The Milky Way stellar halo out to 40 kpc: squashed, broken but smooth

A. J. Deason, V. Belokurov* and N. W. Evans*

Institute of Astronomy, Madingley Road, Cambridge CB3 0HA

Accepted 2011 June 13. Received 2011 June 6; in original form 2011 April 18

ABSTRACT

We introduce a new maximum-likelihood method to model the density profile of blue horizontal branch and blue straggler stars and apply it to the Sloan Digital Sky Survey Data Release 8 photometric catalogue. There are a large number (~20 000) of these tracers available over an impressive 14 000 deg2 in both Northern and Southern Galactic hemispheres, and they provide a robust measurement of the shape of the Milky Way stellar halo. After masking out stars in the vicinity of the Virgo overdensity and the Sagittarius stream, the data are consistent with a smooth, oblate stellar halo with a density that follows a broken power law. The best-fitting model has an inner slope $\alpha_{in} \sim 2.3$ and an outer slope $\alpha_{out} \sim 4.6$, together with a break radius occurring at $\sim 27$ kpc and a constant halo flattening (i.e. ratio of minor axis to major axis) of $q \sim 0.6$. Although a broken power law describes the density fall-off most adequately, it is also well fitted by an Einasto profile. There is no strong evidence for variations in flattening with radius, or for triaxiality of the stellar halo.

Key words: Galaxy: stellar content – Galaxy: structure – galaxies: general – galaxies: haloes – galaxies: individual: Milky Way – galaxies: photometry.

1 INTRODUCTION

The time required for stars in the stellar halo to exchange their energy and angular momentum is very long compared to the age of the Galaxy. Therefore, such stars preserve memories of their initial conditions, and so the structure of the stellar halo is intimately linked to the formation mechanism of the Galaxy itself. This fundamental insight was already noted by Eggen, Lynden-Bell & Sandage (1962). It is the reason why the stellar halo has attracted such interest despite containing only a small fraction of the total stellar mass of the Galaxy. Stars diffuse more quickly in configuration space, as opposed to energy and angular momentum space. Therefore, the spatial structure of the stellar halo may be smooth, even though it is built up from merging and accretion.

The simplest way of studying the stellar halo is through star counts. Typically, RR Lyrae or blue horizontal branch (BHB) stars are used as tracers, as they are relatively bright ($M_g \sim 0.5$–0.7; e.g. Sirko et al. 2004) and can be detected at radii out to $\sim 100$ kpc. The gathering of such data is painstaking work, and carries the price that sample sizes are often small. Such studies are consistent with a stellar halo that is round in the outskirts (with a minor axis to major axis ratio of $q = 1$) and more flattened in the inner parts with $q \sim 0.5$ (e.g. Hartwick 1987; Preston, Shectman & Beers 1991). Rather than selecting typical halo stars, an alternative approach is to model deep star count data in pencil-beam surveys at intermediate and high galactic latitudes, allowing for contamination of the star counts by the thin and thick disc populations. This was attempted by Robin, Reylé & Crézé (2000), who found a best-fitting halo density law with flattening $q \sim 0.76$, together with a power-law fall-off $\alpha$ of 2.4 [i.e. $\rho \sim (\text{distance})^{-\alpha}$]. A similar, slightly later, attempt by Siegel et al. (2002) using data in seven Kapteyn selected areas yielded $q \sim 0.6$ and $\alpha \sim 2.75$.

Efforts to detect variations in the flattening with radius have also been undertaken. Preston et al. (1991) argued that the flattening changes from strongly flattened ($q = 0.5$) at 1 kpc to almost round at 20 kpc. However, work by Sluis & Arnold (1998) using a compilation of 340 RR Lyraes and BHB stars found a constant flattening of $q \sim 0.5$, with no evidence for changes with radius, as well as a power-law index $\alpha \sim 3.2$. The most recent work by De Propris, Harrison & Mares (2010) utilizing 666 BHB stars from the 2dF QSO Redshift Survey find that the halo is approximately spherical with a power-law index $\alpha \sim 2.5$ out to $\sim 100$ kpc. Similarly, Sesar, Juric & Ivezić (2011) studying main-sequence turn-off (MSTO) stars from the Canada–France–Hawaii Telescope Legacy Survey find that the flattening is approximately constant at $q \sim 0.7$ out to 35 kpc.

The Sloan Digital Sky Survey (SDSS) has transformed our knowledge of the stellar halo. Although it had been suspected that the stellar halo is criss-crossed with streams and substructures ever since the discovery of the disrupting Sagittarius (Sgr), the SDSS provided a memorable picture of the debris in the Field of Streams (Belokurov et al. 2006). A wealth of substructure has now been identified, including the Sgr stream, the Virgo overdensity and the
Hercules–Aquila Cloud (e.g. Ibata, Gilmore & Irwin 1995; Belokurov et al. 2006, 2007; Jurić et al. 2008). This has been seen as vindication of modern theories of galaxy formation, which predict that stellar haloes are built up almost exclusively from the debris of disrupting satellites (e.g. Bullock & Johnston 2005; De Lucia & Helmi 2008; Cooper et al. 2010). A number of studies have attempted to model the smooth halo component by avoiding these known substructures (e.g. Jurić et al. 2008). The results are only in very rough agreement, suggesting that the density profile of the Milky Way has a power-law slope in the range $2 < \alpha < 4$ and a flattening varying from $0.4 < q < 0.8$ (e.g. Yanny et al. 2000; Chen et al. 2001; Newberg & Yanny 2006; Jurić et al. 2008; Sesar et al. 2011).

Even though panoramic photometric surveys like SDSS provide a large sample of stellar halo tracers over a considerable portion of the sky, there is no consensus on the flattening and shape of the stellar halo. MSTO stars are commonly used tracers owing to their large numbers and the ease by which they can be photometrically identified (e.g. Bell et al. 2008). The absolute magnitudes of such stars centre around $M_g \sim 4.5$ but have a wide range of values ($\sigma_{M_g} \sim 0.9$ mag), which limits the accuracy to which the density profile can be estimated. BHB stars are superior distance estimators ($\sigma_{M_g} \sim 0.15$), but are significantly scarcer than main-sequence stars. Moreover, they suffer from contamination by blue straggler (BS) stars due to their similar colours. BHB and BS stars can be distinguished by their Balmer line profiles (e.g. Kinman, Suntzeff & Kraft 1994; Yanny et al. 2000; Clewley et al. 2002; Sirko et al. 2004), but this requires spectroscopic information. Whilst spectroscopic samples can cleanly identify BHB stars, the variety of results on flattening (e.g. Preston et al. 1991; Sluis & Arnold 1998; De Propris et al. 2010) and density fall-off (e.g. Xue et al. 2008; Brown et al. 2010) suggests that the completeness biases are difficult to understand and control.

