Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **482,** 4215–4232 (2019)
Advance Access publication 2018 October 30

doi:10.1093/mnras/sty2924

The GALAH survey: velocity fluctuations in the Milky Way using Red Clump giants

Shourya Khanna [1,2★] Sanjib Sharma,[1,2★] Joss Bland-Hawthorn [1,2,3] Michael Hayden,[1,2] David M. Nataf,[4] Yuan-Sen Ting,[5,6,7] Janez Kos,[1,2] Sarah Martell,[8] Tomaž Zwitter,[9] Gayandhi De Silva,[1,10] Martin Asplund,[11] Sven Buder,[12] Ly Duong,[11] Jane Lin,[11] Jeffrey D. Simpson,[10] Borja Anguiano,[13] Jonathan Horner,[14] Prajwal R. Kafle,[15] Geraint F. Lewis,[1] Thomas Nordlander,[2,11] Rosemary F.G. Wyse,[4] Robert A. Wittenmyer[16,17] and Daniel B. Zucker[1,2]

Affiliations are listed at the end of the paper

Accepted 2018 October 22. Received 2018 October 12; in original form 2018 April 20

ABSTRACT

If the Galaxy is axisymmetric and in dynamical equilibrium, we expect negligible fluctuations in the residual line-of-sight velocity field. Recent results using the *APOGEE* survey find significant fluctuations in velocity for stars in the mid-plane ($|z|$ <0.25 kpc) out to 5 kpc, suggesting that the dynamical influence of non-axisymmetric features, i.e. the Milky Way's bar, spiral arms, and merger events extends out to the Solar neighbourhood. Their measured power spectrum has a characteristic amplitude of 11 km s^{-1} on a scale of 2.5 kpc. The existence of such large-scale streaming motions has important implications for determining the Sun's motion about the Galactic Centre. Using Red Clump stars from *GALAH* and *APOGEE*, we map the line-of-sight velocities around the Sun ($d < 5$ kpc), and $|z|$ <1.25 kpc from the mid-plane. By subtracting a smooth axisymmetric model for the velocity field, we study the residual fluctuations and compare our findings with mock survey generated by GALAXIA. We find negligible large-scale fluctuations away from the plane. In the mid-plane, we reproduce the earlier *APOGEE* power spectrum but with 20 per cent smaller amplitude (9.3 km s^{-1}) after taking into account a few systematics (e.g. volume completeness). Using a flexible axisymmetric model the power amplitude is further reduced to 6.3 km s^{-1}. Additionally, our simulations show that, in the plane, distances are underestimated for high-mass Red Clump stars which can lead to spurious power amplitude of about 5.2 km s^{-1}. Taking this into account, we estimate the amplitude of real fluctuations to be <4.6 km s^{-1}, about a factor of three less than the *APOGEE* result.

Key words: stars: distances – stars: fundamental parameters – Galaxy: kinematics and dynamics.

1 INTRODUCTION

The Milky Way is a large late-type disc galaxy. While the Galaxy has had a quiescent accretion history and is not thought to have experienced a major merger in the past 10 Gyr (Stewart et al. 2008; Bland-Hawthorn & Gerhard 2016), it is orbited by nearby dwarf galaxies, some of which can cross the disc and perturb it. Some of these orbiting galaxies become disrupted by these encounters, torn asunder to create streams of material orbiting the galaxy such as the Sagittarius stream (Majewski et al. 2003). In addition, the Milky Way also hosts a central bar extending out to 5 kpc (Wegg, Gerhard & Portail 2015), the dynamical effect of which can also be seen in the Solar neighbourhood as structures in velocity space. Prominent examples of such structures include the Hercules stream (e.g. Dehnen 1998; Bovy 2010; Hunt et al. 2018). While it would seem natural to assume kinematic non-axisymmetry at small radii it is worth investigating whether the dynamical effects of the bar extend out to large radii such as around the Solar neighbourhood. Studying the velocity substructure in the Milky Way thus holds clues to large-scale evolutionary processes in the Galaxy.

★ E-mail: skha2680@uni.sydney.edu.au (SK);
sanjib.sharma@sydney.edu.au (SS)

Over the last two decades local surveys such as *GCS* (Geneva-Copenhagen Survey, Nordström et al. 2004) and *RAVE* (Radial Velocity Experiment, Williams et al. 2013) have mapped the Solar neighbourhood extensively and shown evidence of velocity gradients in the disc. With the advent of large-scale surveys it is now possible to venture out of the Solar neighbourhood. For example, with *APOGEE* (Apache Point Observatory Galactic Evolution Experiment; Majewski, APOGEE Team & APOGEE-2 Team 2016) the Galactic disc in the mid-plane ($|z| < 0.25$ kpc) has been mapped out to 15 kpc. The synergy with other ongoing surveys such as *GALAH* (GALactic Archaeology with HERMES; Martell et al. 2017) and *LAMOST* (Large sky Area Multi-Object fibre Spectroscopic Telescope; Zhao et al. 2012), now allows us to study the region away from the mid-plane and attempt to visualize kinematics and chemistry in 3D.

The limitations of small-scale surveys can be understood in the context of the Local Standard of Rest (LSR)[1] which is generally defined based on results from the *GCS* survey to be $(U, V, W)_{\odot} = (11.1, 12.24, 7.25)$ km s^{-1} (Schönrich et al. 2010). However, it has been suggested that the 'true' LSR may differ from this current standard, most notably through the study of kinematics of *APOGEE* Red Clump stars by (Bovy et al. 2015; hereafter B15). In their analysis they subtract an axisymmetric model for the line-of-sight velocity field and find significant residual bulk motion or streaming of about 11 km s^{-1}. Fourier analysis of the residual motion shows that the scale of fluctuations is about 3 kpc, i.e. much larger than the Solar neighbourhood. B15 suggest that the whole solar neighbourhood is moving with respect to the Galaxy on a non-axisymmetric orbit due to perturbations from the Galactic Bar. Furthermore, taking this large-scale streaming motion into account, they suggest that the value of $V_{\odot}$, the Sun's motion relative to the circular velocity ($V_{\rm circ}$), be revised upwards to 22.5 km s^{-1}, implying that the Solar neighbourhood is moving ahead of the LSR. Although, the proper motion of Sgr A$*$ is well constrained at 6.379 ± 0.024 mas yr^{-1} (Reid & Brunthaler 2004), there is still uncertainty on the distance of the Sun from the Galactic centre ($R_{\odot}$). It is thus important to have multiple methods to constrain the Solar motion about the Galactic centre. Robin et al. (2017) take this forward by making use of highly accurate proper motions from *Gaia* data release 1 (Gaia Collaboration 2016) and *RAVE*-DR4 to model the asymmetric drift and leave the Solar motion as a free parameter. They propose a new, much lower value for $V_{\odot}$ of 0.94 km s^{-1}, but again their analysis is only restricted to the Solar vicinity probing well within 2 kpc of the Sun with mean distances around 1 kpc.

