Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **490,** 3426–3439 (2019)
Advance Access publication 2019 October 7

doi:10.1093/mnras/stz2793

The total stellar halo mass of the Milky Way

Alis J. Deason ,[1★] Vasily Belokurov [2] and Jason L. Sanders [2]

[1] *Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[2] *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

Accepted 2019 October 1. Received 2019 September 30; in original form 2019 August 7

ABSTRACT

We measure the total stellar halo luminosity using red giant branch (RGB) stars selected from *Gaia* data release 2. Using slices in magnitude, colour, and location on the sky, we decompose RGB stars belonging to the disc and halo by fitting two-dimensional Gaussians to the Galactic proper motion distributions. The number counts of RGB stars are converted to total stellar halo luminosity using a suite of isochrones weighted by age and metallicity, and by applying a volume correction based on the stellar halo density profile. Our method is tested and calibrated using *Galaxia* and *N*-body models. We find a total luminosity (out to 100 kpc) of $L_{\rm halo} = 7.9 \pm 2.0 \times 10^8\,{\rm L}_\odot$ excluding Sgr, and $L_{\rm halo} = 9.4 \pm 2.4 \times 10^8\,{\rm L}_\odot$ including Sgr. These values are appropriate for our adopted stellar halo density profile and metallicity distribution, but additional systematics related to these assumptions are quantified and discussed. Assuming a stellar mass-to-light ratio appropriate for a Kroupa initial mass function ($M^\star/L = 1.5$), we estimate a stellar halo mass of $M^\star_{\rm halo} = 1.4 \pm 0.4 \times 10^9\,{\rm M}_\odot$. This mass is larger than previous estimates in the literature, but is in good agreement with the emerging picture that the (inner) stellar halo is dominated by one massive dwarf progenitor. Finally, we argue that the combination of a $\sim 10^9\,{\rm M}_\odot$ mass and an average metallicity of $\langle[{\rm Fe/H}]\rangle \sim -1.5$ for the Galactic halo points to an ancient ($\sim$10 Gyr) merger event.

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

1 INTRODUCTION

The halo of our Galaxy is littered with the stellar debris of destroyed dwarf galaxies. This trash can of material extends out to several hundred kiloparsecs, and gives important insight into the assembly history of the Milky Way and its dark matter potential. Moreover, the remains of the destroyed dwarfs can tell us about the properties of the lowest mass galaxies in the Universe.

The content, size, extent, and kinematics of the stellar halo has been studied extensively over the past few decades (see reviews by Helmi 2008; Belokurov 2013). In particular, the number counts of old and relatively metal-poor stars have revealed that the density profile of the stellar halo approximately follows a power-law with index ~ -2.5 within 20 kpc, and then falls-off more rapidly thereafter, with power-law index ~ -4.0 (e.g. Watkins et al. 2009; Sesar et al. 2010; Deason, Belokurov & Evans 2011; Faccioli et al. 2014; Pila-Díez et al. 2015). Note, however, that the form of the density profile at larger distances (>40–50 kpc) is still highly uncertain (e.g. Deason et al. 2014; Xue et al. 2015; Slater et al. 2016; Hernitschek et al. 2018). The change in density at $\sim$20 kpc profile signifies a transition between the 'inner' and 'outer' halo. Deason et al. (2013) argued that this broken profile is caused by the accretion of a massive dwarf galaxy at early times. In this scenario, the break radius represents the last apocentre of the accreted dwarf, and beyond this furthest point of the orbit, the contribution of the debris from this massive dwarf is significantly diminished. Thus, this picture suggests that the inner stellar halo is dominated by one massive accretion event, while the outer halo is a dusting of several (lower mass) destroyed dwarfs.

The arrival of the *Gaia* (Gaia Collaboration 2016a) data releases (Gaia Collaboration 2016b, 2018a), which provide six-dimensional phase-space measurements for thousands of local halo stars and proper motion measurements for hundreds of thousands of halo stars, reinvigorated our ideas about the structure of the halo and confirmed the insight we gained from the halo number counts. In particular, Belokurov et al. (2018), Haywood et al. (2018), and Helmi et al. (2018) used a combination of kinematical and chemical data from *Gaia*, SDSS, and APOGEE to find that the inner halo is indeed dominated by one massive accretion event, which occurred >8 Gyr ago. This significant event in the history of the Galaxy has been dubbed the *Gaia*-Sausage (aptly named due to its highly eccentric orbit) or *Gaia*-Enceladus. Follow-up studies have added extra weight to the growing consensus that the *Gaia*-Sausage rules the (inner) halo: for example, Deason et al. (2018) and Lancaster et al. (2019) used the kinematics of distant halo stars to dynamically show the transition at $\sim$20 kpc between the 'Sausage' dominated

★ E-mail: alis.j.deason@durham.ac.uk

regime and the outer halo, and Myeong et al. (2018), Myeong et al. (2019), and Massari, Koppelman & Helmi (2019) used the dynamics of the globular cluster population in action space to show that many are likely related to the *Gaia*-Sausage, as expected if this is a massive merger event.

As mentioned above, the density profile of halo stars has proved an invaluable measure to constrain the Galaxy's assembly history. However, the normalization of the density profile, and hence the total stellar halo mass, has proven to be more complicated to measure. This is mainly because the tracers we often use to map the halo star distribution out to large distances, i.e. the RR Lyrae and blue horizontal branch stars, are difficult to relate to the *overall* halo population. This is because the exact broad-brand colour (and hence temperature) distribution of the helium burning stars depends on additional 'hidden' parameters (see e.g. Gratton et al. 2010).

Moreover, it is difficult to provide a robust normalization when using surveys that have non-trivial selection functions and/or are limited in their spatial extent. Most measures of the total stellar halo density are limited to local halo star samples, and a wide range of density normalizations have been quoted in the literature: $\rho_0 = 3.0$–$15.0 \times 10^{-5}\,\mathrm{M_{\odot}\,pc^{-3}}$ (e.g. Morrison 1993; Fuchs & Jahreiß 1998; Gould, Flynn & Bahcall 1998; Digby et al. 2003; Jurić et al. 2008; de Jong et al. 2010). Many of these measures were estimated before the density profile out to large distances was known, and hence relating the local stellar density to the total stellar halo mass is non-trivial. More recently, Bell et al. (2008) estimated the total stellar mass using main-sequence turn-off stars in SDSS, and Deason et al. (2011) used counts of blue horizontal branch stars in SDSS. Both these studies favour relatively low stellar halo masses $M^{\star}_{\rm halo} \sim 3$–$4 \times 10^8\,\mathrm{M_{\odot}}$, but there is sizeable uncertainty relating these tracer populations to the overall stellar halo (see above). In addition, these measures do not include the few $\sim 10^8\,\mathrm{M_{\odot}}$ substructures in the halo, which also contribute to the mass, so the total mass, based on the Bell et al. (2008) and Deason et al. (2011) estimates for the 'field' halo, is in the range $M^{\star}_{\rm halo} \sim 4$–$7 \times 10^8\,\mathrm{M_{\odot}}$ (cf. Bland-Hawthorn & Gerhard 2016).

Currently, different estimates of the Galactic stellar halo mass vary by a factor of 2, but, more worryingly, the uncertainty of these estimates is not robustly quantified. Perhaps more puzzling is that the recent deluge of evidence for a massive accretion event dominating the stellar halo, appears at odds with the rather low value of total stellar halo mass quoted in the literature. Rectifying this apparent conundrum is crucial in order to place the Milky Way in the cosmological context with other, similar mass galaxies. Both simulations and observations show that at fixed galaxy (or halo) mass the range of stellar halo masses is large, reflecting a wide diversity of assembly histories (e.g. Pillepich et al. 2014; Merritt et al. 2016; Harmsen et al. 2017; Elias et al. 2018; Cañas et al. 2019; Monachesi et al. 2019). Moreover, work by Deason, Mao & Wechsler (2016) and D'Souza & Bell (2018) show that the stellar halo mass is critically linked to the most massive dwarf progenitor of the halo. Thus, in order to reconcile several global properties of the Milky Way halo (e.g. density profile, metallicity) with the currently favoured assembly history scenario, it is imperative that we procure a robust total stellar halo mass, complete with a well-defined uncertainty.

In this paper, we utilize the exquisite data from *Gaia* to estimate the total stellar halo luminosity and mass using red giant branch (RGB) stars. Compared to previous work, we take advantage of the full sky coverage of the *Gaia* survey, and use the proper motion distributions to decompose disc and halo populations. In Section 2, we describe the selection of RGB stars, and introduce the models that we use to guide and calibrate our analysis. Number counts of halo stars are estimated in bins of colour, magnitude, and area on the sky, and our process for decomposing the disc and halo populations is described in Section 3. In Section 4, we determine the normalization per halo tracer from stellar population models, and volume correct the number counts in order to estimate the total stellar halo luminosity. We also quantify how well this procedure performs on *N*-body stellar halo models. We discuss our resulting stellar halo mass in Section 5, and summarize our main conclusions in Section 6.

2 HALO RED GIANT BRANCH STARS

Our aim is to use counts of halo RGB stars to estimate the total stellar halo luminosity. RGB stars are ideal tracers for this purpose as they are intrinsically bright, relatively numerous, and are present at all ages and metallicities. Moreover, we are able to cleanly select RGB stars using *Gaia* data (see below). In order to guide us through the stellar halo selection and luminosity estimate, we make use of 'toy' models of the Galaxy, which are tailored towards the *Gaia* data release 2 (GDR2) astrometry and photometry.

2.1 Galaxia and N-body models

We use the *Galaxia* model (Sharma et al. 2011) to create a synthetic survey of the Milky Way. We choose the default (analytical) *Galaxia* model for the disc population (the Besançon model, Robin et al. 2003), and the Bullock & Johnston (2005, hereafter BJ05) *N*-body models for the stellar halo. *Galaxia* employs a scheme to sample the *N*-body models, which ensures that the phase-space density of the generated stars is consistent with that of the *N*-body particles. There are 11 stellar halo models, each representing a different assembly history and stellar halo mass. This suite of simulated stellar haloes have been used extensively in the literature (e.g. Bell et al. 2008; Xue et al. 2011; Deason et al. 2013), and although there may be limitations relative to more sophisticated cosmological simulations, they are an incredibly useful tool for testing and calibrating observational survey data.

The BJ05 models only include halo stars from accreted dwarf galaxies, there are no halo stars born '*in situ*' in the parent halo, as predicted by cosmological hydrodynamic simulations (e.g. Zolotov et al. 2009; Font et al. 2011). However, if this population does exist (this is still not clear in the Milky Way: Deason et al. 2017; Belokurov et al. 2018; Di Matteo et al. 2018; Haywood et al. 2018) it is likely confined to the inner halo and will have similar properties to the thick disc (Zolotov et al. 2009; McCarthy et al. 2012; Pillepich et al. 2014; Belokurov et al. 2019; Gallart et al. 2019). Thus, in our decomposition of halo/disc populations (see Section 3) any *in situ* halo stars will likely be labelled as disc stars. However, we cannot exclude the possibility that some fraction of the stellar halo mass we compute in the *Gaia* data has an *in situ* origin. This is discussed further in Section 5.