An independent constraint on the density profile of the stellar halo is provided by the velocity distribution of the halo stars. A kinematic analysis by Carollo et al. (2007; see also Carollo et al. 2010) suggests that the stellar halo comprises two components with different density profiles and metallicities. The authors find that the density profile becomes shallower beyond $15-20$ kpc. This is in stark contrast to the studies by Watkins et al. (2009) and Sesar et al. (2010), who find a significantly steeper density profile beyond $\sim30$ kpc from the distribution of RR Lyrae stars in SDSS Stripe 82. A caveat to the interpretation of kinematic studies is that the density distribution is not measured directly but rather modelled by assuming a dark matter halo potential.

Therefore, the present state of play is distressingly inconclusive, and a further attack on the problem of the shape of the stellar halo is warranted. In this study, we introduce a new method to model both BHB and BS stars based on photometric information alone. We make use of the SDSS data release 8 (DR8) photometric survey which has now mapped an impressive $\sim14$ 000 deg$^2$ of the sky with both northern and southern coverage. In contrast to previous work, we combine both the wide sky coverage of the SDSS with the accurate distance estimates provided by the BHB stars to model the density profile of the stellar halo out to $\sim40$ kpc.

The paper is arranged as follows. In Section 2.1, we describe the SDSS DR8 photometric data and our selection criteria for A-type stars. The remainder of Section 2 introduces the probability distribution for BHB and BS membership based on colour alone and outlines the absolute magnitude–colour relations for the two populations. In Section 3, we describe our maximum-likelihood method to determine the density profile of the stellar halo, and in Section 4 we present our results. Finally, we draw our main conclusions in Section 5.

## 2 A-TYPE STARS IN SDSS DATA RELEASE 8

### 2.1 DR8 imaging

The SDSS (York et al. 2000) is an imaging and spectroscopic survey covering roughly $\sim1/4$ of the sky. Imaging data is obtained using a CCD camera (Gunn et al. 1998) on a 2.5-m telescope (Gunn et al. 2006) at Apache Point Observatory, New Mexico. Images are obtained simultaneously in five broad optical bands ($ugriz$; Fukugita et al. 1996). The data are processed through pipelines to measure photometric and astrometric properties (Lupton et al. 2001; Smith et al. 2002; Stoughton et al. 2002; Pier et al. 2003; Ivezić et al. 2004; Tucker et al. 2006). The SDSS DR8 contains all of the imaging data taken by the SDSS imaging camera and covers over $\sim14$ 000 deg$^2$ of the sky (Aihaara et al. 2011). We select objects classified as stars with clean photometry. The magnitudes and colours we use in the following sections have been corrected for extinction following the prescription of Schlegel, Finkbeiner & Davis (1998).

In the top panel of Fig. 1, we show the sky coverage of SDSS DR8 in equatorial coordinates. For comparison, the sky coverage of the SDSS DR5 is indicated by the darker grey region. The more recent SDSS DR covers both northern and southern latitudes. We exclude latitudes $|b| > 30^\circ$, so as to concentrate on regions well away from the plane of the galaxy. Over the distance range probed by this work, the SDSS footprint $(|b| > 30^\circ)$ covers approximately $20\per$ of the total volume of the stellar halo. In this study, we use BHB and BS stars to map the density profile of the stellar halo. These A-type stars are selected by choosing stars in the following region in colour–colour space:

$$0.9 < u - g < 1.4$$
$$-0.25 < g - r < 0.0.$$  

This selection is similar to other work using A-type stars (e.g. Yanny et al. 2000; Sirko et al. 2004) and is chosen to exclude main-sequence stars, white dwarfs and quasars. The bottom left-hand panel of Fig. 1 highlights the colour selection box. Whilst we assume that all of our selected stars are BHB or BS stars, there may be a non-negligible contribution by variable stars, such as RR Lyrae. We use multi-epoch Stripe 82 data to estimate the fraction of variable stars in the same magnitude and colour range as our sample. We use the light-curve catalogue compiled by Bramich et al. (2008) and classify variable stars according to the criteria outlined in Sesar et al. (2007). The resulting fraction of variable stars is $\sim5$ per cent. This small fraction of non-BHB and non-BS stars will therefore make little difference to the results of this work.

In the bottom right-hand panel of Fig. 1, we show the error in $u - g$ and $g - r$ colours as a function of $g$-band magnitude. The photometric errors in $u - g$ are larger than in $g - r$, especially at fainter magnitudes. This can be compared with the typical separation between BHB and BS stars in $u - g$ colour (see bottom panel of Fig. 2) which ranges from 0.05 to 0.1 mag. Mean photometric error $\sigma_{(u-g)}$ reaches 0.05 at $g \sim 18.5$, and beyond that the value rapidly increases. Accordingly, in this study we only select stars in the magnitude range $16 < g < 18.5$. This corresponds to a distance range of $\sim4-40$ kpc for typical absolute magnitudes of BHB and BS stars. Note that BHB and BS stars have different absolute magnitudes, and so span separate, but overlapping, distance ranges (see Section 2.3).
2.2 Ridgelines in colour–colour space

We seek to measure the centroids of the BHB and BS loci in colour space as well as their intrinsic widths. Naturally, this can only be done provided there exists a robust classification of the A-type stars according to their surface gravity. It has been shown (e.g. Clewley et al. 2002; Sirko et al. 2004; Xue et al. 2008) that BHB and BS stars can be separated cleanly on the basis of their Balmer line profiles. We proceed by selecting A-type stars from the spectral SDSS DR7 catalogue within the same colour range as our DR8 photometric sample. Restricting the sample to high signal-to-noise ratio (S/N) spectra in the magnitude range 16 < g < 17, we fit the Balmer lines Hγ, Hδ and Hβ with Sérsic profile of the form

\[ y = 1.0 - a \exp \left( \frac{|x - x_0|}{b} \right)^c, \]

where \( x_0 \) and \( a \) give the wavelength and the line depth at the line centre, respectively. The parameters \( b \) (the scale width) and \( c \) are related to the linewidth and line shape, respectively. The relation between combined linewidths and line shapes of the three Balmer lines Hγ, Hδ and Hβ are shown in the top panel of Fig. 2. The BHB (blue points) and BS (red points) stars are clearly separated in this diagram, and the decision boundary is indicated by the dashed black line. We use this spectral classification to pinpoint the loci of the two populations in the \( u-g, g-r \) colour–colour space. In our analysis, these 'ridgelines', i.e. approximate centres in \( u-g \) as a function of \( g-r \), are defined by third-order polynomials:

\[
(u - g)_{\text{BHB}}^0 = 1.167 - 0.775(g - r) - 1.934(g - r)^2 \\
+ 9.936(g - r)^3,
\]

\[
(u - g)_{\text{BS}}^0 = 1.078 - 0.489(g - r) + 0.556(g - r)^2 \\
+ 13.444(g - r)^3,
\]

for \(-0.25 < g - r < 0.0\). The ridgelines are shown by the thick blue and red lines in the bottom panel of Fig. 2. The green line indicates the approximate boundary between BHB and BS stars in \( u-g, g-r \) space. In addition, we calculate the intrinsic spread of the two populations about their ridgelines. We find \( \sigma_{\text{BHB},0}(u - g) = 0.04 \) and \( \sigma_{\text{BS},0}(u - g) = 0.045 \). Fig. 2 makes it apparent that, even for brightest stars, photometric information alone is not enough to separate BHB and BS stars. Therefore, for a given star we define the probability of BHB or BS class membership based on its distance in colour–colour space from the appropriate locus. We assume that both populations are distributed in a Gaussian manner about their ridgelines. We model the conditional probability of measuring \( u-g \) and \( g-r \) colours, given the star of each species, as