Using a purely astrometric sample from *Gaia*-TGAS, Antoja et al. (2017) detect velocity asymmetries of about 10 km s^{-1} between positive and negative Galactic longitudes. Once again, however, their study is only based on proper motions of the stars involved. The detected asymmetry seems directed away from the Solar neighbourhood and towards the outer disc. Similarly, using SDSS-DR12 white dwarf kinematics, Anguiano et al. (2017) find $\partial V_{\rm R}/\partial R = -3 \pm 5$ km s^{-1} and an asymmetry in $\langle V_{\rm Z} \rangle$ between the population above and below the plane of the Galaxy.

Given the compelling evidence of non-equilibrium kinematics shown by a diverse stellar population, it would clearly be interesting to perform a large-scale 3D analysis of the Milky Way. In this paper we examine the line-of-sight kinematics of Red Clump (RC) stars selected from the *GALAH* and *APOGEE* spectroscopic surveys. In general, observed data has a non-trivial selection function, and in some cases leaves a strong signature on the data. Not taking this into account can lead to spurious fluctuations in the velocity distribution of the target stars. Hence it is imperative to check and compare the results to those obtained through use of a synthetic catalogue of stars. To this end we make use of axisymmetric galaxy models using the GALAXIA[2] code (Sharma et al. 2011).

Throughout the paper we adopt a right-handed coordinate frame in which the Sun is at $R_{\odot} = 8.0$ kpc from the Galactic centre and has Galactocentric coordinates $(X_{\rm gc}, Y_{\rm gc}, Z_{\rm gc}) = (-8.0, 0, 0)$ kpc. The cylindrical coordinate angle ϕ increases in the anticlockwise direction. The rotation of the Galaxy is clockwise in the $(X_{\rm gc}, Y_{\rm gc})$ plane. The heliocentric Cartesian frame is related to Galactocentric by $X_{\rm hc} = X_{\rm GC} + 8$, $Y_{\rm hc} = Y_{\rm GC}$, and $Z_{\rm hc} = Z_{\rm GC}$. $X_{\rm hc}$ is negative toward $\ell = 180^{\circ}$ and $Y_{\rm hc}$ is positive towards Galactic rotation. For transforming velocities between heliocentric and Galactocentric frames we use $(X_{\rm gc,\odot}, Y_{\rm gc,\odot}, Z_{\rm gc,\odot}) = (U_{\odot}, \Omega_{\odot} R_{\odot}, W_{\odot})$. Following Schönrich et al. (2010), we adopt $(U, W)_{\odot} = (11.1, 7.25)$ km s^{-1}, while for the azimuthal component we use the constraint of $\Omega_{\odot} = 30.24$ km s^{-1} kpc^{-1} which is set by the proper motion of Sgr A$*$, i.e. the Sun's angular velocity around the Galactic centre (Reid & Brunthaler 2004).

The structure of the paper is as follows: In Section 2.1 we briefly describe our Red Clump selection scheme, which includes using GALAXIA to calibrate de-reddened colours against spectroscopic parameters in order to select a pure Red Clump sample. This is used to derive distances in Section 2.2. Then in Section 3.1 we briefly discuss the observed and simulated datasets used in the paper. Our kinematic model and methods are described in Section 3.3. In Section 4.1 we test our model on a GALAXIA all-sky sample and identify high-mass Red Clump population as a contaminant. Next, in Section 4.2 we analyse observed data in the mid-plane and compare with the *APOGEE* result of B15. Then in Section 4.3 we extend the analysis to the offplane region and compare our results with selection function matched GALAXIA realizations before discussing our findings in Section 5.

2 SELECTING PURE RED CLUMP SAMPLE TO ESTIMATE DISTANCES

2.1 Red Clump calibration and selection

The Red Clump (RC) is a clustering of red giants on the Hertzsprung–Russell diagram (HRD), that have gone through helium flash and now are quietly fusing helium in the convective core. These stars on the helium-burning branch of the HRD have long been considered 'standard candles' for stellar distances as they have very similar core masses and luminosities (Cannon 1970). While many studies use the RC, there is considerable variation in the literature over calibration for the absolute magnitude of these stars. Studies of RC stars using parallaxes from *Hipparcos* have shown

[1]The local standard of rest (LSR) is defined as the frame of reference of a star at the Sun's location that is on a circular orbit in the Galactic gravitational potential. It is thus assumed that the LSR has no radial or vertical motion w.r.t. the Galactic centre, as suggested by the proper motion of Sgr A$*$ which shows that such motion is negligible (Schönrich, Binney & Dehnen 2010).

[2]GALAXIA is a stellar population synthesis code based on the Besancon Galactic model by Robin et al. (2003). GALAXIA uses its own 3D extinction scheme to specify the dust distribution and the isochrones to predict the stellar properties are from the Padova database (Marigo et al. 2008; Bertelli et al. 1994). Full documentation is available at http://galaxia.sourceforge.net/Galaxia3pub.html

that their average absolute magnitude in the K_s (hereafter K) band spans the range $-1.65 < M_{Ks} < -1.50$ (Girardi 2016) and there is ongoing effort to revise this using *Gaia* (Hawkins et al. 2017). The colour dependence of the Red Clump on population parameters is also well known, for example the $(J - Ks)$ colour is predicted to increase by 0.046 mag from ([Fe/H],[α/Fe]) = $(-0.30, +0.10)$ to (0.00, 0.00) (Nataf et al. 2016). Given this variation, in this work we choose not to assume single M_{Ks} value to estimate distances but instead derive an empirical relation between K-band absolute magnitude and metallicity [Fe/H]. We choose the K band for two reasons. While some passbands are more affected than others by metallicity variations within the RC population, such effects seem to be greatly reduced in the K band (Salaris & Girardi 2002). This, combined with the fact that the K band is least affected by extinction, makes it a reliable passband in which to derive fundamental properties of the RC population.

However, we will first need to select a reliable sample of Red Clump stars. Our selection function is based in terms of de-reddened colours $C_{JK} = (J - K)_0$ and the spectroscopic parameters: surface gravity $\log g$, metallicity [Fe/H], and effective temperature $T_{\rm eff}$ as described in Bovy et al. (2014). In the *APOGEE* Red Clump catalog (Bovy et al. 2014) the photometry is corrected for extinction using the Rayleigh Jeans Colour Excess method (RJCE; Majewski, Zasowski & Nidever 2011) which requires photometry in 2MASS and [4.5 μm] bands. However, it is difficult to get de-reddened colours accurately from photometry alone. So, to overcome this, we use pure Red Clump stars from GALAXIA to derive empirical relations expressing C_{JK} in terms of [Fe/H] and $T_{\rm eff}$. This allows us to derive de-reddened colours from spectroscopic parameters, which we can then use to select the Red Clump samples for any given spectroscopic sample. In particular, the GALAXIA Red Clump sample is also used to obtain the aforementioned M_{Ks}–[Fe/H] curve, which is used to estimate distances (see Tables A1-A3). The procedure above is described in full detail in Appendix A.