A synthetic survey is produced from the models in Johnson–Cousins bandpasses and converted to the *Gaia* photometry using the relations given in Jordi et al. (2010). Uncertainties in photometry and astrometry applicable to GDR2 are also included in the model. This is implemented using the PYTHON PYGAIA package.[1] This

[1] https://pypi.org/project/PyGaia/

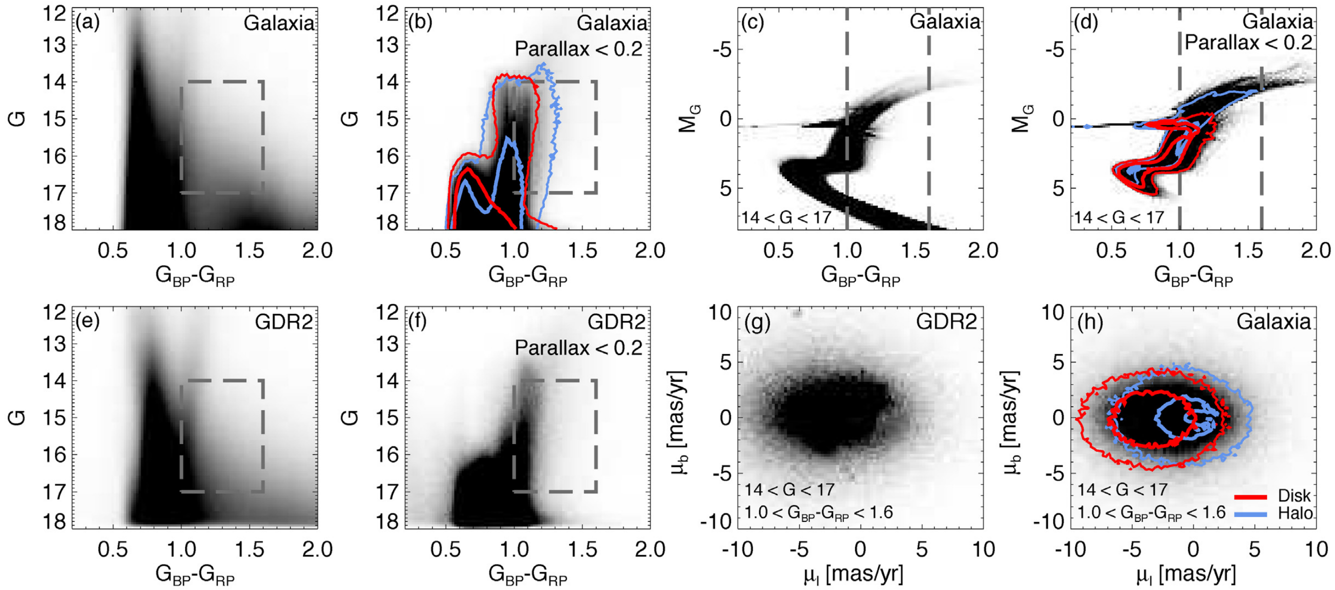

Figure 1. Colour–magnitude diagrams (CMDs) and proper motion distributions for the *Galaxia* models and GDR2. In all panels, only stars with high latitude ($|b| > 30°$) are shown. *Panel (a):* Apparent magnitude versus colour for stars in *Galaxia*. Here, the halo component is from an *N*-body model (Halo-7, see main text). The dashed lines indicate the colour range used in this work to select red giant branch stars. Photometric and astrometric errors applicable to *Gaia* data release 2 have been applied to the model. *Panel (b):* Apparent magnitude versus colour for stars in *Galaxia* with small parallax (parallax <0.2). This cut removes nearby dwarfs. The red and blue contours indicate the disc and halo stars, respectively. *Panel (c):* Absolute magnitude versus colour for stars in *Galaxia*. Here, stars with apparent magnitudes $14 < G < 17$ are shown. *Panel (d):* Absolute magnitude versus colour for stars in *Galaxia* with small parallax. *Panel (e):* Apparent magnitude versus colour for stars in GDR2. Note stars in close proximity to the Magellanic Clouds have been removed. *Panel (f):* Apparent magnitude versus colour for stars in GDR2 with small parallax. *Panel (g):* Proper motions of stars in GDR2 in Galactic coordinates. Here, we only consider stars with parallax <0.2, $1.0 < G_{\rm BP} - G_{\rm RP} < 1.6$ and $14 < G < 17$. *Panel (h):* Proper motions in *Galaxia* (with same selection as GDR2). The disc and halo components have distinct, but overlapping, proper motion distributions. These sequences vary across the sky (see Fig. 3).

module implements the performance models for *Gaia* which are publicly available.[2]

In the top panels (a)–(d) of Fig. 1 we show colour–magnitude diagram (CMDs) for high-latitude ($|b| > 30°$) stars in the *Galaxia* model. Panels (a) and (b) show apparent magnitude versus colour, and panels (c) and (d) show absolute magnitude versus colour (with apparent magnitude restricted to $14 < G < 17$). The dashed lines indicate the colour range we consider for candidate RGB stars. In panels (b) and (d), we exclude stars with parallax >0.2 (approx. $D < 5$ kpc). This cut removes nearby dwarf stars, but there are still disc giants present. We indicate the disc and halo populations with the red and blue contours, respectively. We have checked that the completeness of the halo star sample is not significantly affected by the parallax cut. We find that, for the magnitude and colour range under consideration, the halo stars with $D > 5$ kpc are complete to ≥90 per cent. Our selection of RGB stars, based on magnitude, colour, and parallax, spans the distance range $D \sim$ 5–100 kpc. In panel (h) of Fig. 1 we show the proper motions of the RGB stars in Galactic coordinates (μ_ℓ, μ_b). Here, we only consider stars with parallax <0.2, $1.0 < G_{\rm BP} - G_{\rm RP} < 1.6$ and $14 < G < 17$. The disc and halo components are indicated with the red and blue contours, respectively. The disc and halo components have distinct, but overlapping, proper motion distributions. Here, we are showing all stars across the sky, but these sequences vary depending on the Galactic coordinates (see Fig. 3). This figure shows that the proper motion distributions of RGB stars can be used to disentangle the disc and halo populations. In Section 3, we use the proper motion distributions to estimate the number of RGB stars in the halo in bins of colour, magnitude, and position on the sky.

In panels (e)–(g) of Fig. 1 we show the equivalent CMDs (using only apparent magnitude) and proper motions of the GDR2 data (see below).

2.2 Gaia DR2

The models described in the previous subsection are tailored towards the GDR2 data set. Before going further, we briefly describe our selection of the real *Gaia* data. We select stars from GDR2 with photometry, parallax, and proper motion information. The photometry is corrected for extinction using the Schlegel, Finkbeiner & Davis (1998) dust maps, and we use the relations given in Gaia Collaboration (2018b) to correct the G, $G_{\rm RP}$, and $G_{\rm BP}$ bandpasses. We only include stars with renormalized unit weight error, RUWE < 1.4 (Lindegren 2018), which ensures stars with unreliable astrometry are excluded. In addition, we exclude stars with large BP/RP flux excess using the selection given in Gaia Collaboration (2018b). These cuts remove ∼8 per cent of the sample in the colour, magnitude, and latitude range under consideration (see below). Note most of the star excised are at the fainter, redder region of our selection. We assume that these quality cuts affect the disc and halo populations equally, and thus increase our estimated luminosity (and mass) estimate of the Milky Way (see Section 4) by 8 per cent. From the cleaned sample, we select RGB stars at high latitude ($|b| > 30°$) with parallax <0.2, $1.0 < G_{\rm BP} - G_{\rm RP} < 1.6$ and $14 < G < 17$ (see Fig. 1).

In the following Section, we decompose the disc and halo RGB stars using proper motion information. First, we illustrate this

[2] http://www.cosmos.esa.int/web/gaia/science-performance

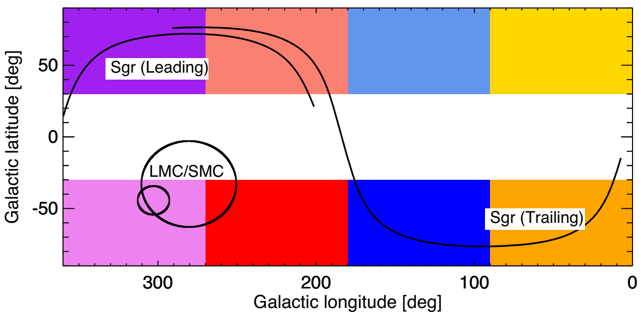

Figure 2. Slices in Galactic longitude and latitude used to fit the disc/halo components. Each bin is fitted separately. The colour coding indicated is adopted throughout the paper. The Sgr leading and trailing arms are shown. When Sgr is excluded, stars lying within 12 deg of these tracks are omitted. Stars in close proximity to the LMC and/or SMC are excluded in our analysis.

process using the *Galaxia* models, and we then apply the technique to our GDR2 sample.

3 DISC–HALO DECOMPOSITION

In Fig. 1, we showed that our selection of RGB stars includes both halo and disc populations. In order to disentangle these populations, we use the two-dimensional proper motion distributions. We assume 2D (for each component of proper motion) Gaussian distributions for both the halo and disc. This Gaussian approximation is reasonable as we (independently) fit in bins of magnitude, colour, and position on the sky, rather than fit the entire distribution with one 2D Gaussian. We use six bins in magnitude (between $14 < G < 17$), six bins in colour (between $1.0 < G_{\rm BP} - G_{\rm RP} < 1.6$), and eight spatial bins. The spatial bins are shown in Fig. 2. When applying this method to the *Gaia* data, we exclude stars within 30 deg of the Large Magellanic Cloud (LMC) and 10 deg of the Small Magellanic Cloud (SMC). We also perform the analysis both with and without stars in the vicinity of the Sagittarius (Sgr) stream. The Sgr stars are selected to lie within 12 deg of the tracks shown in Fig. 2 (see Deason et al. 2012; Belokurov et al. 2014). Note that when we use the BJ05 stellar halo models we do not attempt to excise any streams or satellites, so the biases from unrelaxed substructures are likely more pronounced in the models than the data.

In Fig. 3, we show the true Gaussian parameters for the disc and halo populations in the *Galaxia* model. For this illustration the halo component is Halo-7, although similar trends are seen in all of the haloes. This figure shows that the overlap between the disc and halo components varies as a function of magnitude, colour, and position on the sky. In some cases the overlap is larger, and in others the populations are more distinct. To perform the fits simultaneously (i.e. without knowing which stars belong to disc or halo), we model the proper motion distributions with a mixture of two (halo + disc) multivariate Gaussians using the Extreme Deconvolution algorithm described in Bovy, Hogg & Roweis (2011). In Fig. 4, we show the outcome of these fits for the *Galaxia* model. Note that we initialize the fits using the true Gaussian values for the disc and a halo model (Halo-7 in this case). This step is taken to avoid misclassification of the halo/disc components. However, we check that initializing with different halo profiles or an independent disc model makes little difference to the results (see later). Fig. 4 shows that in some bins the decomposition works well, while in others we are unable to clearly