\[
p(u|g) | \text{BHB} \propto \exp \left( -\frac{(u - g) - (u - g)^0_{\text{BHB}}}{2\sigma^2_{\text{BHB}}} \right),
\]

\[
p(u|g) | \text{BS} \propto \exp \left( -\frac{(u - g) - (u - g)^0_{\text{BS}}}{2\sigma^2_{\text{BS}}} \right).
\]
Note that, in fact, these probabilities are also functions of $g - r$ since the centre of the Gaussian distribution, $(u - g)^0$ varies with the $g - r$ colour (see equation 3). The dispersion about the ridgeline centre depends on the intrinsic width and the photometric errors in $u - g$, $\sigma = \sqrt{\sigma^2_0 + \sigma^2_{u-g}}$. The colour-based posterior probabilities of class membership are then

$$P(\text{BHB} \mid ugr) = \frac{p(ugr \mid \text{BHB})N_{\text{BHB}}}{p(ugr \mid \text{BHB})N_{\text{BHB}} + p(ugr \mid \text{BS})N_{\text{BS}}}.$$  

$$P(\text{BS} \mid ugr) = \frac{p(ugr \mid \text{BS})N_{\text{BS}}}{p(ugr \mid \text{BHB})N_{\text{BHB}} + p(ugr \mid \text{BS})N_{\text{BS}}}.$$  

The total numbers of stars $N_{\text{BHB}}$ and $N_{\text{BS}}$ in a given colour range can then be found iteratively by integrating equations (5). In Table 1, we give the fraction of BHB and BS stars in five $g - r$ colour bins. The fraction ranges from $f_{\text{BHB}}$ ~ 0.15 at redder colours to $f_{\text{BHB}}$ ~ 0.6 at bluer colours. This is in good agreement with the overall BHB-to-BS ratios estimated by Bell et al. (2010) and Xue et al. (2008) who use similar magnitude ranges. Fig. 3 demonstrates the evolution of the separation between the two populations in colour space. For illustration purposes, the SDSS $u - g$ colour is transformed into $(u - g)^*$ using the following relation:

$$(u - g)^* = (u - g) - (u - g)^0_{\text{border}}.$$  

Here, $(u - g)^0_{\text{border}} = 1.223 - 0.632(g - r) - 0.689(g - r)^2 + 11.690(g - r)^3$ is defined by the approximate boundary line between BHB and BS stars shown by the green line in Fig. 2. In these new coordinates, the curved shape of the decision boundary becomes a straight line. In Fig. 3, we fit two Gaussian distributions to the distribution in $(u - g)^*$ for five bins in $g - r$. This two-component model fits the overall distribution very well. The ratio between the amplitudes of the two Gaussians varies with $g - r$ colour. The fraction of BHB stars increases towards bluer colours, in good agreement with our estimates in Table 1.

### Table 1. The fraction of BHB and BS stars in different colour bins. We give the $g - r$ colour range, the total number of stars, the estimated fraction of BHB stars and the estimated fraction of BS stars.

| $g - r$ range | $N_{\text{tot}}$ | $f_{\text{BHB}}$ | $f_{\text{BS}}$ |
|--------------|------------------|-----------------|----------------|
| $-0.05 < g - r < 0.00$ | 5189 | 0.154 | 0.846 |
| $-0.10 < g - r < -0.05$ | 4973 | 0.231 | 0.769 |
| $-0.15 < g - r < -0.10$ | 4151 | 0.346 | 0.654 |
| $-0.20 < g - r < -0.15$ | 3564 | 0.536 | 0.464 |
| $-0.25 < g - r < -0.20$ | 2413 | 0.613 | 0.387 |

2.3 Absolute magnitudes

Let us now derive a relationship between the absolute magnitude and colour of BHB and BS stars. BHB stars are intrinsically brighter, and their absolute magnitude varies little as a function of temperature (colour) or metallicity. In comparison, BS stars are intrinsically fainter and span a much wider range in absolute magnitude.

The absolute magnitudes of BHB stars are calibrated using star clusters with SDSS photometry published by An et al. (2008). 10 star clusters have prominent BHB sequences: M2, M3, M5, M13, M53, M92, NGC 2419, 4147, 5053 and 5466. The density distribution of absolute magnitudes of BHB stars in these clusters is shown as a function of $g - r$ colour in the left-hand panel of Fig. 4. There is little variation of the absolute magnitude with colour for BHB stars. $M_g$ changes by 0.2 (from 0.65 to 0.45) in the $-0.25 < g - r < 0$ range. The inset panel shows the variation of absolute magnitude about the $M_g(\text{BHB})$ versus $g - r$ trend ($M_g^*$). The spread of this distribution is $\sim$0.1, indicating that there is a tight relation describing the BHB absolute magnitude. The star clusters have metallicities typical of halo stars and ranging from $-2.3 < [\text{Fe/H}] < -1.3$, but we find no obvious trend with metallicity.

BS stars are less common in globular clusters than BHB stars. Instead, to calibrate the absolute magnitudes of BS stars, we make use of stars in Stripe 82 belonging to the Sgr stream. The distance to the stream in the right ascension range $25^\circ < \alpha < 40^\circ$ was

$1$ We adopt distance moduli for NGC 2419 and 4147 of 19.8 and 16.32, respectively. These differ from the values given in table 1 of An et al. (2008). We find that these revised values are in better agreement with the colour–magnitude diagrams of the clusters.
The distribution of colours in (\(\cos\)) space. Each row shows the distribution for a different range in \(g-r\) colour. A two-component model is fitted to the data. The overall model is shown by the black line. Individual Gaussians are shown by the red (BS) and blue (BHB) lines, respectively. Stars with \((u-g)^* < 0\) are dominated by BS stars, whilst stars with \((u-g)^* > 0\) are dominated by BHB stars. The ratio between the amplitude of the Gaussians gives an estimate of the overall number ratio between the two populations.

estimated by Watkins et al. (2009) using RR Lyrae stars as \(D_{\text{SP}} = 26.1 \pm 5.6\) kpc. In the middle panel of Fig. 4, we show the absolute magnitude of stars in Stripe 82 between \(25^\circ < \alpha < 40^\circ\) as a function of colour. The density contours are constructed by using stars outside of the range in right ascension as a background and by computing the density contrast. An obvious plume of BS stars is apparent in the density plot. This is shown by the contour levels extending off the main sequence. The solid red line shows the estimated absolute magnitude–colour relation for these BS stars. For comparison, we show the absolute magnitude versus colour relation for BS stars estimated by Kinman et al. (1994). This relation is converted from Johnson–Cousins photometry (\(UBV\)) to Sloan photometry (\(agr\)) using the transformation derived in Jester et al. (2005). The different coloured dots show the relation for different metallicity stars. The absolute magnitude calibration derived for the BS stars in the Sgr stream is almost identical to the Kinman et al. (1994) relation for BS stars with metallicity [Fe/H] = −1.5. This is in good agreement with the metallicity of the stream stars found by Watkins et al. (2009), ([Fe/H] = −1.43).