We now check the accuracy of our selection function in recovering the Red Clump stars. For this we compute precision (fraction of selected stars that are part of the Red Clump) and recall (fraction of Red Clump stars that are selected), which are two commonly used measures of accuracy in the field of information retrieval. We find that 97.6 per cent of our selected stars are part of the Red Clump. Since our selection is based on spectroscopic parameters that have uncertainties associated with them, we also explored the effects of adding Gaussian errors of $(\sigma_{\log T_{\rm eff}}, \sigma_{\rm [Fe/H]}, \sigma_{\log g}) = (0.011, 0.05, 0.1)$ dex. For $T_{\rm eff} = 4700$ K, the typical temperature of a Red Clump star, the uncertainty in temperature is 120 K. In spite of the uncertainties, the precision of the selected stars was found to be 83 per cent (see Table 1 for summary), however, the recall dropped to 69 per cent. If the uncertainty on temperature is reduced by a factor of two, the precision increases to 91 per cent and recall to 85 per cent, suggesting that it is important get precise and accurate temperatures.

2.2 Distances

We now proceed to estimate distances for our Red Clump sample. The distance modulus for a given passband λ corrected for extinction is given by

$$d_{\rm mod} = m_\lambda - M_\lambda - A_\lambda, \tag{1}$$

with apparent magnitude m_λ, absolute magnitude M_λ, and extinction A_λ. For the K band, M_λ is derived using our M_K–[Fe/H] relation (Table A3), while for the extinction we make use of the derived intrinsic colours C_{JK}([Fe/H], $T_{\rm eff}$)[3] i.e.

$$(J - K) - C_{JK} = (A_J - A_K), \tag{2}$$

and this can be related to the reddening $E(B - V)$ using $f_\lambda = A_\lambda)/E(B - V)$ as

$$(A_J - A_K) = (f_J - f_K)E(B - V). \tag{3}$$

After rearranging, we get the general relation

$$A_\lambda = f_\lambda \times \frac{(J - K) - C_{JK}}{f_J - f_K}, \tag{4}$$

with f_λ as in Schlegel, Finkbeiner & Davis (1998) (see Table A2).

To illustrate the accuracy achieved in estimating distances for the GALAXIA sample, we show the residuals in $d_{\rm mod}$ with respect to the true distance modulus in Fig. 1. The residuals lie close to zero, with a typical distance uncertainty of 4 per cent. There is also no significant bias as a function of metallicity [Fe/H]. If uncertainty in spectroscopic parameters are taken into account the dispersion in estimated distance modulus $\sigma_{\rm dmod}$ increases and this is shown in Table 1 for some typical cases. The main reason for the increase in $\sigma_{\rm dmod}$ is the contamination from stars that are not Red Clump, e.g. RGB stars, which can be understood from the quoted precisions in the table. The quantity in brackets denotes $\sigma_{\rm dmod}$ for the actual Red Clump stars, which is not significantly affected by addition of spectroscopic uncertainties.

[3] See equation (A8).

Table 1. Accuracy of the Red Clump selection function as predicted by all-sky $J < 15$ mock samples from GALAXIA. Shown are results for cases with: (1) no uncertainty on spectroscopic parameters; (2) with uncertainty typically expected from high resolution spectra, e.g. *APOGEE* (Holtzman et al. 2015); and (3) same as (2) but with smaller $\sigma_{\log T_{\rm eff}}$ which should be achievable with good quality spectra.

	$\sigma_{\log T_{\rm eff}}$ (dex)	$\sigma_{\rm [Fe/H]}$ (dex)	$\sigma_{\log g}$ (dex)	Recall[a] (per cent)	Precision[b] (per cent)	$\sigma_{\rm dmod}$[c] mag
1	0.0	0.0	0.0	100.0	97.9	0.10 (0.09)
2	0.011	0.05	0.1	77.0	93.6	0.16 (0.12)
3	0.0055	0.05	0.1	87.0	94.8	0.12 (0.11)

[a] We define Red Clump as those stars that satisfy equations (A1)–(A4). So, recall refers to fraction of total number of Red Clump stars that are selected.
[b] Here precision refers to fraction of selected stars that have initial stellar mass $> M_{\rm RGB,tip}$
[c] The quantity in brackets denotes $\sigma_{\rm dmod}$ for the actual Red Clump stars, which is not significantly affected by addition of spectroscopic uncertainties.

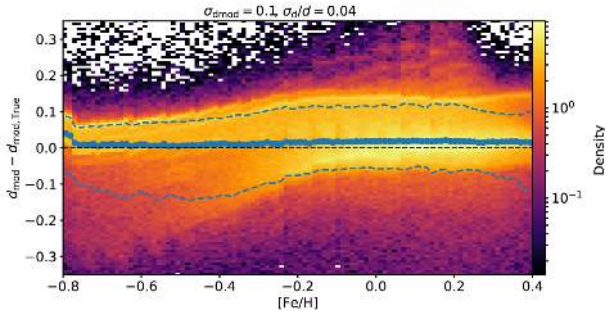


Figure 1. Distance accuracy using Red Clump calibration on GALAXIA: The residuals in the distance modulus are concentrated close to zero (black dotted line) and there is no significant bias with metallicity. The blue dotted lines indicate the 1σ bounds, with the overall distance error being 4 per cent.

3 DATA AND METHODS

3.1 Datasets

In this paper we make use of data from the *APOGEE* and *GALAH* surveys from which Red Clump stars are selected using the selection scheme described in Appendix A unless otherwise specified. Following is a brief overview of the datasets used for our analysis:

We downloaded the Red Clump catalog[4] of *APOGEE* DR12 (Bovy et al. 2014), in order to compare our results directly with B15. This dataset contains 19 937 stars and will be referred to as *ADR12RC*. Similarly we also obtained the latest available RC catalog[5] from *APOGEE* DR14 (Abolfathi et al. 2018). This contains 29 502 stars and will be referred to as *ADR14RC*. In both cases, while we do not apply our Red Clump selection method, we do estimate the distances using the scheme in Section 2.2. Our distances were found to be in excellent agreement with those in the *APOGEE* Red Clump catalog. In our analysis the distances are used to compute velocity maps, and we found that there was no difference between the velocity maps computed using either of the distances.

Where it appears the additional tag 'SF_Bovy' explicitly means that the dataset used has exact selection as in the *APOGEE* Red Clump catalogs.

Next, from the internal release of *GALAH* data up to 2017 October we preselect stars in the magnitude range $9 < V < 14$. The data include fields observed as part of the *K2-HERMES* (Wittenmyer et al. 2018) and *TESS-HERMES* (Sharma et al. 2018) programmes but not the fields observed as part of the pilot[6] survey. Also, data without a proper selection function (field id <-1) were excluded from the analysis. The spectroscopic parameters are from the same pipeline that was used in Sharma et al. (2018) and further details of spectroscopic analysis techniques used can be found there and in Duong et al. (2018). Details on reduction and estimation of radial velocity are in Kos et al. (2017). From the full data we select Red Clump stars using our scheme in Appendix A and obtain 33183 RC stars. This is merged with *ADR14RC* to form a combined observed dataset called *GADR14RC* and again where it appears, the additional tag 'SF_New' signifies that our selection method was employed. This combined set provides a more complete (x, y) spatial coverage as shown in Figs 2 and 3, where only for comparison we show 44166 Red Clump stars from *RAVE*-DR5 (Kunder et al. 2017) using our selection scheme. The combined dataset allows us to explore the region well beyond the Solar neighbourhood.

