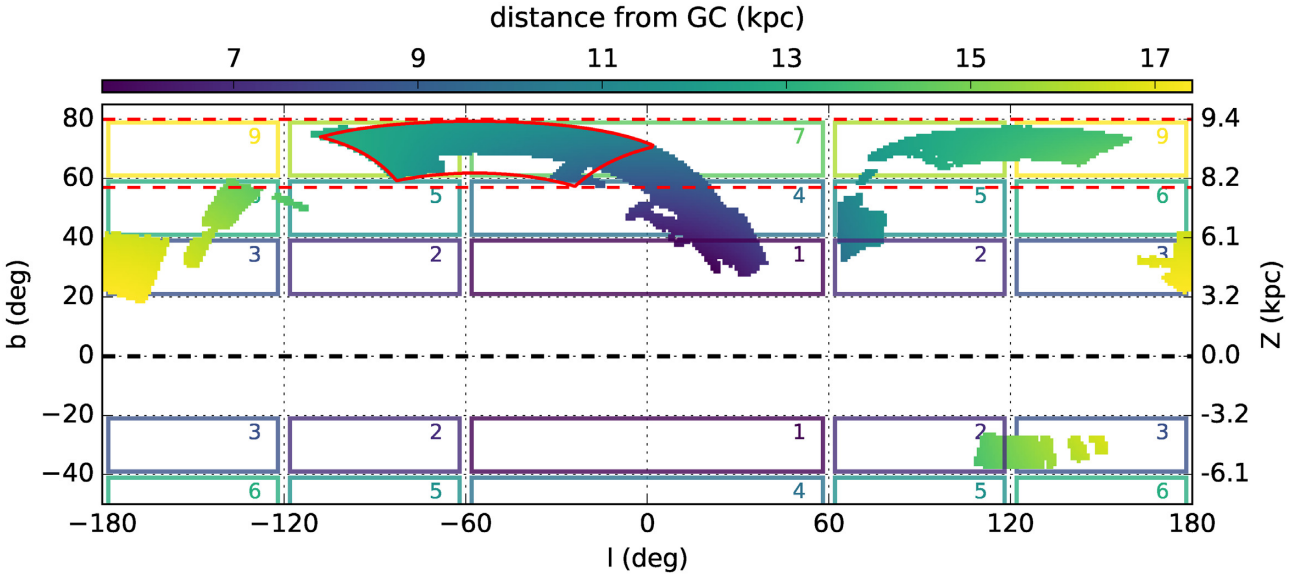

Figure 8. The $\sim 2500\,\mathrm{deg}^2$ *Pristine* footprint, colour coded by the distance to the Galactic Centre for the halo sample, assuming a heliocentric distance of 9.5 kpc. The coloured squares and the corresponding numbers separate the footprint by latitude and longitude, and each coloured region corresponds to the MDF of the same colour in Fig. 10. The region outlined in red shows the region selected to contain the Virgo overdensity, and the red-dashed lines are the equal latitude comparison sample.

of 9.5 kpc, as shown by the black histogram in the left-hand panel of Fig. 4. To convert from heliocentric distances to Galactocentric, we use a solar position of $(X, Y, Z) = (8.2, 0, 0.011)$ kpc. We separate the sky into nine regions, made by the intersection of three regions in longitude and three regions in latitude. The longitude regions are each 60 deg, and are mirrored across the Galactic Centre, and the latitude regions are each 20 deg and are mirrored across the Galactic plane. The nine regions are shown by the coloured squares and labelled by number in the Figure. Fig. 10 shows the corresponding MDFs for the halo sample in each of these regions, coloured according to the corresponding region in Fig. 8 and with the number of stars making up the sample also listed in the top left-hand corner. These are plotted over the grey, filled histogram representing the full halo sample for comparison. We see that for most locations in the footprint, the MDF is invariant, meaning that the relative contribution of the disc and halo does not change with varying latitude or longitude, suggesting that the halo cuts were effective in selecting a pure halo sample. Additionally, it shows that the halo is relatively well mixed in metallicity at these distances. Regions 2, 5, 6, and 9 show some small-scale deviations from the total distribution, but these are all regions that have relatively small numbers of stars and thus these features may be a result of insufficient statistical sampling rather than real substructures. In region 3, however, we see a significantly stronger contribution at the metal-rich end, suggesting that at low galactic latitudes there is an additional contribution from a higher metallicity population.

5.2 The Monoceros ring

Region 3 is at relatively low Galactic latitude and in the direction of the Galactic anticentre. In this observational direction, a substantial amount of substructure has been previously observed, including the Monoceros Ring (Newberg et al. 2002), A13 (Sharma et al. 2010), TriAnd (Majewski et al. 2004; Rocha-Pinto et al. 2004), the Anti-Centre Stream (ACS) and the Easter Banded Structure (EBS; Grillmair 2006; Grillmair, Carlin & Majewski 2008; Li et al. 2012). Although these structures were originally thought to be disrupted debris from accreted satellite galaxies, some recent studies have shown that some or all of these structures may be the result of kicked-up disc material, given that their chemistry and kinematics are consistent with a disc origin (e.g. Li et al. 2017; Bergemann et al. 2018; Sheffield et al. 2018). The peak of the distance distribution of the halo sample coincides with the distance to these structures (d $\sim$10 kpc), suggesting that these are indeed responsible for the observed metal-rich component in the MDF. Despite being able to differentiate this structure from the rest of the halo based on its metallicity distribution, we cannot from the MDF alone constrain whether this indeed is consistent with kicked-up disc material or debris from an accreted satellite.

To further investigate this metal-rich feature, we separate region 3 into high latitude ($|b| > 30$ deg) and low latitude ($|b| < 30$ deg) stars. The MDFs of each of these latitude ranges are shown in the bottom left-hand panel of Fig. 10. From this, we see that for stars with $|b| > 30$ deg, the metal-rich component predominantly disappears, and the MDF is consistent with that of the rest of the footprint regions. We therefore choose to cut out this region from the halo sample and keep only the stars with $|b| > 30$ deg to minimize potential contamination of metal-rich stars from the disc and its associated structures.

5.3 The contribution of the disc

We investigate the possible degree of disc contamination in the halo sample further using the Galaxia mock catalogue described in Section 3.3. With this model, we can identify which component of the Galaxy a given star belongs to. Fig. 9 shows the relative contributions of the thin disc, thick disc, and halo for each of the MSTO-selected samples. For the brighter samples with $g_0 < 19$, the sample is dominated by disc stars, and for the $19 < g_0 < 20$ sample there are slightly more halo stars than disc stars. For the halo sample, the contamination from the disc decreases significantly compared to the faintest MSTO sample. This is due to the greater

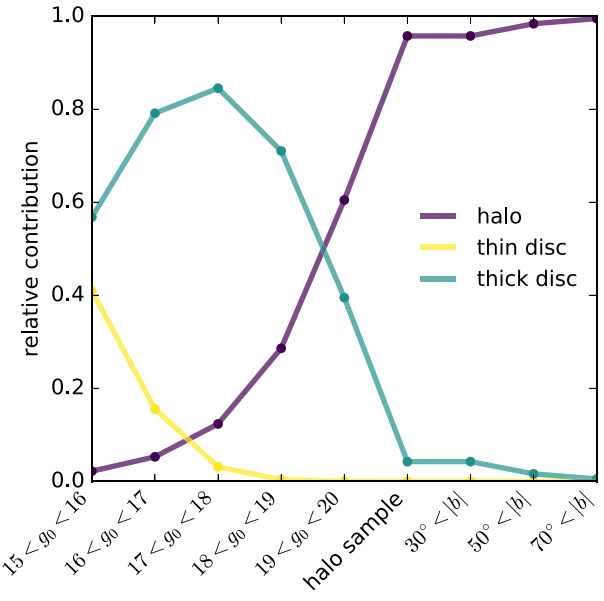

Figure 9. Relative contribution of the thin disc, thick disc, and halo for the different distance samples when selected in the Galaxia mock catalogue. The magnitude ranges refer to the MSTO samples, with the colour cut $0.15 < (g_0 - i_0) < 0.6$, and the halo sample refers to the sample made with $0.15 < (g_0 - i_0) < 0.4$ and $19 < g_0 < 20$, as described in the text. The latitude cuts are applied to the halo sample.

minimum distance and also due to the removal of metal-rich stars with the narrower colour cuts when selecting the MSTO. Adding latitude cuts further increases the purity of the halo selection until at a latitude of $70^\circ < |b|$ the sample consists of almost entirely halo stars. Although this model represents a simplified version of the Galaxy, with an axisymmetric disc devoid of perturbations and substructures from interactions with accreted satellites, this still offers insight into the relative contributions of the selections to the purity of the sample. Therefore, due to the distance range probed by the selected halo sample, the contribution of foreground metal-rich disc stars is likely to be low, although there may still be some component of high-latitude, kicked-up disc material, as discussed in Section 5.2.