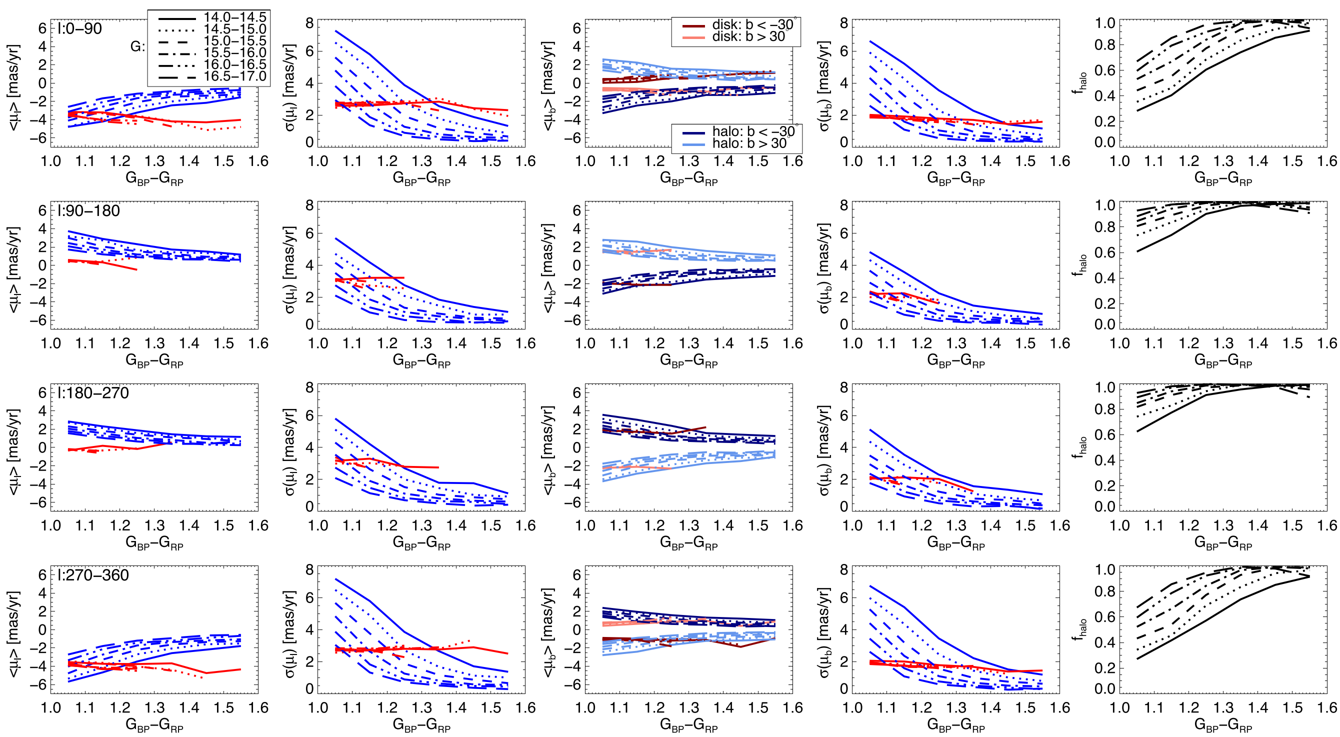

Figure 3. The mean (first and third panels) and dispersion (second and fourth panels) of the *Galaxia* model proper motions in Galactic coordinates as a function of $G_{\rm BP} - G_{\rm RP}$ colour. Blue and red lines indicates the halo and disc components, respectively. Different magnitude bins are shown with different linestyles and each row shows a different bin in Galactic longitude. The sequences are very similar for bins above and below the Galactic plane, except for $\langle\mu_b\rangle$ (third column), which we indicate with different shades of blue and red. The last panel on the right indicates the fraction of halo stars as a function of colour.

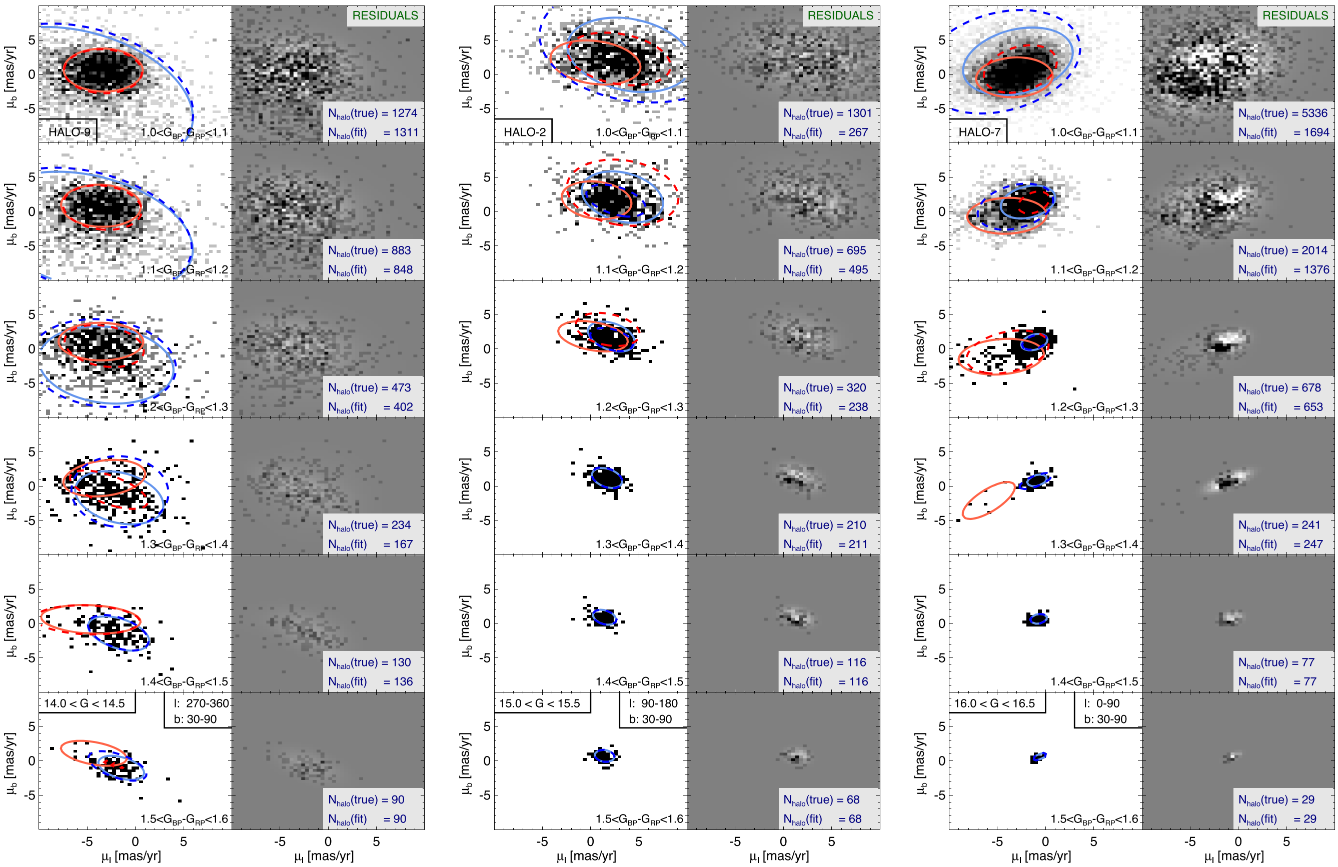

Figure 4. *Left columns:* Extreme deconvolution (XD) fits to the Galaxia proper motion distributions in bins of $G_{BP} - G_{RP}$ colour. The solid red and blue lines show the true disc and halo distributions, and the dashed lines show the XD fits. Here, we use the true Galaxia model values to initialize the XD fit. *Right columns:* The residuals from the fit. Here, the pixel size is 0.4 mas yr^{-1}, and the shading saturates at an excess of $\Delta = \pm 5$. The true and estimated number of halo stars is given in the bottom right hand corner. Three examples are shown for different magnitude ranges and bins on the sky.

distinguish the distinct components. We note the true and fitted halo amplitudes (number of halo stars) in the bottom right corner of the panels. Bins at redder colours and fainter magnitudes have little, if any, disc component so the fits are straightforward. However, even with a significant disc contribution (e.g. at bluer colours and brighter magnitudes) we can sometimes get good estimates of the halo amplitudes.

The reliability of the decomposition for each bin is illustrated in Fig. 5. Here, we show the relative difference between the estimated and true number of halo stars. We have combined results from all 11 BJ05 haloes and show the median and 16/84 percentiles. In certain bins, our estimates are poor (over/underestimate by more than 30 per cent) and these are shown with the black crosses. These are cases where the overlap between disc and halo makes decomposition based on proper motion alone very difficult. However, in most of the bins (70 per cent) we are able to recover the true number of halo stars to within 30 per cent. When we apply this method to the *Gaia* data we can exclude the bins with significant systematics.

In Fig. 6, we show examples of the 2D Gaussian fits to the *Gaia* data. These example bins are the same as in Fig. 4. We show the more general results in Fig. 7. Here, we can see the resulting Gaussian parameters behave similarly to the model predictions (shown in Fig. 3). We note that a noise component becomes apparent in the faintest bins ($16.0 < G < 17$), which is labelled as 'disc'. This component is relatively minor, as the number of stars belonging to the disc in the faint, red bins is very low ($N \lesssim 50$). Moreover, in all bins, the halo component appears to be well behaved, which gives us confidence that our estimated halo amplitudes are reasonable.

Figs 6 and 7 also help us evaluate one of the assumptions we have made in our modelling: that the proper motion is a reliable distance indicator. In essence, we are using proper motion to disentangle distant halo stars and nearby disc stars. However, populations such as the thick disc or *in situ* halo could potentially break this decomposition if their proper motion distributions mimic the (accreted) halo. The results of the decomposition give us confidence that this is not the case. First, the general trends seen in Fig. 7 look similar to the model predictions shown in Fig. 3. Note the agreement is even better when we compare with the 'fitted' values for the model, rather than the 'true' values. This agreement is non-trivial: it shows that the inferred halo population in GDR2 resembles the accreted halo population in the models. If thick disc or *in situ* halo stars were contaminating the results, the proper motions distributions would be inflated (because these stars are closer), and would not necessarily narrow with colour and magnitude, as seen in Fig. 7. Second, the redder bins (see e.g. lower panels of Fig. 6) appear to be almost entirely comprised of very distant stars with small proper motions. If a significant fraction of thick disc or *in situ* halo stars were contaminating these bins, the proper motion distributions would be much broader. However, we caution that we cannot exclude the possibility that our halo sample includes any

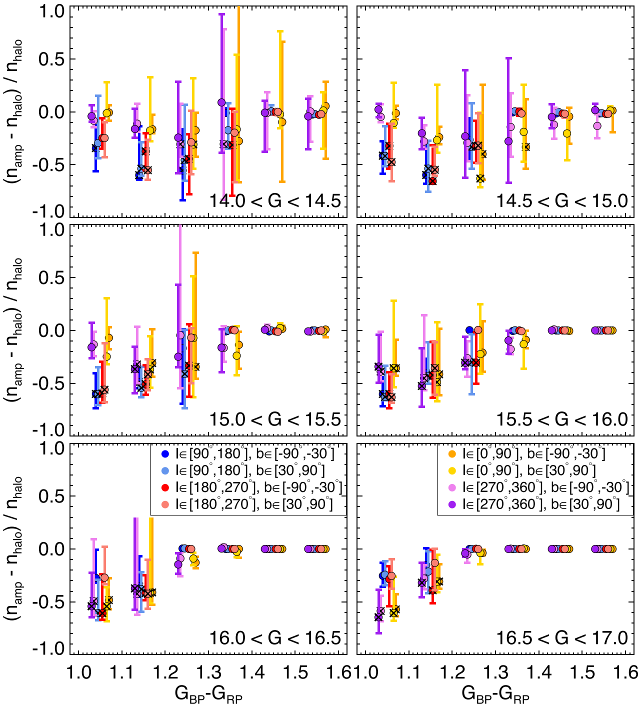

Figure 5. The estimated number of halo RGB stars in the *Galaxia* models from the XD fitting ($n_{\rm amp}$) relative to the true number ($n_{\rm true}$) as a function of colour. Here, we have combined results from all 11 BJ05 haloes and show the median and 16/84 percentiles. Each panel indicates a different magnitude bin. The coloured filled circles indicate the (eight) bins in Galactic longitude and latitude. The colour scheme is given in the legend (also shown in Fig. 2). Bins where the amplitude is underestimated or overestimated by more than 30 per cent, are shown with black crosses. In these cases, the disc and halo are difficult to distinguish, and we can exclude these bins in our analysis. However, in most of the bins (70 per cent) we are able to recover the true number of halo stars to within 30 per cent.

in situ material, particularly if these stars can reach out to large distances. This is discussed further in Section 5.

To provide error estimates on the number of halo RGB stars in each bin, we perform the fits $N = 100$ times. Before each fit, we scatter the parallax according to the error distribution and then make a cut of parallax <0.2. This step essentially adds/removes stars from the analysis with parallax close to the limiting threshold. In addition, we randomly select one of the 11 BJ05 haloes to initialize the fits. As a final check, we initialize the disc parameters using a completely independent model to *Galaxia*. For this we use the disc model described in Sanders & Binney (2015). This model has an action distribution that varies smoothly with age and metallicity using analytic prescriptions for dynamical heating, radial migration, and the radial enrichment of the interstellar medium over time. A mock catalogue of on-sky position, magnitude, age, metallicity, mass, and velocities was generated using Markov Chain Monte Carlo sampling (Foreman-Mackey et al. 2013) of the model combined with a set of PARSEC isochrones. We require samples to have $14 < G < 17$ and $1 < (G_{\rm BP} - G_{\rm RP}) < 1.6$, and convolve the output samples in parallax, proper motion, and magnitudes using nearest neighbours in magnitude and on-sky position from GDR2. We find that, after initializing the disc component using the Sanders & Binney (2015) model, the resulting halo amplitudes are very similar and do not significantly affect our derived luminosity (see following section).