To examine the validity of this analysis, we use GALAXIA to simulate the selection functions of *APOGEE*[7] (Zasowski et al. 2013) and *GALAH* (Martell et al. 2017), and generate a combined Red Clump dataset using our selection schemes for direct comparison with *GADR14RC*. Finally, for Section 4.1 we also generate an all-sky mock Red Clump catalogue to test our kinematical models. All GALAXIA samples were generated with the 'warp' option turned off in order to allow easier interpretation of our experiments.

[4]APOGEE DR12-RC fits files.
[5]APOGEE DR14-RC fits files.
[6]Data collected before 2014 March i.e. with cob id<1403010000 is excluded, where cob id = date × 10 000 + run no.
[7]APOGEE DR14 fields.

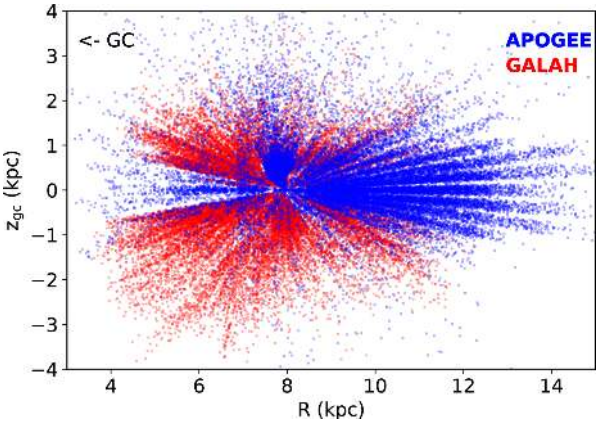


Figure 2. Distribution of the combined Red Clump dataset GADR14RC in Galactocentric $R-z$ plane. While the *APOGEE* coverage dominates in plane and towards the anticentre, *GALAH* surveys the off-plane region more extensively.

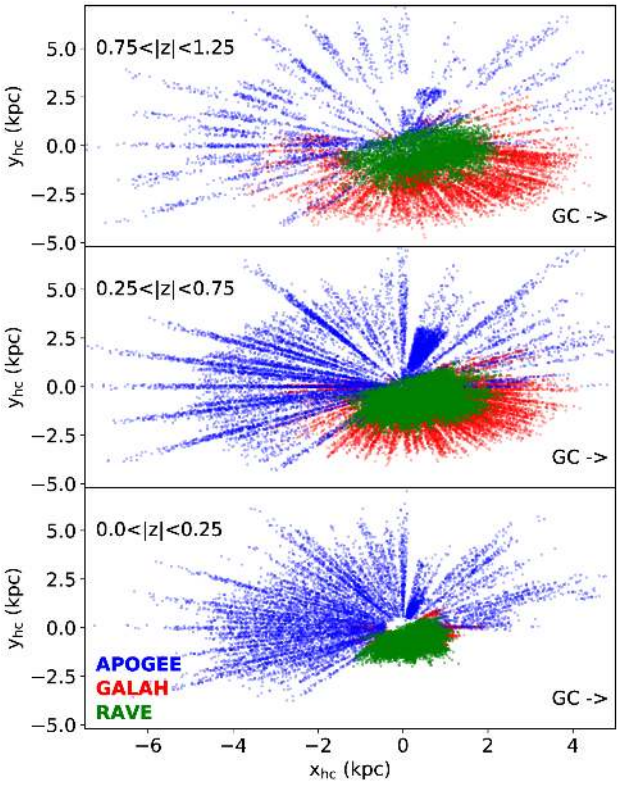


Figure 3. Distribution of GADR14RC in heliocentric (x, y) coordinates where the symmetrical regions above and below the plane have been merged together in slices in z (kpc). *APOGEE* probes deep into the disc while *GALAH* provides good coverage moving away from the plane, and to illustrate this also shown is the coverage of RAVE DR5.

3.2 Proper motions

In order to transform from the heliocentric to Galactocentric frame we require highly accurate proper motions. *Gaia* DR1 has provided high precision parallaxes for about 2 million objects and the DR2 (expected 2018 April) will extend this to nearly a billion objects and will also provide proper motions. In the meantime the two extensively used proper motion catalogues PPMXL (Roeser, Demleitner & Schilbach 2010) and UCAC4 (Zacharias et al. 2013)

have been improved using Gaia DR1 positions to produce UCAC5 (Zacharias, Finch & Frouard 2017) and HSOY (**H**ot **S**tuff for **O**ne **Y**ear, Altmann et al. 2017). Until Gaia DR2 these updated catalogues will provide proper motion with 1–5 mas yr^{-1} precision. For all our observed datasets, where available, we use the average of UCAC5 and HSOY values, and default (UCAC4) proper motions elsewhere. We have checked that this has no impact on our results. Moreover, our main analysis does not make use of proper motions.

3.3 Kinematic model

In this section we will describe the framework of our kinematical modelling. Our goal is to reproduce the observed line-of-sight velocity field ($V_{\rm los}$) using an axisymmetric Galactic model. In our scheme the Galactocentric velocity distribution, $V = (V_R, V_z, V_\phi)$, follows the triaxial Gaussian distribution

$$p(V|r,\tau) = \frac{1}{(2\pi^{3/2})\sigma_R\sigma_\phi\sigma_z}\exp\left\{\frac{V_R^2}{\sigma_R^2}+\frac{V_z^2}{\sigma_z^2}+\frac{(V_\phi-\overline{V}_\phi)^2}{\sigma_\phi^2}\right\}, \tag{5}$$

where we assume that $\overline{V}_R$ and $\overline{V}_z$ are negligible. The mean Galactocentric azimuthal velocity $\overline{V}_\phi$ can be written using Strömberg (1946) as

$$\overline{V}_\phi^2 = V_{\rm circ}^2(R,z) + \overbrace{\sigma_R^2\left(\frac{{\rm d}\ln\rho}{{\rm d}\ln {\rm R}}+\frac{{\rm d}\ln\sigma_{\rm R}^2}{{\rm d}\ln {\rm R}}+1-\frac{\sigma_\phi^2}{\sigma_R^2}+1-\frac{\sigma_z^2}{\sigma_R^2}\right)}^{V_{asym}}, \tag{6}$$

where $V_{\rm circ}$ is the Galactocentric circular velocity, and $V_{\rm asym}$ is the asymmetric drift. Assuming exponential density profiles for the Galactic disc ($\rho \propto {\rm e}^{-R/R_{\rm d}}$, Sharma & Bland-Hawthorn 2013) and velocity dispersion ($\sigma \propto {\rm e}^{\frac{R-R_\odot}{R_\sigma}}$) we get

$$\overline{V}_\phi^2 = V_{\rm circ}^2(R,z) + \sigma_R^2\left(-\frac{R}{R_{\rm d}}-\frac{2R}{R_\sigma}+1-\frac{\sigma_\phi^2}{\sigma_R^2}+1-\frac{\sigma_z^2}{\sigma_R^2}\right). \tag{7}$$

