university of groningen / faculty of science and engineering

Bachelor Research Project
Kapteyn Astronomical Institute

Tracing the outer galactic halo with Blue Horizontal Branch stars

Author:
Lei Titulaer
(s3765679)

Supervisor:
dr. Else Starkenburg
Second reader:
prof. dr. Eline Tolstoy

July 1, 2021

Abstract

We present a new method to select Blue Horizontal Branch (BHB) stars from a sample of A-stars. BHB stars are used as standard candles to trace outward into the galactic halo, but in order to use them as such they need to be distinguished from contaminating Blue Straggler (BS) stars that have similar temperatures. Firstly, the method presented by Starkenburg et al. (2019) is applied to a new, larger sample of stars. The wide-field SDSS ugr-bands and the narrow-band Pristine $CaHK$-band are used to separate the BHB sequence from the BS sequence in the $u-CaHK$, $g-r$ color-color space. We then extend the method and use the deeper CFIS u-band in combination with Pan-STARRS gr-bands and Pristine $CaHK$-band in order to be able to improve the selection of stars at distances further out.

From the inferred heliocentric distances of the sample of BHB stars we are able to identify multiple known substructures and globular clusters in the halo. The Sagittarius stream is analysed and we are able to detect a recently discovered outward 'spur' feature by Sesar et al. (2017). The density profile of the 'smooth' component of the galactic halo is modelled by fitting a negative power law. This improves the result by Starkenburg et al. (2019) for the 'smooth' halo component. Using the new method utilising the CFIS u-band, we are not able to necessarily probe deeper into the halo, but we demonstrate that the completeness at distances over ~ 75 kpc is largely improved, providing a clearer picture of the halo substructure in the $75 - 120$ kpc regime.

Contents

1 Introduction

The galactic halo has been a topic of interest for astronomers for multiple decades (Helmi, 2020). This diffuse, far stretching Milky Way component enclosing all other constituents of the Milky Way can tell us a lot about the history of the Milky Way and its evolution over time. It is generally accepted that the Milky Way has formed through mergers with smaller galaxies and that these mergers explain the structures present in the Galaxy's halo (Helmi, 2020). Simulations have shown that the profile of the galactic halo is highly dependent on the accretion of these smaller galaxies. Accretions form substreams in the halo which become less apparent over time and eventually dissolve and become part of the 'smooth' halo component (e.g. Bullock & Johnston, 2005). The halo can thus tell a lot about how active the accretion history of the Milky Way has been.

A very interesting property of the halo is its density profile. We know that galaxies with steep profiles have experienced a much more active accretion history than galaxies of the same mass with less steep profiles (e.g. Thomas et al., 2018). Constraining the density profile of the Milky Way's halo is of great importance to infer how and how long ago the galaxy formed. Understanding the formation of the Milky Way halo in its turn allows to give further constraints on models explaining the formation of the Galaxy, ranging from understanding dark matter to constraining the mass of the Milky Way (Helmi, 2020).

The galactic halo contains many irregularities. One of the most prominent examples of the substructures in the outer galactic halo is the Sagittarius stream, a remnant of the merger of the Sagittarius dwarf galaxy ~ 8 Gyr ago (Helmi, 2020). In order to map the smooth halo component, the Sagittarius stream, and other structures in the galactic halo, an ongoing effort is made to trace out the halo as far as possible using standard candles like blue horizontal branch stars, RR Lyrae (e.g. Deason et al., 2011; Sesar et al., 2017; Fukushima et al., 2018) and other tracers (Helmi, 2020). Using standard candles is useful as their consistent absolute magnitude can be used to infer distances to the stars. However, we are constrained by sensitivity of surveys at faint magnitudes and effectiveness of selection mechanisms to select the tracers. In this work we will apply the method presented by Starkenburg et al. (2019) to a larger sky coverage, and select blue horizontal branch stars to trace the galactic halo. Additionally, we will adapt the method also to select blue horizontal branch stars but while using a different, deeper filter system.

1.1 BHB stars

Blue horizontal branch (BHB) stars are bright A-type stars that can be used as standard candles. BHB stars differ from main-sequence A-stars by having longer lifetimes and being less massive (Yanny et al., 2000). They are ideal tracers for reaching far out in the galactic halo, as they are present in old populations and have bright absolute magnitudes ($M_g \approx 0.5$), allowing them to be identified at large distances (e.g. Thomas et al., 2018). The color space of BHB stars is not contaminated by many other types of stars at large distances, as main sequence stars do not have long enough life times to be present in the galactic halo. There has not been evidence that significant star formation has taken place in the galactic halo for the past billion years, which is longer than average A-star main sequence life time (Yanny et al., 2000). However, BHB stars do have close spectroscopic resemblance to fainter blue stragglers (BS). BS are just like BHB stars present in old populations, which means that they can contaminate a sample of BHB stars. It is of importance to clean a selection of BHB stars from BS stars, as BS stars have intrinsically fainter magnitudes than BHB stars (Starkenburg et al., 2019), which would give a crooked view of the galactic halo density. BHB stars are giant-helium burning stars, whereas BS stars originate from mergers of hydrogen-burning stars on the main sequence (e.g Yanny et al., 2000; Starkenburg et al., 2019). Hence, BS stars differ from BHB stars by having larger gravities, $\log g = 2.8 - 3.75$ for BHB stars and $\log g = 3.75 - 5.0$ for BS stars (e.g. Starkenburg et al., 2019; Thomas et al., 2018; Vickers et al., 2012). This difference can be detected through the Balmer line profiles, which are narrower for BHB stars than for BS (Vickers et al., 2012, and bottom panel of figure 1). Spectroscopic

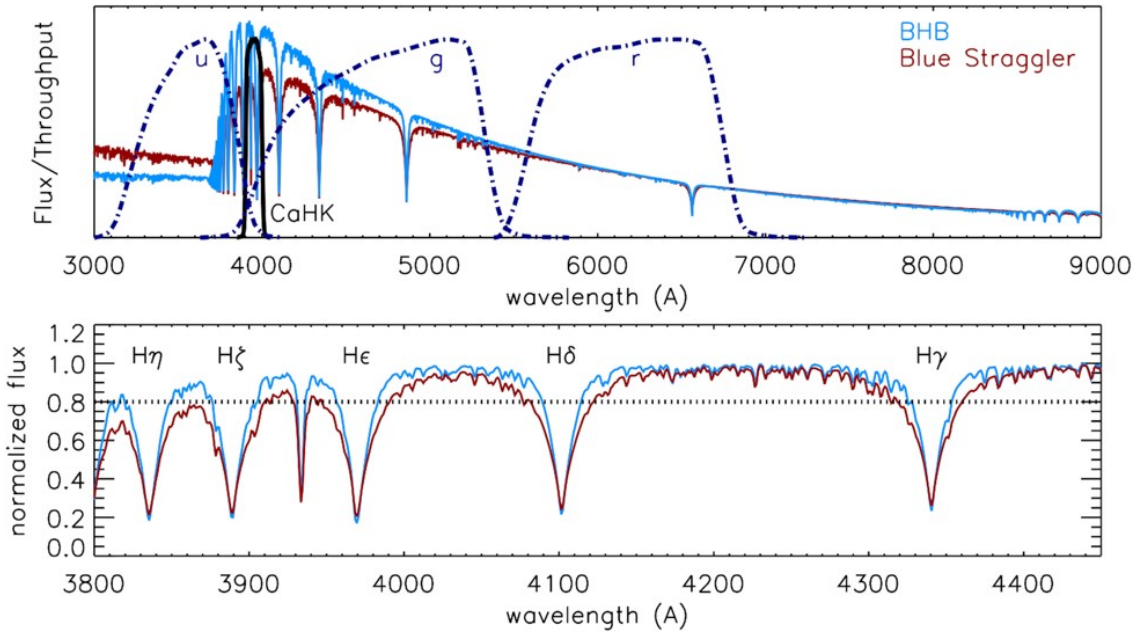

Figure 1: The top frame shows synthetic spectra from the database by Munari et al. (2005). The transmission curves for the SDSS *ugr* filter bands are plot in blue dotted curves and the black solid curve corresponds to the transmission curve of the CaHK filter from the Pristine survey. The bottom frame shows the Balmer absorption lines and the synthetic spectra for BHB and BS stars. Image taken from Starkenburg et al. (2019).

data can be used to distinguish the two stars using this feature, but this is costly in terms of telescope time, which provides difficulties at faint apparent magnitudes. Therefore, methods have been developed to separate BHB and BS stars using broad-band photometric data.

The top frame of figure 1 shows the overall spectra for BHB and BS stars. Next to the different widths of the Balmer line profiles, it is clear that the intensity of the BHB spectrum is higher than that of the BS spectrum above ~ 3800 Å, but that below that wavelength range the BS spectrum dominates. This difference can photometrically be detected when using the right filters.

1.2 Previous work

Fukushima et al. (2018) select BHB stars from the wide field Sloan Digital Sky Survey (SDSS) and Hyper-Suprime-Cam (HSC) survey solely by using various color cuts. First a set of A-stars is selected in the $g-r$, $i-z$ color-color space, after which BS stars are filtered out using the $u-g$, $g-r$ bands. This method creates a sample of BHB stars which reaches a completeness of 67% and purity of 62%.

In the work of Starkenburg et al. (2019), blue horizontal branch (BHB) stars are selected using the Pristine survey CaHK narrow-band photometry combined with *ugr* photometry from the SDSS survey. The work adapts a method by Deason et al. (2011), who apply a method to select BHB stars using only SDSS photometry. Deason et al. (2011) separate BHB stars in the $u-g$, $g-r$ color-color space using the SDSS *ugr* bands from data release 8 (DR8). Starkenburg et al. (2019) use a similar method using the SDSS *ugr* bands, but differ from Deason et al. (2011) by also using the $u-CaHK$, $g-r$ space to select BHB stars, taking the *ugr* bands from SDSS DR14 and the $CaHK$ magnitude from the narrow-band Pristine survey.

The different intensities between the BHB and BS spectrum above ~ 3800Å result in in a different

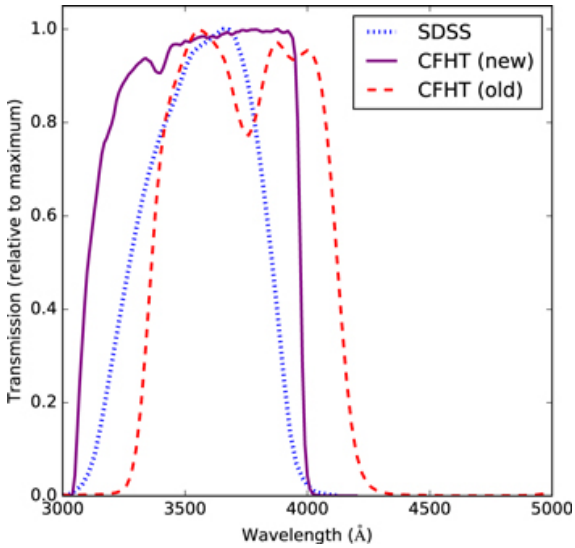

Figure 2: Transmission curves for the SDSS u-band filter and the CFIS u-band filter. The label 'CFHT (new)' corresponds to the u-band filter from CFIS DR3, used in this work (the survey used to have a different name). Image taken from Ibata et al. (2017)

flux in the filter bands. As can be seen in the top frame of figure 1, the SDSS u-band is sensitive towards the wavelengths < 3800Å and the SDSS g-band collects information above 3800Å. By combining these contradicting properties, the difference between BHB stars and BS can be maximised and the two type of stars can be separated from each other. It also shows in figure 1 that the difference between the BHB and BS intensity is at its maximum within the range of the Pristine $CaHK$ filter, making this filter perfect to differentiate between these stars. This is very well established in Starkenburg et al. (2019). The large difference between the two spectra in the Pristine $CaHK$ filter results in a large improvement in the completeness of the BHB star selection. Starkenburg et al. (2019) improve the completeness from 46 % using the SDSS ugr-bands to a spectacular 91 % when including the Pristine $CaHK$ filter. The purity in both methods is 93 %. Still, the analysis is constrained by the sensitivity of the rather shallow SDSS u-band filter. To get a deeper view of the galactic halo, Starkenburg et al. (2019) propose to use the u-band data from the much deeper Canada-France Imaging Survey (CFIS).

The CFIS data is obtained using the Megacam wide-field imager on the 3.6m wide Canada-France Hawaii Telescope (Boulade et al., 2003), allowing CFIS-u to detect ~ 3 magnitudes fainter than SDSS-u (Starkenburg et al., 2019). Because the u-band is the limiting band of the SDSS survey, CFIS-u can be used to fully make use of the other SDSS bands. Furthermore, the deeper u-band data can be used to infer for example metallicity, which allows for tracing the chemical components of the galactic halo (Ibata et al., 2017). The CFIS u-band was already used in Thomas et al. (2018) in combination with the $griz$-bands from the Pan-STARRS survey to make a selection of BHB stars. The CFIS data allow Thomas et al. (2018) to reach out as far as 220 kpc, reaching outward significantly further than previous methods which often did not reach much further than 100 kpc (e.g. Hernitschek et al., 2018; Sesar et al., 2017; Deason et al., 2014).