4 TOTAL STELLAR HALO LUMINOSITY

In the previous Section, we calculated the number of halo RGB stars in bins of colour, magnitude, and regions on the sky. We now want to convert these numbers into an estimate of the total stellar halo luminosity (and hence stellar mass). We provide a luminosity estimate for each bin, by applying the following two corrections:

(i) *Stellar population correction:* We use isochrones to relate the number of RGB stars in a given colour bin to the total luminosity. Here, we use the PARSEC isochrones (Bressan et al. 2012), with metallicities in the range $-2.5 <$ [M/H] <0.0 and ages 10–14 Gyr. These isochrones are solar scaled, but halo stars are alpha enhanced with [α/Fe] $\sim$ 0.3 (e.g. Venn et al. 2004). Hence, we use the relation given by Salaris & Cassisi (2005) to relate [M/H] to [Fe/H]: [M/H] = [Fe/H] + 0.2 for [α/Fe] = 0.3. For each isochrone, we calculate the number of RGB stars per unit luminosity as a function of colour. We adopt the PARSEC isochrones as our 'fiducial' stellar population model (these are also the models used in *Galaxia*) and we comment on the changes to our results when other models are used in Section 5. For each of our six colour bins (with 0.1 dex width) we calculate $N_{\rm RGB}/{\rm L}_\odot$. This procedure requires us to assume an initial mass function (IMF)

$$\frac{N_{{\rm RGB},i}}{{\rm L}_\odot} = \frac{\int_{m_1}^{m_2} \xi(m)\,{\rm d}m}{\int_{m_{\rm min}}^{m_{\rm max}} L\,\xi(m)\,{\rm d}m}, \qquad (1)$$

where $\xi(m)$ is the IMF and i denotes the isochrone. The limits m_1 and m_2 give the mass range probed by a particular colour bin and $m_{\rm min}, m_{\rm max}$ denotes the full range of masses probed by the isochrone. Note that the luminosity estimate is only weakly dependent on the IMF, as most of the commonly used IMF parametrizations are very similar for the high-mass stars, which dominate the stellar light. In comparison, the stellar mass strongly depends on the adopted IMF, as the uncertainty of the mass function for low-mass stars, which dominate the mass, is significant. It is for this reason that we provide a robust estimate of total stellar luminosity, rather than mass. This luminosity can later be converted to stellar mass using the appropriate stellar mass-to-light ratio for a given IMF (see Section 5). For the *Galaxia* models we use a Chabrier IMF (Chabrier 2003, as assumed for the halo's *N*-body component in this model), and we use the Kroupa IMF (Kroupa 2001) when estimating the Milky Way halo luminosity using *Gaia* data. In practice, these IMFs are comparable and give very similar luminosity (and mass) estimates.

We next convert $N_{{\rm RGB},\,i}/{\rm L}_\odot$ for each isochrone into an overall estimate by weighting the isochrones using a metallicity distribution function (MDF) and age distribution. For the *Galaxia* models, we fit a Gaussian to the true MDF of the halo, and for the *Gaia* data we adopt an MDF from the literature with $\langle$[Fe/H]$\rangle = -1.5$, σ([Fe/H]) = 0.5 (An et al. 2013; Zuo et al. 2017). For the ages, we assume a uniform age distribution in the range 10–14 Gyr. The top panel of Fig. 8 shows the resulting (weighted) $N_{\rm RGB}/{\rm L}_\odot$ for each colour bin. The error bars indicate the 16/84 percentiles given the adopted MDF and age distribution. We now have a way to relate total number of RGB stars in a colour bin to the luminosity. However, our estimates from the previous section are in bins of magnitude and area on the sky, and thus each probe a different volume of the halo. Thus, the final correction is to volume correct each bin.

(ii) *Volume correction:* Each bin in magnitude, colour, and region of the sky probes a different volume of the halo. Thus, to convert our estimated number of halo RGB stars to total number of RGB stars we need to volume correct. This requires adopting a density profile

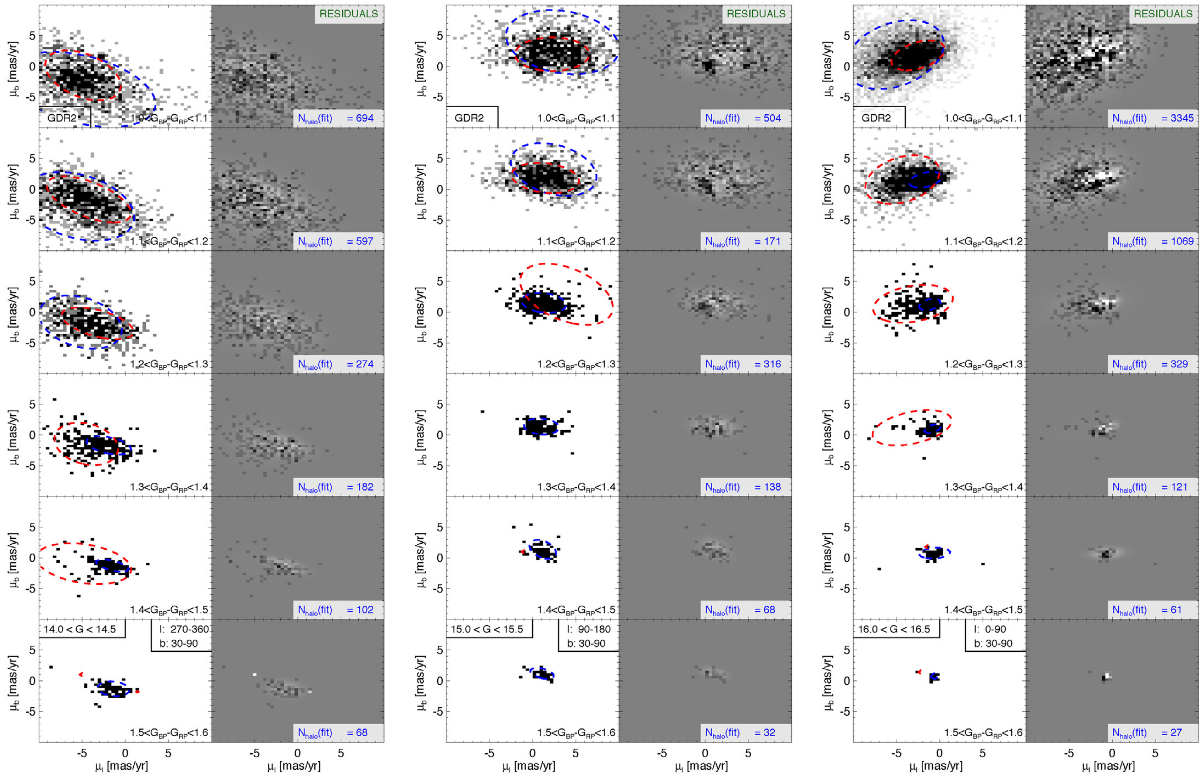

Figure 6. *Left columns:* Extreme deconvolution (XD) fits to the GDR2 proper motion distributions in bins of $G_{\rm BP} - G_{\rm RP}$ colour. The dashed red and blue lines show the estimated disc and halo distributions. Here, we use the Galaxia model values to initialize the XD fit. *Right columns:* The residuals from the fit. Here, the pixel size is 0.4 mas yr^{-1}, and the shading saturates at an excess of $\Delta = \pm 5$. The estimated number of halo stars is given in the bottom right hand corner. Three examples are shown for different magnitude ranges and bins on the sky.

for the stellar halo. This has been measured for the Milky Way in previous work, and we adopt an Einasto profile when applying to the *Gaia* data, with $n = 1.7$, $R_{\rm e} = 20$ kpc and minor-to-major axis ratio $q = 0.6$ (Deason et al. 2011). For the *Galaxia* models we, fit an Einasto profile directly to the halo stars out to 100 kpc. For all 11 haloes, the values typically lie in the range: $n = 1$–5, $R_{\rm e} = 15$–40 kpc, and $q = 0.4$–0.8. Our volume correction relates the volume probed by each bin to the total volume, which we assume goes out to 100 kpc. Hence, our luminosity estimates are within 100 kpc, although this is more or less identical to the total luminosity as there is very little stellar halo mass beyond 100 kpc. We use the PARSEC isochrones to calculate the distance range probed in each bin and by adopting a density profile this can be converted into a volume

$$\frac{\text{Vol, i}}{\text{Total Vol}} = \frac{\int_{D_1}^{D_2}\int_{\ell_1}^{\ell_2}\int_{b_1}^{b_2} \rho(D,\ell,b)\, D^2 \cos(b)\, {\rm d}D{\rm d}\ell{\rm d}b}{\int_{D=0{\rm kpc}}^{D=100{\rm kpc}}\int_{\ell=0^\circ}^{\ell=360^\circ}\int_{b=-90^\circ}^{b=90^\circ} \rho(D,\ell,b)\, D^2 \cos(b)\, {\rm d}D{\rm d}\ell{\rm d}b}, \quad (2)$$

where i denotes an individual isochrone and D_1, D_2, ℓ_1, ℓ_2, b_1, b_2 denote the range in distance and area probed by each bin (where the minimum value of $D_1 = 5$ kpc). The combined estimates are then calculated by weighting the isochrones by an MDF and age distribution. In the bottom panel of Fig. 8, we show this volume correction for one bin in ℓ and b as a function of magnitude and colour.

After applying the corrections outlined above we can estimate the total stellar halo luminosity. First, we test the method on the *Galaxia* models, for which we know the true halo luminosity. In Fig. 9, we show the estimated luminosity for every bin in colour (x-axis), magnitude (panel), and area on the sky (coloured symbols) for three example haloes. The light grey region indicates the combined estimate for all bins, and the dark grey region indicates the combined estimate for selected bins. These selected bins are identified in the previous section, and exclude bins where the overlap between disc and halo prevents a good estimate of the number of halo RGB stars. Here, approximately 30 per cent of the bins are excluded and these are indicated with the black crosses in the figure. The black dashed line indicates the true halo luminosity (out to 100 kpc). The luminosity estimates in each bin have large error bars, but the combination of a large number of these bins can give a ~5 per cent measure (but note this error is just statistical!). Reassuringly, the estimates in different bins generally agree very well, apart from the bins that we have already identified as having systematic differences (black crosses).