However, equation (7) is valid only for the case where the principle axis of the velocity ellipsoid is aligned with the spherical coordinate system (r, θ, ϕ) centred on the Galactic centre, i.e. $\overline{V_R V_z} = (V_R^2 - V_z^2)(z/R)$ (Binney & Tremaine 2008). There is however, evidence to suggest that the ellipsoid is aligned with the cylindrical system (R, ϕ, z) (e.g. Binney et al. 2014), in which case $\partial\overline{V_R V_z}/\partial z = 0$ and the term $1 - \frac{\sigma_z^2}{\sigma_R^2}$ drops out from equation (7). Since the actual answer probably lies in between the two alignments, we instead take into account the contribution of dispersion terms ($\sigma_{\phi,R,z}$) as a new parameter $c_{\rm ad}$,

$$\overline{V}_\phi^2 = V_{\rm circ}^2(R,z) + \sigma_R^2\left(-\frac{R}{R_{\rm d}}-\frac{2R}{R_\sigma}+c_{\rm ad}\right). \tag{8}$$

Using the above framework we can now describe the individual models employed:

(i) *Bovy1*: The model used by B15 is derived in Bovy et al. (2012, B12 hereafter). Essentially they assume $\partial\overline{V_R V_z}/\partial z = 0$, exponential surface density profile, exponential velocity dispersion profile, and a constant circular velocity, and then use the distribution function from Dehnen (1999) to model the asymmetric drift. Sharma et al. (2014) fitted the B12 model to *RAVE* data and showed that the B12 model can be approximated by setting $c_{\rm ad} = 0.28$ in equation (8). In order to reproduce the results of B15 we adopt this value for $c_{\rm ad}$. Furthermore, in accordance with B15, we set $R_d = 3$ kpc, $R_\sigma = \infty$ kpc, $\sigma_R = 31.4$, and assume a flat profile for the circular velocity $V_{\rm circ} = 220\,{\rm km\,s^{-1}}$, $V_{\rm circ} + V_\odot = 242.5\,{\rm km\,s^{-1}}$. We use *Bovy1* only for the mid-plane ($|z| < 0.25$ kpc) as was the case in B15.

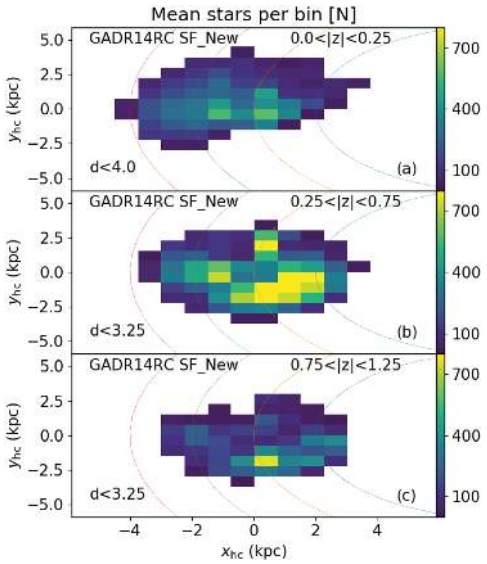


Figure 4. Stellar number density of the *GADR14RC* dataset for the three vertical slices used in our analysis. Each pixel has a minimum of 20 stars. The radial cuts applied here correspond to completeness in magnitude limited selection (see Fig. 11 for details).

(ii) *globalRz*: B15 model requires making a number of assumptions, e.g. about the circular velocity profile, the σ_ϕ/σ_R ratio as well as the $\sigma_{\rm R}$ profile. Typically, $\sigma_{\rm R}$ in the disc lies around 20–40 km s^{-1} (Bovy et al. 2012), however the vertical variation in dispersion requires proper modelling of the AVR and hence a good handle on stellar ages. Moreover, $V_{\rm circ}$ itself has a non-trivial profile, as for example was found with kinematic analysis of *RAVE* where gradient in both radial ($\propto \alpha_R(R - R_\odot)$) and vertical directions ($\propto \alpha_z|z|^{-1.34}$) was reported (Sharma et al. 2014). Furthermore, if we compute $\overline{V}_\phi$ using proper motions, we see that the profiles are not flat in R (Fig. 4).

Given that some of the assumptions might not be correct, for our analysis we adopt a flexible model for $\overline{V}_\phi$, that is a second degree multivariate polynomial in cylindrical Galactocentric coordinates R and z, more specifically,

$$\overline{V}_\phi = \sum_{i=0}^{2}\sum_{j=0}^{2} a_{ij}(R-R_\odot)^i z^j. \tag{9}$$

The model prediction in Galactocentric coordinates can be transformed to heliocentric coordinates assuming $(U_\odot, \Omega_\odot R_\odot, W_\odot)$ for the solar motion and fitted to observed the line-of-sight velocity $V_{\rm los,mod}$. The $\Omega_\odot$ is given by the proper motion of Sgr A*, and hence this approach does not require us to assume a value for $V_\odot$ or $V_{\rm circ}$. In order to fit for the coefficients a_{ij}, we assume that the observed $V_{\rm los}$ is a Gaussian, $\mathcal{N}(.||{\rm mean, dispersion})$, centred at $V_{\rm los,mod}$ with dispersion $\sigma_{\rm los} = 31.4\,{\rm km\,s^{-1}}$ (similar to B15). This can be summarized as

$$p(V_{\rm los}|a_{ij}, l_{\rm gc}, z_{\rm gc}, R) = \mathcal{N}(V_{\rm los}|V_{\rm los,mod}, \sigma_v), \tag{10}$$

and we call this model *globalRz*. The *MCMC* fitting is carried out using the *bmcmc* package (Sharma 2017).

Table 2. Parameters to model the velocity dispersion σ_R as a function of height $|z|$.