We compute the spread of absolute magnitudes for each colour bin as \(\sigma_{\text{dist}} \sim 0.5\). This dispersion takes into account the distance errors (\(D_{\text{SP}} = 26.1 \pm 5.6\) kpc) and encompasses a range of metallicities (see dashed red lines in Fig. 4). Hence, we conclude that our calibration for BS absolute magnitudes does not have a strong metallicity bias. Note that the ‘average’ BS absolute magnitude is \(\sim 2.5\), approximately 2 mag fainter than the BHB stars. This is entirely consistent with the absolute magnitudes of halo BS stars found by Yanny et al. (2000) and Clewley et al. (2004).

The resulting absolute magnitudes of BHB and BS stars as a function of \(g-r\) colour are

\[
M_g(\text{BHB}) = 0.434 - 0.169(g - r) + 2.319(g - r)^2 \\
+ 20.449(g - r)^3 + 94.517(g - r)^4,
\]

which are valid over the colour range \(-0.25 < g-r < 0\). For BHB stars, this allows us to estimate accurately the distances of BHB candidates. For BS stars, the relation is only approximate and the scatter for each colour interval needs to be taken into account for all distance estimates.

In the right-hand panel of Fig. 4, we show the estimated radial distributions for the BHB and BS populations. We select high-probability BHB and BS stars by using the membership probabilities defined in equation (5). ‘High’-probability BHB/BS stars are defined as those for which we are 68 per cent (or 1\(\sigma\)) confident of BHB/BS membership. The BHB stars probe a radial range from \(r \sim 10\) to \(\sim 45\) kpc. BS stars, which have fainter absolute magnitudes, only probe out to \(\sim 30\) kpc. However, there is a large degree of overlap between the two populations between 10 and 30 kpc.

3 MAXIMUM-LIKELIHOOD METHOD

In this section, we outline the maximum-likelihood method used to constrain the density profile of the stellar halo. An important assumption in the modelling is that the BHB and BS stars follow the same density distribution, modulo an overall scaling. The number of BHB and BS stars in a given increment of magnitude and area on the sky is then described by

\[
\Delta N_{\text{BHB}} \left( m_g - M_{\text{BHB}}^g, \ell, b \right) = \rho_{\text{BHB}}(m_g - M_{\text{BHB}}^g, \ell, b) D_{\text{BHB}}^3 \\
\times \frac{1}{5} \ln 10 \Delta m_g \cos \ell \Delta \ell \Delta b
\]

\[
\Delta N_{\text{BS}} \left( m_g - M_{\text{BS}}^g, \ell, b \right) = \rho_{\text{BS}}(m_g - M_{\text{BS}}^g, \ell, b) D_{\text{BS}}^3 \\
\times \frac{1}{5} \ln 10 \Delta m_g \cos \ell \Delta \ell \Delta b.
\]

Here, the distance increment \(\Delta D\) has been converted into the apparent magnitude increment via the relation \(\Delta D = \frac{1}{5} \ln 10 D \Delta m_g\). The normalizing factors, \(\rho_{\text{BHB}} = N_{\text{BHB}}V_{\text{BHB}}\) and \(\rho_{\text{BS}} = N_{\text{BS}}V_{\text{BS}}\), are found by performing volume integrals over the SDSS DR8 sky.
coverage and over the required magnitude range:

\[
V_{\text{BHB}}(M_g^{\text{BHB}}) = \int \int \int \left[ \rho \left( m_g - M_g^{\text{BHB}} \right), \ell, b \right] D_{\text{BHB}}^3 \times \frac{1}{5} \ln 10 \, dm_g \cos b \, d\ell \, db .
\]

\[
V_{\text{BS}}(M_g^{\text{BS}}) = \int \int \int \left[ \rho \left( m_g - M_g^{\text{BS}} \right), \ell, b \right] D_{\text{BS}}^3 \times \frac{1}{5} \ln 10 \, dm_g \cos b \, d\ell \, db .
\]

These normalizing integrals depend on the absolute magnitude of the stars and hence on the \( g - r \) colour \( (M_g = M_g(g - r)) \) from equation 7. Note that the values of \( V_{\text{BHB}} \) and \( V_{\text{BS}} \) play a minor role in identifying the maximum-likelihood model given the choice of parameters, but are important when evaluating the performances of different model families. We also assume that our sample consists only of BHB and BS stars so the total number of stars is the sum of these two populations, \( N_{\text{tot}} = N_{\text{BHB}} + N_{\text{BS}} \), where \( N_{\text{BHB}} = f_{\text{BHB}}N_{\text{tot}} \) and \( N_{\text{BS}} = f_{\text{BS}}N_{\text{tot}} \). The overall fraction of BHB to BS stars varies as a function of \( g - r \) colour, as shown in the previous section.

Combining equations (4) and (8) gives the number of stars in a cell of colour, magnitude, longitudinal and latitudinal space:

\[
\Delta N = p(\text{ugr} \mid \text{BHB}) \Delta N_{\text{BHB}} + p(\text{ugr} \mid \text{BS}) \Delta N_{\text{BS}}
\]

\[
= N_{\text{tot}} v(\text{ugr}, l, b) \Delta (g - r) \Delta m_g \Delta \ell \Delta b \cos b \frac{1}{5} \ln 10 .
\]

where the stellar density is

\[
v(\text{ugr}, l, b) = p(\text{ugr} \mid \text{BHB}) \frac{f_{\text{BHB}}}{V_{\text{BHB}}} \left( m_g - M_g^{\text{BHB}} \right), \ell, b \right] D_{\text{BHB}}^3 \]

\[
+ p(\text{ugr} \mid \text{BS}) \frac{f_{\text{BS}}}{V_{\text{BS}}} \left( m_g - M_g^{\text{BS}} \right), \ell, b \right] D_{\text{BS}}^3 .
\]

Here, each star is assigned a ‘BHB distance’ \( (D_{\text{BHB}}) \) as well as a ‘BS distance’ \( (D_{\text{BS}}) \). The colour probability functions weight the contribution of each star to the BHB density or the BS density. For simplicity, we group all the stars into five bins in \( g - r \) of width 0.05 mag. Thus, stars in each \( g - r \) bin have the same normalizations

\[(V_{\text{BHB}}, V_{\text{BS}}), \text{fraction of BHB/BS stars } (f_{\text{BHB}}, f_{\text{BS}}) \text{ and absolute magnitudes } (M_{g^{\text{BHB}}}, M_{g^{\text{BS}}}) \]. We take into account the uncertainty in the BS absolute magnitude by convolving the BS number density with a Gaussian magnitude distribution. This distribution is centred on the estimated absolute magnitude \( (M_g^{\text{BS}} = M_{g^{\text{BS}}}(g - r)) \) and has a dispersion of \( \sigma_M = 0.5 \) (see Fig. 4).