The results for all 11 of the BJ05 haloes are shown in Fig. 10. Here, we show the estimated luminosity relative to the true luminosity. The grey filled circles show the combined estimates from all bins, and the blue filled circles show the combined estimates from a subset of 'robust' bins. The luminosity is typically underestimated by 20 per cent when all bins are used. This is because for certain bins the halo and disc populations cannot be properly decomposed. However, if we disregard these bins we are able to recover the true value to within 25 per cent. Note the scatter across all 11 haloes is larger than the individual statistical error bars (~5 per cent). This

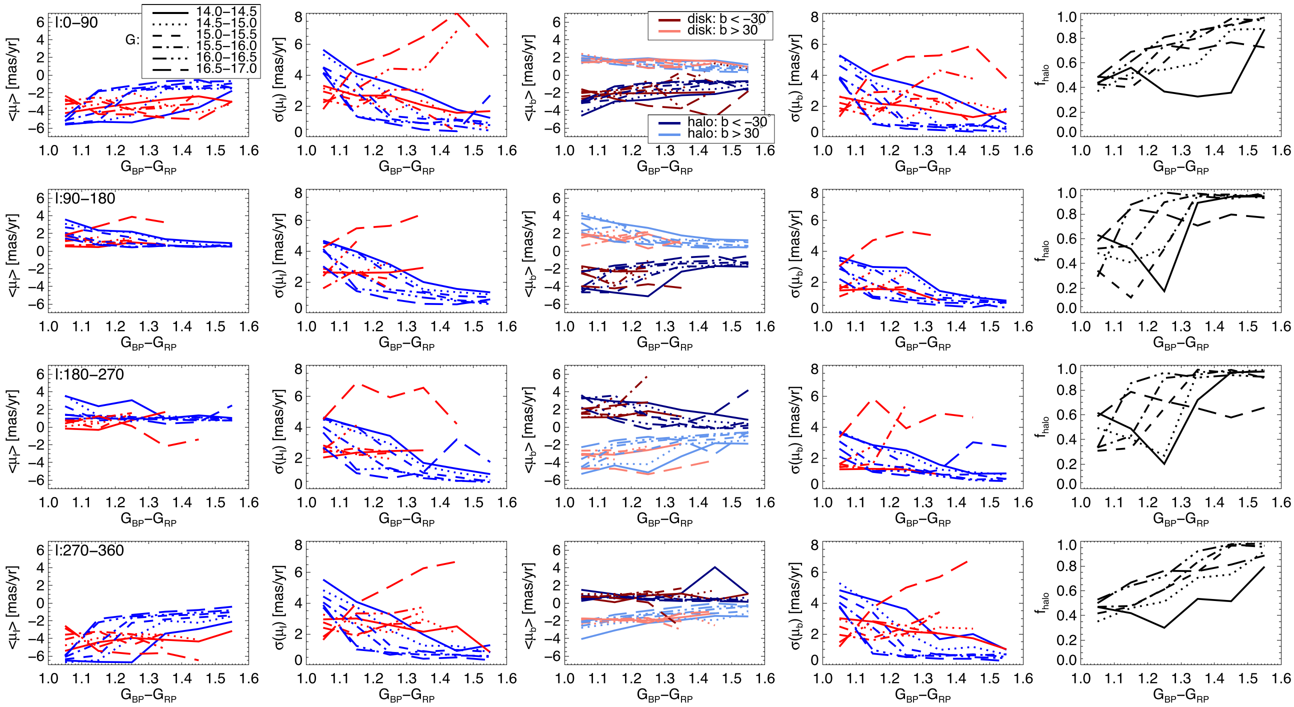

Figure 7. The mean (first and third panels) and dispersion (second and fourth panels) of the GDR2 model proper motions in Galactic coordinates as a function of $G_{\rm BP} - G_{\rm RP}$ colour. Blue and red lines indicate the estimated halo and disc components, respectively. Different magnitude bins are shown with different linestyles, and each row shows a different bin in Galactic longitude. The sequences are very similar for bins above and below the Galactic plane, except for $\langle\mu_b\rangle$ (third column), which we indicate with different shades of blue and red. The last panel on the right indicates the fraction of halo stars as a function of colour. The sequences follow roughly the expected trends (see Fig. 3). However, the 'disc' component in the faintest magnitude bin appears to be dominated by noise.

is due to systematic effects, such as substructure, non-Gaussian MDFs, non-Einasto density profiles etc. So, this exercise gives us a more robust estimate of the error of our estimated luminosity.

We now apply our procedure to the *Gaia* data and show the results for the luminosity estimate in Fig. 11. Here, we have performed the analysis both with and without the Sgr stream. When the Sgr stream is included, the estimated luminosity increases by 15 per cent. It is clear that including Sgr enhances the halo luminosity, particularly in the fainter, redder bins. This is particularly evident in the $\ell \in [270, 360]$, $b \in [30, 90]$ bin, which is where the apocentre of the Sgr leading arm (at $D \sim 50$ kpc) is dominant. These results give a rough estimate of the Sgr luminosity of $L_{\rm Sgr} \sim 1.5 \times 10^8\,{\rm L}_\odot$, in good agreement with the value derived by Niederste-Ostholt, Belokurov & Evans (2012). Owing to the systematics we deduced in the previous section, we use a subset of bins to calculate our best luminosity estimate. We find $L_{\rm halo} = 7.9 \pm 2.0 \times 10^8\,{\rm L}_\odot$ excluding Sgr, and $L_{\rm halo} = 9.4 \pm 2.4 \times 10^8\,{\rm L}_\odot$ including Sgr. Here, we have assumed, based on comparison to *N*-body models, that this estimate is accurate to 25 per cent. Note that if we had used all available bins, our estimates are reduced by ∼10 per cent and the statistical error is smaller. However, as shown in Fig. 10, the systematic error increases and the mass is likely underestimated when all bins are used.

5 DISCUSSION

5.1 A relatively high Galactic stellar halo mass?

In the preceding Section(s), we have used counts of RGB stars in GDR2 to estimate the total luminosity of the Galactic halo. This can be converted to a stellar mass by adopting an appropriate stellar mass-to-light ratio. Using the (weighted) suite of PARSEC iscohrones described earlier (with uniform ages between 10–14 Gyr and an MDF with $\langle[{\rm Fe/H}]\rangle = -1.5$), we estimate stellar mass-to-light ratios of 1.3, 1.5, and 2.8 for a Chabrier, Kroupa, and Salpeter (Salpeter 1955) IMF, respectively. We adopt the Kroupa IMF as our fiducial model, which gives: $M^\star_{\rm halo} = 1.2 \pm 0.3 \times 10^9\,{\rm M}_\odot$ (exc. Sgr) and $M^\star_{\rm halo} = 1.4 \pm 0.4 \times 10^9\,{\rm M}_\odot$ (inc. Sgr). Alternatively, we can express these values in terms of the local stellar halo density: $\rho_0 = 6.9 \times 10^{-5}\,{\rm M}_\odot\,{\rm pc}^{-3}$ (exc. Sgr), $\rho_0 = 8.1 \times 10^{-5}\,{\rm M}_\odot\,{\rm pc}^{-3}$ (inc. Sgr). These values can be multiplied by factors of 1.3/1.5 or 2.8/1.5 if Chabrier or Salpeter IMFs are preferred.

Our estimated stellar mass is significantly higher than recent values in the literature. For example, Bell et al. (2008) and Deason et al. (2011) find masses $M^\star_{\rm halo} = 3\text{–}4 \times 10^8\,{\rm M}_\odot$, which, even with the additional few $\times 10^8\,{\rm M}_\odot$ of substructures that are likely not accounted for in these models, is a factor of 2–3 lower than our estimate. However, it is worth pointing out that both of these estimates rely on an approximate relation between number of blue horizontal branch or main-sequence turn-off stars and luminosity. These works use globular clusters to calibrate this relation, but there is no simple way to quantify the sources of systematic errors in this approach. Indeed, although the low mass quoted by Bell et al. (2008) and Deason et al. (2011) are often cited in the literature, the estimates are relatively 'back of the envelope' and were not the main focus of the papers. All studies estimating the stellar halo luminosity or mass (including this one) face the difficult problem of converting number counts of (tracer) stars to a luminosity. The main advantages of our new estimate are (1) the uninterrupted all-sky, large volume probed by *Gaia* and (2) a thorough exploration of the

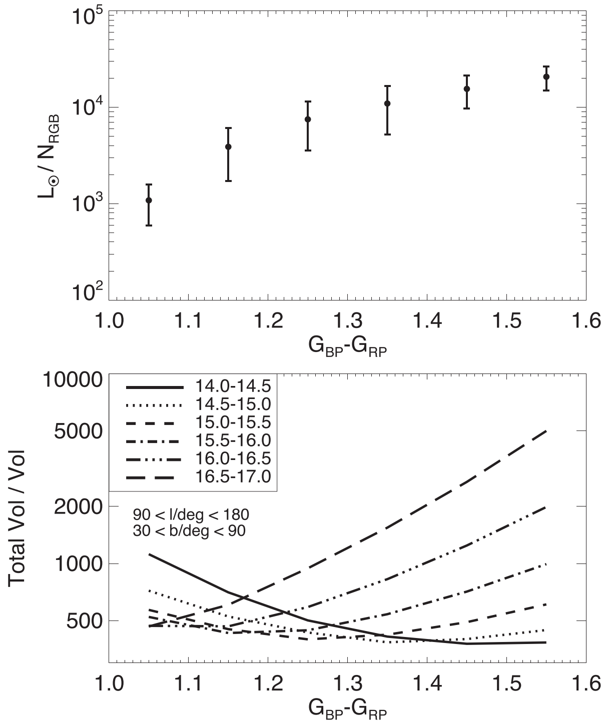

Figure 8. *Top:* The relation between total luminosity and number of RGB stars per colour bin. Here, we have used a set of weighted PARSEC isochrones assuming uniform ages in the range 10–14 Gyr and a metallicity distribution with $\langle[\mathrm{Fe/H}]\rangle = -1.5$, $\sigma([\mathrm{Fe/H}]) = 0.5$. *Bottom:* The total volume (out to 100 kpc) relative to the volume probed by a colour bin. Different linestyles correspond to different magnitude bins. Here, we used the weighted isochrones to estimate the distance range probed by a specific colour, magnitude bin, and we use the stellar halo density profile to relate the volume probed to the total volume. For the Milky Way, we assume an Einasto profile with $n = 1.7$, $R_e = 20$ kpc, and a minor-to-major axial ratio $q = 0.6$ (Deason et al. 2011).

various systematic uncertainties, including using simulations to test the method, the influence of the adopted IMF and stellar isochrones, and the influence of the adopted stellar density profile and MDF (see following subsection). It is intriguing that our estimate is in better agreement with results deriving from relatively nearby halo star counts (e.g. Morrison 1993; Gould et al. 1998; de Jong et al. 2010), but these require significant extrapolation to convert to a *total* stellar halo mass. Importantly, our estimated mass is in good agreement with the recent result posted by Mackereth & Bovy (2019). These authors find $M^{\star}_{\rm halo} = 1.3^{+0.3}_{-0.2} \times 10^9\,\mathrm{M}_\odot$ using RGB star counts in APOGEE DR14 data.

It is worth remarking that the low (few $\times 10^8\,\mathrm{M}_\odot$) stellar halo mass often quoted for the Milky Way is also at odds with recent results from *Gaia* suggesting the (inner) halo is dominated by an ancient, massive accretion event with $M^\star \sim 0.5$–$1 \times 10^9\,\mathrm{M}_\odot$ (Belokurov et al. 2018; Helmi et al. 2018). Moreover, analyses of the kinematics and ages of the Galactic globular cluster populations point to a small number of massive ($\sim 10^9\,\mathrm{M}_\odot$) Milky Way progenitors (Myeong et al. 2018; Kruijssen et al. 2019). While it is feasible that some of the stars from these massive progenitors end up in the stellar disc (and thus avoid being accounted for in analysis given the $|b| > 30^\circ$ cut), the majority of the debris should be in the halo. Thus, our new estimate of $M^{\star}_{\rm halo} \sim 10^9\,\mathrm{M}_\odot$ agrees with the emerging picture of a massive progenitor dominating the stellar halo mass and, importantly, provides a direct accounting of the debris from this event.

5.2 Model assumptions and systematic uncertainties

In this subsection, we explore the systematic uncertainties related to our model assumptions. First, we consider the halo density profile. We adopt a flattened Einasto stellar halo density profile from Deason et al. (2011) to volume correct the RGB star counts in magnitude, colour, and spatial bins. The form of the density profile of halo stars within 50 kpc from various sources in the literature are in broad agreement (e.g. Faccioli et al. 2014; Pila-Díez et al. 2015; Xue et al. 2015), but they differ in detail. In the left-hand panel of Fig. 12, we compute the total halo luminosity for various different density profiles. Note, here for ease of computation, and as we are interested in relative differences, we adopt a single isochrone model with age $T = 10$ Gyr and metallicity $[\mathrm{Fe/H}] = -1.5$. The filled points use the density profiles relevant for the 11 BJ05 halo models. These points are shown to illustrate the range of values that can be found if there is little knowledge about the halo density profile. In this case, the dispersion in the luminosity estimates (neglecting the obvious outlier) is ∼35 per cent of the mean. Note we checked that the outliers in the BJ05 haloes have rather extreme density profile parameters (at least relative to the MW). The lines indicate various density profiles in the literature (Deason et al. 2011; Sesar et al. 2011; Faccioli et al. 2014; Pila-Díez et al. 2015; Xue et al. 2015). These observed profiles have been computed using a range of tracers (blue horizontal branch, RR Lyrae, main sequence, and giant stars) and data sources. This figure illustrates that the derived luminosity can vary significantly with the adopted density profile. In general, profiles that are steeper (at large distances) lead to higher luminosity estimates, as the volume correction factor is larger. The grey region in Fig. 12 indicates the approximate 1σ dispersion in luminosity for the range of observed density profiles, which is ∼30 per cent of our fiducial result using the Deason et al. (2011) profile. We note that it is reassuring that the profile by Xue et al. (2015), which uses RGB stars as tracers, gives a similar answer to our fiducial result. We keep the Deason et al. (2011) profile to give our main result, but note that an additional systematic error (of 30 per cent) can be included in order to account for uncertainties in the stellar halo density profile.