$\|z\|$	0.0	0.5	1.0	2.0
δ_{σ_0}	0.0	5.0	7.0	7.0

(iii) *Strom_z*: Finally, we will now describe the model for our GALAXIA simulations. While we could just use the *globalRz* model to approximate kinematics in GALAXIA, however, flexible models like *globalRz* with many free parameters run the risk of overfitting the data. Hence we devise a more realistic model. Note, our aim here is to generate a simple and realistic null hypothesis case, i.e. a smooth axisymmetric model that has no velocity fluctuations. The default model in GALAXIA is based on the Strömberg relation with parameters from the RAVE-GAU kinematic model from Sharma et al. (2014) (S14; their last column of table 6). This model is able to describe the z variation in velocity dispersions (≈AVR), but it requires stellar ages as input. Since for observed data, ages are not available, instead of using the default model, we modify it take the variation of σ_R with height z into account. For this we adopt the following form for σ_R,

$$\sigma_R(R, z) = (\sigma_0 + \delta_{\sigma_0}(|z|)) \exp(-R/R_\sigma). \quad (11)$$

and fit for σ_0 and δ_{σ_0} using a mock GALAXIA realization with RAVE-GAU model. We find $\sigma_0 = 30.7\,\mathrm{km\,s^{-1}}$ and δ_{σ_0} for three different values of $|z|$ is given in table 2. To obtain δ_{σ_0} for any arbitrary value of z we use linear interpolation. For the thick disc we assume a mono-age population (11 Gyr old) and we assume that the thick disc obeys the AVR of the thin disc. If this is not done the velocity distribution in the upper slices might deviate strongly from a Gaussian distribution and this will lead to velocity fluctuations and will make the simulation unsuitable for out null hypothesis test.

Finally, we fit equation (8) to the mock GALAXIA realization and find the best match for $c_{\rm ad} = 0.77$. The disc and velocity dispersion scale lengths are adopted directly from S14, i.e. $(R_d, R_\sigma) = (2.5, 13.7)$ kpc. Lastly, in GALAXIA we use $V_{\rm circ} + V_\odot = 226.84 + 12.1\,\mathrm{km\,s^{-1}}$ and the circular velocity profile is from Sharma et al. (2011) and is not flat.

3.4 Fourier analysis of velocity fluctuations

Each dataset is divided into three slices in z (as in Fig. 3), and further binned into (x, y) space with bins of size 0.75×0.75 kpc^2. The resulting stellar density map for each vertical slice of the *GADR14RC* dataset is shown in Fig. 5. For each bin, we calculate the residual $\Delta V_{\rm los} = V_{\rm los} - V_{\rm los,mod}$, to produce a 2D velocity fluctuation image h. To reduce the contribution from Poisson noise we set $h = 0$ for bins than have less than 20 stars. Next we perform Fourier analysis of the image h and calculate the 2D power spectrum of fluctuations as

$$P_{kl} = \frac{1}{N_{\rm eff}} |A_{\rm kl}|^2 \Delta x \Delta y, \quad (12)$$

where $A_{\rm kl}$ is the 2d Fast Fourier Transform of the image h and Δx and Δy are the size of the bins along x and y directions. $N_{\rm eff}$ is the effective number of bins in the image and is given by $\sum_i \sum_j H(n_{ij} - 20)$, where H is the Heaviside step function and n_{ij} is the number of stars in the (i, j)-th bin. Next we average P_{kl} azimuthally in bins of $k = \sqrt{k_x^2 + k_y^2}$ to obtain the 1D power spectrum $P(k)$. The $P(k)$ as defined above satisfies the following normalization condition given by the *Parseval's* theorem,

$$\int_0^\infty P(k) 2\pi k \, \mathrm{d}k = \sum_k \sum_l P_{kl} \Delta k_x \Delta k_y = \frac{\sum_i \sum_j H(n_{ij} - 20) h_{ij}^2}{N_{\rm eff}}. \quad (13)$$

We present $\sqrt{P(k)}$ that has the dimensions of $\mathrm{km\,s^{-1}}$ as our final result. The presented formalism to compute the power spectrum is slightly different that of B15, but it matches the results of B15 and importantly ensures that the estimated power spectrum $P(k)$ is invariant to changes in size of the bin, the overall size of the image box, and bins with missing data. The noise for the power spectrum is calculated in the same manner except that for the input signal we use normally distributed data with zero mean and dispersion equal to the standard deviation of $\Delta V_{\rm los}$.

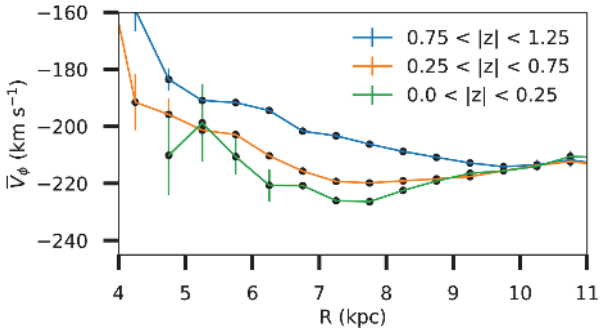


Figure 5. Mean Galactocentric rotation ($\overline{V}_\phi$) derived using proper motions for the combined dataset GADR14RC and shown for different z slices. The profiles look parabolic in nature with a steepening gradient as we move away from the plane.

4 RESULTS

The observed data have complicated selection functions in terms of magnitude and spatial coverage. Therefore before we study the observed data, we will first consider a much simpler dataset using GALAXIA that has uniform spatial coverage. This will allow us to test the method described in Section 3.3 and explore any selection function related biases.

4.1 GALAXIA all-sky sample: high-mass Red Clump stars

Using GALAXIA we generate an all-sky sample that has $H < 13.8$, the magnitude boundary of the *APOGEE* data set in the mid-plane, and select Red Clump stars using the scheme in Appendix A. We make three versions of this dataset, one with true distances ($d = d_{\rm True}$), one with Red Clump-derived distances ($d = d_{\rm RC}$), and one with Red Clump-derived distances but only for stars with $d_{\rm RC}/d_{\rm True} > 0.8$. The last of these is chosen to provide a control sample to check for systematic errors in distances. For each data set we fit the *globalRz* kinematic model to $V_{\rm los}$ data and derive the $\overline{V}_\phi$ profile and then construct the $\Delta V_{\rm los}$ map (Section 3.3). In Fig. 6, we only show results for the mid-plane region with $|z| < 0.25$ kpc. The panels in first column compare the derived $\overline{V}_\phi$ profile with the actual $\overline{V}_\phi$ profile, computed directly using line-of-sight motion, proper motions and true distances. The $\overline{V}_\phi$ profile computed using Red Clump distance is also shown alongside. The panels in second column show the map of velocity fluctuations $\Delta V_{\rm los}$, while their power spectrum is shown in panels of the third column. The median power spectrum expected