As can be seen from Figure 2, the CFIS-u band filter is not identical to the SDSS u-band filter in shape but is a bit wider, extending to redder and bluer wavelengths. This affects any color spaces including the u-band, and it will affect the effectiveness of our newly proposeed method as well.

1.3 This work

In this work we will follow the method presented by Starkenburg et al. (2019) for the SDSS ugr-bands and the Pristine $CaHK$-band. We will use more recent data from the Pristine survey, covering a larger area than before. This allows us to observe new structures and obtain a better view on the 'smooth' halo component. The method will further be adjusted to also include the data from the deeper CFIS u-band. We will use the gr-bands from Pan-STARRS DR-2 and the $CaHK$-band from the Pristine survey. We aim to improve the selection of BHB stars with the CFIS-u filter and hope to reach out much further than established in Starkenburg et al. (2019).

1.4 Data

As mentioned, data from various surveys will be used in this work. The various surveys will be shortly introduced in the next four subsections.

1.4.1 SDSS

The SDSS survey (York et al., 2000) is a wide field survey, using a 2.5m telescope (Gunn et al., 2006) located at the Apache Point Observatory in Mexico, observing images in five broad optical bands ($ugriz$) (e.g. Xue et al., 2011; Deason et al., 2011). SDSS data has been utilised widely for studying multiple phenomenon in the galaxy, being used in multiple thousand refereed papers thus far (Aihara et al., 2011). SDSS data thus provides a good starting point for our analysis.

1.4.2 The Pristine survey

The Pristine survey (Starkenburg et al., 2017) employs a narrow-band filter, initially intended to observe the metallicity-sensitive Ca-H and Ca-K absorption features (Starkenburg et al., 2017). The filter band, in combination with u-band data, turns out to be incredibly useful to distinguish BHB stars from BS even though not initially having been designed for it (Starkenburg et al., 2019).

1.4.3 CFIS-u

The CFIS u-band has already been introduced shortly in section 1.2. Imaging data for CFIS-u were obtained using the Canada-France Hawaii Telecope (Ibata et al., 2017). The survey was intended to improve imaging data towards the bluer wavelengths, in which for example SDSS data is often limiting (Ibata et al., 2017).

1.4.4 Pan-STARRS

For the gr-bands, we use imaging data from the Pan-STARRS survey. The Pan-STARRS survey is just like the SDSS survey a wide field survey, and data is obtained using a 1.8m telescope (e.g. Hodapp et al., 2004), located at the Haleakala Observatories on the island of Maui (Chambers et al., 2016). The Pan-STARRS data has some advantages over the SDSS survey. For our purpose it is mainly advantageous that the Pan-STARRS survey is deeper and has more coverage compared to SDSS (Chambers et al., 2016). The combination with the deeper CFIS-u data will provide a detailed view of the galactic halo stretching out as far as possible.

2 Methodology

2.1 Stars selection

In order to start the analysis, a selection of cuts needs to be made in the data files with stars in order to have a set with as many BHB stars as possible but removing already a large fraction of the contamination.

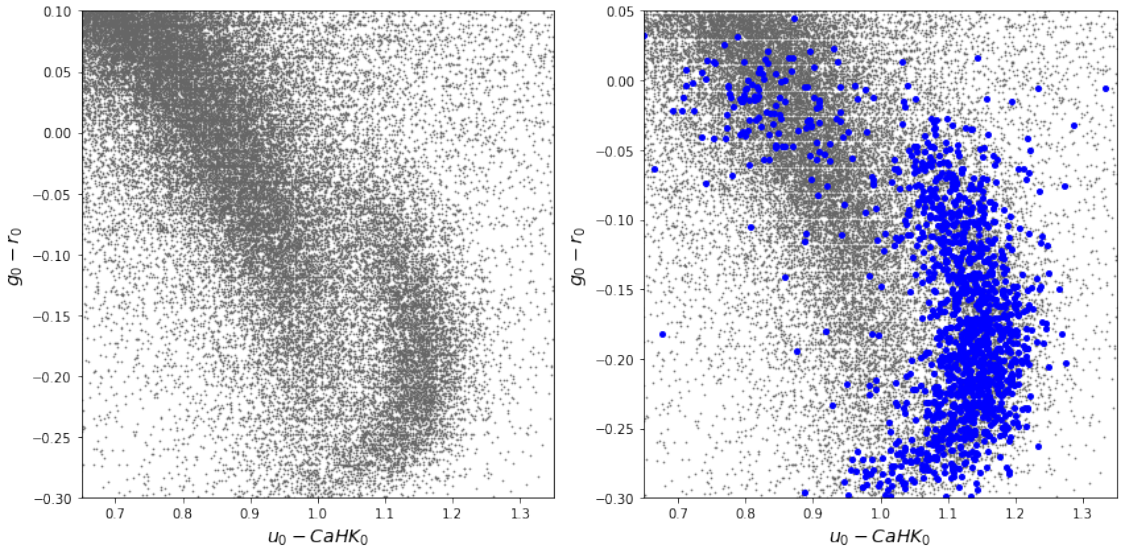

Figure 3: The left frame shows the selection of stars in the SDSS/Pristine color-color space. The right frame shows the confirmed BHB stars from the Xue et al. (2011) database (blue dots) which are as well present in the SDSS/Pristine set.

2.1.1 SDSS/Pristine

We select stars by using magnitudes from the stars present both in the Pristine set and the SDSS DR-14 survey. We have a file with the de-reddened *ugriz* magnitudes from the SDSS survey and the de-reddened $CaHK$ magnitude from the Pristine survey. For the SDSS bands, we use the de-reddened magnitudes as provided by the SDSS collaboration in DR-14. For the CaHK magnitudes from the Pristine survey we use the Schlegel dust map and a specifically calculated extinction coefficient for the $CaHK$ filter (see Starkenburg et al., 2017). A-stars are selected in the set using specific color spaces. We start with the same color selection as in Starkenburg et al. (2019), in which the color bands $0.9 < u_0 - g_0 < 1.4$ and $-0.25 < g_0 - r_0 < 0.0$ are applied. This ensures that contamination by white dwarfs becomes negligible and that the sample is cleaned from contamination by quasars (Starkenburg et al., 2019). Subsequently, only the stars which are flagged in the SDSS set as having clean photometry are kept. We also clean out objects which are not sufficiently point sources thus cannot be classified as stars. The selection of stars is plot in the $u_0 - CaHK_0$, $g_0 - r_0$ color space in the left frame of figure 3. Two distinct curves can be identified, one corresponding to BS stars and one to BHB stars.

A check is made on whether the selection is done correctly by over plotting the BHB stars selected by Xue et al. (2011) that are overlapping with our own sample. These stars have a high probability of being BHB stars and were spectroscopically confirmed using a method that utilises measurements of the Balmer line profiles directly. In the right frame of figure 3 we can clearly see that the right curve corresponds to the BHB sample. The other curve corresponds to BS stars, which we want to clean out from our sample. It can be noted that in the color range $g_0 - r_0 > -0.05$ a substantial fraction of stars from Xue et al. (2011) seem to fall on the BS sequence. From Starkenburg et al. (2019) we know that these stars are deemed likely to be misclassified BHB stars. Due to the lower temperatures in the region $g_0 - r_0 > -0.05$, the Balmer line profiles are more narrow making Xue's classification method more difficult in that region. Therefore we further reduce our sample by discarding stars with color $g_0 - r_0 > -0.05$, in line with what was done in Starkenburg et al. (2019).

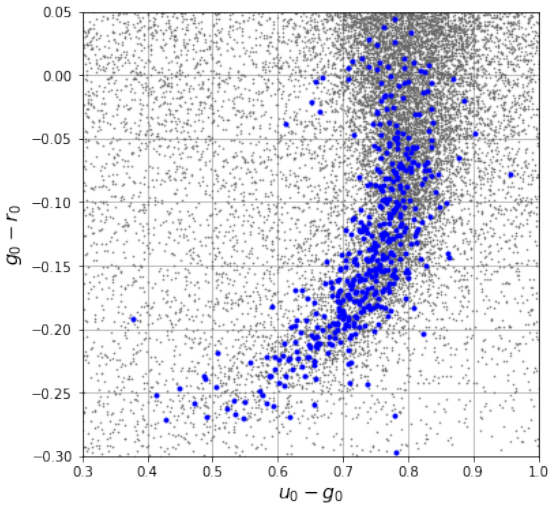

Figure 4: Stars from the CFIS/Pristine/Pan-STARRS set are plot in the $u_0 - g_0$, $g_0 - r_0$ color-color space. The confirmed BHB stars from Xue et al. (2011) are over plot in blue.

2.1.2 CFIS/Pristine/Pan-STARRS

We follow a similar method to select stars present in simultaneously CFIS internal DR-3, Pan-STARRS DR-2, and Pristine. We have a file with fluxes for each star in each band which we first convert to magnitudes. We take the u-band flux from CFIS and the g-band and r-band fluxes from Pan-STARRS. The fluxes are converted to magnitudes using (G. Thomas, priv. comm.)

$$m = -2.5 \log(\Phi_m / 3631.0), \tag{1}$$

$$\sigma_m = -2.5 \log((\Phi_m + \sigma_{\Phi_m})/3631.0) + 2.5 \log(\Phi_m / 3631.0), \tag{2}$$

where Φ_m is the flux in band m. These magnitudes are not yet corrected for extinction. We calculate for extinction using

$$m_0 = m - A_V \cdot E(B - V), \tag{3}$$

where $E(B-V)$ is the color excess for each star, given in our data file and A_V is the extinction correction for each band which is taken from the appendix of Schlafly & Finkbeiner (2011), where we take the values corresponding to a standard reddening of $R_V = 3.1$. For the CFIS u-band, the extinction correction is not given in the appendix from Schlafly & Finkbeiner (2011), so we use the extinction correction for SDSS-u given by $A_V = 4.239$, as also used in Thomas et al. (2018).

Due to the different filters in this set compared to the SDSS/Pristine set, we need to define our own color space to select the A-stars. We plot the stars in the $u_0 - g_0$, $g_0 - r_0$ color-color space and overplot the overlapping confirmed BHB stars from Xue et al. (2011) (see figure 4). We see that most BHB stars fall on an explicit curve. We select the stars in the color cuts $-0.25 < g_0 - r_0 < 0.0$ and $0.50 < u_0 - g_0 < 0.9$ in order to have as little contamination as possible from non-BHB stars while at the same time keeping the completeness as large as possible.

Again, stars which are not flagged as having clean photometry are deselected. Then the selection of stars is also plot in the $u_0 - CaHK_0$, $g_0 - r_0$ color space (left panel of figure 5). Right away we see that the BS and BHB sequences are less distinct as in the SDSS color space. This is due to the different filter shape of the CFIS u-band filter compared to SDSS-u (see also figure 2), resulting in the sequences being closer together.

Similar to before, the overlapping stars from Xue et al. (2011) are shown in the right frame of figure 5 and once more we see that the BHB stars fall onto an own sequence but again above

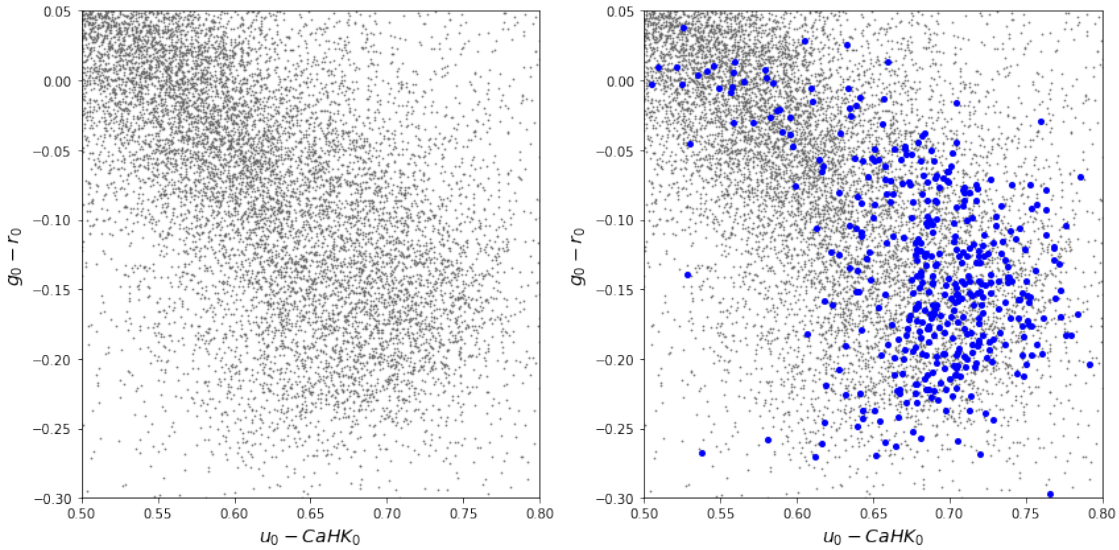

Figure 5: The left frame shows the selection of stars in the CFIS/Pristine/Pan-STARRS color-color space. The right frame shows the confirmed BHB stars from the Xue et al. (2011) database (blue dots) which are as well present in the CFIS/Pristine/Pan-STARRS set.