The log-likelihood function can then be constructed from the density distribution,

\[
\log L = \sum_{i=1}^{N_{\text{tot}}} \log \left[ v \left( m_g^{i}, \text{ugr}^i, \ell^i, b^i \right) \right] \cos b .
\]

The number of free parameters constrained by the likelihood function depends on the complexity of the model stellar halo density profile. The log-likelihood is maximized to find the best-fitting parameters using a brute-force grid search.

4 RESULTS

In this section, we outline the results of applying our maximum-likelihood method to the sample of A-type stars selected from the SDSS DR8. We consider in turn a number of simple density profiles with constant flattening – single power law, broken power law and Einasto – before examining the case for refinements, such as triaxiality, radial variations in shape and substructure.

4.1 Single power-law profile

First, let us consider a simple power-law density model of the form

\[
\rho(r_g) \propto r_g^{-\alpha} , \quad r_g^2 = x^2 + y^2 + z^2 q^{-2} .
\]

The parameters \( q \) and \( \alpha \) describe the halo flattening and the power-law fall-off in stellar density, respectively. Oblate density distributions have \( q < 1 \), spherical \( q = 1 \) and prolate \( q > 1 \).

We summarize the best-fitting single power-law model parameters in Table 2. Using the entire sample, we find the maximum-likelihood model parameters of \( \alpha = 2.6 \) and \( q = 0.65 \). We repeat the analysis by excising stars in the regions of the Virgo overdensity

© 2011 The Authors, MNRAS 416, 2903–2915
Monthly Notices of the Royal Astronomical Society © 2011 RAS
Table 2. A summary of our best-fitting oblate power-law models with and without the Virgo overdensity and the Sgr stream. We give the total number of stars, the model parameters, the average log-likelihood value for the model and $\sigma_{\text{tot}}$.

| $N_{\text{tot}}$ | $\alpha$ | $q$ | $\ln(L) \times 10^4$ | $\sigma_{\text{tot}}$ | With V and S? |
|------------------|----------|-----|---------------------|------------------------|---------------|
| 20 290           | 2.60$^{+0.05}_{-0.06}$ | 0.65$^{+0.02}_{-0.02}$ | $-17.117$ | 0.38 $\pm$ 0.01 | Yes           |
| 15 403           | 2.90$^{+0.05}_{-0.01}$ | 0.53$^{+0.02}_{-0.01}$ | $-12.2917$ | 0.22 $\pm$ 0.02 | No            |

and Sgr stream, which amounts to removing stars in the regions defined by

$$0 < X < 30, \quad X = 63.63961\sqrt{2(1 - \sin b)}.$$  (14)

This is the mask introduced by Bell et al. (2008) to remove stars belonging to the Virgo overdensity (as well as parts of the leading tail of the Sgr stream). Another portion of the Sgr stream is located in the southern part of the sky (see Fig. 5) and is removed by

$$0' < \alpha < 50', \quad -30' < \delta < 0'. $$  (15)

By discarding these large overdensities, we find a slightly steeper power law and a more flattened shape with $\alpha = 2.9$ and $q = 0.53$. There is also a substantial increase in the maximum likelihood, implying that the fit is indeed affected by the presence of these large overdensities. In Fig 6, we show the maximum-likelihood contours for the model parameters $q$ and $\alpha$ for both cases. The contours encompass the 1$\sigma$, 2$\sigma$ and 3$\sigma$ regions, respectively. We find that our likelihood function has a well-defined peak. The maximum-likelihood model parameters ($\alpha = 2.9$, $q = 0.53$, excluding overdensities) are in good agreement with some of the previous work assuming a single power-law model for the stellar halo (e.g. Yanny et al. 2000; Newberg & Yanny 2006; Jurić et al. 2008). This exercise indicates that the two large overdensities induce a bias in deduced model parameters, and, therefore, in the subsequent sections, we remove stars in the vicinity of the Virgo overdensity and Sgr stream in all our calculations.

We use our maximum-likelihood smooth oblate halo model ($\alpha = 2.9$, $q = 0.53$) to show the residuals of the data minus model on the sky. In Fig. 5, we compare the data and model in both equatorial and galactic coordinates. The top and middle panels show the data and model on the sky (the magnitude distribution has been collapsed), whilst the bottom panels show the residuals of the model. The Virgo overdensity is the most obvious feature located at ($\alpha \sim 190^\circ$, $\delta \sim 30^\circ$). This overdensity covers a substantial fraction of the North Galactic cap. A portion of the Sgr stream can be seen at ($\alpha \sim 20^\circ$, $\delta \sim -25^\circ$) in the South Galactic cap. The residuals for these features reach up to approximately three times the model values. We can estimate the total fraction of stars residing in these overdensities from the excess numbers of stars in these regions of the sky. Approximately $\sim$5.5 per cent of our total sample ($N_{\text{tot}} = 20 290$) reside in the northern Virgo+Sgr overdensity, whilst <1 per cent occupy the southern portion of the Sgr stream. The fraction of

![Figure 5](https://example.com)
stars in these overdensities is relatively small as they are located in the vicinity of the Northern and Southern Galactic caps where the density of stars is small. However, as the relative difference [i.e. (data-model)/model] between the data and model is large in these regions; these overdensities can influence the likelihood values.

4.2 Broken power-law profile

We relax our models to allow a change in the steepness of the density fall-off and consider broken power-law models of the form

\[ \rho(q) \propto \begin{cases} r_q^{-\alpha_{\text{in}}} & r_q \leq r_b \\ r_q^{-\alpha_{\text{out}}} & r_q > r_b \end{cases} \]

In Fig. 7, we show the maximum-likelihood contours for the inner and outer power laws (left-hand panel) and break radii and flattening (right-hand panel). A model with break radius \( r_b = 27 \text{kpc} \) is preferred with slopes of \( \alpha_{\text{in}} = 2.3, \alpha_{\text{out}} = 4.6 \). A steeper power law is favoured at larger radii, while the power law within the break radius is shallower. Note that the break radius is an ellipsoidal fall-off and consider broken power-law models of the form

\[ \rho(q) \propto \begin{cases} r_q^{-\alpha_{\text{in}}} & r_q \leq r_b \\ r_q^{-\alpha_{\text{out}}} & r_q > r_b \end{cases} \]

\[ \ln \frac{\rho(q)}{\rho(q_{\text{eff}})} = -d_q \left( \frac{r_q}{r_{\text{eff}}} \right)^{1/n} - 1. \]

The shape of the density profile is described by the parameter \( n \). Density distributions with steeper inner profiles and shallower outer profiles are generated by large values of \( n \). For example, dark matter haloes with 'cuspy' inner profiles typically have values of \( n \sim 6 \). The parameter \( d_q \) is a function of \( n \). For \( n \geq 0.5 \), a good approximation is given by \( d_q = 3n - 1/3 + 0.0079/n \) (Graham et al. 2006).

This profile allows for a non-constant fall-off without the need for imposing a discontinuous break radii. We find the maximum-likelihood solutions for \( q \) and \( r_{\text{eff}} \). Our best-fitting Einasto model has parameters \( q = 0.58, n = 1.7 \) and \( r_{\text{eff}} = 20 \text{kpc} \). The slope of the density profile varies rapidly with radius as indicated by the relatively small value of \( n \).