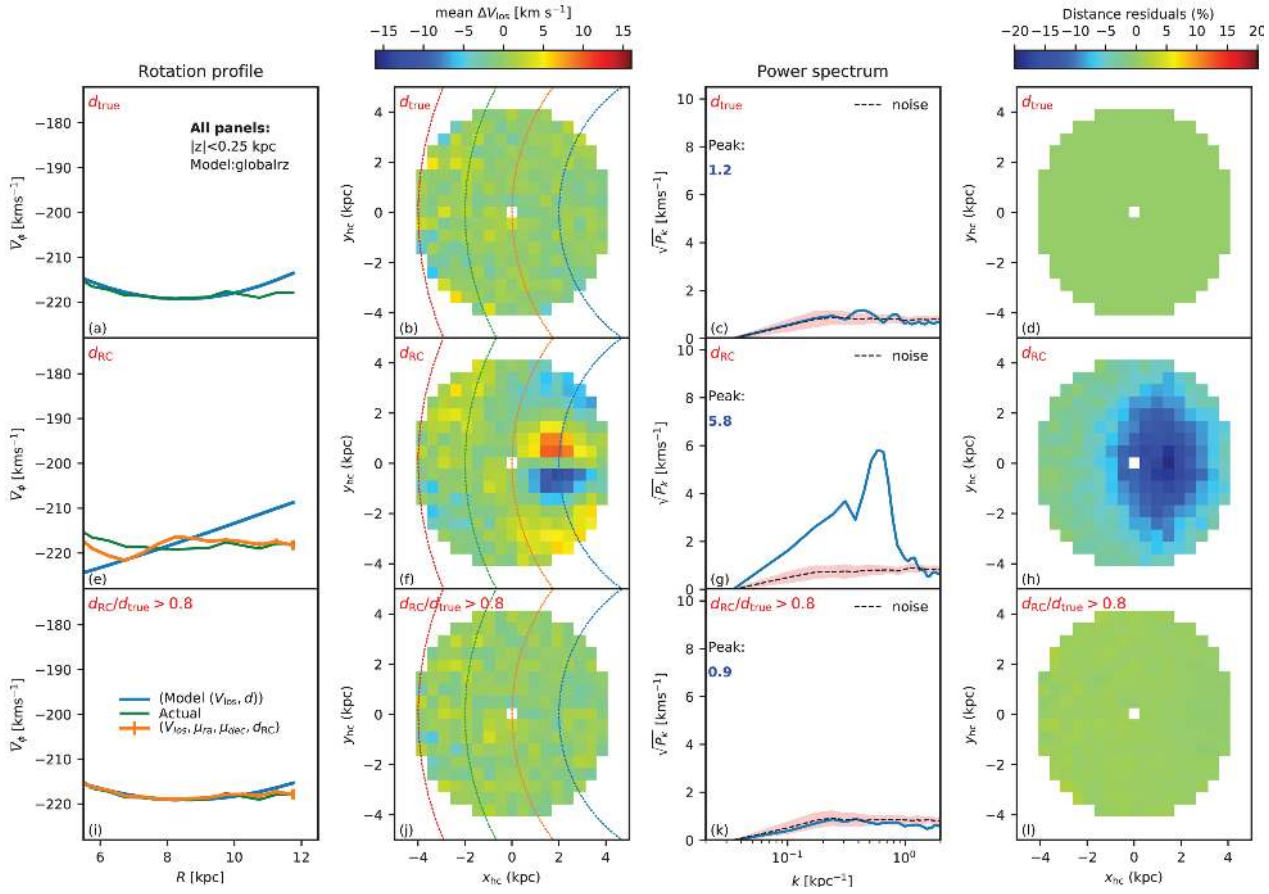


Figure 6. Kinematic modelling of GALAXIA all-sky sample. Results with three different choices of distance are shown, true distances (top panels), Red Clump distances (middle panels), and Red Clump distances but with stars restricted to $d_{RC}/d_{True} > 0.8$ (bottom panels). Each column shows a different aspect of the kinematics. (a, e, f): The $\overline{V}_\phi$ as function of R obtained by fitting the *globalRz* model to V_{los}. The actual $\overline{V}_\phi$ profile and the profile obtained with proper motion and $d = d_{RC}$ is also plotted alongside. (b, f, j) The line-of-sight residual velocity map obtained after subtracting the best-fitting *globalRz* model (also overplotted are curves of R = [6, 8, 10, 12] increasing towards negative X_{hc}). (c, g, k) Power spectrum of the residual velocity map. (d, h, l) Map of distance residuals computed with respect to d_{True}.

due to Poisson noise is shown in dotted black and 68 percentile spread around it based on 20 random realizations is shown in pink. Finally, in the fourth column we show the map of distance residuals. The results for each case are summarized below.

(i) True distances $d = d_{True}$: It is clear that for the true distances we are able to recover the profile by fitting *globalRz* model to V_{los}. This is also reflected in the map of ΔV_{los}, where we obtain a smooth map with negligible residuals. Furthermore, the 1D power spectrum also has amplitude consistent with noise of about 2 km s^{-1}. This scenario is as would be expected of a perfectly axisymmetric galaxy.

(ii) RC distances $d = d_{RC}$: The results are more interesting for the Red Clump derived distances case. Here, the actual $\overline{V}_\phi$ profile (green line) is not reproduced accurately by the *globalRz* model (blue line) unlike the previous case. The model overestimates the profile beyond the Solar circle ($R_\odot$) and underestimates it towards the Galactic centre. The $\overline{V}_\phi$ profile computed using proper motions also does not match the actual profile. The ΔV_{los} residual map shows a peculiar dipole along the y axis for $x > 0$. This feature gives rise to a sharp peak in the power spectrum with amplitude of 5.9 km s^{-1} at a physical scale of $k^{-1} = 1.6$ kpc. Exactly at the location where we see high residuals in V_{los} we also see high residual in distances.

(iii) RC distances but only for $d_{RC}/d_{True} > 0.8$: The results of this case are very similar to that for case where we use true distances.

For the first case with true distances the residuals in both V_{los} and distance are zero by definition. For the second case with RC distances, we see significant residuals. It is clear that the region corresponding to the high ΔV_{los} residual also corresponds to high distance residual, i.e. distance errors. This suggests the cause of high residuals is systematic errors in distances. This is further confirmed by the results of the third case, where we restrict the analysis to stars with $d_{RC}/d_{True} > 0.8$ and find no residuals in V_{los} or distances.

For the second case, the distance residuals are negative which means that the distances are underestimated. This would have the effect of bringing stars closer to us than in reality, more importantly, their kinematics would be inappropriate for their inferred location. This is why we see a dipole in the V_{los} maps. Since the velocity field is incorrect, the best-fitting *globalRz* model fails to reproduce the actual $\overline{V}_\phi$ profile. Due to systematics in distances the velocity profile inferred using proper motion would also be wrong, and this is the reason for the mismatch of the orange line with the green line in Fig. 6(e). Again this is confirmed in Fig. 6(i), where we restrict

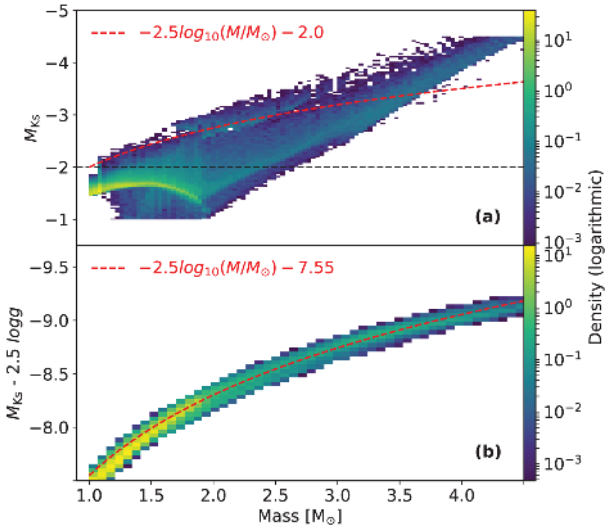


Figure 7. Mass distribution of Red Clump stars for a $H < 13.8$ sample simulated with GALAXIA. Panel (a) shows the distribution of Red Clump stars in the (M_{K_s}, Mass) plane. It is clear that the luminous Red Clump stars also have higher mass. Stars above the black dashed roughly correspond to where $d_{\rm RC}/d_{\rm True} > 0.8$. Panel (b) shows the distribution of Red Clump stars in the ($M_{K_s} - 2.5 \log g$, Mass) plane. The tight relation is because the Red Clump stars lie in a narrow range of $T_{\rm eff}$. The red dotted line is $=-2.5\log M$ in both the panels.

