Draft version August 17, 2023
Typeset using LaTeX default style in AASTeX62

MKY Project Outline

Jiani Ding

(Dated: November 2020)

1. INTRODUCTION

We attempt to use a new approach to constrain the Milky Way accretion history. Due to the long dynamical time, the velocity of the majority of the stars in the outer halo substructure will remain coherent. The upcoming large redshift survey DESI will provide a largest number of 'tracers' (giant and horizontal branch stars) in the outer halo (beyond $\sim$ 20 kpc). Thus, we establish a quantity (the velocity difference between stellar pairs) to investigate the substructure in the outer halo. The properties in the distributions of the velocity difference are able to trace the significance of the substructure as a function of the radius. We will first test our method on the simulated halos from Bullock & Johnson simulation, and then apply it on the observation data.

2. SIMULATION DATA

We are using Bullock & Johnston (2005) Halo simulation model with the Galaxia galaxy simulation to produce the simulated galaxy and stars for different accretion history. Take halo05,12 and 17 as examples, the diagram that illustrates the accretion history of halo 05, halo 12 and halo 17 is showed in Fig. 1. We test our method on the halo05, halo12 and halo 17 (halo with recent accretion history) and all other halos in the simulation to investigate the correlation between the quantity defined by our method and the quantities given by the simulation that are related to the accretion history of the different halos.

3. GENERATING SIMULATED STARS

Based on the Bullock & Johnston (2005) simulation, we first simulate the halo data for eleven halos and then use the code in the Sharma et al. (2011) to generate a synthetic survey with continuous and smooth star distributions. For the galactic distance from 20 kpc to 80 kpc, the red giant and BHB stars are ideal for the substructure study with the DESI Milkyway Survey depth. We make a gravity cut for selecting the red giant in our study and randomly select 10000 stars for each halo from the simulation, in order to simulate testing our method in the DESI survey team but with simulation data.

4. VELOCITY DIFFERENCE BETWEEN STAR PAIRS

We compute the pairwise velocity difference for stellar pairs with a limited spatial scale separation in the distance 20-76 kpc. The reasons we choose the galactic distance ranging from 20 kpc to 70 kpc are as follows. 1. The outer halo ($>$ 20 kpc) has longer dynamical time and higher signal for the substructure. 2. The limitation of the stars number in the distant halo $>$ 76 kpc limit the statistic for the pairwise velocity of the star pairs.

We will change the limited spatial scale separation for those stellar pairs in order to see how the velocity difference distribution varies with respect to different spatial scale. We will test our theory on three different spatial scales (6,12 kpc). Take the 6 kpc scale case as an example. For the first four shells, we use a 6 kpc shell sizes. For the last two shells, we use a 16 kpc shell size. Then we calculate the velocity difference between the stars inside the shell and outside the shell with different spatial distance separations (6,12 kpc). As for the detailed process, we first select a subsample ($\sim$ 60000) for all the stars from the halo and then select all the stars inside a 6/16 kpc shell (galactic distance 20-26 kpc for the first shell, 26-32 kpc for the second...). Then we calculate velocity difference between the stars inside the shell and all the stars within a spatial separations (6,12 kpc). Finally, we plot the velocity difference distributions. The pairwise velocity difference different star pairs for halo05 is showed in Fig.2. The pairwise velocity difference different star pairs for halo12 is showed in Fig.3. The pairwise velocity difference different star pairs for halo17 is showed in Fig.4.

However, due to the limitation of the DESI survey, the majority of our sample may be the stars in the red giant branch. From Fig. 5, the cut for surface gravity of the red giant branch is around 2.5. After making the surface gravity cut, we calculate the velocity difference.

5. UNBIN STATISTICS FITTING FOR THE TWO GAUSSIAN/ONE GAUSSIAN MODEL

We also try to apply a MCMC fitting for the original velocity difference data without any binning. We use a two Gaussian function (narrow Gaussian plus broad Gaussian function) for modeling the data by assuming a ratio R between the integration of the area under the narrow and broad Gaussian distribution. After normalization, the two Gaussian model can be written by

$$R \times \frac{exp(-(x-center)^2}{(2*sigma*sigma))sigma\sqrt{(2\pi)}} + (1-R) \times \frac{exp(-(x-center2)^2}{(2*sigma2*sigma2))sigma2\sqrt{(2\pi)}} \quad (1)$$

, where x is the data, center1 and center2 is the centers for the narrow and broad Gaussian distribution and sigma1 and sigma2 is the σ for the narrow and broad Gaussian distribution. The data for fitting will be the velocity difference in each galactic radial distance shells (20-26 kpc, 26-32 kpc, 32-38 kpc, 38-44 kpc, 44-60 kpc, 60-76 kpc) for different halos (halo1-halo11).