$g_0 - r_0 > -0.05$ a significant part of the BHB stars confirmed by Xue et al. (2011) seem to fall onto the BS curve. Because of the likelihood of these stars being misclassified, the stars in this range are discarded and only stars with $g_0 - r_0 < -0.05$ are considered.

2.2 Ridgelines & probabilities

In order to determine whether each star in the selection is either a BHB star or a BS, we want to define ridgelines that go through the middle of the BHB sequence and BS sequence and describe these sequences. With these ridgelines we can assign a probability to each star either being a BHB or a BS. The procedure on how this is done is first described for the SDSS/Pristine set, after which the procedure for the CFIS/Pristine/Pan-STARRS set is explained.

2.2.1 Ridgelines SDSS/Pristine

In the work of Starkenburg et al. (2019), ridgelines are fit to synthetic stellar spectra from Munari et al. (2005). Munari et al. (2005) supplies a database of synthetic spectra for stars with parameters covering the whole of the HR diagram. In order to fit the ridgelines, spectra for stars with typical parameters for A-stars are selected from the database. For BHB stars the typical parameters $7500 \leq T_{\rm eff} \leq 9500$, $3.0 \leq \log g \leq 3.5$, and $-2.5 \leq$ [Fe/H] ≤ -1.5 are taken. For BS stars the parameters are $7500 \leq T_{\rm eff} \leq 9500$, $4.0 \leq \log g \leq 4.5$, and $-1.0 \leq$ [Fe/H] ≤ -0.5. These spectra are integrated under the SDSS and Pristine filters and synthetic magnitudes are inferred. The resulting colors are then computed and plot in the SDSS/Pristine color space after which ridgelines can be fit through the spectra. The spectra do not fit directly, as a shift of a total of 0.08 has to be applied. An initial shift of 0.04 is necessary to correct for the offset between the SDSS u-band and its AB magnitude as integrated under the filter curve. An additional shift of 0.04 is necessary because this is the mean shift between SDSS DR-8 and SDSSS DR-14 magnitudes (Starkenburg et al., 2019) (the Munari spectra are integrated under the filter from SDSS DR-8, in this analysis SDSS DR-14 is used). The spectra are shown in the left frame of figure 6. A third order power

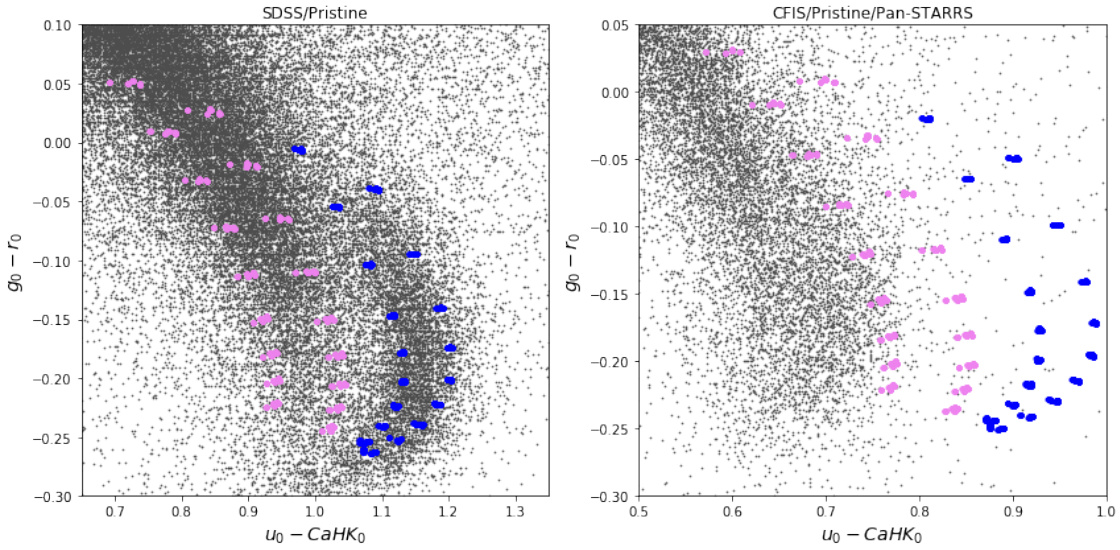

Figure 6: The synthetic stellar spectra as computed by Munari et al. (2005) are plot on top of the sequences in the SDSS/Pristine and CFIS/Pristine/Pan-STARRS color space (BHB = blue, BS = violet). It is clear that the synthetic spectra do not explain the BHB and BS sequences well in the CFIS/Pristine/Pan-STARRS space. Also note that the range on the x-axis for that space is wider than in for example figure 5.

law is fit through the spectra. Starkenburg et al. (2019) finds the following ridgelines:

$$(u_0 - CaHK_0)^0_{\rm BHB} = 0.997 - 1.465\,(g_0 - r_0) + 0.411\,(g_0 - r_0)^2 + 18.531\,(g_0 - r_0)^3\,, \quad (4)$$

$$(u_0 - CaHK_0)^0_{\rm BS} = 0.832 - 1.222\,(g_0 - r_0) - 2.094\,(g_0 - r_0)^2 + 1.046\,(g_0 - r_0)^3\,. \quad (5)$$

If we follow the same procedure as above, the fitted curves are somewhat off to the curves found in Starkenburg et al. (2019). This is probably due to a different fitting mechanism. Because the curves from Starkenburg et al. (2019) seem to fit better towards the lower $g_0 - r_0$ regime (see left frame of figure 9), we proceed with using these ridgelines for the SDSS/Pristine color space.

2.2.2 Probabilities SDSS/Pristine

With these ridgelines a probability can be assigned to each star being either a BHB star or a BS. We here follow the method presented by Deason et al. (2011) for the SDSS $u_0 - g_0$, $g_0 - r_0$ space, which was applied and adjusted for the $u_0 - CaHK_0$, $g_0 - r_0$ space in Starkenburg et al. (2019). The probability for having color $u_0 - CaHK_0$ for a star of each sequence is given by

$$p(ugr\,CaHK \mid \text{BHB}) \propto \exp\left(-\frac{\left[(u_0 - CaHK_0) - (u_0 - CaHK_0)^0_{\rm BHB}\right]^2}{2\sigma^2_{\rm BHB}}\right), \quad (6)$$

$$p(ugr\,CaHK \mid \text{BS}) \propto \exp\left(-\frac{\left[(u_0 - CaHK_0) - (u_0 - CaHK_0)^0_{\rm BS}\right]^2}{2\sigma^2_{\rm BS}}\right). \quad (7)$$

Here $(u_0 - CaHK_0)^0$ is the center of the sequences, which is described by the ridgelines in equations 4 and 5. We note that this probability thus is also dependent on the $g_0 - r_0$ color. The standard deviation depends on the intrinsic width of the distribution and the photometric error on the color

and is given by

$$\sigma_{\mathrm{BHB}} = \sqrt{(\sigma^{0}_{\mathrm{BHB}})^2 + (\sigma_{u_0 - CaHK_0})^2}, \tag{8}$$

$$\sigma_{\mathrm{BS}} = \sqrt{(\sigma^{0}_{\mathrm{BS}})^2 + (\sigma_{u_0 - CaHK_0})^2}. \tag{9}$$

Here $\sigma^{0}_{\mathrm{BHB}}$ and σ^{0}_{BS} is the error on the intrinsic width of the sequences. For the SDSS/Pristine color-color space, these values are taken from Starkenburg et al. (2019) and are given by $\sigma^{0}_{\mathrm{BHB}} = 0.04$, $\sigma^{0}_{\mathrm{BS}} = 0.045$. We apply another cut to our sample and remove any object that lies beyond 3σ from the two ridgelines, where we use σ_{BHB} and σ_{BS} from equations 8 and 9. This removes outliers from the sample.

The posterior probabilities of being in one of either classes is described by

$$P(\mathrm{BHB} \mid ugr\,CaHK) = \frac{p(ugr\,CaHK \mid \mathrm{BHB}) N_{\mathrm{BHB}}}{p(ugr \mid \mathrm{BHB}) N_{\mathrm{BHB}} + p(ugr\,CaHK \mid \mathrm{BS}) N_{\mathrm{BS}}}, \tag{10}$$

$$P(\mathrm{BS} \mid ugr\,CaHK) = \frac{p(ugr\,CaHK \mid \mathrm{BS}) N_{\mathrm{BS}}}{p(ugr\,CaHK \mid \mathrm{BHB}) N_{\mathrm{BHB}} + p(ugr\,CaHK \mid \mathrm{BS}) N_{\mathrm{BS}}}. \tag{11}$$

Here N_{BHB} and N_{BS} describe the number of BHB stars and BS stars in a color band. Deason et al. (2011) finds these values by iteratively integrating equations 10 and 11 for the $u-g$, $g-r$ SDSS color space. In Starkenburg et al. (2019) it is argued that the sequences in the $u-CaHK$, $g-r$ color space are well enough seperated that no prior is needed on the ratios of BHB and BS stars as a function of color, as this allows to be more agnostic about the fraction of BHB and BS stars as a function of color and distance. The values of N_{BHB} and N_{BS} are both taken to be equal to 1. Since we are following the method by Starkenburg et al. (2019), we will stick to these values for the calculation of our probabilities. Nevertheless, the values for N_{BHB} and N_{BS} are computed for completeness (see table 1).

2.2.3 Ridgelines CFIS/Pristine/Pan-STARRS

In the CFIS/Pristine/Pan-STARRS color space, fitting the ridgelines took a more extensive effort than for the SDSS/Pristine color space. The attempts made to find the ridgelines and the corresponding failures are described in the next three subsections.

2.2.3.1 Using Munari et al. (2005) synthetic spectra

A first attempt was made to integrate the synthetic spectra from Munari et al. (2005) under the various filter curves and fit ridgelines to these spectra, as done for SDSS/Pristine in Starkenburg et al. (2019). However, when computed it was quickly clear that the spectra were far off in comparison to the actual stars (see right frame of figure 6). The thought arose that a zero-point shift was needed as for SDSS/Pristine, but no evidence to substantiate this claim could be found. The suggestion was made that the theoretical model would be inaccurate in the near-UV, as the models would not have sufficient information in that wavelength (G. Thomas, priv. comm.). This could be (part of) the reason, as shown in Figure 2, the CFIS u-band filter has a higher throughput to bluer wavelengths when compared to the SDSS u-band. An attempt is made to manually shift the ridgelines and fit them over the sequences in the CFIS/Pristine/Pan-STARRS color space. In order to get a remotely reasonable looking fit, both curves fitted to the Munari spectra need a multiplication of a factor of 0.75, and thereafter the BS curve needs another shift of 0.04. This shift works but is purely empirical and not rooted in any theoretical expectations.

2.2.3.2 Using Xue et al. (2008) BHB & BS sample

A second attempt is made to fit ridgelines to a sample of BHB and BS stars which have already been spectroscopically confirmed through another method. Xue et al. (2008) supplies a database

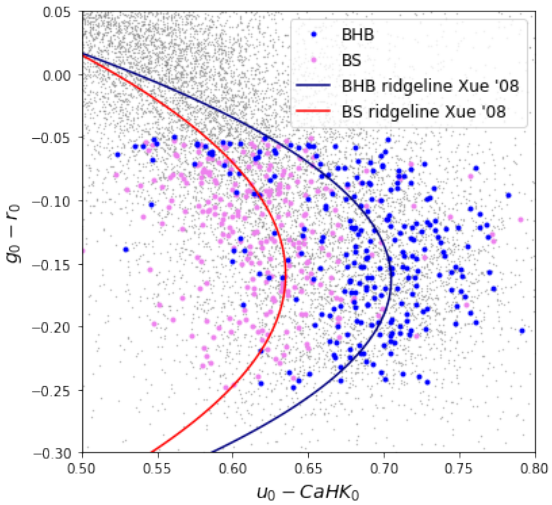

Figure 7: Overlapping stars in the Xue et al. (2008) database and the CFIS/Pristine/Pan-STARRS set are plot over all the stars in the CFIS/Pristine/Pan-STARRS data set (grey points). The Xue et al. (2008) stars are only selected in the color range in which the further analysis is done. Second order polynomials are fit through the BHB and the BS sequences of the overlapping stars (dark blue and red curves, respectively).

of both BHB or BS stars which have been spectroscopically confirmed using a method similar to Xue et al. (2011). We find the overlapping coordinates for the stars in the CFIS/Pristine/Pan-STARRS set and plot the confirmed stars in the $u_0 - CaHK_0$, $g_0 - r_0$ space (figure 7). We then fit a second order polynomial to the BHB sample and the BS sample. in figure 7 it can be seen that the curves do not seem to represent the actual sequences very well, with in particular the BHB sequence. In the higher $g_0 - r_0$ region, there seem to be multiple misclassified BHB stars which deflect the BHB ridgeline in that region. Furthermore there is only a relatively small sample of stars in the CFIS/Pristine/Pan-STARRS space to which the ridgelines are fit.