5.4 The Virgo overdensity

In addition to investigating the effect of direction in Figs 8 and 10, we also take a closer look at the largest Galactic substructure in our footprint that is probed by our distance range: the Virgo overdensity. Although the specific nature and origin of this structure is still not well understood, it is known that the Virgo overdensity is located at a distance of 8–25 kpc (e.g. Lokhorst et al. 2016), and that it is likely comprised of overdense regions of halo stars. Located at approximately RA = 190 deg and Dec. = −5 deg and spanning a large region of sky of over 1000 deg^2, a significant region of the overdensity is located within the *Pristine* footprint. The current data set is therefore well suited to exploring the MDF of this structure. Although the Sagittarius stream is also prominent over our footprint, its vast leading and trailing arms are too distant to be sufficiently seen in the sample presented in this work (see, for example, Law & Majewski 2010). We therefore select a region around the Virgo overdensity, and compare its metallicity distribution to another region of the *Pristine* footprint at the same Galactic latitude that does not include Virgo. The region selected to contain Virgo is 180 deg < RA < 210 deg and −1 deg < Dec. < 20 deg, and is demarcated in red in Fig. 8, and the comparison sample is shown by the red-dashed lines. The MDFs of these regions are shown in the bottom middle and bottom right-hand panels of Fig. 10. Although there are minor fluctuations from the underlying halo sample, these are not significant enough to draw any conclusions. Therefore, this indicates that the Virgo overdensity metallicity is not very different from that of the general stellar halo at these Galactic latitudes.

6 DISCUSSION

So far, we have produced an MDF for a carefully selected halo sample, with $0.15 < (g_0 - i_0) < 0.4$, $19 < g_0 < 20$, $|b| > 30^\circ$ and a cut in $(u_0 - g_0)$ versus $(g_0 - i_0)$ space to remove young disc stars. We calculated with isochrones and a Galaxia model that this sample spans ∼6−20 kpc, a distance range that nicely probes the inner Galactic halo. This sample is not free of selection biases, as we discussed in Section 4.2, but these are mostly present at the metal-rich end, and constitute moderate corrections that we have applied at [Fe/H] < −1.5.

The slope of the metal-poor regime, particularly at [Fe/H] < −1.5, is important in that it constrains the relative ratios of the most metal-poor stars in the Galaxy, which in turn encode the rate of chemical enrichment at early times. This is useful for building chemical enrichment models and for constraining the physical processes driving them. This also helps to better understand the role of metal-poor stellar populations in the formation of the Galactic halo, such as accreted dwarf galaxies and GCs.

6.1 Biases in the photometric metallicities

The *Pristine* survey is fairly unbiased in the stars that it observes. However, there are still various interdependencies on the colour, magnitude, and derived metallicities. We list the most important effects next, and discuss their possible effects on the MDF in the following section:

(i) *Dependence of the metallicity sensitivity on colour:* Fig. 11 shows the overlapping sample of SDSS/SEGUE stars with the *Pristine* sample (using the same quality cuts on the *Pristine* photometry and the SDSS/SEGUE spectra as described in Section 3 of Starkenburg et al. 2017). The colour coding by metallicity nicely illustrates the metallicity sensitivity of the survey. This figure also shows that at bluer colours the range of metallicities spans a narrower range in $(CaHK_0 - g_0)$ than at redder colors, meaning that for the same uncertainties in photometry there will be a larger uncertainty in metallicity, and consequently a larger contamination across metallicity bins. In addition, the histogram in the bottom panel of Fig. 11 shows that most of the stars are concentrated at $(g_0 - i_0) < 1.0$. The population redder than that is mostly dominated by metal-rich disc dwarfs.

(ii) *Dependence on the depth of the survey on metallicity:* Due to the nature of the narrow-band *Pristine* Ca H&K filter, stars that are metal-poor will have smaller absorption lines and thus appear brighter in the photometry. This results in a slight difference in brightness for the most metal-poor stars compared to the metal-rich ones. The middle panel of Fig. 11 shows that on average, the more metal-poor stars will be 0.2−0.5 mag brighter in the *Pristine* narrow-band filter than metal-rich stars. This will lead to the metal-poor stars probing a slightly larger volume. However, imposing the colour cut of $0.15 < (g_0 - i_0) < 0.4$ used to select the halo