Using the statistical method illustrated in appendix C.2. (Kepner et al. 1999), we assume virtual bins that are sufficiently small so no bin contains more that one velocity difference data. The sum of the log probability L = $-\sum_{empty} n^j_{model} - \sum_{filled} n^j_{model} + \sum_{filled} lnn^j_{model}$. In this equation, the likelihood only depends on the value of the data of the velocity difference. The first two terms are the sum over all the empty bins and all the filled bins. The first two terms are just the total number of data points, which is a constant, so for finding the optimize parameters (width, center for the narrow and broad Gaussian distribution and the ratio between the narrow and broad Gaussian distribution), we just need to maximize the last term of this likelihood function. We use the uniform distributions -50 $<$ center1 $<$ 50 and 0 $<$ sigma1 $<$ 70 and -50 $<$ center2 $<$ 50 and 100 $<$ sigma2 $<$ 500 and -1e-5 $<$ R $<$ 1-1e-5 as the prior distribution and using MCMC to search for the best fitted parameters, especially the ratio R between the narrow and broad Gaussian distribution. We use the emecee MCMC package and the above likelihodd in the fitting process. The ratio R will indicate the significance of the substructure since it reflect the integration of the area under the narrow and broad Gaussian distribution. We use a 6000 iterations with a 1500 burn iterations and the final corner plots indicate a converged trace and result for all the halos.

6. ERROR OF THE GAUSSIAN FITTING

We have tested the convergence of the traces in the MCMC unbin fitting. We follow Goodman & Weare (2010) and use the integrated autocorrelation time to quantify the effects of sampling error on MCMC results. The corner plot made from the converged trace show the confidence level for each fitted parameters for all the halos. The 1-sigma, 2-sigma, 3-sigma errors for those parameters could be extracted from the corner plots. Take the corner plot for the seven fitted parameters from halo17 as an example, in Fig. 6, the trace for the seven parameters in the corner plot for the halo17 have converged and the corner plot showing less correlation between those parameters, the 1-sigma, 2-sigma errors are revealed in the plot too.

7. RATIO OF DETECTABLE SUBSTRUCTURE AND THE MINIMUM NUMBER TO ATTAIN THE RATIO

We first plot ratio R as a function of the galactic radial distance for all the halos. We then explore the minimum number of stars we need for each shell (each shell corresponds to a Ratio R) to attain a converged ratio R in the MCMC fitting for all the halos. The minimum number is defined as the number of stars required to get a converged result for R within 2-sigma error of the R attained for the MCMC run with the full sample in each shell. The final results (minimum number of stars required to attain a converged trace for the best-fitted parameters ratio R) for (halo02, halo15, halo17 and halo20) are showed below.

8. MEASUREMENT RESULTS

We finally choose the two Gaussian (narrow Gaussian component + broad Gaussian component) and one Gaussian model fitting to analyze our results. We apply this approach (two and one Gaussian fitting and errors of the best fitted parameters attained from hessian matrix) to all the halos in Bullock & Johnston (2005). We have seven parameters (amplitude, width, center of the narrow Gaussian, the amplitude, width, center of the broad Gaussian and the offset

of the resultant distribution) for the two Gaussian model and four parameters (amplitude, width, center of the broad Gaussian and the offset of the resultant distribution) for the one Gaussian model. We also attain the errors for each of the parameter of each shell in all the halos. The results (using a 12 kpc spatial separation) for the ratio between the narrow and broad Gaussian peak as a function of radius is in Fig. 7. We also explore the minimum number of stars required to attain a converged result for the ratio measurement in each shell of all the halos. The final results are showed in the Fig. 8. This result indicate that to achieve a converged results for some of the outer halos, we only require < 1000 stars.

9. TESTING VELOCITY DIFFERENCE MEASUREMENT METHOD IN OBSERVATION DATA

After testing our method in the simulation data, we also apply the method in the observation data. We apply our technique on the 5000 Blue Horizontal Branch (BHB) stars in sample from (Xue et al. 2011). The final distribution of the velocity difference measurement with a 12 kpc scale separation constraint for the 5000 BHB stars sample is showed in 9, with a 6 kpc scale separation constraint is in 10. The 3-sigma region of the GMM fitting on the BHB star sample distribution are showed in Fig. 9 and Fig. 10.