4.4 Model comparisons

We now test how accurately our models represent the observed magnitude distribution of the data. Using Monte Carlo methods, we create a distance distribution according to the model density profile. The fake data is given a \( g - r \) colour distribution drawn from the real data, which can then be converted into absolute magnitudes (see equation 7). The ratio of BHB to BS stars in each colour bin is chosen to match the values given in Table 1. In the case of the BS stars, absolute magnitudes are determined from the \( g - r \) colour by drawing randomly from a Gaussian distribution centred on the estimated value (from equation 7) with a dispersion of 0.5 mag. The resulting (apparent) magnitude distributions for our models are compared with the data in Fig. 8.
In the top left-hand panel, we show the magnitude distribution for our single power-law oblate model with the solid red line. The error bars take into account the Poisson uncertainty as well as the uncertainty spread of absolute magnitudes for BS stars. The distribution of magnitudes for our DR8 A-type star sample is shown by the black points. There is reasonably good agreement, but there is a notable deviation at fainter magnitudes. For comparison, we show the maximum-likelihood spherical model by the red dashed line (with $\alpha \sim 2.7$), which provides a very poor fit to the data. The top right-hand panel shows the magnitude distributions for the single power-law, broken power-law and Einasto profiles by the solid red, blue and green lines, respectively. The broken power-law and Einasto models provide a better representation of the data than a single power-law model, although there are still discrepancies at the faintest magnitudes.

In the bottom panels of Fig. 8, we show the magnitude distributions for the most probable BHB and BS stars. We select stars with $P_{\text{BHB}} > 0.7$ as 'BHB' stars and stars with $P_{\text{BS}} > 0.7$ as 'BS' stars (see Section 2.3). Of course, this is for illustration purposes only, as a clean separation of BHB and BS stars requires spectroscopic classification. While the BS stars are adequately described by a single power-law model, this is a poor description of the BHB stars. This is unsurprising, as the BS stars cover a smaller distance range and barely populate distances beyond the break radius. The bottom panels of Fig. 8, we show the magnitude distributions for the most probable BHB and BS stars. We select stars with $P_{\text{BHB}} > 0.7$ as 'BHB' stars and stars with $P_{\text{BS}} > 0.7$ as 'BS' stars (see Section 2.3). Of course, this is for illustration purposes only, as a clean separation of BHB and BS stars requires spectroscopic classification. While the BS stars are adequately described by a single power-law model, this is a poor description of the BHB stars. This is unsurprising, as the BS stars cover a smaller distance range and barely populate distances beyond the break radius. The bottom panels of Fig. 8, we show the magnitude distributions for the most probable BHB and BS stars. We select stars with $P_{\text{BHB}} > 0.7$ as 'BHB' stars and stars with $P_{\text{BS}} > 0.7$ as 'BS' stars (see Section 2.3). Of course, this is for illustration purposes only, as a clean separation of BHB and BS stars requires spectroscopic classification. While the BS stars are adequately described by a single power-law model, this is a poor description of the BHB stars. This is unsurprising, as the BS stars cover a smaller distance range and barely populate distances beyond the break radius. The bottom panels of Fig. 8, we show the magnitude distributions for the most probable BHB and BS stars. We select stars with $P_{\text{BHB}} > 0.7$ as 'BHB' stars and stars with $P_{\text{BS}} > 0.7$ as 'BS' stars (see Section 2.3). Of course, this is for illustration purposes only, as a clean separation of BHB and BS stars requires spectroscopic classification. While the BS stars are adequately described by a single power-law model, this is a poor description of the BHB stars. This is unsurprising, as the BS stars cover a smaller distance range and barely populate distances beyond the break radius. The bottom panels of Fig. 8, we show the magnitude distributions for the most probable BHB and BS stars. We select stars with $P_{\text{BHB}} > 0.7$ as 'BHB' stars and stars with $P_{\text{BS}} > 0.7$ as 'BS' stars (see Section 2.3). Of course, this is for illustration purposes only, as a clean separation of BHB and BS stars requires spectroscopic classification. While the BS stars are adequately described by a single power-law model, this is a poor description of the BHB stars. This is unsurprising, as the BS stars cover a smaller distance range and barely populate distances beyond the break radius. The bottom panels of Fig. 8, we show the magnitude distributions for the most probable BHB and BS stars. We select stars with $P_{\text{BHB}} > 0.7$ as 'BHB' stars and stars with $P_{\text{BS}} > 0.7$ as 'BS' stars (see Section 2.3). Of course, this is for illustration purposes only, as a clean separation of BHB and BS stars requires spectroscopic classification. While the BS stars are adequately described by a single power-law model, this is a poor description of the BHB stars. This is unsurprising, as the BS stars cover a smaller distance range and barely populate distances beyond the break radius. The bottom panels of Fig. 8, we show the magnitude distributions for the most probable BHB and BS stars. We select stars with $P_{\text{BHB}} > 0.7$ as 'BHB' stars and stars with $P_{\text{BS}} > 0.7$ as 'BS' stars (see Section 2.3). Of course, this is for illustration purposes only, as a clean separation of BHB and BS stars requires spectroscopic classification. While the BS stars are adequately described by a single power-law model, this is a poor description of the BHB stars. This is unsurprising, as the BS stars cover a smaller distance range and barely populate distances beyond the break radius.
left-hand panel clearly shows the need for a more steeply declining density law at larger radii.

In Fig. 9, we show the distribution of (probable) BHB stars in spherical shells with the solid black line. Here, we only consider the most likely BHB stars (with \(P_{\text{BHB}} > 0.7\), as for these we can accurately estimate their distance. The red, blue and green lines show the radial distributions for our best-fitting single power-law, broken power-law and Einasto models, respectively. The single power-law model is a poorer description of the data, whilst the broken power-law and Einasto models both provide better representations of the radial distribution.

4.5 Refinements: triaxiality and a radially dependent shape

A natural question to ask is whether further refinements might provide a still more accurate description of the data. Up to now, we have assumed spheroidal halo models with a constant flattening with radius.

First, we consider whether triaxiality makes any improvement. The definition of \(r_q\) is modified to

\[
r_q^2 = x^2 + y^2 p^{-2} + z^2 q^{-2}.
\]

We fit single power-law triaxial models to the data and obtain maximum-likelihood parameters of \(q = 0.50\) and \(p = 0.71\). The likelihood value increases relative to an oblate model with the inclusion of this extra parameter \([-2\ln(\mathcal{L}_{\text{oblate}}/\mathcal{L}_{\text{triaxial}}) \sim 200]\). The magnitude distribution of the triaxial model is largely the same as the oblate model. We inspect the difference between these two models in equatorial coordinates in Fig. 10. The dot–dashed lines indicate the regions of the sky with low galactic latitudes \(|b| < 40^\circ\). Regions with latitudes below \(|b| < 30^\circ\) have been removed. The triaxial model is notably overdense relative to the oblate model in the regions centred on \((\alpha, \delta) = (125^\circ, 50^\circ)\) and \((\alpha, \delta) = (60^\circ, 10^\circ)\). The latter region, located in the southern part of the sky, is close to the portion of the Sgr stream excised in our best-fitting model. The former is coincident with the Monoceros ring (Newberg et al. 2002), which can be identified, for example, in fig. 1 of Belokurov et al. (2006). We suggest that the apparent triaxiality may well be caused by the presence of these overdensities. The increased flexibility of the model can cause it to ‘fit’ to such substructure, and thus the increase in likelihood may be an artefact.