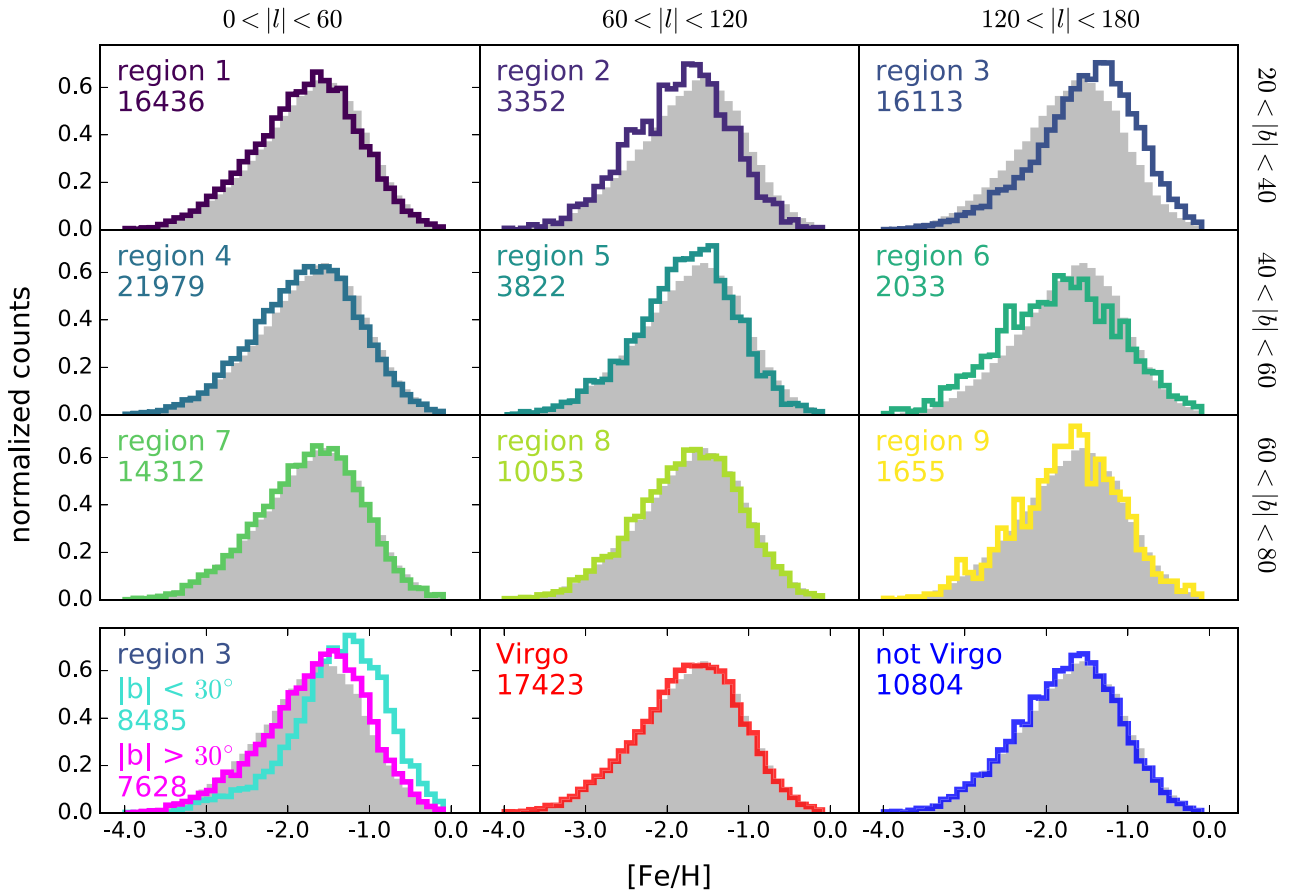

Figure 10. MDFs for the nine different regions described in the text, with colours that correspond to the numbered regions depicted in Fig. 8. The grey-filled histogram is the MDF of the total, GMM corrected halo sample for comparison. The bottom left-hand panel shows the MDFs for region 3, separated into a high- and low-latitude sample (fuchsia and cyan, respectively), and the other two panels in the bottom row show the region selected to contain Virgo in red (the selected region is outlined in Fig. 8), and the rest of the stars at the same Galactic latitude as a comparison are shown in blue.

sample reduces this metallicity-dependant magnitude discrepancy to 0.1−0.2 mag, which is on the order of the uncertainty inherent in determining these distances with isochrones.

(iii) *Dependence of measurement uncertainties on magnitude:* Another effect results from the uncertainties on the photometric flux being larger at fainter magnitudes. At the bright end, the uncertainties in both the Ca H&K and the SDSS broad-band magnitudes are small, but at the faint end $\sigma_{\rm Ca\,H\&K} >> \sigma_{\rm SDSS}$, such that it is largely the Ca H&K uncertainty that drives the overall metallicity uncertainty. The photometric uncertainties as a function of magnitude for the sample are summarized in Table 2. The larger uncertainties at the faint end are partially mitigated with the GMM rescaling, especially since we construct the GMM based on a fainter sample, resulting in a model that takes into account the larger uncertainties in the spectroscopic data set.

(iv) *Dependence of contamination fraction as a function of metallicity:* As previously mentioned in the introduction, there are far more metal-rich stars than metal-poor stars in the halo MDF, meaning that more metal-rich stars will scatter into the metal-poor regime than the other way around, resulting in a relatively higher contamination fraction at low metallicities. This is largely corrected for by the GMM rescaling, but only in so far as metal-rich stars that scatter into the metal-poor end can be identified by medium-resolution spectroscopy.

6.2 Previous determinations of the halo MDF

The MDF of the halo produced from *Pristine* data is shown in Fig. 12. The grey histogram is the raw *Pristine* MDF, the cyan is after rescalling with the GMM, and the black histrogram is corrected with both the GMM and the correction factors for the colour cut from Table A2. Table A3 shows the counts and corresponding bins used to produce each of the Pristine histograms in Fig. 12. Application of the GMM correction to the halo sample has two main effects on the MDF, namely that it shifts the peak of the distribution to slightly higher metallicities, and removes contaminants from the metal-poor tail and shifts them to higher metallicities. The corrections for the colour cut do not significantly change the slope of the metal-poor tail of the MDF ([Fe/H] < -2), but modestly enhance the relative numbers of stars in the bins at higher metallicities.

The right-hand panel of Fig. 12 shows a comparison of several MDFs from the literature that are presented in this paper, and the left-hand panel shows the same but with a log scale to better emphasize the differences in the slopes of the metal-poor regime. Each distribution is scaled to the *Pristine* distribution at [Fe/H] $= -2$, to allow for an easier comparison of the slopes. The purple, yellow and green histograms are the MDFs presented in An et al. (2015), Allende Prieto et al. (2014), and Schörck et al. (2009), respectively, and the blue-dashed line is the MDF expected

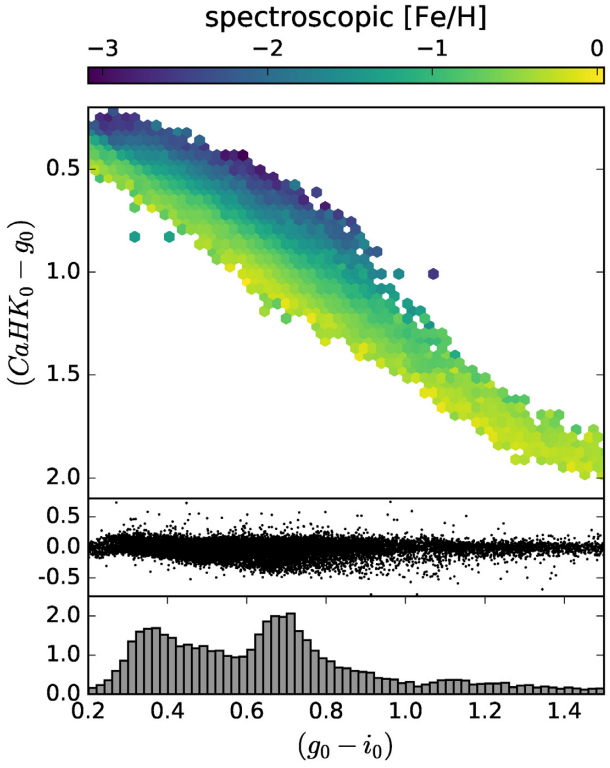

Figure 11. Top panel: Colour–colour plot of the SEGUE spectroscopic sample, colour coded by SEGUE metallicities. Middle panel: $(CaHK_0 - g_0)$ colour subtracted by the median value per $(g_0 - i_0)$ colour bin. Bottom panel: histogram of the distribution of the top panel.

Table 2. Uncertainties as a function of magnitude for the *Pristine* Ca H&K and SDSS broad-band magnitudes.

	median uncertainties		
Magnitude	g	i	Ca H&K
$g = 15$	0.003	0.004	0.004
$g = 16$	0.004	0.004	0.006
$g = 17$	0.005	0.005	0.01
$g = 18$	0.006	0.006	0.02
$g = 19$	0.01	0.01	0.05
$g = 20$	0.02	0.02	0.1

from the simple chemical enrichment model presented in Hartwick (1976), assuming an effective yield of $y_{\rm eff} = -1.6$. The MDF of Ryan & Norris (1991) is not shown in this Figure, but is in good agreement with the Hartwick (1976) model. In this section, we first describe each MDF and how it was produced, and then discuss how it compares to our results.

6.2.1 *The chemical evolution model of Hartwick (1976)*

Using a sample of 60 halo GCs, Hartwick (1976) modifies a simple chemical evolution model (Searle & Sargent 1972; Pagel & Patchett 1975) to describe the MDF. Starting from a closed box of primordial gas, this model is based on a simple parametrization of the rate at which the interstellar medium (ISM) is enriched in heavy elements by exploding stars. The model assumes instantaneous recycling and a constant IMF, and can be described using a single free parameter known as the effective yield, $y_{\rm eff}$. This parameter represents the ratio of the mass of heavy elements ejected to the mass locked up in long-lived stars or remnants, and also sets the location of the peak of the distribution. Following Ryan & Norris (1991), we choose $y_{\rm eff} = -1.6$. Under the assumption that the halo GC population traces the field star population, we can use this model to approximate the idealized MDF of the Milky Way halo.