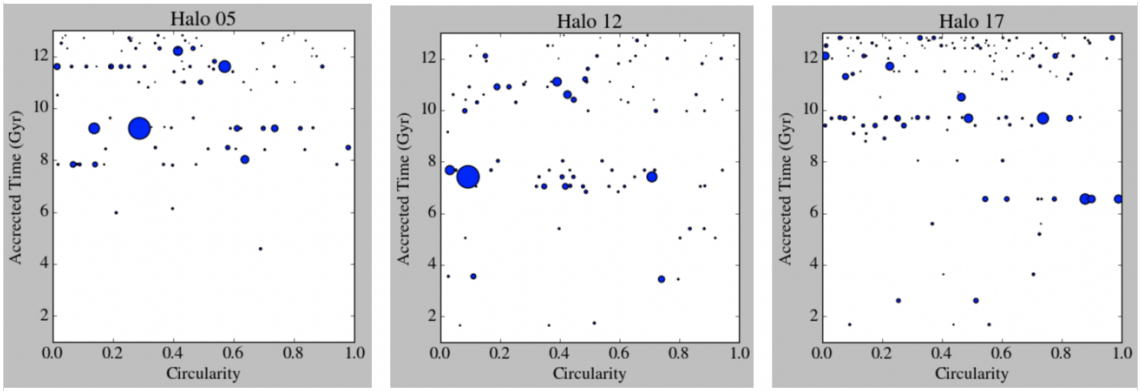

Figure 1. Accretion History plot for halo05, halo12 and halo17. From left to right panel is the accretion time versus circularity for halo05, halo12 and halo17. The size of the blue data points stands for the stellar mass of the satellites.

REFERENCES

Bullock, J. S., & Johnston, K. V. 2005, ApJ, 635, 931, doi: 10.1086/497422

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65, doi: 10.2140/camcos.2010.5.65

Kepner, J., Fan, X., Bahcall, N., et al. 1999, ApJ, 517, 78, doi: 10.1086/307160

Sharma, S., Bland-Hawthorn, J., Johnston, K. V., & Binney, J. 2011, ApJ, 730, 3, doi: 10.1088/0004-637X/730/1/3

Xue, X.-X., Rix, H.-W., Yanny, B., et al. 2011, ApJ, 738, 79, doi: 10.1088/0004-637X/738/1/79