An additional model assumption is the adopted metallicity distribution function. In our fiducial results, we adopt a MDF for the halo with $\langle[\mathrm{Fe/H}]\rangle = -1.5$. This value is motivated by results in the literature (An et al. 2013; Zuo et al. 2017), but lower and higher average metallicities have also been reported (e.g. Xue et al. 2015; Conroy et al. 2019). To explore the effect of the MDF on our results, we show in the right-hand panel of Fig. 12 the derived luminosity as a function of average metallicity. Here, we keep the same dispersion in the MDF ($\sigma = 0.5$ dex) but vary the mean ($\langle[\mathrm{Fe/H}]\rangle$). Note that the Deason et al. (2011) density profile is adopted, but the same relative trend is seen with different stellar halo density profiles. The relation is shown relative to the fiducial luminosity assuming $\langle[\mathrm{Fe/H}]\rangle = -1.5$. The figure shows that the luminosity estimate is dependent on the adopted MDF. The adopted metallicity affects the derived total luminosity in two main ways: (1) higher metallicity isochrones have lower luminosity per unit number of RGB stars ($\mathrm{L}_\odot / N_{\rm RGB}$) and (2) the distances, in a given colour and magnitude bin are lower at higher metallicity, and hence the volume correction (Total Vol/Vol) is typically smaller. These two effects both lead to a reduction in total luminosity at higher metallicities (and an increase at lower metallicities). To account for the variation with metallicity,

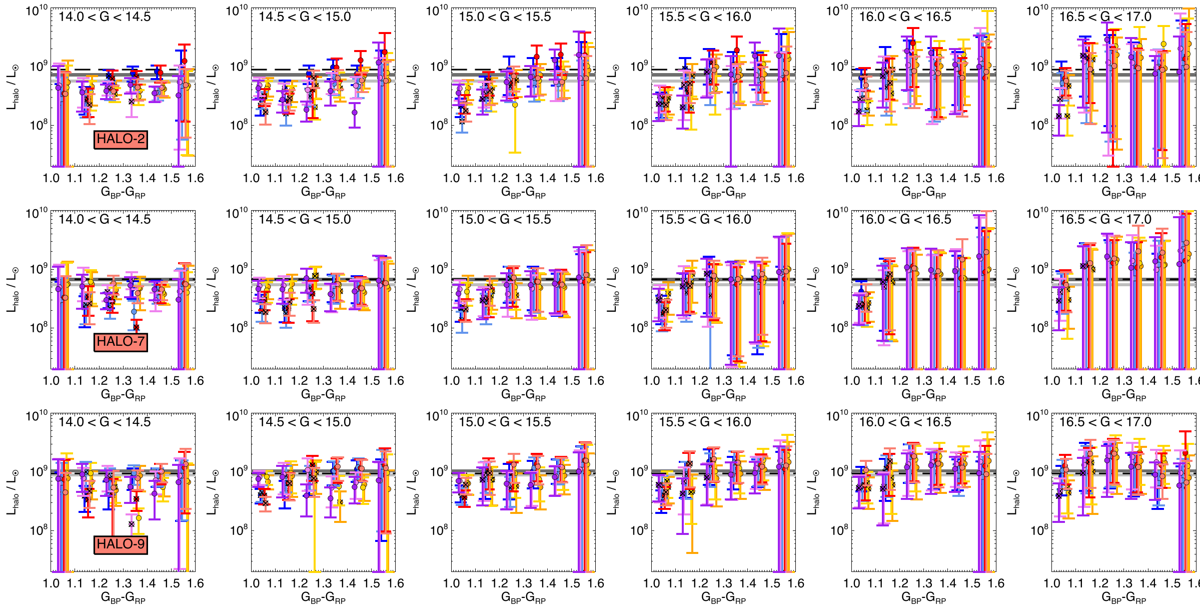

Figure 9. The estimated (total) stellar halo luminosity as a function of colour. Each panel shows a different magnitude bin. For each colour, magnitude bin, there are eight bins on the sky. The colour coding is the same as in Fig. 2. We show three example haloes from the *Galaxia* + *N*-body models. The dashed black line shows the true value, and the (light) shaded grey region indicates the combined estimate from all of the colour, magnitude, and (ℓ, b) bins. The dark shaded grey region indicates the combined estimated when 30 per cent of the bins (shown with black crosses) are excluded.

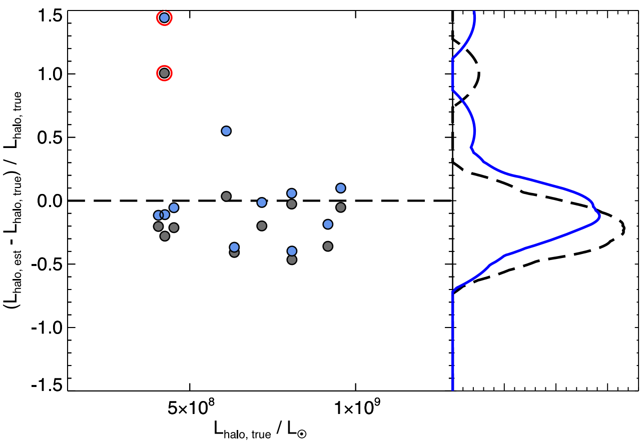

Figure 10. The estimated luminosity for the *Galaxia* + *N*-body haloes relative to the true values as a function of stellar halo luminosity. Here, the 'total' luminosity is defined within 100 kpc. The right-inset panel shows the PDF for the $(L_{\rm halo,\,est} - L_{\rm halo,\,true})/L_{\rm halo,\,true}$ values. The grey points are the estimates when all bins are used, and the blue points are when bins with high levels of contamination are excluded. For the majority of haloes, we can recover the true value to within ∼25 per cent. An outlier (halo-10) is indicated with a red circle; this halo has significant contribution from unrelaxed substructure.

we compute a quadratic relation between luminosity and the average metallicity

$$\frac{L_{\rm halo}}{L_{\rm halo,\langle[\rm Fe/H]\rangle=-1.5}} = 1.0 - 0.9851\left(\langle[\rm Fe/H]\rangle + 1.5\right) + 0.2670\left(\langle[\rm Fe/H]\rangle + 1.5\right)^2. \quad (3)$$

Here, the halo luminosity can be adjusted from the fiducial estimate (assuming $\langle[\rm Fe/H]\rangle = -1.5$) using the above relation. Note that this equation is only valid for average metallicities in the range $-2.0 < \langle[\rm Fe/H]\rangle < -1.0$. For completeness, we also provide a conversion formula for the total stellar halo mass assuming a Kroupa IMF. Note the relation is not identical to the luminosity conversion (modulo a factor conversion) as the stellar mass-to-light ratio depends on metallicity. For example, for a Kroupa IMF (and assuming old ages) the stellar mass-to-light ratio is 1.6(1.4) for $\langle[\rm Fe/H]\rangle = -1.0(-2.0)$.

$$\frac{M^\star_{\rm halo}}{M^\star_{\rm halo,\langle[\rm Fe/H]\rangle=-1.5}} = 1.0 - 0.9104\left(\langle[\rm Fe/H]\rangle + 1.5\right) + 0.2473\left(\langle[\rm Fe/H]\rangle + 1.5\right)^2. \quad (4)$$

Conroy et al. (2019) recently reported that the average stellar halo metallicity is higher than previously thought, with $\langle[\rm Fe/H]\rangle = -1.2$. If we use this increased metallicity in the formula given above then the derived stellar halo mass is $M^\star_{\rm halo} = 1.05 \times 10^8\,{\rm M}_\odot$, i.e. 25 per cent lower than our fiducial estimate (see following subsection for further discussion).

Finally, our stellar halo mass (and luminosity) estimate is also dependent on the suite of isochrones used in the analysis, as the predictions for the RGB can vary between different stellar population models (see e.g. Hidalgo et al. 2018). If we repeat our analysis (assuming our fiducial density profile, MDF, IMF assumptions) with the MIST (Choi et al. 2016) or BaSTI (Hidalgo et al. 2018) models, we find (total) stellar masses of $M^\star_{\rm halo} = 0.85 \times 10^9\,{\rm M}_\odot$ and $M^\star_{\rm halo} = 1.1 \times 10^9\,{\rm M}_\odot$, respectively. These masses are slightly lower than our fiducial results (based on the PARSEC isochrones), but still consistent within the uncertainties. The complexities of modelling the RGB in isochrone libraries is beyond the scope of this paper, but this, in addition to the systematic effects mentioned above, is an important consideration for stellar halo mass estimates and will need close attention in future work to achieve both precise and accurate measures.

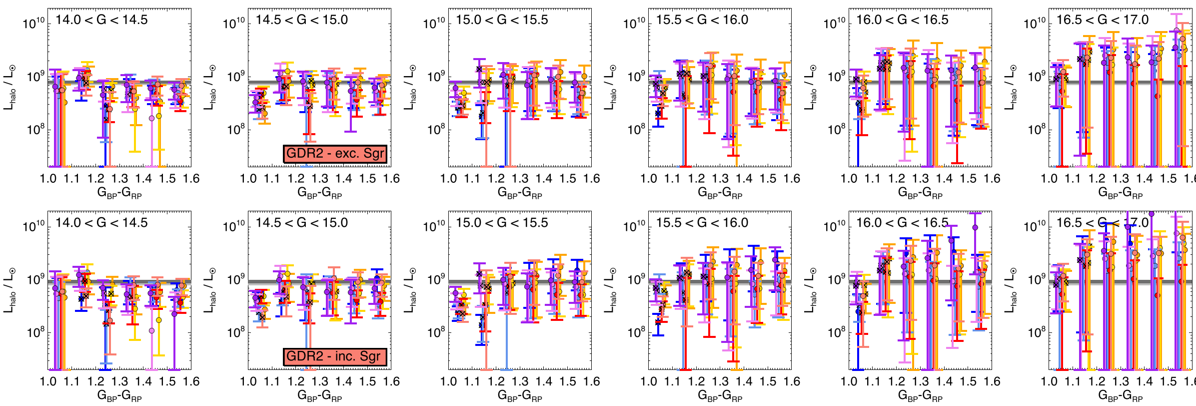

Figure 11. The estimated (total) stellar halo luminosity as a function of colour. Each panel shows a different magnitude bin. For each colour, magnitude bin, there are eight bins on the sky. The colour coding is the same as in Fig. 2. The top and bottom rows show cases with Sgr excluded (top, $L_{\rm halo} = 7.9 \times 10^8\,{\rm L}_\odot$) and included (bottom, $L_{\rm halo} = 9.4 \times 10^8\,{\rm L}_\odot$). Including Sgr increases the total luminosity by ~15 per cent. The (light) shaded grey region indicates the combined estimate from all of the colour, magnitude, and (ℓ, b) bins. The dark shaded grey region indicates the combined estimated when 30 per cent of the bins (shown with black crosses) are excluded.