2.2.3.3 Using probabilities SDSS/Pristine

A final attempt made to fit the ridgelines is by using the confirmed BHB and BS stars from the SDSS/Pristine filter. We take the stars which are present both in the SDSS/Pristine set and the CFIS/Pristine/Pan-STARRS set, and select the stars which have posterior probability $P(\text{BHB}) > 0.8$ and $P(\text{BS}) > 0.8$ in the SDSS/Pristine set. These are plot in the CFIS/Pristine/Pan-STARRS space and third order polynomials are fit through the sequences (see figure 8). These ridgelines already provide a much better fit to the sequences than the other two proposed methods.

In figure 9 the ridelines following the three different methods are plot next to each other in the CFIS/Pristine/Pan-STARRS space but also in the SDSS/Pristine space for comparison to the original curve by Starkenburg et al. (2019). Within the range $-0.25 < g_0 - r_0 < -0.05$ the curves computed using the overlapping stars provides the best resemblance to the original curve from Starkenburg et al. (2019) in the SDSS/Pristine color-color space (we do show the Munari curves in both frames for completeness, but since this method is not rooted in any theoretical expectations in the CFIS/Pristine/Pan-STARRS color-color space, these curves are not considered for the further analysis). Therefore the curves following the overlapping stars method will be used in the remainder of this paper for the CFIS/Pristine/Pan-STARRS set. These ridgelines are given by

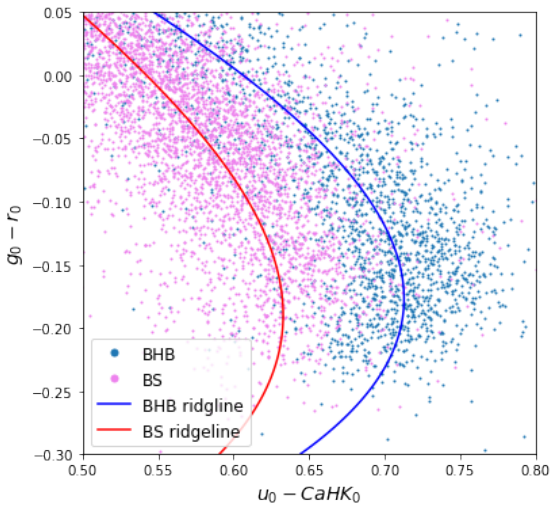

Figure 8: Stars present in both the SDSS/Pristine set and CFIS/Pristine/Pan-STARRS are plot in the CFIS/Pristine/Pan-STARRS color-color space. The stars which have a high probability of being BHB or BS according to the SDSS/Pristine color-color space ($P(BHB) > 0.8$ and $P(BS) > 0,8$) are plot in blue and pink respectively. Third order polynomials are fit trough the sequences and are described by the blue and red curves.

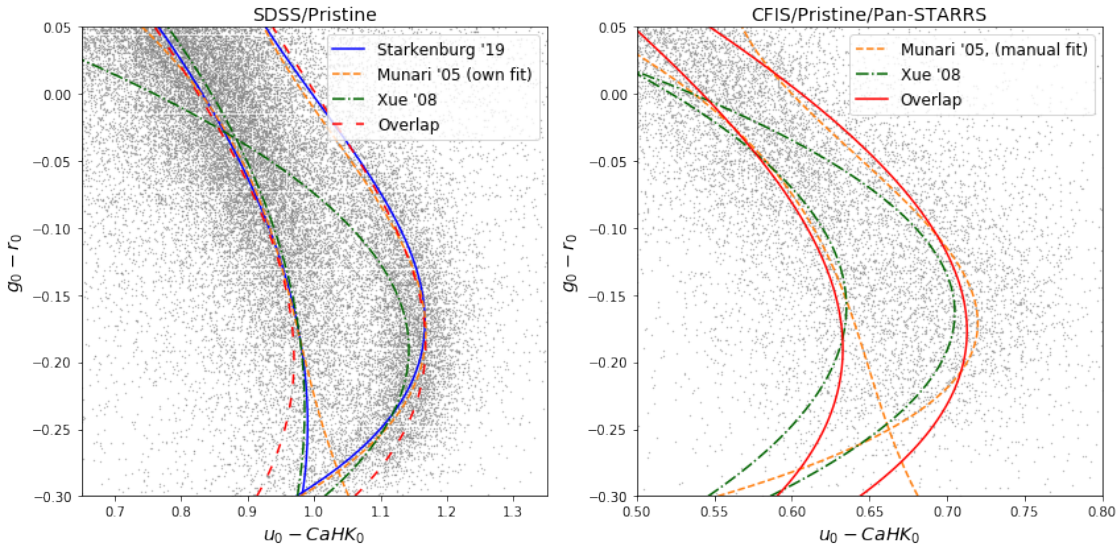

Figure 9: Ridgelines computed with different methods plotted in one figure. The left figure shows the stars in the SDSS/Pristine color space and the right figure shows the stars in the CFIS/Pristine/Pan-STARRS color space. The solid lines represent the curves which are used in the further analysis.

Table 1: The fraction of BHB stars and BS in the SDSS/Pristine set and the CFIS/Pristine/Pan-STARRS set in different color bins. The values are used in equations 10 and 11. The columns represent the color bin, the total number of stars in that bin, the fraction of BHB stars and the fraction of BS.

	SDSS/Pristine			CFIS/Pristine/Pan-STARRS		
	$N_{\rm tot}$	$f_{\rm BHB}$	$f_{\rm BS}$	$N_{\rm tot}$	$f_{\rm BHB}$	$f_{\rm BS}$
$-0.5 < g_0 - r_0 < -1.0$	5875	0.270	0.730	2261	0.364	0.636
$-1.0 < g_0 - r_0 < -1.5$	4707	0.391	0.609	1840	0.468	0.532
$-1.5 < g_0 - r_0 < -2.0$	3948	0.543	0.457	1543	0.527	0.473
$-2.0 < g_0 - r_0 < -2.5$	3086	0.658	0.342	901	0.490	0.510

$$(u_0 - CaHK_0)^0_{\rm BHB} = 0.606 - 1.091\,(g_0 - r_0) - 2.091\,(g_0 - r_0)^2 + 3.752\,(g_0 - r_0)^3\,, \quad (12)$$

$$(u_0 - CaHK_0)^0_{\rm BS} = 0.542 - 0.837\,(g_0 - r_0) - 1.268\,(g_0 - r_0)^2 + 3.283\,(g_0 - r_0)^3\,. \quad (13)$$

2.2.4 Probabilities CFIS/Pristine/Pan-STARRS

We use the above ridgelines to assign a probability to each star being either a BHB star or BS in a similar manner as in section 2.2.2. The probability of a star of each sequence having color $u_0 - CaHK_0$ are described by equations 6 and 7, but now $(u_0 - CaHK_0)^0$ is described by the ridgelines from equations 12 and 13. The posterior probabilities are described by equations 10 and 11.

Since the sequences in the CFIS/Pristine/Pan-STARRS color space have different widths than in the SDSS/Pristine color space, we have to adjust our values for $\sigma^0_{\rm BHB}$ and $\sigma^0_{\rm BS}$ in equations 12 and 13. When comparing the right frames of figures 3 and 5, it can be noted that the width of the sequence of the Xue et al. (2011) stars in the bottom panel is less than in the top panel (note the different scales on the x-axis). We approximate that this difference is about 0.5, which is why we take the values for $\sigma^0_{\rm BHB}$ and $\sigma^0_{\rm BS}$ to be half of that in the SDSS/Pristine filter. We get

$$\sigma^0_{\rm BHB} = 0.02,$$
$$\sigma^0_{\rm BS} = 0.025.$$

Then similar to the SDSS/Pristine set, a cut is made in the sample to remove any object which is beyond 3σ from the two ridgelines.

The sequences for the BS and BHB stars are not as well separated in the CFIS/Pristine/Pan-STARRS color space as in the SDSS/Pristine space, so now we do want to interatively integrate equations 10 and 11 in order to compute values for $N_{\rm BHB}$ and $N_{\rm BS}$. These values are described in table 1 for various color bands.

2.3 Obtaining distances

Because BHB stars are standard candles, we can use their magnitude to obtain their distances and we can compute a profile of the galactic halo. For the SDSS filter curve, we follow the equation by Deason et al. (2011), who give the following relation for the absolute magnitude for BHB stars in the SDSS color space:

$$M_{g_{\rm SDSS},{\rm BHB}} = 0.434 - 0.169\,(g_0 - r_0) + 2.319\,(g_0 - r_0)^2 + 20.449\,(g_0 - r_0)^3 + 94.517\,(g_0 - r_0)^4\,. \quad (14)$$

With the absolute magnitude and the observed SDSS g_0-band magnitude we can compute the distance using the distance modulus (e.g. Kutner, 2003)

$$d = 10^{(0.2(g_{0_{\rm SDSS}} - M_{g_{\rm SDSS}} + 5)}. \quad (15)$$

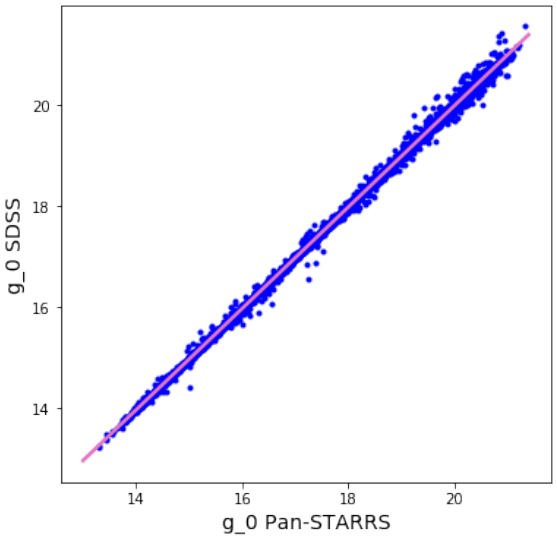

Figure 10: $g_{0_{\rm PS}}$ on the horizontal axis versus $g_{0_{\rm SDSS}}$ on the vertical axis for the overlapping stars in the SDSS/Pristine set and the CFIS/Pristine/Pan-STARRS set. The pink line represents a linear fit through the data.

Deason et al. (2011) only supply a relation for the absolute magnitude in the SDSS color space. Therefore, we need a relation to convert the $g_0 - r_0$ Pan-STARRS color to the $g_0 - r_0$ SDSS color in the CFIS/Pristine/Pan-STARRS set. Thomas et al. (2018) provide this conversion, given by

$$(g_0 - r_0)_{\rm SDSS} = 1.18\,(g_0 - r_0)_{\rm PS} + 0.02. \tag{16}$$

We can now compute the absolute g-band magnitude for the SDSS filter, but we can only obtain the distance if we have a means to convert the Pan-STARRS g-band magnitude to the SDSS g-band magnitude. To obtain this relation specifically for A-stars, the g_0(Pan-STARRS) magnitudes are plot versus the g_0(SDSS) magnitudes in figure 10. A linear fit can be made to convert the magnitudes. The conversion is given by

$$g_{0_{\rm SDSS}} = 1.006\,g_{0_{\rm PS}} - 0.130 \tag{17}$$

With equations 16 and 17 we can convert the Pan-STARRS magnitudes to SDSS values and use equation 15 to compute the distances for the BHB stars in the CFIS/Pristine/Pan-STARRS set as well.

3 Results

3.1 Probabilities

With equations 10 and 11 each star is assigned a probability of being a BHB star. We plot the probabilities in a histogram for each filter set, once using the computed values for $N_{\rm BHB}$ and $N_{\rm BS}$, and once while using $N_{\rm BHB} = N_{\rm BS} = 1$. In figure 11 we see that for both the SDSS/Pristine set as the CFIS/Pristine/Pan-STARRS set, the histogram shows peaks at $P(\rm BHB) \to 1$ and $P(\rm BHB) \to 0$. This shows that there is a clear distinction that stars are assigned either a probability of being BHB or a probability of being BS. Furthermore, we see that the fraction of stars with high BHB probabilities drops slightly when the N-values computed in table 1 are used.