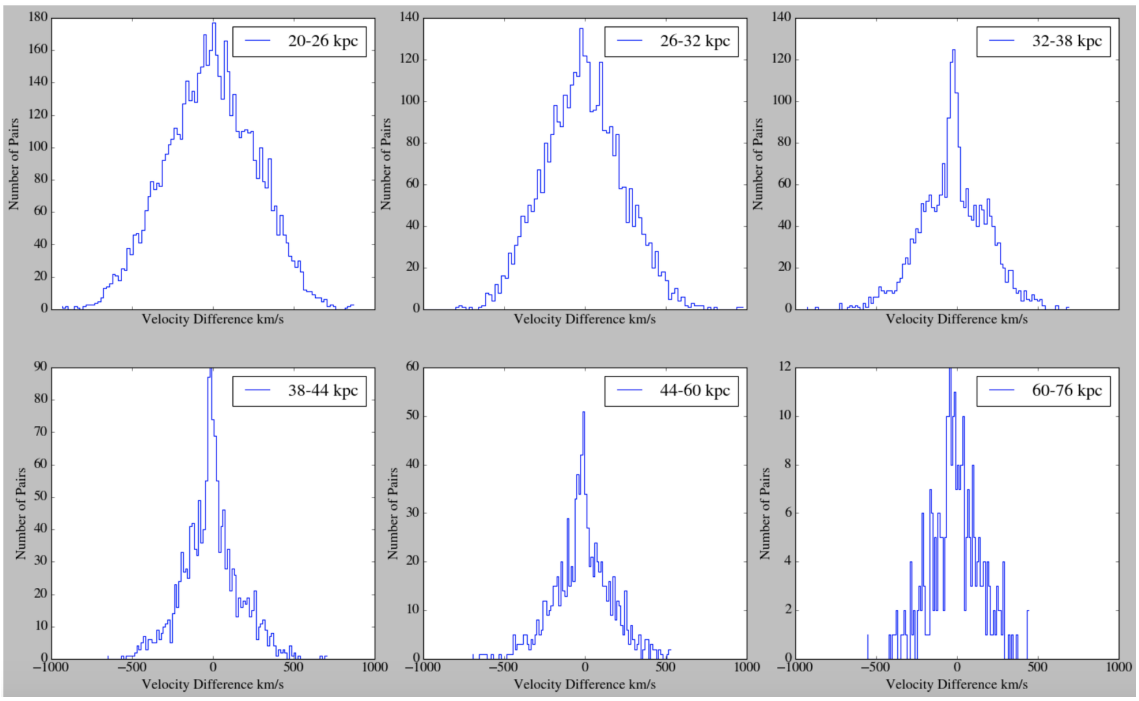

Figure 2. Pairwise velocity difference between different star pairs for halo05

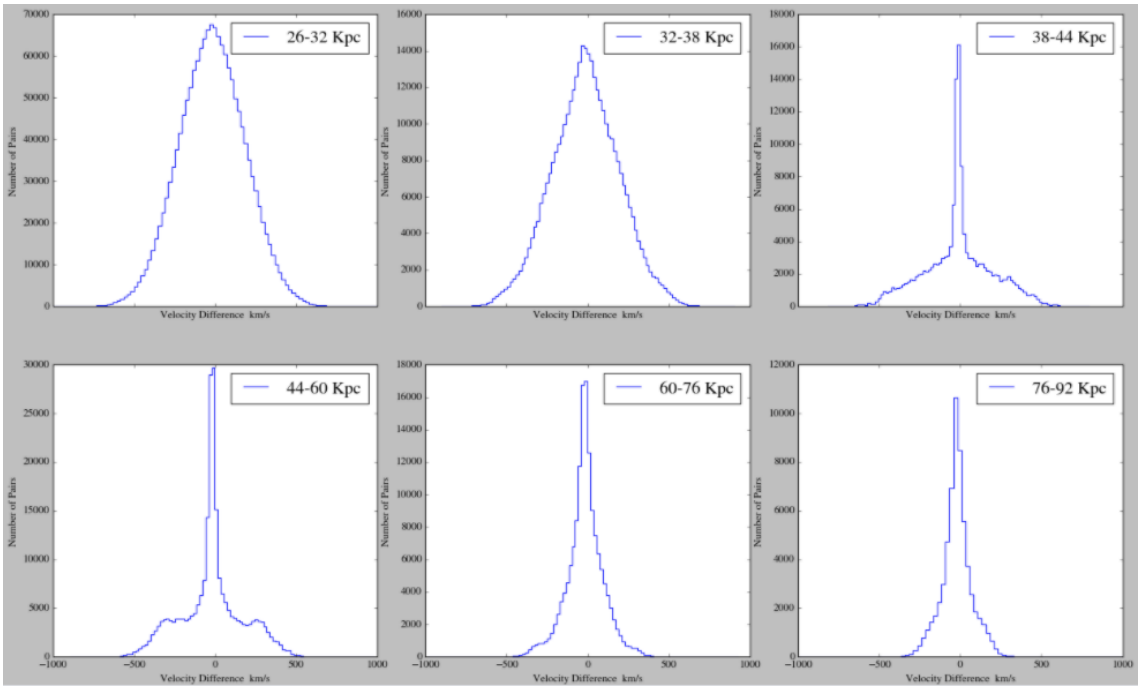

Figure 3. Pairwise velocity difference between different star pairs for halo12