6.2.2 *SDSS sample (Allende Prieto et al. 2014)*

Allende Prieto et al. (2014) use a sample of spectroscopically selected F-type stars from the SDSS survey (in particular those targeted in the BOSS survey as flux calibrators) to determine a metallicity distribution of the Milky Way halo. The MDF shown by the yellow histogram in Fig. 12 is produced using a sample of ∼16 000 F-type MSTO stars selected with $2.5 < \log g < 4.4$ and $g_0 > 17$. Most of the observed stars are at high Galactic latitudes, and they estimate them to probe a distance range from 5 kpc out to a few tens of kpc. They note that this sample is biased due to colour cuts used to select the BOSS spectrophotometric stars (summarized in their table 1). To correct for these, they compute photometric colours for a grid of spectral energy distribution models covering a range of stellar parameters and identify that ones make it through the colour cut and are in the temperature range of $5600 < T_{\rm eff} < 6500$ K. They then rescale the MDF to account for the missing populations, making the simplifying assumption that all stars are evenly distributed over the stellar parameter space. Although this slightly boosts the number of stars at the peak of the distribution at [Fe/H] = −1.6, it has very little effect on the tails of the distribution. The corrections are expected to be most significant for higher metallicities, as at these metallicities few stars with typical turn-off temperatures and gravities are selected in BOSS (as is also apparent from the results in table 1 of Allende Prieto et al. 2014).

6.2.3 *HES sample (Schörck et al. 2009)*

Schörck et al. (2009) estimate the slope of the MDF at the metal-poor end by correcting the observed MDF from the Hamburg ESO Survey (HES, Christlieb, Wisotzki & Graßhoff 2002) for selection effects. For the correction, they create an MDF following Hartwick (1976), replace all the [Fe/H] values with pairs in their observables KP (an index on the Ca K absorption line following Beers et al. 1999) and colour $(B - V)_0$, convolve them with Gaussian uncertainties, and assess the completeness in [Fe/H] versus $(B - V)_0$ space after applying the HES selection cuts. The result of this exercise is that they find HES to be essentially complete below [Fe/H] = −3.5, but the corrections become quite significant at higher metallicities. For example, at [Fe/H] = −2.05 they apply corrections of a factor of 10−80 (depending on colour). They build the MDF based on a sample of 1638 HES stars with available spectroscopic follow-up, and scale it up to include all 3439 stars from the total candidate sample.

6.2.4 *The sample of Ryan & Norris (1991)*

Ryan & Norris (1991) present the metallicity distribution of the local halo from 372 subdwarfs selected in proper motion studies with space velocities relative to the LSR > 250 km s^{-1}. They find a very good fit to the Hartwick (1976) model using $y_{\rm eff} = -1.6$. The excellent fit does not include the 10 per cent most metal-rich stars in

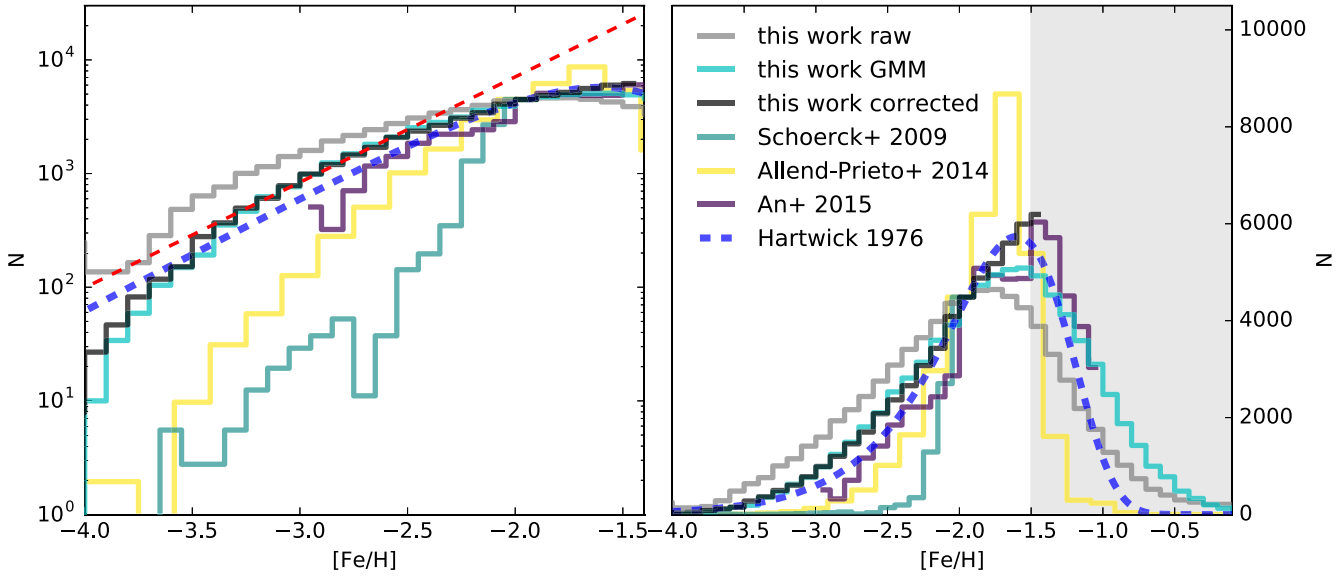

Figure 12. Comparison of the *Pristine* MDF to previous works. The right-hand panel shows the MDFs over the full metallicity range to illustrate the global shapes and location of the peaks, while the left-hand panel shows only the metal-poor end and is plotted with a log scale to emphasize the differences in the metal-poor slopes. The grey shaded region shows where the corrections to the *Pristine* MDF become large and uncertain, at [Fe/H] > -1.5. The raw MDF of the *Pristine* halo sample is shown in grey, the GMM corrected in cyan, and the GMM/colour corrected in black. The MDFs from An et al. (2015), Allende Prieto et al. (2014), and Schörck et al. (2009) are shown in purple, yellow, and green, respectively, and the Hartwick (1976) model is shown as the blue-dashed line. The red-dashed line is an exponential fit to the *Pristine* corrected sample in the metallicity range of $-3.4 <$ [Fe/H] < -2.5.

the sample, but they speculate that this might be contamination from the disc. The model would predict 13 stars with [Fe/H] < -3, whereas 6 are observed, this difference is, however, too small to be significant for their sample size.

6.2.5 The sample of An et al. (2015)

An et al. (2015) derive a halo MDF photometrically from main-sequence turn-off stars observed in the extra deep SDSS stripe 82, and estimate that their sample covers a distance range of 5−10 kpc. They find excellent agreement with the work of Ryan & Norris (1991) mentioned above. In addition to a single Hartwick (1976) or Gaussian model fit, they also provide a double Gaussian fit to their derived MDF, finding a peak at [Fe/H] $= -1.4$ and a second one at [Fe/H] $= -1.9$. The extra Gaussian component is motivated by a difference in kinematics for the stars in both populations, where the second population shows a larger number of retrograde orbiting stars.

6.3 Comparison to the literature

6.3.1 The number of metal-poor stars

Fig. 12 shows that there are distinct differences between the various MDFs in the relative numbers of stars they contain at [Fe/H] < -2. Most notably, the *Pristine* halo sample shows the largest number of stars, even more than the Hartwick (1976) model, in the range $-3.4 <$ [Fe/H] < -2.2. The photometric sample of An et al. (2015) shows a pronounced bimodality, with the metal-poor component showing a similar abundance of stars to *Pristine* for [Fe/H] < -2.3, but it begins to drop off at [Fe/H] < -2.6, presumably as the metallicity calibration becomes less reliable. Both of these samples are in reasonably good agreement with the model. The SDSS/BOSS spectroscopic sample of Allende Prieto et al. (2014) shows a smaller contribution with respect to the photometric samples at [Fe/H] < -2, and the HES sample of Schörck et al. (2009) shows the lowest relative contribution of metal-poor stars of all of the shown MDFs.