For the SDSS/Pristine set, we obtain 5928 stars with $P(\rm BHB) > 0.8$ when using $N_{\rm BHB} = N_{\rm BS} = 1$ and 5829 stars with $P(\rm BHB) > 0.8$ when using $N_{\rm BHB}$ and $N_{\rm BS}$ from table 1. There is a total of

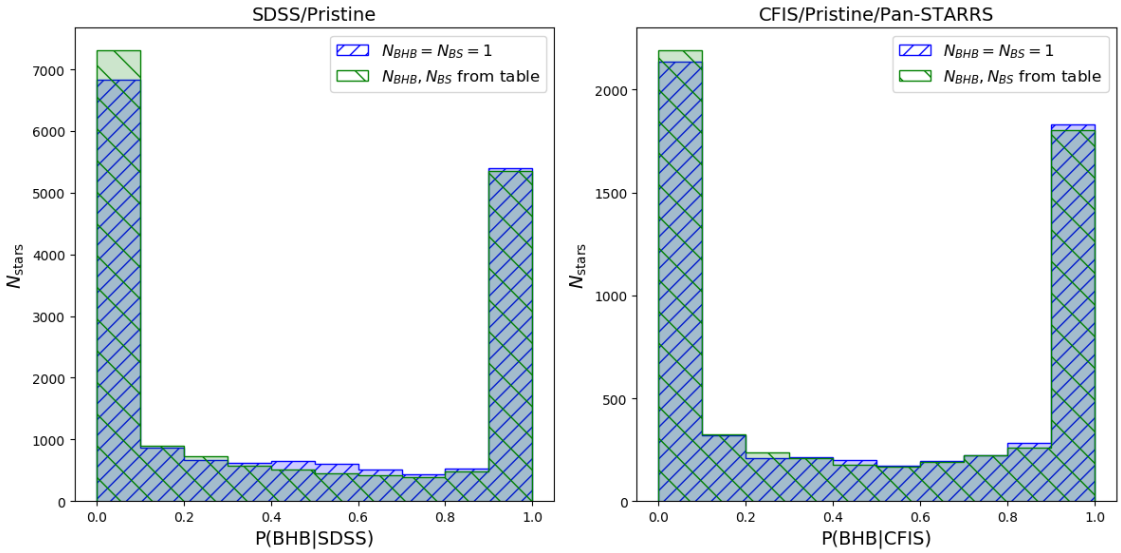

Figure 11: A histogram of the probabilities of each star being BHB. The historgrams are plot for the stars in the SDSS/Pristine set (left frame) and for the stars in the CFIS/Pristine/Pan-STARRS set (right frame). For both sets, the values computed with $N_{\rm BHB} = N_{\rm BS} = 1$ (blue-striped columns) and with the values for $N_{\rm BHB}$ and $N_{\rm BHB}$ as computed in table 1 (green striped columns) are shown.

5686 stars which have $P(\rm BHB) > 0.8$ in both methods. This means that for the SDSS/Pristine set, only 1.7% less stars are identified as BHB when using $N_{\rm BHB}$ and $N_{\rm BS}$ from table 1, and only 4.1% of stars which were initially identified as BHB are not identified as BHB anymore.

For the CFIS/Pristine/Pan-STARRS set, 2116 stars that have $P(\rm BHB) > 0.8$ are identified when using $N_{\rm BHB} = N_{\rm BS} = 1$, and 2060 stars with $P(\rm BHB) > 0.8$ are identified when using $N_{\rm BHB}$ and $N_{\rm BS}$ from table 1. There are 2049 stars which have probabilities $P(\rm BHB) > 0.8$ in both methods. In the CFIS/Pristine/Pan-STARRS case, only 2.6% less stars are identified when using $N_{\rm BHB}$ and $N_{\rm BS}$ from table 1 opposed to taking $N_{\rm BHB} = N_{\rm BS} = 1$, and only 3.1% of stars previously assigned $P(\rm BHB) > 0.8$ are assigned a probability $P(\rm BHB) < 0.8$.

We see that the effect of taking the different N-values is rather small. As already argued by Starkenburg et al. (2019), not assigning a prior allows us to be more agnostic about the probabilities. In this way the selection and size of the sample will not affect the probability of a star coming out as BHB star or not, and the result will be consistent with rerunning the experiment with more data or when using different color cuts or selection criteria. Therefore we will use the values $N_{\rm BHB} = N_{\rm BS} = 1$ for the probabilities in both sets in the remainder of this paper.

3.2 Completeness

In order to know how good our set of BHB stars represents the actual population of BHB stars in the halo, we want to compare the selection of BHB stars with a set of known BHB stars. We again consider the spectroscopically confirmed BHB stars by Xue et al. (2011) in the range $-0.25 < g_0 - r_0 < -0.05$. We check how many of these BHB stars are assigned probability $P(\rm BHB) > 0.8$, using both methods. We find that for the SDSS/Pristine method, the completeness is 86.5% and for the CFIS/Pristine/Pan-STARRS method a completeness of 73.7% is reached.

Starkenburg et al. (2019) reach a completeness of 91% which means that our implemented method does a slightly worse job. This can be due to the larger coverage of the sample, or due to the fact that some slightly different color cuts were made than in Starkenburg et al. (2019) (see section

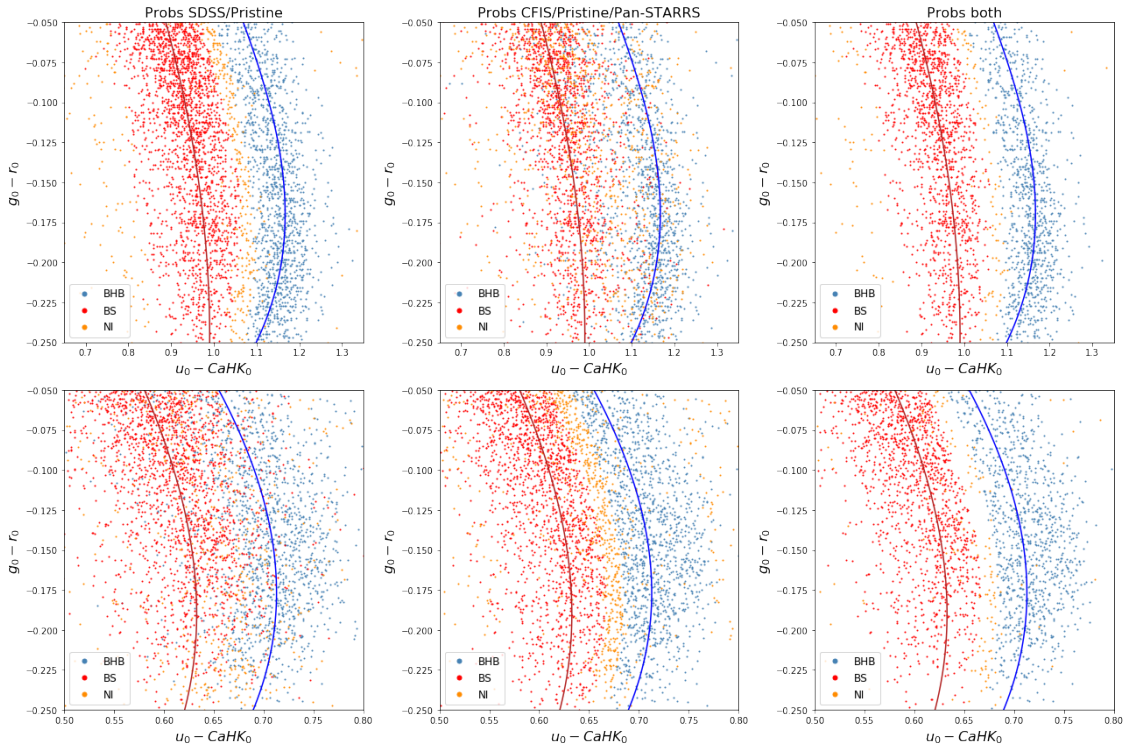

Figure 12: The overlapping stars in the SDSS/Pristine and CFIS/Pristine/Pan-STARRS plot in both color-color spaces (SDSS/Pristine top frames, CFIS/Pristine/Pan-STARRS bottom frames). The stars are marked for having probability $P(\text{BHB}) > 0.8$ (blue), $P(\text{BS}) > 0.8$ (red), and $0.2 < P(\text{BHB}), P(\text{BS}) < 0.8$ (yellow (not identified, NI)). In the left frames, the probabilities taken are computed using the SDSS/Pristine color-color space, the middle frames show the probabilities assigned with the CFIS/Pristine/Pan-STARRS color-color space, and the right frame shows the stars which have $P(\text{BHB}) > 0.8$ and $P(\text{BS}) > 0.8$ in both sets. The ridgelines which represent the center of each sequence are also plot in blue for BHB and dark red for BS.

4.2.2).

3.3 Overlap in data sets

We computed $P(\text{BHB})$ in both the SDSS/Pristine set and the CFIS/Pristine/Pan-STARRS set seperately, but it is as well interesting to see how well the probabilities from both sets relate to each other. BHB stars and BS stars are selected by taking stars with probabilities $P(\text{BHB}) > 0.8$ and $P(\text{BS}) > 0.8$. In figure 12 we plot the overlapping stars from both sets in the two color-color spaces. In the top left and middle bottom frame, we see that the selected BHB and BS stars nicely follow the ridgelines and a fairly abrupt border marks the outsides of the sequences. In the top middle frame and bottom left frame, the stars are plot with the BHB and BS probabilities computed using the opposite filters. Reassuringly we find that the BHB stars and BS stars in both color-color spaces aggregate around the corresponding ridgelines. The sequences in these frames are more diffuse towards the edges, but it is clear that the general selection of the stars went well. As a final check, the stars are plot which have BHB and BS probabilities in both methods. These plots show a very similar behaviour as the other plots, confirming that the methods did a similar job. From the set of overlapping stars, there is a total of 1263 stars identified as BHB when using the SDSS/Pristine method, and using the CFIS/Pristine/Pan-STARRS method we identify 1201. From these stars there is a total of 948 that have probabilities $P(\text{BHB}) > 0.8$ in both sets.

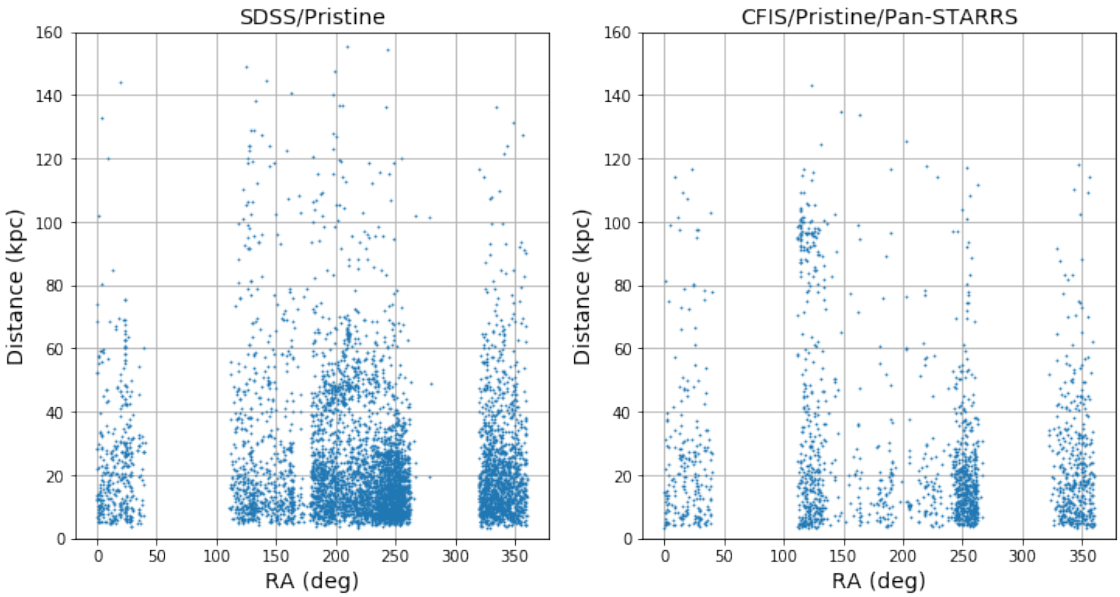

Figure 13: Right ascension versus distance for the stars in the SDSS/Pristine set and the CFIS/Pristine/Pan-STARRS set.

3.4 Distance profile of the halo

Using the property that BHB stars are standard candles, distances to each BHB star are computed using the procedure as described in section 2.3. In figure 13 the right ascension of each star with $P(\text{BHB}) > 0.8$ is plot versus the distance. In figure 14 the coverage of the two sets is shown, together with the known stellar streams and substructures in the halo. Combining figure 13 and 14 some interesting objects can be identified.