Some earlier investigations have found evidence that the shape of the stellar halo changes from a flattened distribution at smaller radii to an almost spherical distribution at larger radii (e.g. Preston et al. 1991). We can test this claim by allowing the flattening \(q\) to vary with radius. Following the reasoning of Sluis & Arnold (1998), we make the following substitution:

\[
q \rightarrow q \sqrt{\frac{r^2 + r_0^2}{q^2 r^2 + r_0^2}}. \tag{19}
\]

The halo flattening is still \(q\) at small radii but tends to sphericity at large radii. The scale radius \(r_0\) determines the radial range over which this change occurs. For example, for large values of \(r_0\) (e.g. larger than the most distant stars) the flattening is approximately constant over the applicable radial range. We fit single power-law models with a varying flattening to the data and find that large-scale radii are preferred \((r_0 > 10^3 \text{kpc})\), with an inner flattening of \(q = 0.53\). This indicates that the flattening is approximately constant over the radial range of the data (out to \(\sim 40\text{kpc}\)). This is in agreement with the deductions of Sluis & Arnold (1998) and Sesar et al. (2011), who also found no real evidence for a varying shape with radius.

In summary, we find that the data are well described by an oblate density distribution with a constant flattening of \(q \sim 0.6\) and a more steeply declining profile at larger radii. We summarize our maximum-likelihood models by giving a ‘side view’ of the profiles in Fig. 11. A spherical model is strongly disfavoured, whereas oblate broken power-law and Einasto models provide a good representation.

Our best-fitting density distribution can be used to estimate the total stellar mass. We find the total number of BHB stars by integrating the BHB density profile over all space in the distance range 1–40 kpc. The number of BHB stars can be converted into a total luminosity using the relation derived in Deason, Belokurov &
The Milky Way stellar halo

The scale dependence of substructure. We show the

Figure 11. A side view of our maximum-likelihood models. Grey-scale shows density of stars in plane of Galactocentric (x, z) at y = 0 in four maximum-likelihood models. The blue, green, yellow and red contours show density levels corresponding to the 50th, 90th, 95th and 99th percentiles of the spherical model. The white dot marks the location of the Sun.

Evans (2011) using globular clusters: \( N_{\text{BHB}}/L \sim 10^{-3} \). Assuming a mass-to-light ratio \( M/L \sim 1–5 \), the stellar mass is approximately \( 2–10 \times 10^8 M_\odot \). Bell et al. (2008) calculate a total stellar mass of \( \sim 3.7 \times 10^8 M_\odot \) using main-sequence stars, in good agreement with our estimate. Note that the total stellar halo mass is believed to be \( \sim 10^9 M_\odot \) (e.g. Morrison 1993), so we are probing a significant fraction of the stellar halo (\( \sim 20–100 \) per cent).

4.6 The amount of substructure

A rough idea of the amount of substructure present in the data can be attained by computing the root mean square (rms) deviation of the models about the data (\( \sigma/\text{tot} \)) as defined by equations (2) and (3) in Bell et al. (2008):

\[
\langle \sigma^2 \rangle = \frac{1}{n} \sum_k (D_k - M_k)^2 - \frac{1}{n} \sum_k (M_k - \bar{M})^2, \tag{20}
\]

\[
\sigma/\text{tot} = \sqrt{\frac{\langle \sigma^2 \rangle}{n}} \tag{21}
\]

Here, \( D_k \) is the number of stars in bin \( k \), \( M_k \) is the expected number from the model, \( \bar{M} \) is a Poisson random deviate from the model value and \( n \) is the total number of bins (in \( l, b \) and \( m_g \)). However, with a total of \( \sim 20000 \) stars we cannot afford to use a fixed bin size over the entire sky: pixels must be simultaneously large enough to contain ample stars and small enough to adequately sample the substructure. To circumvent this problem, we use the Voronoi binning method of Cappellari & Copin (2003) to partition the data in pixels on the sky. This adaptive binning method groups pixels together, forming ‘superpixels’, with the objective of obtaining a constant S/N per bin. We choose the S/N \( \sim 4 \) (assuming Poisson noise) to ensure that the mean number of stars in each 2D bin is \( \gtrsim 10 \). We also check that our results are not affected by the choice of S/N. The stars are split into three g-band magnitude ranges, binned into \( \sim 65000 \) \( 1^\circ \times 1^\circ \) pixels for full sky coverage which are then combined into \( \sim 750 \) superpixels using the procedure described above to estimate the overall \( \sigma/\text{tot} \) values. These are given in the last two columns of Table 3. As expected, our \( \sigma/\text{tot} \) fraction is significantly reduced when the stars in the region of the Virgo overdensity and Sgr stream are removed: the estimated fraction of substructure reduces from 40 to 20 per cent. However, further refinement of our model makes little difference to the \( \sigma/\text{tot} \) fraction, even if the likelihood is increased. This is indicative of the resilience of our maximum-likelihood method to relatively minor amounts of substructure present in the data.

We illustrate the dependence of \( \sigma/\text{tot} \) on spatial scale in Fig. 12 by grouping together superpixels with similar spatial scales and computing \( \sigma/\text{tot} \) for each group. This shows how the fraction of substructure depends on the superpixel size (in \( \text{deg}^2 \)) for different maximum-likelihood models. \( \sigma/\text{tot} \) is computed both with (black lines) and without (red lines) the stars in the vicinity of the Virgo overdensity and the Sgr stream. The spherical model has relatively large values of \( \sigma/\text{tot} \) over all scales, illustrating the poor fit of this model to the data. The single and broken power-law oblate models show a scale dependence, whereby larger scales (larger number of pixels) have an increased \( \sigma/\text{tot} \) fraction relative to smaller scales. Similar behaviour is also true for the triaxial model, but the \( \sigma/\text{tot} \) fractions are slightly lower. As alluded to earlier, we suspect that this may be caused by the triaxial model describing some of the large-scale substructure (namely the Monoceros ring or parts of the Sgr stream which have not been excised).

Different lines in Fig. 12 show the dependence of the \( \sigma/\text{tot} \) measure on the apparent magnitude of the tracers and hence correspond to different distances probed. The slight rise of the rms with magnitude indicates that higher fractions of substructure may be present at larger distances. This can be simply explained with the increase

Figure 12. The scale dependence of substructure. We show the \( \sigma/\text{tot} \) fraction as a function of superpixel size. Individual panels show the relation for our various maximum-likelihood models. The different line styles indicate different magnitude bins. The black lines show the relation for all the stars. The red lines show the relation when stars in the vicinity of the Virgo overdensity and Sgr have been removed.
of the substructure lifetime with radius as governed by the orbital period of the infalling Galactic fragments.