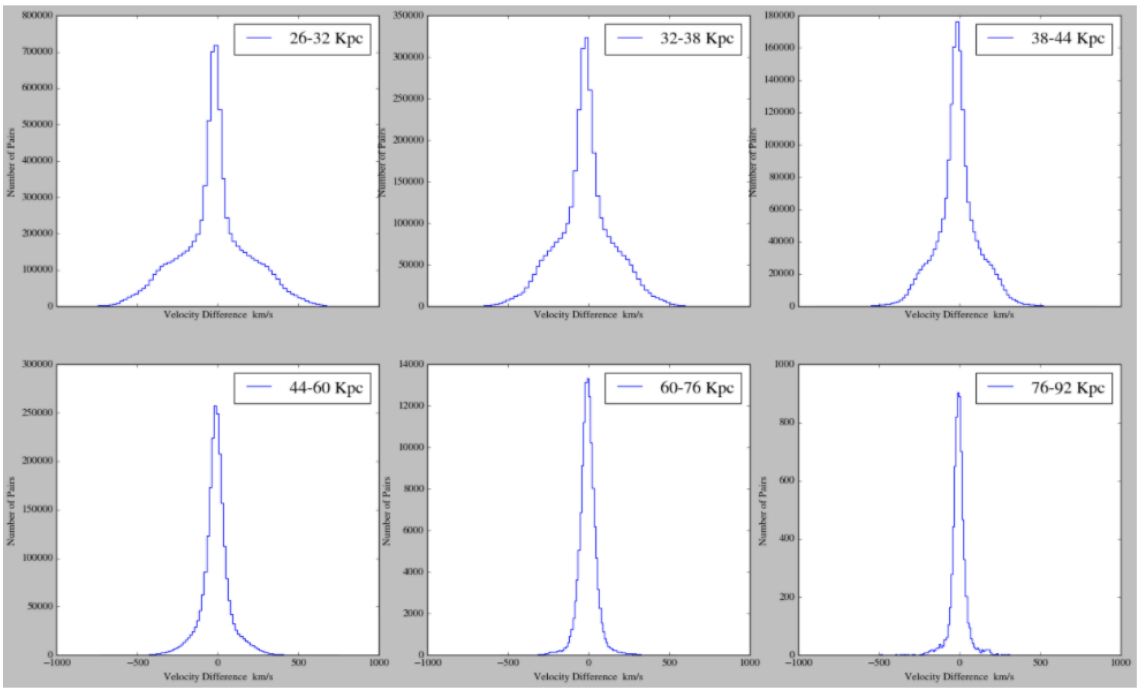

Figure 4. Pairwise velocity difference between different star pairs for halo17

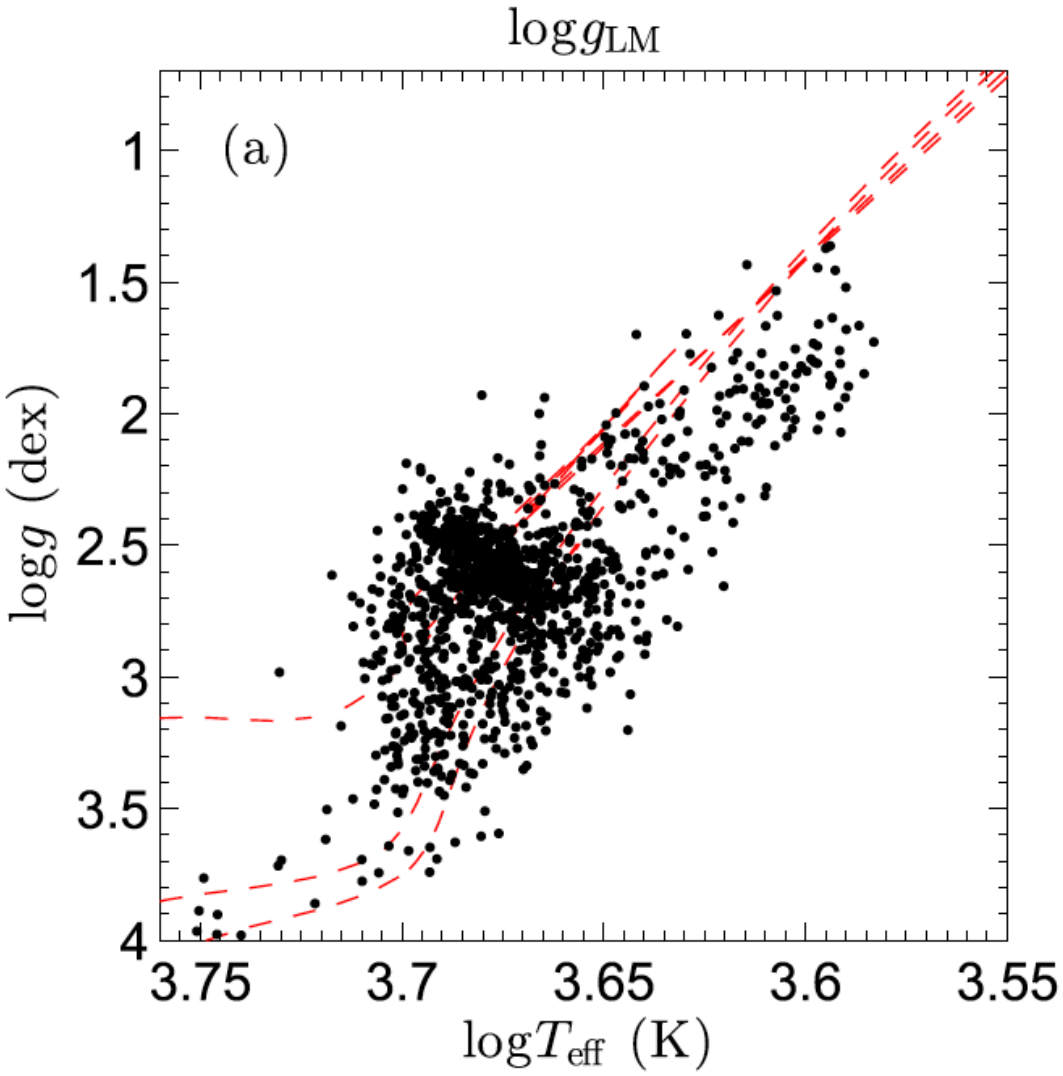

Figure 5. Teff–log g diagrams from Liu et al. 2015

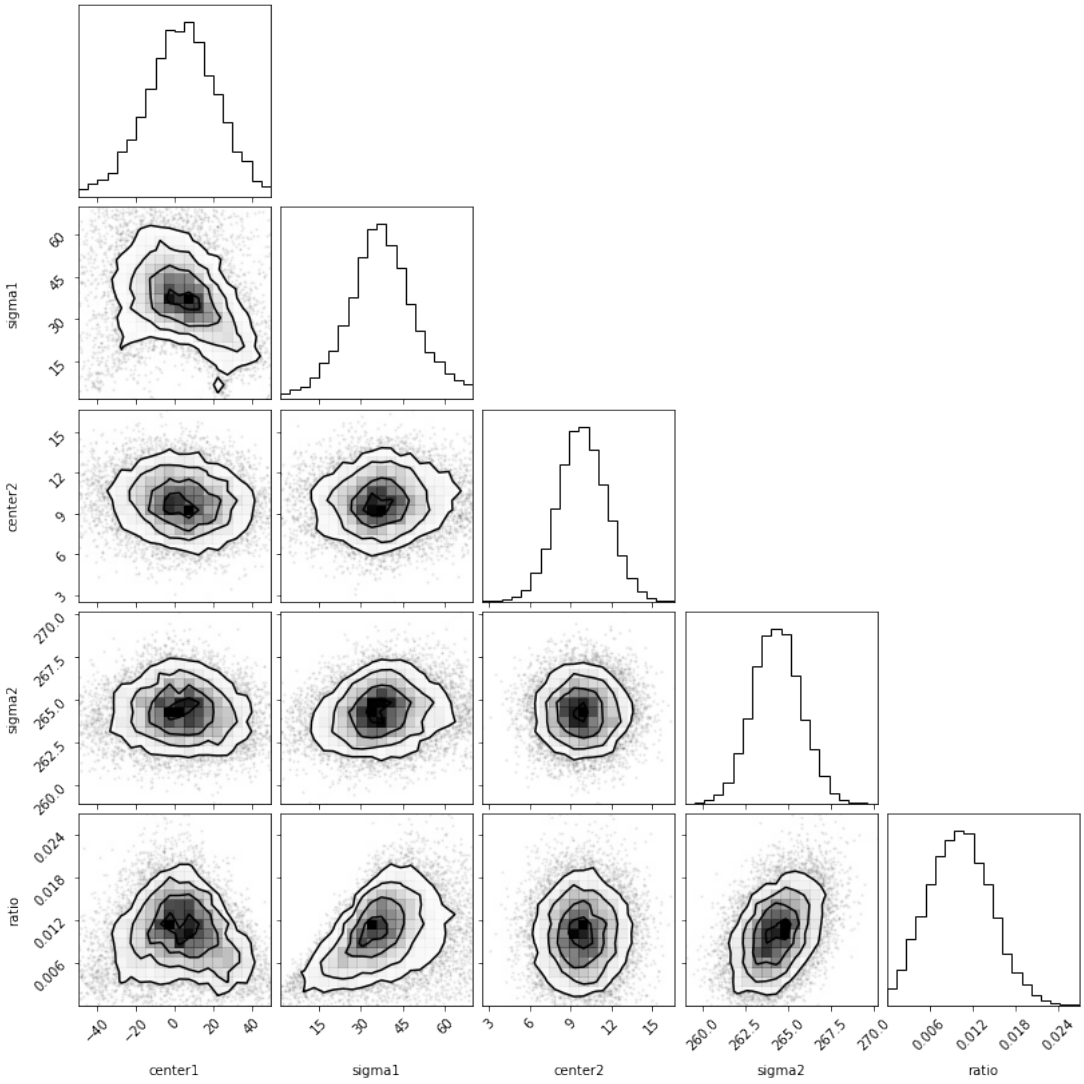