3.4.1 M33

The first time computing figure 13, a line of stars is identified in the CFIS/Pristine/Pan-STARRS frame at RA=23.5, stretching the whole distance range from 40 kpc to 140 kpc. Plotting in equatorial coordinates we find that these stars are centered at (RA, Dec) = (23.5, 31) degrees. We find that at this position the galaxy Messier 33 can be found (e.g. Corbelli et al., 1989). This galaxy is not part of the galactic halo, but located at a distance of $\sim$ 800 kpc (e.g. Corbelli et al., 1989). We consider it unlikely that the observed overdensity of stars at this position is a structure of BHB stars in the halo, but we expect it to rather be a consequence of bad photometry due to the underlying galaxy. Therefore the stars in the range $23 < \alpha < 24$, $30 < \delta < 32$ are removed from the CFIS/Pristine/Pan-STARRS sample.

3.4.2 Substructures

In the left frame of figure 13 we see a clump of stars which can be identified as the leading arm of the Sagittarius stream at RA= $\pm$200 degrees, at around 50 kpc. We can also spot the Bootes dwarf galaxy (Belokurov et al., 2006) in the SDSS/Pristine frame at RA $\approx$ 210 degrees at around 65 kpc, located just a bit farther than the Sagittarius leading arm.

The coverage of the CFIS/Pristine/Pan-STARRS set is not as good in that area, which is why the Sagittarius stream and Bootes dwarf galaxy are not as well identified in that set. Interestingly, at RA $\approx$ 120 degrees at around 90 kpc, we do identify the other trailing arm of the Sagittarius stream in the CFIS/Pristine/Pan-STARRS set, although it is not as well noticeable in the SDSS/Pristine

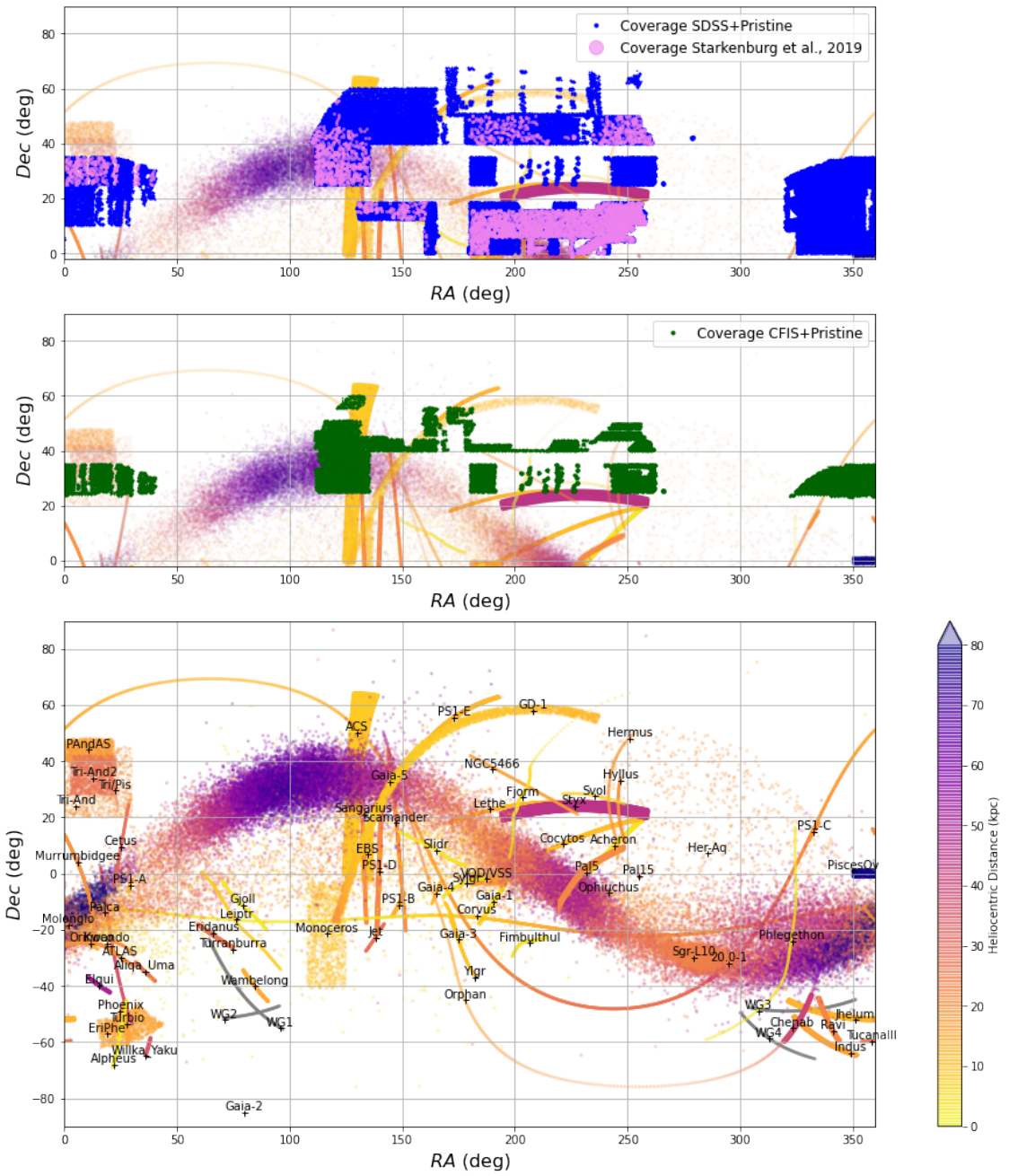

Figure 14: In the top two frames the coverage of the SDSS/Pristine set is shown in blue. The coverage of Starkenburg et al. (2019) is over plot in pink. In the middle frame we see the footprint of the CFIS/Pristine/Pan-STARRS set. Known stellar streams and substructures in the Milky Way are vaguely plot int eh background. The bottom frame shows the full set of known stellar streams and substructures, as given by the `galstreams` package by Mateu et al. (2018).

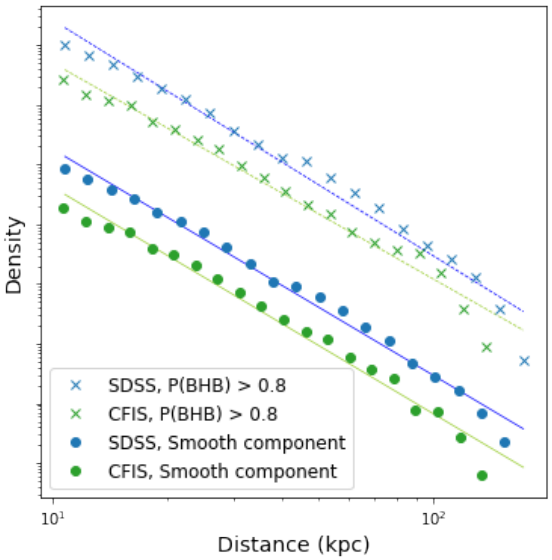

Figure 15: Density versus distance in the galactic halo. Only stars with distance > 10 kpc have been considered. The top curves show the density computed with all stars having P(BHB) > 0.8 in both sets. The bottom densities represent the smooth part of the halo and areas with overdensities are filtered out. The areas filtered out for the bottom sets have values 180 < RA < 220 for the SDSS/Pristine set (Sagittarius stream) and 110 < RA < 140 (Sagittarius stream and NGC 2419) and 205 < RA < 206, 28 < Dec < 29 (Globular cluster M3) for the CFIS/Pristine/Pan-STARRS set.

set. Another interesting feature at RA ≈ 115 degrees is a line of stars stretching from about 80 kpc to 105 kpc. We use the catalog for Milky Way globular clusters by Harris (1996, (2010 edition)) to identify the structure. We identify the line of stars as globular cluster NGC 2419, located at 82.6 kpc at (RA, Dec) = (114.6, 38.5). Although the stars form a line from about 80 kpc to 105 kpc and are not clustered at the distance of the globular cluster, we can explain this through photometry being less accurate for stars in crowded regions such as dense globular clusters. In the CFIS/Pristine/Pan-STARRS frame, a structure can also be identified at RA ≈ 205 degrees. We identify this structure as globular cluster M3 (or NGC 5272), located at (RA, Dec) = (205, 28.2) at 10.2 kpc (Harris, 1996, 2010 edition).

3.5 Density profile

The selected BHB stars can be used to derive a density profile of the galactic halo. The density profile can be modeled by a negative power law of the form (Deason et al., 2011)

$$\rho(r) \propto r^{-\alpha}. \tag{18}$$

In figure 15 the distance is plot versus the density and the power law is fit through. If we fit a power law through the densities derived from all the confirmed BHB stars in the set, a negative power law with a slope of $\alpha = 3.95 \pm 0.60$ is established for the SDSS/Pristine set, and a slope of $\alpha = 3.62 \pm 0.55$ is found for the CFIS/Pristine/Pan-STARRS set. The uncertainties given are the standard deviations of the fit. The values obtained include the substructures in the halo, which can skew the fit. As seen in figure 15, there are bumps in the density profiles due to the presence of the Sagittarius stream in our set, clearly visible at ~ 50 kpc for SDSS/Pristine amd at ~ 100 kpc for the CFIS/Pristine/Pan-STARRS set (also figures 13 and 14). Therefore the stars in the coverage of the sets in which overlap exists with the Sagittarius stream are filtered out and a new fit is made. The resulting negative power law has a slope of $\alpha = 3.79 \pm 0.36$ for the SDSS/Pristine set and $\alpha = 3.80 \pm 0.49$ for the CFIS/Pristine/Pan-STARRS set, a remarkably consistent result!

3.6 Sagittarius stream

In figure 16 we reproduce figure 11 from Starkenburg et al. (2019) in order to observe how well the features of the Sagittarius stream are present in our sample. Of specific interest is the distant 'spur' feature, first found by Sesar et al. (2017) who used a sample of RR Lyrae to trace the halo. The area in which the feature exists is marked by a red rectangle in figure 16. We select the stars within the plane of the Sagittarius stream, taking $-30° < B_{\mathrm{Sgr}} < 30°$, where B_{Sgr} corresponds to latitude in Sagittarius Stream coordinates as defined by Belokurov et al. (2014) (see appendix A.1 for a definition). For both sets the stars with $P(\mathrm{BHB}) > 0.8$ as well as stars with $P(\mathrm{BHB}) > 0.5$ are shown.

Clearly the SDSS/Pristine set provides a much better footprint in the Sagittarius plane than the CFIS/Pristine/Pan-STARRS set. We see that in the SDSS/Pristine set, the Sagittarius stream arms as defined by Hernitschek et al. (2017) are well traced by the sample. Also the distant 'spur' feature can be identified. The CFIS/Pristine/Pan-STARRS sample clearly has much less coverage within the Sagittarius stream plane. Nevertheless, we are very lucky to cover exactly the area where the distant 'spur' feature exists. The feature can very clearly be identified in this set, although the sample does not reach as deep as in the SDSS/Pristine set.

As already introduced by Starkenburg et al. (2019), there have been two recent modelling attempts by Dierickx & Loeb (2017) and Fardal et al. (2019) that managed to reproduce the observed distances as well as the distant 'spur' feature in the Sagittarius stream. The two works differently interpret the feature, where Fardal et al. (2019) expect the spur to extend to 140 kpc and Dierickx & Loeb (2017) expect the spur to extend to 250 kpc (Starkenburg et al., 2019). In the CFIS/Pristine/Pan-STARRS set, the sample reaches out to a maximum distance of ~ 150 kpc, which does not allow us to say anything about which model is preferred. On the contrary, in the SDSS/Pristine set the spur clearly seems to extend for over 200 kpc. Caution should be taken as most stars reaching out that far have probabilities $0.5 < P(\mathrm{BHB}) < 0.8$, thus the deeper view goes at the expense of most likely a higher contamination. Nevertheless, the spur seems to be clearly distinct towards 200 kpc, and we are tempted to favor the model by Dierickx & Loeb (2017) over that by Fardal et al. (2019).

4 Discussion

4.1 Evaluating new method

With the above retrieved results an evaluation can be made of the new procedure utilising the CFIS-u filter compared to the method using SDSS-u. The aim has been to trace out the halo as far as possible. With the CFIS-u band, fainter A-stars should be detectable compared to SDSS-u, which should give a better view of the halo towards the outskirts.

4.1.1 Deeper into the halo?

When comparing the SDSS/Pristine method to the CFIS/Pristine/Pan-STARRS method, we note that when selecting stars with $P(\mathrm{BHB}) > 0.8$, the new method does not trace out the halo significantly further than the old method. In figure 13 this becomes clear, and one could even argue that for some values of right ascension, the SDSS/Pristine method finds more stars at large distances than the CFIS/Pristine/Pan-STARRS method.