From the left-hand panel of Fig. 12, we can see that the shape of the *Pristine* halo sample MDF is not linear in log space, and therefore is not well described by a single exponential curve. Instead, the observed distribution is much more suited to a fit with three separate exponentials, in the metallicity ranges of [Fe/H] < -3.5, $-3.4 <$ [Fe/H] < -2.5, and $-2.5 <$ [Fe/H] < -1.5. The best-fitting line in the metallicity range of $-3.4 <$ [Fe/H] < -2.5 is shown as the red-dashed line in the left-hand panel of Fig. 12, and has a slope of $\Delta(\mathrm{LogN})/\Delta[\mathrm{Fe/H}] = 1.0 \pm 0.1$. Below this metallicity range, the slope becomes substantially steeper with a value of 2.0 ± 0.2.

Using a spectroscopic sample of metal-poor halo giants selected from the *SkyMapper* survey, Da Costa et al. (2019) derive a slope of the MDF of $\Delta(\mathrm{LogN})/\Delta[\mathrm{Fe/H}] = 1.5 \pm 0.1$ for a metallicity range of $-2.75 <$ [Fe/H] < -4. The MDF they derive is well fit by a single exponential for the entire metallicity range, but drops off steeply at [Fe/H] $= -4$. Fitting a single exponential in the same way naively to the *Pristine* MDF also yields a slope of 1.5 ± 0.1, but as discussed above, this does a poor job of describing the distribution. Therefore, the slope in the metal-poor regime from the *Pristine* halo sample is somewhat flatter, consists of more components and has a drop off at a higher metallicity ([Fe/H] < -3.5) than the MDF described in Da Costa et al. (2019). The slopes of both of the spectroscopic samples are also steeper than the slope of the *Pristine* halo sample for [Fe/H] < -2, particularly for the Schörck et al. (2009) sample that shows the steepest slope in the metallicity range of $-2.5 <$ [Fe/H] < -2 of all the MDFs.

Interestingly, the slope of the *Pristine* MDF between $-3.4 <$ [Fe/H] < -2.5 matches very well to the slope of the Hartwick (1976) simple chemical evolution model at [Fe/H] < -2. This

breaks down at [Fe/H] < -3.5, however, where the *Pristine* MDF drops steeply. This also happens to be the metallicity below which the *Pristine* photometric metallicity determination becomes less reliable, and where there are very few stars in the spectroscopic sample with which to fit the GMM correction. Nevertheless, this steep decrease is seen in the other MDFs as well, which all show a drop between $-4 <$ [Fe/H] < -3.5, whereas the model does not. At these very low metallicities – and at the earliest epoch of star formation – chemical enrichment is dominated by a small number of supernovae explosions. Therefore, star formation is likely to proceed in a very inhomogeneous and stochastic way, deviating more strongly from the assumptions of the model that the ISM is well mixed and homogeneous. Additionally, the model does not account for channels of enrichment from external sources, which could very quickly enrich the interstellar medium and result in a steeper slope at the lowest metallicities. In Schörck et al. (2009), they note that despite the discrepancy between the slope of their metal-poor MDF and the Hartwick (1976) model, their MDF better matches theoretical models that include a critical metallicity for low-mass star formation at $Z_{cr} = 10^{-3.4} Z_{\odot}$ (Salvadori, Schneider & Ferrara 2007). It should be noted, however, that Schörck et al. (2009) rely on rescaling the models to match their selection function. Nevertheless, this motivates the exploration of further modifications of these simplified models by incorporating more complex physics (e.g. gas accretion, outflow, and inhomogeneous mixing) that may improve the agreement between the models and the data.

6.3.2 *Selection effects and metallicity bias*

Each of the MDFs discussed above are selected in different ways and from different samples. Ryan & Norris (1991) and Hartwick (1976) describe the local halo, whereas An et al. (2015), Schörck et al. (2009), Allende Prieto et al. (2014), and this work all target the 'distant' halo, by selecting stars at larger distances from the Galactic plane, albeit each exploring a slightly different distance range and thus different Galactic environments. More importantly than just the differences of environments probed, is that each of these samples come with their own selection functions and biases. Despite all of these works applying careful corrections to the MDFs, they still do not agree fully with one another, particularly at the metal-poor end.

As previously discussed, the greatest difficulty in producing an MDF is to correct for metallicity-dependant biases and selection effects incurred from the selection of a halo sample. Spectroscopic samples are typically more affected by completeness issues since stars are specifically selected for follow-up observations, whereas photometric samples observe stars more ubiquitously. In addition, spectroscopic samples may be designed to target metal-poor stars, and thus have an implicit bias towards stars of low metallicity that must be corrected (such as is the case for HES, for example). This correction process is highly non-trivial, and must be based on some models or assumed knowledge of the underlying stellar populations, and an overcorrection could lead to an underestimate of the number of metal-poor stars. In contrast, a photometric approach is likely much more complete (in so far as it is complete within the probed magnitude range), but is more susceptible to contamination due to the larger uncertainties associated with photometric metallicity determinations. These uncertainties may cause a broadening of the distribution, and an increase in the number of stars in the wings (i.e. at the metal-poor and metal-rich ends). We have made an effort to correct for this contamination by applying a statistical rescaling of the MDF with a GMM fit to a spectroscopic sample, but this is only effective in correcting for erroneous photometric metallicity determinations that are caught by the spectroscopy. Stars that have peculiar features in their spectra that affect their metallicities but that cannot be detected with low-/medium-resolution spectroscopy (e.g. certain binaries, blue stragglers) will go undetected and thus uncorrected by this approach, and result in an overestimation of the number of metal-poor stars. We have discussed the potential influence of these contaminants in Section 6.1, and although their contribution is likely to be small, this remains a caveat of this work.

Nevertheless, a photometric sample has other advantages, such as larger sample sizes. Despite the many cuts imposed to make the *Pristine* halo sample, it still consists of ∼80 000 stars. The next largest sample from the ones discussed above is from Allende Prieto et al. (2014), and consists of ∼16 000 stars, although the $\log g$ cuts imposed for that sample are more comparable to the faintest MSTO sample presented in this work, which consists of ∼230 000 stars. The narrower $\log g$ sample from Allende Prieto et al. (2014) that is comparable to our halo sample has <1000 stars. This naturally allows for an MDF that is statistically more robust than those made with the somewhat smaller spectroscopic samples.

Ultimately, none of the MDFs presented here are completely devoid of bias or selection effects. It is only through multiple, repeated efforts using a diversity of independent samples and methods, each with their own selection effects and biases, that we may hope to converge on a true determination of the halo MDF. This is particularly promising as we continue to collect large amounts of spectra with large-scale surveys, and continue to improve the quality of photometric metallicity samples and distance determinations.

6.4 Comparison with *Gaia*–Enceladus

The MDF presented for the halo sample, with its peak [Fe/H] $= -1.6$ is qualitatively consistent with the MDF presented for the *Gaia*–Enceladus structure in Helmi et al. (2018; the histogram in their fig. 2), although the MDF presented in this work has a much stronger tail of metal-poor stars then the one presented for *Gaia*–Enceladus. It should be noted that the MDF presented in Helmi et al. (2018) is produced using a sample of APOGEE stars, and no attempt is made to correct for selection effects. Another study conducted by Gallart et al. (2019), select halo stars by taking all stars with large tangential velocities ($>200\,\mathrm{km\,s^{-1}}$) from a local sample ($\lesssim$2 kpc). They further separate these stars based on their metallicity distributions into a blue sequence that is consistent with the *Gaia*–Enceladus population, and a red sequence that they suggest is representative of an '*in situ*' halo population. They determine the metallicity and age distributions for these population using a linear combination of synthetic isochrones, and find that they are both consistent with very old stellar populations. The photometric MDFs for these two sequences match reasonably well with spectroscopic MDFs produced by cross-matching with the LAMOST and GALAH spectroscopic surveys. The MDF of the blue sequence alone has a peak of [Fe/H] ~ -1.4 and loosely follows the distribution provided in Helmi et al. (2018), but the MDF of the red sequence is much more metal-rich with a peak at [Fe/H] ~ -0.7. If we consider these two populations together to make up the halo, then there is a much stronger metal-rich population than is shown in the MDF of the halo sample that we present in this work. However, this may simply indicate differences that result from selection of halo samples in both works as well as the limitations at high metallicity

in this work. If the inner halo is indeed dominated by the stars from a single massive accretion event, then the MDF presented here with the distance range of our halo sample should consist mostly of those stars. However, despite qualitative similarities between the distributions, we cannot with the current data corroborate or reject this hypothesis.