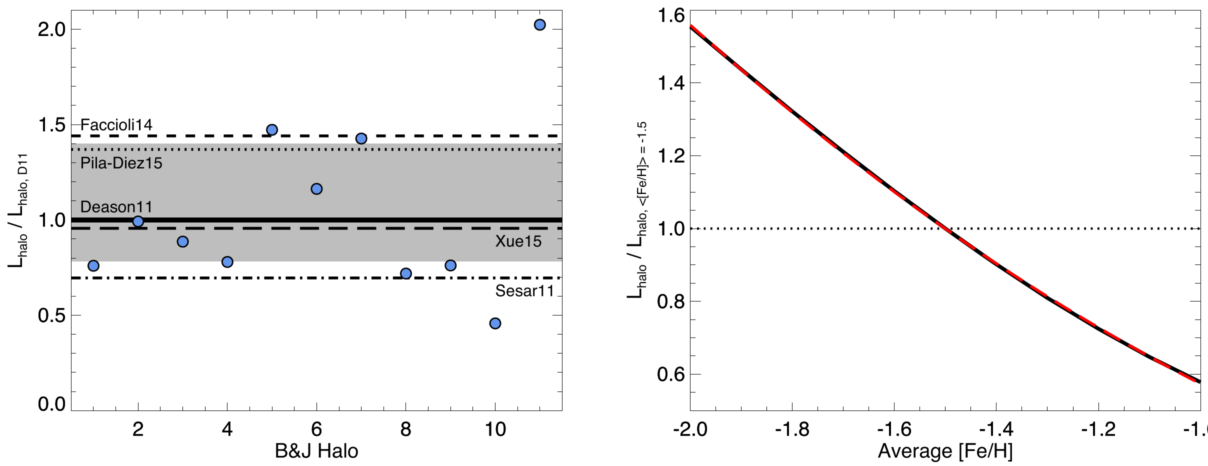

Figure 12. *Left:* The total halo luminosity derived with various stellar halo density profiles relative to the fiducial density profile assumption. Note here we adopt a single isochrone model with age $T = 10$ Gyr and metallicity [Fe/H] $= -1.5$. We use the range of density profiles seen in the BJ05 haloes (filled blue points) and indicate the results for a range of observed profiles in the literature (Deason et al. 2011; Sesar, Jurić & Ivezić 2011; Faccioli et al. 2014; Pila-Díez et al. 2015; Xue et al. 2015). Our fiducial density profile assumption (Deason et al. 2011) lies in the middle of the estimates, but the various profiles have a 1σ dispersion of ~30 per cent around the fiducial value. *Right:* The halo luminosity relative to the fiducial luminosity (with $\langle$[Fe/H]$\rangle = -1.5$) as a function of average metallicity. The derived luminosity strongly depends on the MDF. The dashed red line shows a quadratic fit that can be used to approximately convert our fiducial luminosity estimate to a different MDF.

5.3 Tension between stellar halo mass and metallicity?

Dwarf galaxies follow a fairly tight (~0.2 dex scatter) stellar mass–metallicity relation (Kirby et al. 2013). Following the relation derived by Kirby et al. (2013) based on Local Group galaxies, dwarfs with masses in the range 0.5–1 $\times$ $10^9\,{\rm M}_\odot$ have average metallicities of $\langle$[Fe/H]$\rangle \sim -0.9$ to -0.8 dex. The average metallicity of halo stars is $\langle$[Fe/H]$\rangle \sim -1.5$ (An et al. 2013; Zuo et al. 2017), which is seemingly at odds with a stellar halo mass of $\sim 10^9\,{\rm M}_\odot$ dominated by one massive progenitor. However, this simple exercise ignores two important factors: (1) we are using the $z = 0$ mass–metallicity relation and the stellar halo was built up in the past and (2) the inner halo, within ~20 kpc is likely dominated by a massive progenitor, but the outer parts are likely biased towards lower mass contributors (Deason et al. 2018; Lancaster et al. 2019).

Deason et al. (2016) use cosmological *N*-body simulations to explore the relation between accreted stellar mass and metallicity. They used empirical stellar mass–halo mass relations, and redshift-dependent stellar mass–metallicity relations, to map accreted dark matter subhaloes to stellar halo progenitors. In their fig. 7, they show the relation between the average metallicity of the accreted stellar material and the typical destroyed dwarf mass. For progenitors of $M^\star \sim 0.5$–$1 \times 10^9\,{\rm M}_\odot$, the average metallicity varies between $\langle$[Fe/H]$\rangle \sim -1.0$ and -1.5. The lower metallicities are only obtained when the progenitor is destroyed at very early times, when the average metallicity of the dwarf galaxies (at fixed mass) is lower (Ma et al. 2016) (see also Fattahi et al., in preparation). Thus, in order to reconcile the stellar halo metallicity with a massive progenitor (and hence relatively massive stellar halo), this event must have occurred $\gtrsim$10 Gyr ago. This is exactly the scenario that has been proposed in the *Gaia*-Sausage or *Gaia*-Enceladus discovery papers: an *ancient*, massive accretion event (Belokurov et al. 2018; Haywood et al. 2018; Helmi et al. 2018).

Recently, Conroy et al. (2019) suggested that the average stellar halo metallicity should be revised upwards to ⟨[Fe/H]⟩ = −1.2. In this case, as discussed in the previous section, our estimated stellar halo mass is slightly lower ($\sim 1.0 \times 10^9\,\mathrm{M}_\odot$ rather than $1.4 \times 10^9\,\mathrm{M}_\odot$). The argument above – that this metallicity–stellar halo mass combination favours an ancient accretion event – still holds, but the disparity with the $z = 0$ stellar mass–metallicity relation is less severe.

Finally, as mentioned earlier, although the bulk of the (inner) stellar halo mass may be contributed by the *Gaia*-Sausage, there is still a sprinkling of lower mass, $\sim 10^7$–$10^8\,\mathrm{M}_\odot$ progenitors (e.g. Sgr, Sequoia), with lower average metallicities that also contribute to the total stellar halo mass (and average metallicity). Moreover, there could also be a contribution from *in situ* halo stars to the total stellar halo mass. We discussed in Section 2 that the *Galaxia* + *N*-body models do not include *in situ* halo stars. Belokurov et al. (2019) recently showed evidence for an *in situ* halo population, dubbed 'Splash', in the inner Milky Way halo (see also Gallart et al. 2019). The Splash is kinematically hot and has chemical and kinematic features that are intermediate between halo and thick disc populations. However, importantly, Belokurov et al. (2019) show that the Splash is confined to the inner halo. Indeed, they find at heights of $|z| \sim 10$ kpc the Splash drops to a meagre 5 per cent of the halo density. As our analysis is mainly concerned with distant stars ($D \sim 5$–100 kpc) at high Galactic latitude ($|b| > 30°$), we do not expect our halo mass estimate to be contaminated by more than 5 per cent from Splash stars. However, we do caution that the cosmological simulations do predict a significant amount of distant *in situ* halo material (see e.g. Monachesi et al. 2019). If this is true in the Milky Way, then our total stellar halo mass estimate is a combination of accreted halo stars and any *in situ* material that manages to make it out to significant distances in the halo.

5.4 The Milky Way in context

At fixed galaxy (or halo) mass, the stellar halo mass can vary significantly: this has been seen both in simulations and observations (e.g. Pillepich et al. 2014; Merritt et al. 2016; Elias et al. 2018; Monachesi et al. 2019). Thus, the stellar halo mass is intimately linked to the assembly history of the halo. For example, if a halo is dominated by one accretion event, then the stellar halo mass will reflect the mass of this progenitor (see e.g. Deason et al. 2016; D'Souza & Bell 2018).

In the left-hand panel of Fig. 13, we show the ratio of stellar masses of the accreted (halo) and the galaxy (host) populations against the galaxy's stellar mass. Here, we show the values from the $N = 30$ AURIGA simulations in grey. This is a suite of high-resolution cosmological hydrodynamic simulations of Milky Way mass haloes (see Grand et al. 2017 for more details). The stellar halo masses we show here only include the 'accreted' stellar halo mass. As noted by Monachesi et al. (2019), the stellar halo masses in AURIGA are significantly overestimated if we do not excise the halo stars born *in situ*. We also show observational measurements in the left-hand panel from the Ghosts (black filled circles, Harmsen et al. 2017) and Dragonfly (pale blue filled circles, Merritt et al. 2016) surveys. Our estimate for the Milky Way is shown with the orange star symbol (assuming a Kroupa IMF). Here, we use total Galactic stellar mass derived in Licquia & Newman (2015). Even though our stellar halo mass is larger than some previous estimates in the literature, the Milky Way stellar halo mass fraction is relatively low compared to both external galaxies and the AURIGA simulations.

In the right-hand panel we use the AURIGA simulations to show $M^\star_{\mathrm{halo}}/M^\star_{\mathrm{gal}}$ against the average merger time of the destroyed dwarf galaxies that build-up the stellar halo ($\langle T_{\mathrm{merge}} \rangle$: computed by averaging over all star particles within 100 kpc). The points are coloured (and scaled) according to the average progenitor mass. Note that the quantities in Fig. 13 (e.g. merger times, accreted stellar mass) for the AURIGA simulations are derived in the works by Fattahi et al. (2019) and Monachesi et al. (2019). There is a clear trend between *the epoch* of a dwarf accretion and the fraction of stellar mass in the halo: earlier accretion events lead to a lower fraction of stars in the halo (see also Elias et al. 2018). Early mergers truncate the star formation activity in the progenitor dwarfs, while the dwarfs accreted later were able to continue to form stars. Note that haloes with a large number of progenitors (e.g. Halo-17, blue point in top left of right-hand panel) do not follow this trend as closely as the 'average' progenitor mass and merger time are more ill defined. For illustration, we indicate the Milky Way with the orange star. Here, we have assumed the typical merger time for the *Gaia*-Sausage is 10 ± 2 Gyr ago (4 Gyr since the Big Bang). Even

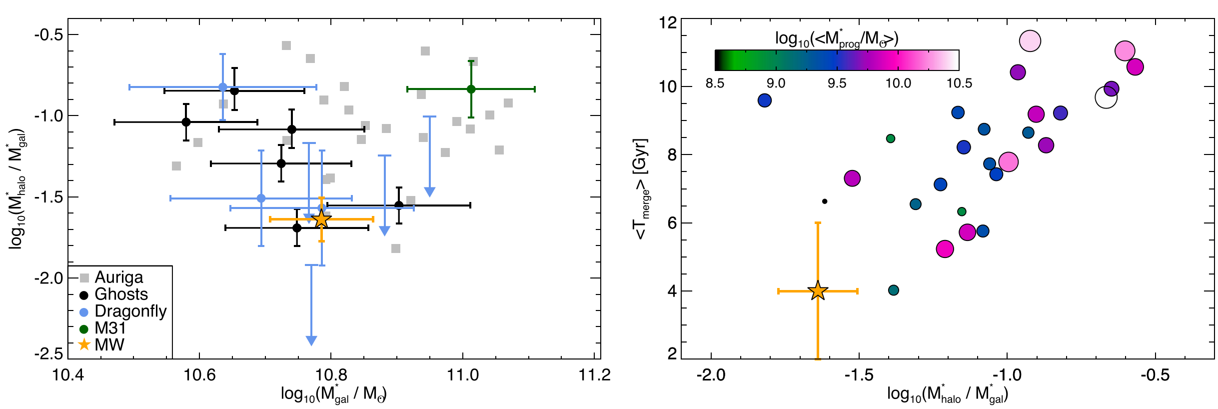

Figure 13. *Left:* The stellar halo mass fraction ($M^\star_{\mathrm{halo}}/M^\star_{\mathrm{gal}}$) as a function of galaxy mass. We show the simulated Auriga galaxies, and observational estimates from Ghosts (Harmsen et al. 2017), Dragonfly (Merritt et al. 2016), and M31 (Sick et al. 2015; Harmsen et al. 2017). The yellow star indicates our measure for the Milky Way assuming a Kroupa IMF (note we use the Galaxy mass from Licquia & Newman 2015). *Right:* The average merger time of Milky Way halo progenitors against stellar halo mass fraction for the AURIGA haloes. The points are coloured (and scaled) according to the average progenitor mass. The Milky Way is indicated with the orange star.

though most haloes in AURIGA experience more recent accretion events, it is clear that an ancient merger event, with little activity after said event, can adequately explain the stellar halo mass fraction of the Milky Way.