From figures 3 and 5 we know that the BHB and BS sequences are much more distinct in the SDSS/Pristine color-color space than in the CFIS/Pristine/Pan-STARRS color-color space. Certainly, the sample size of the CFIS/Pristine/Pan-STARRS set is much smaller than that of the SDSS/Pristine set. With more data, the curves in the CFIS/Pristine/Pan-STARRS color-color space might become more distinct. Nevertheless, solely looking at the spectroscopically confirmed

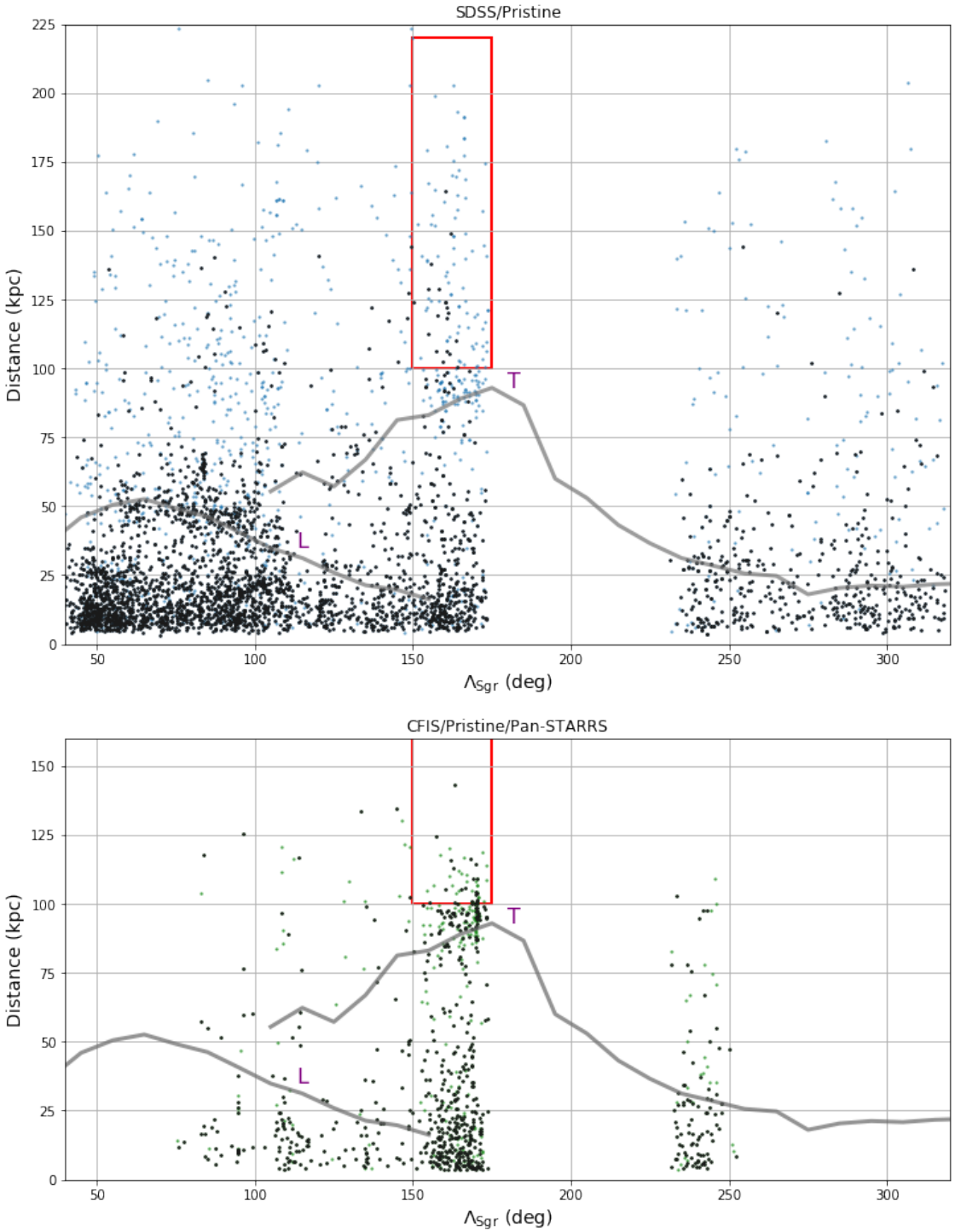

Figure 16: For both SDSS/Pristine and CFIS/Pristine/Pan-STARRS, longitude along the Sagittarius plane is plot versus heliocentric distance for a selection of stars. The longitude along the Sagittarius plane is computed as defined by Belokurov et al. (2014) (for a definition see appendix A.1). The stars with $P(\mathrm{BHB}) > 0.8$ are plot in black, and the stars with $P(\mathrm{BHB}) > 0.5$ are plot in color. The grey curves illustrate the Sagittarius stream components as defined by Hernitschek et al. (2017), where 'L' and 'T' correspond to respectively the leading and trailing arm of the Sagittarius stream (Starkenburg et al., 2019). The red box marks the area of the distant 'spur' feature, identified by Sesar et al. (2017).

BHB stars by Xue et al. (2011), there seems to be more dispersion and more overlap through the BS curve in the CFIS/Pristine/Pan-STARRS space than in the SDSS/Pristine color space. Since the transmission curve of the CFIS-u filter collects more information towards the redder part of the spectrum (see figure 2) than the SDSS-u filter, and since there is less spectroscopically interesting information towards those wavelengths for the separation of BHB stars and BS, the CFIS u-band is more dominated by non-interesting information. This can explain the less distinct sequences in the $u_0 - CaHK_0$, $g_0 - r_0$ color-color space.

4.1.2 Effectiveness at faint magnitudes

As a result of the lesser separation between the ridgelines in the CFIS/Pristine/Pan-STARRS space, more accurate photometry is needed to get similar results as in the SDSS/Pristine space. The top frame of figure 17 shows the average separation between the ridgelines in the SDSS/Pristine space and in the CFIS/Pristine/Pan-STARRS space together with the photometric error in the $u_0 - CaHK_0$ color versus magnitude for all the stars in both bands. Although the photometric error in the CFIS/Pristine/Pan-STARRS band is a lot lower for all $CaHK_0$ magnitudes than the photometric error in the SDSS/Pristine band, the error approaches the scale of the separation between the ridgelines in the color-color space at no significantly different magnitude than for SDSS/Pristine.

We do however notice that the quality of the selection of stars further out seems to have improved when using the CFIS u-band. In the bottom panel of figure 17 we see all the stars with $P(\text{BHB}) > 0.8$ in the magnitude-error space. The errors in the CFIS/Pristine/Pan-STARRS set still do not seem to approach the scale of the separation between the ridgelines at significantly fainter magnitudes than the SDSS/Pristine stars, but the BHB sample at faint magnitudes seems to be much more complete for the CFIS/Pristine/Pan-STARRS set, especially for stars further out than ~ 75 kpc

Although the limiting magnitude for both sets is similar, the uncertainty in SDSS-u has a larger spread at a given $CaHK$ magnitude. This might be related to spatial variations in the SDSS u-band over the footprint (e.g. Finkbeiner et al., 2016). The result is that at that given magnitude, more stars might drop out of our selection. In contrast, the CFIS-u uncertainties are much better behaved which results in a much higher completeness at a given magnitude. This shows by the higher completeness at faint magnitudes in figure 17, providing a much more complete view of the structures in the halo in the range $75 - 120$ kpc.

This is also very well established in figure 13, where the CFIS/Pristine/Pan-STARRS method does a significantly beter job in identifying the Sagittarius stream overdensity at RA ~ 120 at ~ 90 kpc, even though the SDSS/Pristine set and CFIS/Pristine/Pan-STARRS set have similar coverage in that range. Furthermore globular clusters are better identifiable in CFIS/Pristine/Pan-STARRS, as clusters like NGC 2419 and M3 are detectable as opposed to in the SDSS/Pristine filter. This is an interesting result as this means that the new method provides a means to more accurately and more completely map substructures in the galactic halo, although not necessarily further out.

4.1.3 Density profile

In section 3.5 a negative power law with slope $\alpha = 3.8 \pm 0.5$ was found to fit the density profile of the smooth component of the galactic halo. Starkenburg et al. (2019) established a negative slope of 3.5 ± 0.1 for the full sample and a negative slope of 4.0 ± 0.1 for the 'smooth' component. From figure 14 we know that the work of Starkenburg et al. (2019) was constrained by the coverage of the Pristine survey at that point in time, as not much area outside of the Sagittarius stream plane was included. With the new data, the coverage over the 'smooth' component of the halo was largely improved. Since in this work the same method as Starkenburg et al. (2019) was used, it

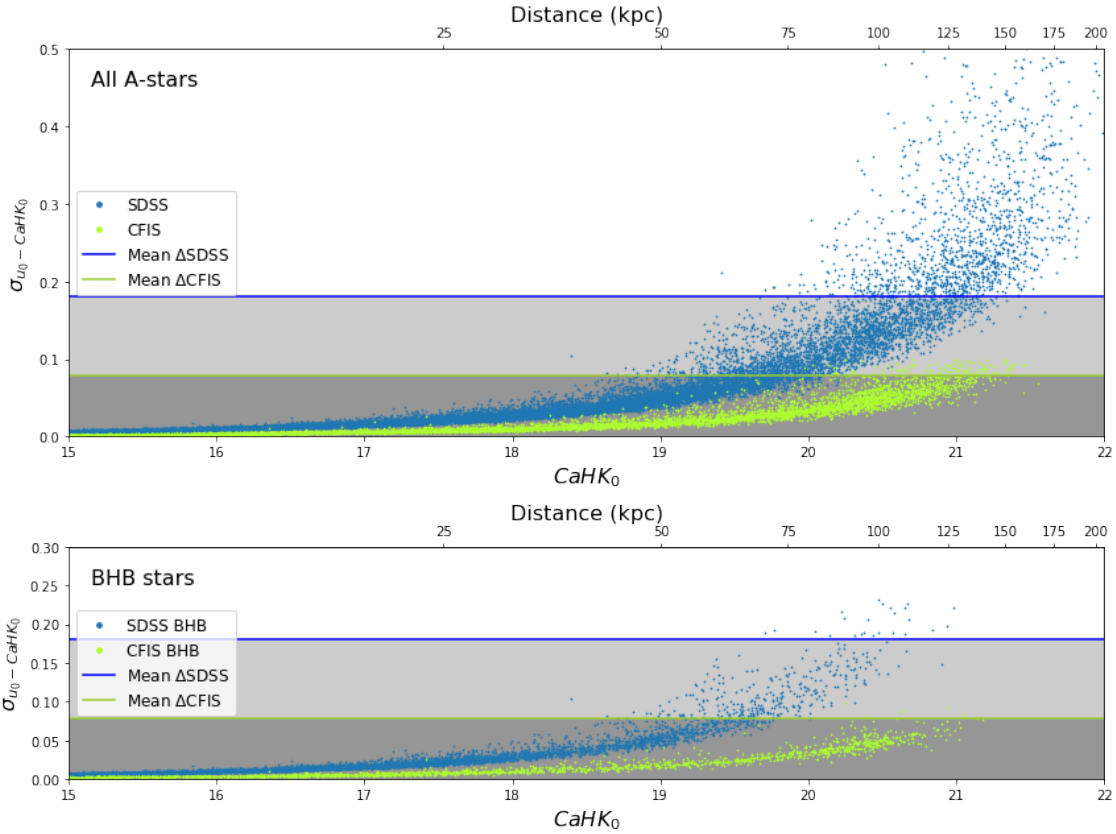

Figure 17: The dots represent the Pristine $CaHK_0$ magnitude against the photometric error in the $u_0 - CaHK_0$ color in both color-color frames. The top frame shows this for all selected A-stars, the bottom panel only shows the stars which have $P(\text{BHB}) > 0.8$. The blue and green horizontal lines represent the mean separation between the BHB and BS ridgeline in respectively the SDSS/Pristine and CFIS/Pristine/Pan-STARRS color space.

The top x-axis corresponds to the average distance to a BHB star with given $CaHK_0$ magnitude. To get the values for the axis the CFIS/Pristine/Pan-STARRS $CaHK_0$-magnitude was first converted to g_0 and $g_0 - r_0$ in a similar manner as for $g_{0_{\text{PS}}}$ to $g_{0_{\text{SDSS}}}$ in section 2.3, after which equations 14 and 15 were used to convert to distance.