6.5 The duality of the stellar halo

It remains an open question in the field of Galactic archaeology of whether the Galactic halo is composed of a single or multiple components. As discussed in Section 6.2.5, An et al. (2015) derive a halo MDF that is consistent with having two distinct components, and suggest that this supports the dual halo hypothesis. The purple histogram in Fig. 12 shows a fairly clear bimodality, to which An et al. (2015) fit two Gaussians with peaks at [Fe/H] $= -1.4$ and -1.9. They go on to estimate that the local population of halo stars is made up of 35–55 per cent outer halo stars. The *Pristine* MDF does not show such a clear bimodality, but is also clearly not well fit by a single exponential in the metal-poor regime, which may indicate some more complex underlying populations. Similarly, neither of the spectroscopic MDFs from Schörck et al. (2009) or Allende Prieto et al. (2014) display an obvious bimodality, and neither of those authors discuss a potential bimodality in the MDFs as evidence for a dual halo. Interestingly, however, the halo MDF produced in Zuo et al. (2017) with SDSS photometry also does not show a clear bimodality, but they nevertheless decompose it into three underlying Gaussian components, one quite metal-rich and consistent with a disc population, and two that they identify as the inner and outer halo. They show this for two samples, one with $-8 < Z < -4$ and $10 < R < 14$ that shows peaks of [Fe/H] $= -1.43$ and -1.92, and a more distant sample with $-12 < Z < -8$ and $12 < R < 16$ for which they derive peaks of [Fe/H] $= -1.51$ and -2.25. With our current data, we do not have precise enough distances, or a clean enough selection of halo giant samples to go further out into the halo and see if a well-defined break radius between these two-halo populations can be identified, although this should be the topic of future work. Therefore, we do not draw a firm conclusion as to whether the MDF presented in this work is consistent with a multi- or single-component halo.

6.6 Implications for the formation of globular clusters

In the Milky Way, there is an apparent paucity of GCs at low metallicities. More specifically, of the ∼160 GCs observed to date, there are none with [Fe/H] < -2.5 (Harris 2010), which is seemingly at odds with the old ages derived for many GCs. In order to gain insight into this discrepancy, we can compare the metal-poor MDF of Milky Way GCs to the MDF of halo field stars to differentiate whether this lack of observed GCs could be the result of a real physical process or simply due to statistical undersampling. From the catalogue of Harris (2010), there are 56 GCs with [Fe/H] ≤ -1.5, 13 with [Fe/H] ≤ -2, and 0 with [Fe/H] ≤ -2.5. Using the relative number of stars at these metallicities from the scaled MDF of this work, we calculate an expected number of GCs to be 56, 26, 10, and 2.8 for [Fe/H] ≤ -1.5, -2, -2.5, and -3, respectively. Taking only the GCs in the Galactocentric distance range of the selected halo sample ($d_{\rm GC} \sim 5{-}20$ kpc), the counts from the Harris (2010) catalogue are 30, 10, 0, and 0, and from *Pristine* 30, 14, 5.4, and 1.5 for [Fe/H] ≤ -1.5, -2, -2.5, -3, respectively. Taking a Poisson distribution with a mean of 5.4, the expected probability of finding no clusters at [Fe/H] < -2.5 would be 0.5 per cent, while a Poisson distribution with a mean of 1.5 gives a probability of 22 per cent of finding no clusters at [Fe/H] < -3. Therefore, we see that if the GC population were to follow the same MDF as the halo field stars, there is a significant discrepancy between the observed and expected number of GCs at [Fe/H] < -2.5. This is suggestive that the perceived lack of metal-poor GCs is due to a real truncation of the GC MDF, or alternatively that there are underlying differences in formation mechanisms of GCs with respect to the field stars. Some mechanisms for this truncation of the GC MDF have been proposed in the literature (e.g. Kruijssen 2019 and references therein), but a detailed discussion of these is beyond the scope of this work. A more robust statistical analysis to put these results on a firmer footing, as well as a more in-depth discussion will be presented in Gieles et al. (in preparation).

Some recent observational efforts have identified low metallicity objects in the faint, low-mass regime where GCs and dwarf galaxies are difficult to differentiate. For example, the Sylgr stream (Ibata, Malhan & Martin 2019) contains at least two stars with [Fe/H] $= -2.92 \pm 0.06$ (Roederer & Gnedin 2019), although it cannot yet be determined whether the progenitor of this stream was a dwarf galaxy or a GC. Another object, Draco II (Laevens et al. 2015), has a low metallicity ([Fe/H] $= -2.7 \pm 0.1$), and shows both dwarf galaxy and GC like characteristics. In an in-depth study of this system, Longeard et al. (2018) favour the scenario that Draco II is a disrupting dwarf galaxy, although they leave open the possibility that it could be a GC observed at the very end of its disruption process. Therefore, it remains to be seen whether the perceived lack of low-metallicity GCs will be solved by detailed observations in the near future, or if this needs to be explained by theories of GC formation.

6.7 Future work

The halo MDF we present here relies on a spectroscopic sample to rescale the photometric metallicities. The *Pristine* spectroscopic follow-up sample, totalling ∼1000 stars, was obtained over a three-year period using a total of ∼165 nights of observing time on single-slit spectrographs mounted on 2−4 m class telescopes. Although this approach was successful at obtaining a sample with which to explore the data set, it is unfeasible to follow-up the full number of *Pristine* candidates, particularly at fainter magnitudes. The William Herschel Telescope Enhanced Area Velocity Explorer (WEAVE; Dalton et al. 2018) is a multi-object spectrograph (MOS), and starting this year, will begin collecting a large number of spectra to survey the northern sky. In a Memorandum of Understanding between the surveys, it is agreed that WEAVE will allocate ∼6 fibres per deg^2 to follow-up *Pristine* metal-poor star candidates as part of its Galactic archaeology program. As the *Pristine* survey footprint currently spans over ~ 5000 deg^2 and is still growing, we estimate that the WEAVE survey will observe a total of $\sim 30\,000$ *Pristine* candidates during its operation. Apart from providing an excellent sample of metal-poor stars in the halo with which to conduct a multitude of Galactic archaeology studies, this will also provide a much larger spectroscopic sample with which to extend this work. This will allow for an even more detailed investigation of the MDF in various galactic environments as we will have much more information on the dependence of contamination in the samples as a function of colour and magnitude.

Other upcoming MOS facilities and large spectroscopic surveys will continue to provide large numbers of spectra of halo stars that can be used to determine MDFs when their selection functions are

not too complex, or with which to correct photometric surveys as in this work. In combination with the next data releases from the *Gaia* mission including more precise and accurate parallaxes, they will provide an excellent opportunity to study the MDF extending further out into the halo and in narrower distance ranges (see e.g. Conroy et al. 2019, for a recent example of this).