In summary, our estimated stellar halo mass supports a scenario whereby the Milky Way experienced an early (∼10 Gyr ago), massive ($M^{\star} \sim 0.5 - 1.0 \times 10^9\,\mathrm{M}_{\odot}$) merger event, and had only relatively minor mergers thereafter.[3]

6 CONCLUSIONS

In this work, we have used counts of RGB stars from GDR2 to estimate the total stellar luminosity of the Milky Way's halo. Using slices in colour, magnitude, and position on the sky, we decompose the disc and halo RGB populations using 2D Gaussian fits to the proper motion distributions. The resulting counts of halo stars are converted into a stellar luminosity using a suite of (weighted) PARSEC isochrones. Our analysis is tested and calibrated on the *Galaxia* model, using the BJ05 *N*-body models for the stellar halo component. Our main results are summarized as follows:

(i) In the majority (70 per cent) of bins in magnitude, colour, and area on the sky we are able to recover the true number of halo RGB stars to $\lesssim$30 per cent. Tests with the *Galaxia* + BJ05 models show that we are able to recover the true total luminosity to within 25 per cent if the metallicity distribution and density profile of the halo stars are known. This confidence interval takes into account realistic systematic uncertainties, such as the presence of substructure and non-Gaussian proper motion distributions.

(ii) After applying our method to GDR2, and assuming an Einasto density profile (Deason et al. 2011) and MDF with $\langle[\mathrm{Fe/H}]\rangle = -1.5$ for the stellar halo, we find a total luminosity of $L_{\mathrm{halo}} = 7.9 \pm 2.0 \times 10^8\,\mathrm{L}_{\odot}$ excluding Sgr, and $L_{\mathrm{halo}} = 9.4 \pm 2.4 \times 10^8\,\mathrm{L}_{\odot}$ including Sgr. The difference when Sgr is included or excluded gives a rough estimate of the total luminosity of the Sgr progenitor: $L_{\mathrm{Sgr}} \sim 1.5 \times 10^8\,\mathrm{L}_{\odot}$, in good agreement with the value derived by Niederste-Ostholt et al. (2012).

(iii) We explore additional systematic uncertainties from our adopted MDF and density profile for the halo. In particular, we find the metallicity strongly influences the derived luminosity and we provide an approximate conversion formula to infer luminosity (and mass) for a different MDF. Moreover, differences in the literature regarding the halo density profile leads to an additional systematic uncertainty of ∼30 per cent in our derived luminosity and mass.

(iv) Assuming a stellar mass-to-light ratio appropriate for a Kroupa IMF ($M^{\star}/L = 1.5$) and our fiducial halo density profile and MDF, we estimate a stellar halo mass of $M^{\star}_{\mathrm{halo}} = 1.4 \pm 0.4 \times 10^9\,\mathrm{M}_{\odot}$. This mass is larger than estimates in the literature using different stellar halo tracers (main-sequence turn-off stars, blue horizontal branch stars) and different methods. However, a mass of $\sim 10^9\,\mathrm{M}_{\odot}$ confirms the emerging picture that the (inner) stellar halo is dominated by one massive dwarf progenitor.

(v) We show that haloes in the AURIGA simulations that have similar stellar halo mass fractions ($M^{\star}_{\mathrm{halo}}/M^{\star}_{\mathrm{gal}} \sim 0.02$) to the Milky Way are likely formed by ancient (∼10 Gyr) mergers. Indeed, the relatively low stellar halo mass fraction and average metallicity of the stellar halo can only be reconciled with a massive progenitor if this was a very early merger event.

[3] At least until the LMC is digested, see Cautun et al. (2019).

ACKNOWLEDGEMENTS

We thank an anonymous referee for providing a thorough and insightful report, which greatly improved the quality of this manuscript.

AD thanks Azi Fattahi for help with simulation data, and Russell Smith for useful discussions regarding stellar population models. We thank the Auriga team for allowing us to use their data in Fig. 12.

AD is supported by a Royal Society University Research Fellowship. AD also acknowledges the support from the STFC grant ST/P000541/1. JLS acknowledge the support of the Leverhulme and Newton Trusts.

This work has made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

We acknowledge the Gaia Project Scientist Support Team and the DPAC.

This work used the DiRAC Data Centric system at Durham University, operated by ICC on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, STFC DiRAC Operations grant ST/K003267/1, and Durham University. DiRAC is part of the National E-Infrastructure.

AD thanks the staff at the Durham University Day Nursery who play a key role in enabling research like this to happen.

REFERENCES

An D. et al., 2013, ApJ, 763, 65
Bell E. F. et al., 2008, ApJ, 680, 295
Belokurov V., 2013, New Astron. Rev., 57, 100
Belokurov V. et al., 2014, MNRAS, 437, 116
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Belokurov V., Sanders J. L., Fattahi A., Smith M. C., Deason A. J., Evans N. W., Grand R. J. J., 2019, preprint (arXiv:1909.04679)
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Bovy J., Hogg D. W., Roweis S. T., 2011, Ann. Appl. Stat., 5, 1657
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Bullock J. S., Johnston K. V., 2005, ApJ, 635, 931 (BJ05)
Cañas R., Lagos C. d. P., Elahi P. J., Power C., Welker C., Dubois Y., Pichon C., 2019, preprint (arXiv:1908.02945)
Cautun M., Deason A. J., Frenk C. S., McAlpine S., 2019, MNRAS, 483, 2185
Chabrier G., 2003, PASP, 115, 763
Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, ApJ, 823, 102
Conroy C., Naidu R. P., Zaritsky D., Bonaca A., Cargile P., Johnson B. D., Caldwell N., 2019, preprint (arXiv:1909.02007)
D'Souza R., Bell E. F., 2018, MNRAS, 474, 5300
de Jong J. T. A., Yanny B., Rix H.-W., Dolphin A. E., Martin N. F., Beers T. C., 2010, ApJ, 714, 663
Deason A. J., Belokurov V., Evans N. W., 2011, MNRAS, 416, 2903
Deason A. J. et al., 2012, MNRAS, 425, 2840
Deason A. J., Belokurov V., Evans N. W., Johnston K. V., 2013, ApJ, 763, 113
Deason A. J., Belokurov V., Koposov S. E., Rockosi C. M., 2014, ApJ, 787, 30
Deason A. J., Mao Y.-Y., Wechsler R. H., 2016, ApJ, 821, 5
Deason A. J., Belokurov V., Koposov S. E., Gómez F. A., Grand R. J., Marinacci F., Pakmor R., 2017, MNRAS, 470, 1259

Deason A. J., Belokurov V., Koposov S. E., Lancaster L., 2018, ApJ, 862, L1

Di Matteo P., Haywood M., Lehnert M. D., Katz D., Khoperskov S., Snaith O. N., Gómez A., Robichon N., 2018, preprint (arXiv:e-print)

Digby A. P., Hambly N. C., Cooke J. A., Reid I. N., Cannon R. D., 2003, MNRAS, 344, 583

Elias L. M., Sales L. V., Creasey P., Cooper M. C., Bullock J. S., Rich R. M., Hernquist L., 2018, MNRAS, 479, 4004

Faccioli L., Smith M. C., Yuan H.-B., Zhang H.-H., Liu X.-W., Zhao H.-B., Yao J.-S., 2014, ApJ, 788, 105

Fattahi A. et al., 2019, MNRAS, 484, 4471

Font A. S., McCarthy I. G., Crain R. A., Theuns T., Schaye J., Wiersma R. P. C., Dalla Vecchia C., 2011, MNRAS, 416, 2802

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fuchs B., Jahreiß H., 1998, A&A, 329, 81

Gaia Collaboration, 2016a, A&A, 595, A1

Gaia Collaboration, 2016b, A&A, 595, A2

Gaia Collaboration, 2018a, A&A, 616, A1

Gaia Collaboration, 2018b, A&A, 616, A10

Gallart C., Bernard E. J., Brook C. B., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019, Nat. Astron., 3, 932

Gould A., Flynn C., Bahcall J. N., 1998, ApJ, 503, 798

Grand R. J. J. et al., 2017, MNRAS, 467, 179

Gratton R. G., Carretta E., Bragaglia A., Lucatello S., D'Orazi V., 2010, A&A, 517, A81

Harmsen B., Monachesi A., Bell E. F., de Jong R. S., Bailin J., Radburn-Smith D. J., Holwerda B. W., 2017, MNRAS, 466, 1491

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., 2018, ApJ, 859, 31

Hidalgo S. L. et al., 2018, ApJ, 856, 125

Jordi C. et al., 2010, A&A, 523, A48

Jurić M. et al., 2008, ApJ, 673, 864

Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, ApJ, 779, 102

Kroupa P., 2001, MNRAS, 322, 231

Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2019, MNRAS, 486, 3180

Lancaster L., Koposov S. E., Belokurov V., Evans N. W., Deason A. J., 2019, MNRAS, 486, 378

Licquia T. C., Newman J. A., 2015, ApJ, 806, 96

Lindegren L., 2018, Re-normalising the astrometric chi-square in Gaia DR2, GAIA-C3-TN-LU-LL-124. Available at: http://www.rssd.esa.int/doc_fetch.php?id = 3757412

Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140

Mackereth J. T., Bovy J., 2019, MNRAS, preprint (arXiv:1910.03590)

Massari D., Koppelman H. H., Helmi A., 2019, A&A, 630, L4

McCarthy I. G., Font A. S., Crain R. A., Deason A. J., Schaye J., Theuns T., 2012, MNRAS, 420, 2245

Merritt A., van Dokkum P., Abraham R., Zhang J., 2016, ApJ, 830, 62

Monachesi A. et al., 2019, MNRAS, 485, 2589

Morrison H. L., 1993, AJ, 106, 578

Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018, ApJ, 863, L28

Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235

Niederste-Ostholt M., Belokurov V., Evans N. W., 2012, MNRAS, 422, 207

Pila-Díez B., de Jong J. T. A., Kuijken K., van der Burg R. F. J., Hoekstra H., 2015, A&A, 579, A38

Pillepich A. et al., 2014, MNRAS, 444, 237

Robin A. C., Reylé C., Derrière S., Picaud S., 2003, A&A, 409, 523

Salaris M., Cassisi S., 2005, Evolution of Stars and Stellar Populations. Wiley, Chichester, UK

Salpeter E. E., 1955, ApJ, 121, 161

Sanders J. L., Binney J., 2015, MNRAS, 449, 3479

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Sesar B. et al., 2010, ApJ, 708, 717

Sesar B., Jurić M., Ivezić Ž., 2011, ApJ, 731, 4

Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, ApJ, 730, 3

Sick J., Courteau S., Cuilland re J.-C., Dalcanton J., de Jong R., McDonald M., Simard D., Tully R. B., 2015, in Cappellari M., Courteau S., eds, Proc. IAU Symp. 311, Galaxy Masses as Constraints of Formation Models. Kluwer, Dordrecht, p. 82

Slater C. T., Nidever D. L., Munn J. A., Bell E. F., Majewski S. R., 2016, ApJ, 832, 206

Venn K. A., Irwin M., Shetrone M. D., Tout C. A., Hill V., Tolstoy E., 2004, AJ, 128, 1177

Watkins L. L. et al., 2009, MNRAS, 398, 1757

Xue X.-X. et al., 2011, ApJ, 738, 79

Xue X.-X., Rix H.-W., Ma Z., Morrison H., Bovy J., Sesar B., Janesh W., 2015, ApJ, 809, 144

Zolotov A., Willman B., Brooks A. M., Governato F., Brook C. B., Hogg D. W., Quinn T., Stinson G., 2009, ApJ, 702, 1058

Zuo W., Du C., Jing Y., Gu J., Newberg H. J., Wu Z., Ma J., Zhou X., 2017, ApJ, 841, 59

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