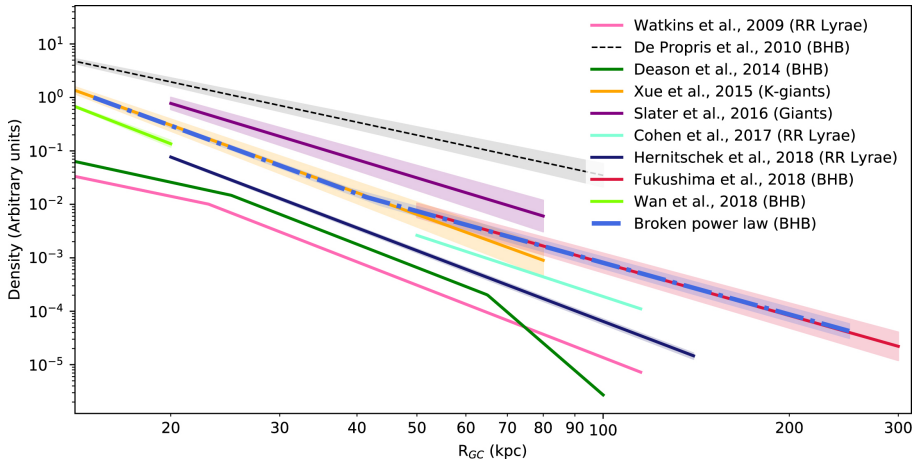

Figure 18: Best-fit models of the density profile of the galactic halo in various work. The blue line ('Broken power law (BHB)') corresponds to the best-fit model obtained by Thomas et al. (2018). Image taken from Thomas et al. (2018).

can be fairly said that especially the power law fit for the 'smooth' component was highly improved.

In the work of Fukushima et al. (2018) a slope with index $\alpha \sim 3.5$ is found for the single power law using a selection of BHB stars. Thomas et al. (2018) find a best fit model of a broken power law with inner slope $\gamma = 4.24 \pm 0.08$, and an outer slope $\beta = 3.21 \pm 0.07$. Both papers use different methods, but we do see that our results fall relatively well in line with these previous results. We note that results of previous analyses have often differed from each other as much as our result differs from other results. In figure 18 a summary is given of results of previous work. Although difficult to observe exactly which power law value each curve corresponds to, we do observe that the various results still show a fair amount of variation. Our result fits well within these results. Nevertheless, the procedure should be further worked out to reduce the uncertainty on the fit, as this is currently still relatively high.

4.2 Further research

The current analysis has still been largely constrained by the coverage of the surveys. In future, more and more area will be covered allowing evaluation of different regions in the halo and eventually coming to a profile of the full galactic halo, while using the same method as applied in this work.

Only a small share of the possible uses of the selection of BHB stars have been addressed in this work. More data will allow for a more thorough analysis of the substructures in the halo, possibly discovering more structures. A better analysis could be made of the Sagittarius stream, trying to understand its formation and substructures even better. Deriving the density profile is only the first step in analysing the galactic halo. Only a single power law was fit to the density profile, but it would be interesting to see the effect of fitting a broken power law in order to get a good comparison with previous results by for example Thomas et al. (2018) or Hernitschek et al. (2018). Furthermore, by combining the data of BHB stars with relative motions of the stars in our set, eventually a constraint could be made on the gravitational potential of the halo, ultimately obtaining a value for the mass of our Milky Way.

4.2.1 Different uncertainties

From the top frame of figure 17 it becomes clear that the photometric errors in the $u_0 - CaHK_0$ color of the CFIS/Pristine/Pan-STARRS set do not exceed ~ 0.1 magnitude. In our data files,

no stars are considered with photometric errors higher than 0.1 on the $CaHK$ magnitude. This means that for detecting stars farthest out we are constrained by the $CaHK$ magnitude. Therefore it would be interesting to see whether more and deeper BHB stars could be identified when considering stars with higher errors on the $CaHK$-magnitude, or that we have already reached the maximum distance we can probe outward to with this method.

Furthermore, we know from figure 17 that the errors on the magnitudes of the SDSS/Pristine sample seem to be much higher than for the CFIS/Pristine/Pan-STARRS space in general. This might explain the less distinct substructures in figure 13 and also the deeper view into the halo for stars with $0.5 < P(\text{BHB}) < 0.8$ in figure 16 can be explained in this way. This is again a reason to also consider stars with higher uncertainties on the magnitudes for the CFIS/Pristine/Pan-STARRS sample, as this would allow for a deeper, although less accurate, view into the halo. By eventually cross checking SDSS/Pristine results with the CFIS/Pristine/Pan-STARRS sample, we could hopefully confirm or disprove the finding on the distant 'spur' feature in section 3.6, in which a light preference was given for the model by Dierickx & Loeb (2017) as opposed to that by Fardal et al. (2019).

There are some areas in around Milky Way satellites and dwarf galaxies which are observed in the Pristine $CaHK$-band, but with longer integration times, providing us with deeper observations (e.g. Longeard et al., 2020). We could try to utilise the data from the excess area around the dwarf galaxies to see what the effect of deeper Pristine data will be on our results. Although we will be highly constrained by the amount of data, it would be interesting to see how well our methods would do with deeper data.

Differentiating BHB stars from BS in other color-color spaces should also be explored. The z-band in companion with the Pristine $CaHK$-band might for example be suited as well to select BHB stars. Combining this band with the Pristine $CaHK$-filter or with other filter bands might provide new means to select BHB stars. Furthermore, better constraints should be set on the completeness of the found set, eventually approximating the completeness as a function of magnitude, as done in the work of Thomas et al. (2018). We could with this information extrapolate the density at far distances and more completely evaluate the density profile far out.

4.2.2 Methodology

At last, some improvements could still be made to the applied method. The method to fit the ridgelines could be improved, and there should be better constraints on the width of the sequences especially in the CFIS/Pristine/Pan-STARRS space. We note that the uncertainty on the width of the sequence has a large effect on the selection especially of the faintest stars in a set. Equations 8 and 9 tell us that the uncertainty on the width of the sequence and the photometric uncertainty of the star accumulate, and the higher the total uncertainty, the lower the chance that a star is given a large probability of being BHB. Therefore, more adequately evaluating the sequence widths will allow for a better completeness and lower contamination of the sample towards the ourskirts.

Starkenburg et al. (2019) also apply a couple of extra cuts in the data which were not applied in our own analysis, which would further narrow down the sample. Starkenburg et al. (2019) apply a cut of 0.2 on the uncertainty of the SDSS g-band magnitude. Also the distances obtained with equation 15 are cross-checked with parallax distances from Gaia DR-2. Although having only little effect ($\sim$ 1 percent less stars), some outliers could be filtered out.

We provide the code developed for the analysis of the data in a directory for which the link is provided in appendix A.2. Anyone interested can use this in order to further develop the method or use the selection of BHB stars for further analysis.

5 Conclusion

In this analysis we used photometric data from SDSS DR-14, the Pristine survey, CFIS intermediate DR-3 and Pan-STARRS DR-2. We explored the behaviour of the A-stars in the survey in the $u_0 - CaHK_0$, $g_0 - r_0$ color-color space in both the SDSS/Pristine as well as in the CFIS/Pristine/Pan-STARRS space. Ridgelines were fit through the BHB sequences and the BS sequences in the two spaces, and probabilities were assigned to each star being either BHB or BS. By selecting stars with $P(\text{BHB}) > 0.8$ and inferring distances to these stars utilising the standard candle property of BHB stars, a profile of the galactic halo could be made. Multiple substructures could be identified and a density profile could be fit to the data. A negative power law with value $\alpha = 3.8 \pm 0.5$ was found for the smooth component of the halo.

When we compare our results to that of Starkenburg et al. (2019), we can notice that various improvements have been made. Due to the significantly larger coverage in the SDSS/Pristine set, we were able to put a better constraint on the density profile of the smooth component of the halo and we were able to better qualify the effectiveness of the method.

By implementing the method for the CFIS u-band as well, we were able to improve the capability to detect substructures far out, although not necessarily probing out further into the halo compared to using SDSS-u. The forecast is that by considering stars with higher errors on the magnitude, we will be able to probe outward into the halo even further. Although the quality of the resulting findings will be lower, this might still result in obtaining new insights on for example the distant 'spur' feature of the Sagittarius stream, the profile of the smooth halo component and possibly detecting new objects. The new method shows a promising means to trace out the substructures inthe halo up to ~ 150 kpc, with more data in future providing us many new opportunities.

Acknowledgements

I would like to gratefully thank Else Starkenburg for the enormous help and support during this research project. Our weekly meetings on Thursday were always very helpful and very much worth looking forward to! It is a shame that we have only been able to meet each other in an online environment, but I am looking forward to meeting you in person in future. Thank you for providing the topic and making me familiar with the interesting field of galactic archaeology. I am looking forward to hopefully extend the results on this research in future!

Also very many thanks to Guillaume Thomas for providing the CFIS/Pristine/Pan-STARRS data set and providing the first steps in analysing it. The insights towards the end of the research on the results were of great help as well and very much appreciated!

References

Aihara H., et al., 2011, Astrophysical Journal, Supplement, 193, 29

Belokurov V., et al., 2006, Astrophysical Journal, Letters, 647, L111

Belokurov V., et al., 2014, Monthly Notices of the Royal Astronomical Society, 437, 116

Boulade O., et al., 2003, in Iye M., Moorwood A. F. M., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. pp 72–81, doi:10.1117/12.459890

Bullock J. S., Johnston K. V., 2005, Astrophysical Journal, 635, 931

Chambers K. C., et al., 2016, arXiv e-prints, p. arXiv:1612.05560

Corbelli E., Schneider S. E., Salpeter E. E., 1989, Astronomical Journal, 97, 390

Deason A. J., Belokurov V., Evans N. W., 2011, Monthly Notices of the Royal Astronomical Society, 416, 2903

Deason A. J., Belokurov V., Koposov S. E., Rockosi C. M., 2014, Astrophysical Journal, 787, 30

Dierickx M. I. P., Loeb A., 2017, Astrophysical Journal, 836, 92

Fardal M. A., van der Marel R. P., Law D. R., Sohn S. T., Sesar B., Hernitschek N., Rix H.-W., 2019, Monthly Notices of the Royal Astronomical Society, 483, 4724

Finkbeiner D. P., et al., 2016, Astrophysical Journal, 822, 66

Fukushima T., et al., 2018, Publications of the Astronomical Society of Japan, 70, 69

Gunn J. E., et al., 2006, Astronomical Journal, 131, 2332

Harris W. E., 1996, The Astronomical Journal, 112, 1487

Helmi A., 2020, Annual Review of Astron and Astrophysics, 58, 205

Hernitschek N., et al., 2017, Astrophysical Journal, 850, 96

Hernitschek N., et al., 2018, Astrophysical Journal, 859, 31

Hodapp K. W., et al., 2004, Astronomische Nachrichten, 325, 636

Ibata R. A., et al., 2017, Astrophysical Journal, 848, 128

Kutner M. L., 2003, Astronomy: A Physical Perspective, second edn. Cambridge University Press

Longeard N., et al., 2020, Monthly Notices of the Royal Astronomical Society, 491, 356

Mateu C., Read J. I., Kawata D., 2018, Monthly Notices of the Royal Astronomical Society, 474, 4112

Munari U., Sordo R., Castelli F., Zwitter T., 2005, Astronomy and Astrophysics, 442, 1127

Schlafly E. F., Finkbeiner D. P., 2011, Astrophysical Journal, 737, 103

Sesar B., Hernitschek N., Dierickx M. I. P., Fardal M. A., Rix H.-W., 2017, Astrophysical Journal, Letters, 844, L4

Starkenburg E., et al., 2017, Monthly Notices of the Royal Astronomical Society, 471, 2587

Starkenburg E., et al., 2019, Monthly Notices of the Royal Astronomical Society, 490, 5757

Thomas G. F., et al., 2018, Monthly Notices of the Royal Astronomical Society, 481, 5223

Vickers J. J., Grebel E. K., Huxor A. P., 2012, Astronomical Journal, 143, 86

Xue X. X., et al., 2008, Astrophysical Journal, 684, 1143

Xue X.-X., et al., 2011, Astrophysical Journal, 738, 79

Yanny B., et al., 2000, Astrophysical Journal, 540, 825

York D. G., et al., 2000, Astronomical Journal, 120, 1579

A Appendix

A.1 Sagittarius Stream coordinate system

The Sgr stream coordinate system as given by Belokurov et al. (2014). α and δ correspond to equatorial coordinates (α = Right Ascension and δ = Declination).

$$\Lambda_{\rm Sgr} = \mathrm{atan2}(-0.93595354\cos(\alpha)\cos(\delta) - 0.31910658\sin(\alpha)\cos(\delta) + 0.14886895\sin(\delta),\\ 0.21215555\cos(\alpha)\cos(\delta) - 0.84846291\sin(\alpha)\cos(\delta) - 0.48487186\sin(\delta))$$

$$B_{\rm Sgr} = \arcsin(0.28103559\cos(\alpha)\cos(\delta) - 0.42223415\sin(\alpha)\cos(\delta) + 0.86182209\sin(\delta)),$$

where $\tan(\mathrm{atan2}(x, y)) = y/x$.

A.2 Code

The developed code is provided in a zipfile in a personal directory on the Kapteyn cloud service. The directory can be accessed using the link below. A password is necessary which can be retrieved by sending an email to titulaer@astro.rug.nl.

The link:
https://astrodrive.astro.rug.nl/index.php/s/lI6yw4ISCsEQZHb