7 CONCLUSIONS

In this paper, we built an MDF of the Galactic halo using a carefully selected halo sample from the *Pristine* survey. The *Pristine* survey currently covers an area of $\sim 2500\,\mathrm{deg}^2$ and provides photometric metallicities to ∼4 million FGK-type stars. From this sample, we applied colour and magnitude cuts to select MSTO samples spanning various heliocentric distance ranges between 0.6 and 25 kpc. We found that there is a gradient showing decreasing metallicity with increasing distance, with a clear transition from a strong disc component at [Fe/H] ~ -0.7 in the brighter samples to a metal-poor peak at [Fe/H] ~ -1.6, where the halo dominates in the fainter samples. We then selected a halo sample consisting of ∼80 000 stars, selected with $0.15 < (g_0 - i_0) < 0.4$, $19 < g_0 < 20$, $|b| > 30\,\mathrm{deg}$, and a cut in $(u_0 - g_0)$ versus $(g_0 - i_0)$ space to remove young foreground disc stars. This produced a sample covering a heliocentric distance range of ∼6−20 kpc, which we confirmed to have a Galactocentric radius of $5 < R_{\mathrm{GC}} < 20\,\mathrm{kpc}$ and height above the disc $|Z| > 4\,\mathrm{kpc}$ using a mock *Pristine* sample generated with the GALAXIA code. We applied a statistical rescaling based on a sample of *Pristine* follow-up spectroscopy and SDSS/SEGUE stars, to account for contamination of metal-rich stars at the metal-poor end, and a correction to account for the preferential selection of metal-poor stars by the colour cut used to select the distance limited halo sample. At low galactic latitudes ($20\,\mathrm{deg} < |b| < 30\,\mathrm{deg}$) and in the direction of the Galactic anticentre, we identified a metal-rich population at the disc–halo interface, which we concluded to correspond to the Monoceros/TriAnd/ACS/EBS/A13 structure. We also selected a region containing the Virgo overdensity but found no difference in the MDF of this region when compared to other regions of equal Galactic latitude. Comparing the MDF from this work with others derived in the literature, we find a larger fraction of stars at [Fe/H] < -2 than in previously published halo MDFs. In particular, we found a slope of $\Delta(\mathrm{LogN})/\Delta[\mathrm{Fe/H}] = 1.0 \pm 0.1$ in the metallicity range of $-3.4 <$ [Fe/H] < -2.5, which is substantially shallower than previous observational MDFs from spectroscopic samples, but agrees well with the simple closed-box chemical enrichment model of Hartwick (1976), although the model overestimates the number of stars at [Fe/H] < -3.5. We compare the MDF to that provided in Helmi et al. (2018) and Gallart et al. (2019) for the *Gaia*–Enceladus structure, but despite some qualitative similarities we cannot conclude whether or not the halo MDF we present is consistent with a single large merger event dominating the stellar content in the inner halo. We see no clear evidence of duality in the halo MDF, but conclude that we need better distances and cleaner samples of distant giants to probe the inner and outer halo in order to properly address this. Finally, we compare the MDF for Milky Way GCs and the MDF for the halo field stars presented in this work, and find that if these two populations follow the same metallicity distribution, the expected number of observed GCs is 10 at [Fe/H] < -2.5 and 2.8 at [Fe/H] < -3. We therefore suggest that the lack of GCs at [Fe/H] < -2.5 in the Milky Way is due to physical processes, rather than statistical undersampling.

ACKNOWLEDGEMENTS

We would like to thank Søren Larsen and Mark Gieles for useful discussions that motivated the development of this work, as well as Johannes Schuster for his help in improving the manuscript. We gratefully thank the Canada–France–Hawaii Telescope (CFHT) staff for performing the observations in queue mode, for their reactivity in adapting the schedule, and for answering our questions during the data-reduction process. We thank Nina Hernitschek for granting us access to the catalogue of Pan-STARRS variability catalogue.

KY, ES, and AA gratefully acknowledge funding by the Emmy Noether program from the Deutsche Forschungsgemeinschaft (DFG). This work has been published under the framework of the IdEx Unistra and benefits from a funding from the state managed by the French National Research Agency (ANR) as part of the investments for the future program. NFM gratefully acknowledges support from the ANR funded project 'Pristine' (ANR-18-CE31-0017) along with funding from CNRS/INSU through the Programme National Galaxies et Cosmologie and through the CNRS grant PICS07708. KY, ES, AA, PJ, VH, and NFM benefitted from the International Space Science Institute (ISSI) in Bern, CH, thanks to the funding of the Teams 'The Formation and Evolution of the Galactic Halo' and 'Pristine'. JIGH acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities (MICIU) under the 2013 Ramón y Cajal program MICIU RYC-2013-14875, and also from the Spanish Ministry project MICIU AYA2017-86389-P. CL thanks the Swiss National Science Foundation for supporting this research through the Ambizione grant PZ00P2_168065.

Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the CFHT that is operated by the National Research Council (f Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. SDSS-IV is managed by the Astrophysical Research Consortium for the The Pan-STARRS1 Surveys have been made possible through contributions of the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max-Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, Queen's University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under grant NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation under grant AST-1238877, the University of Maryland, and Eotvos Lorand University.

Funding for the SDSS IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard–Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins

University, Kavli Institute for the Physics and Mathematics of the Universe/University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam, Max-Planck-Institut für Astronomie (Heidelberg), Max-Planck-Institut für Astrophysik (Garching), Max-Planck-Institut für Extraterrestrische Physik, National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group,Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

REFERENCES

Abadi M. G., Navarro J. F., Steinmetz M., 2006, MNRAS, 365, 747
Aguado D. S. et al., 2019, MNRAS, 490, 2241
Allende Prieto C. et al., 2014, A&A, 568, A7
An D. et al., 2013, ApJ, 763, 65
An D., Beers T. C., Santucci R. M., Carollo D., Placco V. M., Lee Y. S., Rossi S., 2015, ApJ, 813, L28
Arentsen A. et al., 2020, MNRAS, 491, L11
Beers T. C., Rossi S., Norris J. E., Ryan S. G., Shefler T., 1999, AJ, 117, 981
Beers T. C. et al., 2012, ApJ, 746, 34
Belokurov V., 2013, New Astron. Rev., 57, 100
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Bergemann M. et al., 2018, Nature, 555, 334
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Bonifacio P. et al., 2019, MNRAS, 487, 3797
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Caffau E. et al., 2017, Astron. Nachr., 338, 686
Carollo D. et al., 2007, Nature, 450, 1020
Casagrande L., Wolf C., Mackey A. D., Nordland er T., Yong D., Bessell M., 2019, MNRAS, 482, 2770
Christlieb N., Wisotzki L., Graßhoff G., 2002, A&A, 391, 397
Conroy C., Naidu R. P., Zaritsky D., Bonaca A., Cargile P., Johnson B. D., Caldwell N., 2019, ApJ, 887, 237
Da Costa G. S. et al., 2019, MNRAS, 489, 5900
Dalton G. et al., 2018, in Evans C. J., Simard L., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII. SPIE, Bellingham, p. 107021B
Deason A. J., Mao Y.-Y., Wechsler R. H., 2016, ApJ, 821, 5
D'Souza R., Bell E. F., 2018b, MNRAS, 474, 5300
D'Souza R., Bell E. F., 2018a, Nat. Astron., 2, 737
Frebel A., Norris J. E., 2015, ARA&A, 53, 631
Gallart C., Bernard E. J., Brook C. B., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019, Nat. Astron., 3, 932
Grillmair C. J., 2006, ApJ, 651, L29
Grillmair C. J., Carlin J. L., Majewski S. R., 2008, ApJ, 689, L117
Harmsen B., Monachesi A., Bell E. F., de Jong R. S., Bailin J., Radburn-Smith D. J., Holwerda B. W., 2017, MNRAS, 466, 1491
Harris W. E., 2010, preprint (arXiv:1012.3224)
Hartwick F. D. A., 1976, ApJ, 209, 418
Haywood M., Di Matteo P., Lehnert M. D., Snaith O., Khoperskov S., Gómez A., 2018, ApJ, 863, 113
Helmi A., 2008, A&AR, 15, 145
Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
Hernitschek N. et al., 2016, ApJ, 817, 73
Ibata R. A., Gilmore G., Irwin M. J., 1994, Nature, 370, 194
Ibata R. A. et al., 2014, ApJ, 780, 128
Ibata R. A. et al., 2017, ApJ, 848, 129
Ibata R. A., Malhan K., Martin N. F., 2019, ApJ, 872, 152
Ivezić, Ž. et al., 2008, ApJ, 684, 287
Keller S. C. et al., 2007, Publ. Astron. Soc. Aust., 24, 1
Laevens B. P. M. et al., 2015, ApJ, 813, 44
Koppelman H., Helmi A., Veljanoski J., 2018, ApJ, 860, L11
Kruijssen J. M. D., 2019, MNRAS, 486, L20
Law D. R., Majewski S. R., 2010, ApJ, 714, 229
Li J., Newberg H. J., Carlin J. L., Deng L., Newby M., Willett B. A., Xu Y., Luo Z., 2012, ApJ, 757, 151
Li T. S. et al., 2017, ApJ, 844, 74
Lokhorst D. et al., 2016, ApJ, 819, 124
Longeard N. et al., 2018, MNRAS, 480, 2609
Majewski S. R., Ostheimer J. C., Rocha-Pinto H. J., Patterson R. J., Guhathakurta P., Reitzel D., 2004, ApJ, 615, 738
Malhan K., Ibata R. A., Martin N. F., 2018, MNRAS, 481, 3442
McConnachie A. W. et al., 2018, ApJ, 868, 55
Myeong G. C., Evans N. W., Belokurov V., Sand ers J. L., Koposov S. E., 2018, MNRAS, 478, 5449
Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235
Newberg H. J. et al., 2002, ApJ, 569, 245
Oman K., Starkenburg E., Navarro J., 2017, Galaxies, 5, 33
Pagel B. E. J., Patchett B. E., 1975, MNRAS, 172, 13
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Robertson B., Bullock J. S., Font A. S., Johnston K. V., Hernquist L., 2005, ApJ, 632, 872
Robin A. C., Reylé C., Derrière S., Picaud S., 2003, A&A, 409, 523
Rocha-Pinto H. J., Majewski S. R., Skrutskie M. F., Crane J. D., Patterson R. J., 2004, ApJ, 615, 732
Roederer I. U., Gnedin O. Y., 2019, ApJ, 883, 84
Ryan S. G., Norris J. E., 1991, AJ, 101, 1865
Salvadori S., Schneider R., Ferrara A., 2007, MNRAS, 381, 647
Schönrich R., Asplund M., Casagrande L., 2011, MNRAS, 415, 3807
Schönrich R., Asplund M., Casagrande L., 2014, ApJ, 786, 7
Schörck T. et al., 2009, A&A, 507, 817
Searle L., Sargent W. L. W., 1972, ApJ, 173, 25
Sharma S., Johnston K. V., Majewski S. R., Muñoz R. R., Carlberg J. K., Bullock J., 2010, ApJ, 722, 750
Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, ApJ, 730, 3
Sheffield A. A., Price-Whelan A. M., Tzanidakis A., Johnston K. V., Laporte C. F. P., Sesar B., 2018, ApJ, 854, 47
Starkenburg E. et al., 2017, MNRAS, 471, 2587
Venn K. et al., 2020, MNRAS, 492, 3241
York D. G. et al., 2000, AJ, 120, 1579
Youakim K. et al., 2017, MNRAS, 472, 2963
Zuo W., Du C., Jing Y., Gu J., Newberg H. J., Wu Z., Ma J., Zhou X., 2017, ApJ, 841, 59