When interpreting these plots, it is important to bear in mind that the smallest superpixels are biased towards the densest parts of the halo, while the emptiest regions are sampled by the largest agglomerations of 2D bins. This bias could potentially explain at least some of the trend with size. On the scales smaller than several tens of degrees, \( \sigma/\text{tot} \) in our models is in the range \( 0.05 < \sigma/\text{tot} < 0.2 \), indicating that the halo is not dominated by substructure and is relatively smooth. While there is evidence that the fraction of substructure increases with scale, this apparent increase in \( \sigma/\text{tot} \) could also be caused by spurious pixels. In sparse regions, the grouping of many pixels, each containing very few stars, poses a limitation to this method.

5 CONCLUSIONS

We have developed a new method to simultaneously model the density profile of both BHB and BS stars based on their broad-band photometry alone. The probability of BHB or BS membership is defined by the locus of these stars in \( u - g, g - r \) space. We use these colour-based weights to construct the overall probability function for the density of the two populations. When applied to the A-type stars selected from the SDSS DR8 photometric catalogue, the best-fitting stellar halo models are identified by applying a maximum-likelihood algorithm to the data.

Based on the data from regions with \( 30^\circ < |b| < 70^\circ \), we find that the stellar halo is not spherical in shape, but flattened. Spherical models cannot reproduce the distribution of the A-type stars in our sample and are discrepant with the data at both bright and faint magnitudes. Our best-fitting models suggest that the stellar halo is oblate with a flattening (or minor-axis-to-major-axis ratio) of \( q = 0.59 \) with a typical uncertainty of \( \sigma_q \sim 0.1 \). As a simple representation of the stellar halo, we advocate the use of a broken power-law model with an inner slope \( \alpha_{\text{in}} = 2.3 \) and an outer slope \( \alpha_{\text{out}} = 4.6 \), the break radius occurring at about 27 kpc. This gives the formula

\[
\rho(r) \propto \begin{cases} r_q^{-2.3} & r_q \leq 27 \text{ kpc}, \\ r_q^{-4.6} & r_q > 27 \text{ kpc}, \end{cases}
\]

with \( r_q^2 = R^2 + z^2/0.59^2 \). For those who prefer Einasto models, an equally good law for the stellar halo density is

\[
\rho(r) \propto \exp \left[ -4.77 \left( r_q/20\text{kpc} \right)^{1/1.7} - 1 \right].
\]

These two formulae are fundamental results of our paper, and can be summarized as the Milky Way stellar halo is \textit{squashed and broken}. The results are qualitatively similar to those obtained by Sesar et al. (2011) using main-sequence stars. There is no evidence in the data for variation of the flattening with radius. There is some mild evidence for a triaxial shape, but the apparent triaxiality might be due to the presence of the Monocerus ring at low latitudes (\( |b| < 40^\circ \)) and regions of the southern Sgr stream.

The rms deviation of the data around the maximum-likelihood model \( \sigma/\text{tot} \) typically ranges between 5 and 20 per cent. This indicates that the Milky Way stellar halo, or at least the component traced by the A-type stars in the SDSS DR8, is \textit{smooth} and not dominated by unrelaxed substructure.

This finding is discrepant with the conclusions of Bell et al. (2008). These authors use MSTO stars selected from SDSS DR5 to model the stellar halo density. They argue that no smooth model can describe the data and conclude that the stellar halo is dominated by substructure (with \( \sigma/\text{tot} \geq 0.33 \)). We suggest that these contrasting results may be due to the different methods and tracers used. Bell et al. (2008) search for the lowest \( \sigma/\text{tot} \) fraction models. However, we find that the \( \sigma/\text{tot} \) values vary little between different density models even if the likelihood is substantially increased. Furthermore, a lower \( \sigma/\text{tot} \) could also indicate that a model is ‘fitting’ to any substructure (e.g. triaxial models both in this work and in Bell et al. 2008). Bell et al. (2008) also use main-sequence stars as tracers which are more numerous than A-type stars, but are much poorer distance indicators. The adopted absolute magnitude scale for main-sequence stars is dependent on metallicity and colour. Bell et al. (2008) assume a median absolute magnitude of \( M_r \sim 4.5 \) with a scatter of \( \sigma_{M_r} = 0.9 \); an appropriate choice for a halo population with metallicity [Fe/H] \sim −1.5. The scatter, even with an assumed metallicity, is greater than the spread in absolute magnitude for BS stars (\( \sigma_{M_r} \sim 0.5 \) per colour bin). Uncertainties in the distances to stars could lead to compact substructures becoming more blurred out over a wider range of distance. It is possible that Bell et al. (2008) see a somewhat younger and/or metal-richer component of the stellar halo with the main-sequence tracers. This component could be more unrelaxed than the one traced by A-type stars.

Our conclusion that the stellar halo is composed of a smooth underlying density, together with some additional substructures such as the Virgo overdensity and the Sgr stream, is very reassuring. If the stellar halo were merely a hotchpotch of tidal streams and unrelaxed substructures, then modelling and estimation of total mass and potential would be much more difficult. Many of the commonly used tools of stellar dynamics – such as the steady-state Jeans equations – implicitly assume a well-mixed and smooth equilibrium. This raises the hope that a full understanding of the spatial and kinematic properties of stars in the smooth, yet squashed and broken, stellar halo can yield the gravitational potential and dark matter profile of the Galaxy itself.

ACKNOWLEDGMENTS

AJD thanks the Science and Technology Facilities Council (STFC) for the award of a studentship, whilst VB acknowledges financial support from the Royal Society. We thank the anonymous referee for many helpful suggestions.

REFERENCES

Aihara H. et al., 2011, ApJS, 193, 29
An D. et al., 2008, ApJS, 179, 326
Bell E. F. et al., 2008, ApJ, 680, 295
Bell E. F., Xue X. X., Rix H., Rahlan C., Hogg D. W., 2010, AJ, 140, 1850
Belokurov V. et al., 2006, ApJ, 642, L137
Belokurov V. et al., 2007, ApJ, 657, L89
Bramich D. M. et al., 2008, MNRAS, 386, 887
Brown W. R., Geller M. J., Kenyon S. J., Diaferio A., 2010, ApJ, 139, 59
Bullock J. S., Johnston K. V., 2005, ApJ, 635, 931
Cappellari M., Copin Y., 2003, MNRAS, 342, 345
Carollo D. et al., 2007, Nat, 450, 1020
Carollo D. et al., 2010, ApJ, 712, 692
Chen B. et al., 2001, ApJ, 553, 184
Clewley L., Warren S. J., Hewett P. C., Norris J. E., Evans N. W., 2002, MNRAS, 337, 87
Clewley L., Warren S. J., Hewett P. C., Norris J. E., Evans N. W., 2004, MNRAS, 352, 285
Cooper A. P. et al., 2010, MNRAS, 406, 744
De Lucia G., Helmi A., 2008, MNRAS, 391, 14
De Propris R., Harrison C. D., Mares P. J., 2010, ApJ, 719, 1582