Figure 6. MCMC Corner plot for the seven fitted parameters of the halo17.

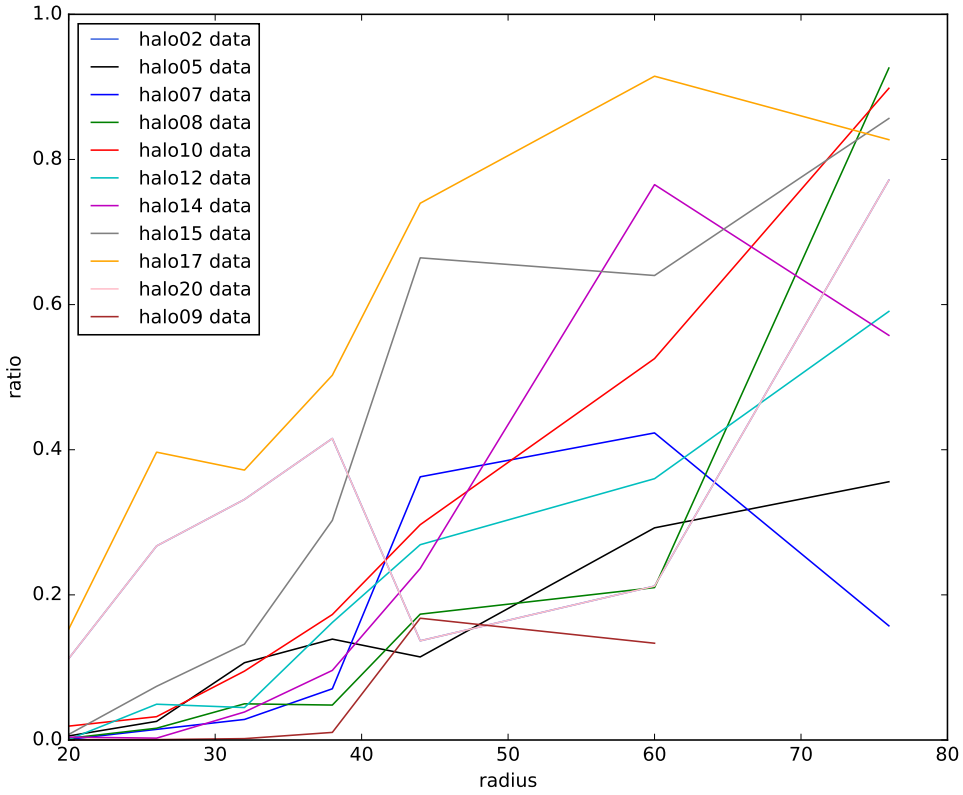

Figure 7. The ratio between the integration area under the narrow and broad Gaussian distribution as a function of the radius.