APPENDIX A: TABLES OF THE MDFS AND CORRECTIONS

Table A1. Best-fitting parameters for the GMM shown in the right-hand panel of Fig. 6.

Component	Mean (x, y)	Weight	Covariance matrix $\begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$
1	(−1.320, −1.501)	0.28	$\begin{bmatrix} 0.183 & 0.113 \\ 0.113 & 0.188 \end{bmatrix}$
2	(−1.893, −2.121)	0.39	$\begin{bmatrix} 0.232 & 0.197 \\ 0.197 & 0.222 \end{bmatrix}$
3	(−2.448, −2.815)	0.33	$\begin{bmatrix} 0.251 & 0.033 \\ 0.033 & 0.127 \end{bmatrix}$

Table A2. Correction factor applied for the colour cut to select the halo sample. The rightmost columns is the correction factor applied to produce the black histogram in Fig. 12 and is the mean of the 10 and 12 Gyr columns.

[Fe/H]	8 Gyr	10 Gyr	12 Gyr	14 Gyr	Correction
−3.95	1.000	1.000	1.000	1.000	1.000
−3.85	1.000	1.000	1.000	1.000	1.000
−3.75	1.000	1.000	1.000	1.000	1.000
−3.65	1.000	1.000	1.000	1.000	1.000
−3.55	1.000	1.000	1.000	1.000	1.000
−3.45	1.000	1.000	1.000	1.000	1.000
−3.35	1.000	1.000	1.000	1.000	1.000
−3.25	1.000	1.000	1.000	1.000	1.000
−3.15	1.009	1.000	1.011	1.031	1.006
−3.05	1.007	1.003	1.003	1.027	1.003
−2.95	1.006	1.012	1.000	1.011	1.006
−2.85	1.009	1.009	1.015	1.018	1.012
−2.75	1.000	1.016	1.024	1.006	1.020
−2.65	1.005	1.007	1.037	1.000	1.022
−2.55	1.007	1.017	1.037	1.014	1.027
−2.45	1.010	1.034	1.048	1.041	1.041
−2.35	1.012	1.049	1.079	1.059	1.064
−2.25	1.024	1.058	1.085	1.112	1.072
−2.15	1.050	1.069	1.096	1.141	1.083
−2.05	1.083	1.090	1.131	1.191	1.111
−1.95	1.111	1.127	1.180	1.258	1.154
−1.85	1.129	1.160	1.242	1.330	1.201
−1.75	1.156	1.193	1.303	1.429	1.248
−1.65	1.212	1.239	1.356	1.594	1.298
−1.55	1.267	1.297	1.492	1.835	1.395
−1.45	1.315	1.403	1.679	2.054	1.541
−1.35	1.391	1.531	1.932	2.338	1.732
−1.25	1.479	1.689	2.137	5.174	1.913
−1.15	1.635	2.029	2.869	–	2.449
−1.05	1.841	2.411	5.600	–	4.006
−0.95	2.209	2.658	–	–	–
−0.85	2.471	10.235	–	–	–
−0.75	3.101	–	–	–	–
−0.65	20.823	–	–	–	–
−0.55	–	–	–	–	–
−0.45	–	–	–	–	–
−0.35	–	–	–	–	–
−0.25	–	–	–	–	–
−0.15	–	–	–	–	–
−0.05	–	–	–	–	–

Table A3. MDFs for the raw *Pristine* halo sample, the sample corrected with the GMM, and the sample corrected with the GMM and colour cuts, plotted in Fig. 12 as the grey, cyan, and black histograms, respectively.

[Fe/H]	Raw *Pristine*	GMM	Corrected
−3.95	395	10	10
−3.85	143	34	34
−3.75	170	59	59
−3.65	294	104	104
−3.55	500	149	149
−3.45	657	192	192
−3.35	787	353	353
−3.25	1037	466	466
−3.15	1189	625	628
−3.05	1458	765	767
−2.95	1650	984	989
−2.85	1999	1241	1255
−2.75	2235	1491	1520
−2.65	2516	1808	1847
−2.55	2845	2098	2154
−2.45	3186	2531	2634
−2.35	3521	2816	2996
−2.25	3770	3142	3366
−2.15	4116	3596	3892
−2.05	4503	3910	4342
−1.95	4635	4486	5174
−1.85	4778	4730	5680
−1.75	4798	4951	6178
−1.65	4650	5054	6557
−1.55	4409	5086	7092
−1.45	4011	4925	7589
−1.35	3415	4529	7841
−1.25	2865	4124	7889
−1.15	2259	3588	8787
−1.05	1816	3101	12421
−0.95	1326	2477	–
−0.85	1051	1875	–
−0.75	763	1426	–
−0.65	538	997	–
−0.55	415	779	–
−0.45	310	483	–
−0.35	264	331	–
−0.25	217	200	–
−0.15	218	130	–
−0.05	–	63	–

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