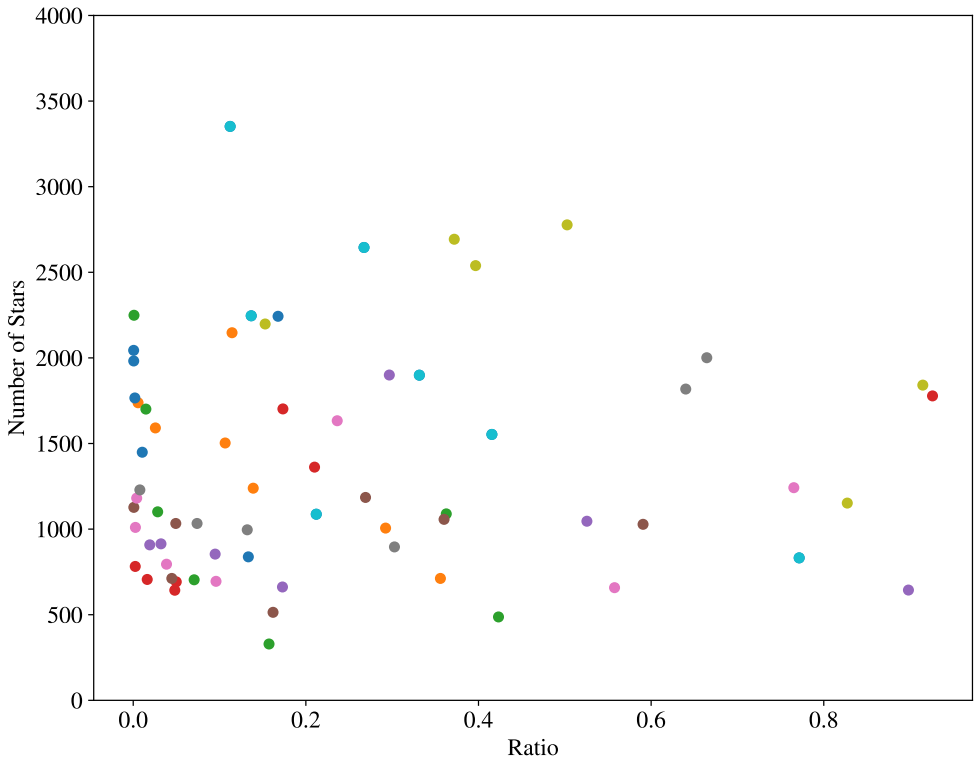

Figure 8. The minimum number of stars required to attain a converged result for the measurements of the seven parameters in each shell of all the halos.

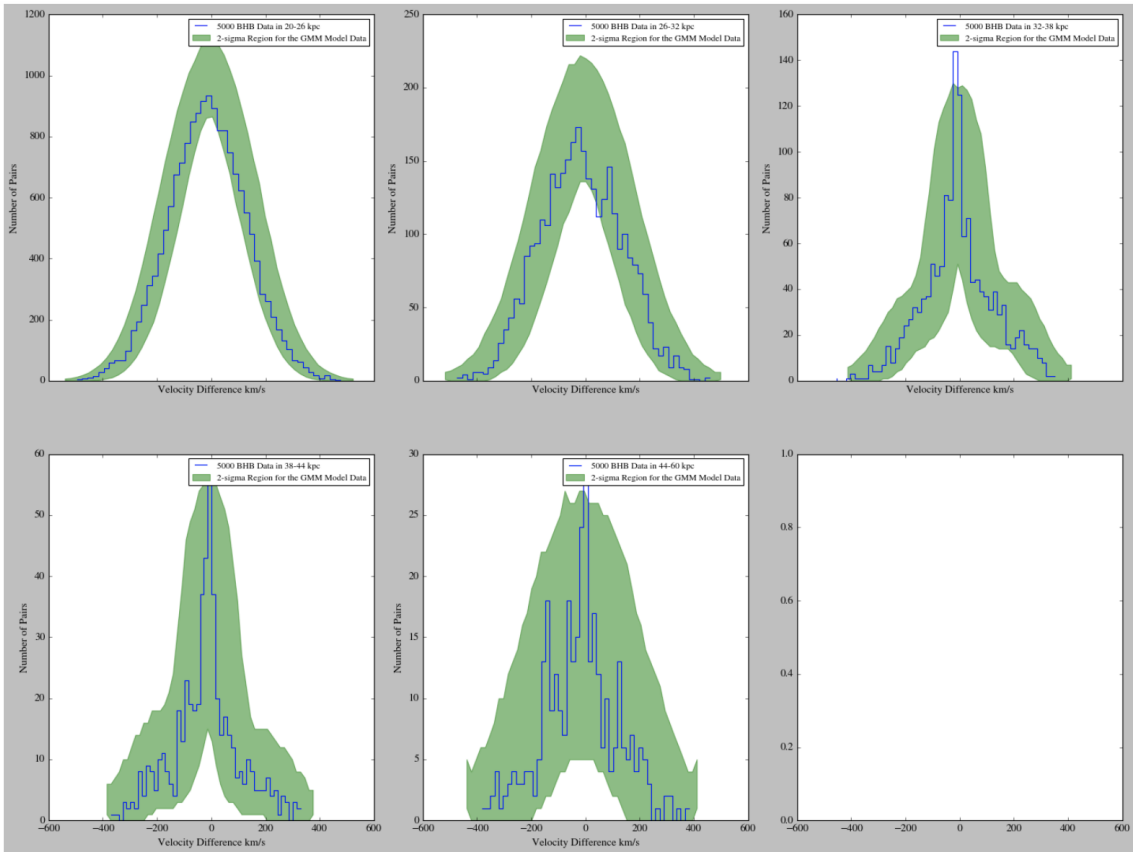

Figure 10. The 3-sigma region of the GMM fitting on the BHB star sample distribution with a 6 kpc scale constraint

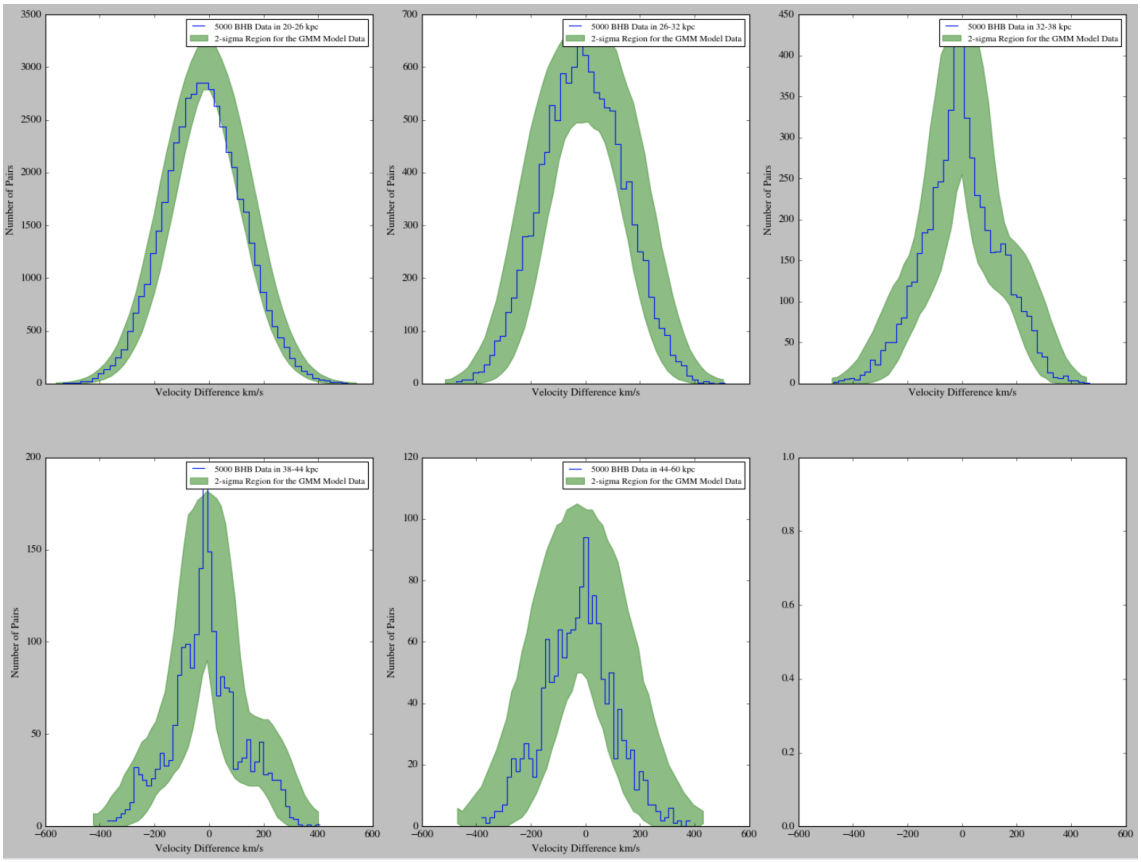

Figure 9. The 3-sigma region of the GMM fitting on the BHB star sample distribution with a 12 kpc scale constraint