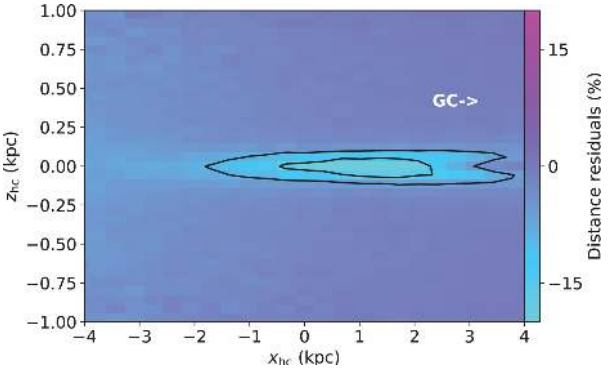


Figure 8. Properties of Red Clump stars for a $H < 13.8$ sample simulated with GALAXIA. We show the map of mean distance residuals ($d_{\rm RC} - d_{\rm True})/d_{\rm True}$. It can be seen that close to the plane and towards the Galactic Centre, the distance is underestimated. The contours indicate the fraction of Red Clump stars (10 per cent and 40 per cent levels shown) that are unusually luminous, identified using $M_{\rm Ks} < -2.0$. This suggests that distance errors are due to luminous Red Clump stars.

stars to $d_{\rm RC}/d_{\rm True} > 0.8$ and there is no mismatch between any of the $\overline{V}_\phi$ profiles.

We now investigate the cause of systematic errors in distances of Red Clump stars. We generate an all sky $H < 13.8$ sample with GALAXIA, identify Red Clump stars in it, and then study their properties. Fig. 7(a) shows the distribution of Red Clump stars in the plane of M_{Ks} and stellar mass M. Typically, Red Clump stars have $M_{Ks} \approx -1.60$, however Fig. 7(a) shows that there is a tail extending down to much brighter magnitudes. Stars with $d_{\rm RC}/d_{\rm True} < 0.8$ that were responsible for strange features in residual velocity maps in Fig. 6 correspond to $M_{Ks} < -2$ and this is shown as the black dashed line in the panel. In the tail below the line, brightness is strongly correlated with stellar mass, which extends up to 4 $M_\odot$. We know that mass of a red giant star is anticorrelated with age (e.g. Sharma, Stello & Bland-Hawthorn 2016; Miglio et al. 2017), with massive stars being in general younger. So the cause for the systematic errors in the Red Clump distances is the presence of young Red Clump stars that have high mass and luminosity.

The anticorrelation of absolute magnitude with mass is easy to understand. Red Clump stars lie in a narrow range of $T_{\rm eff}$. Hence their luminosity L is proportional to R^2. Given that surface gravity $g = M/R^2$, and since M_K represents the luminosity L well, we have

$$M_K \propto -2.5 \log L \propto -2.5(\log M - \log g) \quad (14)$$

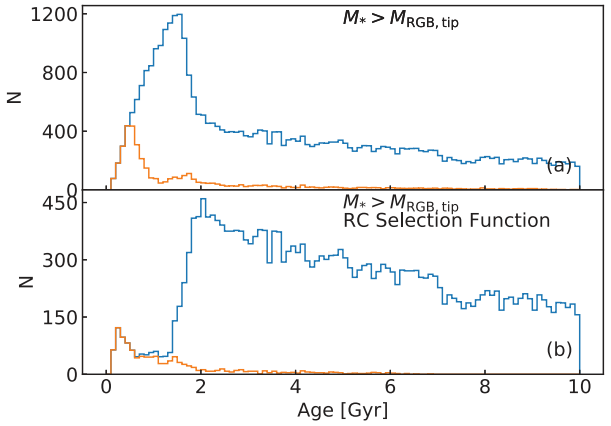
$$M_K - 2.5 \log g \propto -2.5 \log M. \quad (15)$$

For a given $\log g$, the magnitude decreases with mass and the expected trend is shown in Fig. 7(a). For Red Clump stars $\log g$ is not constant, to take this into account in Fig. 7(b), we show stars in the ($M_K - 2.5\log g$, Mass) space. The stars now perfectly follow the predicted relation of equation (15).

We now investigate as to where we expect to find such high-mass stars and in which regions do we expect significant errors in distances. Fig. 8 shows the map of distance residual in the (x, z) plane. We see that the distance residuals are high in the mid-plane of the Galaxy and towards the Galactic Centre. The contours overplotted on Fig. 8, show the fraction of Red Clump stars that have $M_{\rm Ks} < -2$, i.e. very luminous. Close to the plane and towards the Galactic Centre in certain areas the fraction is higher than 0.3. The regions of high distance residuals correspond to region with higher fraction of high-mass Red Clump stars, this provides a causal link for the high distance residuals.

Why is the contamination from young, high-mass RC so prominent close to the plane and towards the Galactic Centre? This is due to a combination of four different effects. First, due to the age scale height relation in the Galaxy, younger stars have smaller scale height and are closer to the plane. Secondly, the surface density profile of stars in the Galaxy falls off exponentially with distance from the Galactic Centre, which means there are more such stars towards the Galactic Centre. Thirdly, along any given line of sight the volume of a cone around it increases as square of the distance. So more stars from far away with larger true distances are displaced to regions with smaller apparent distances. Finally, the spectroscopic selection function designed to select RC stars also plays a role in making the high mass stars appear more prominently. For constant star formation rate the number of Red Clump stars shows a sharp peak around an age of 1.5 Gyr (Girardi 2016). But our contaminant bright stars having $M_{Ks} < -2$, peak at 0.5 Gyr and are not associated with the peak at 1.5 Gyr. The age distribution of RC stars in GALAXIA is shown in Fig. 9(a), also shown are the contaminant bright stars. Fig. 9(b) shows the age distribution after applying our RC selection function. The peak at 1.5 Gyr vanishes but not the one at 0.5 Gyr. It is clear that the selection function introduces a strong age bias rejecting a significant fraction of young stars, but the young contaminant bright stars are not rejected, instead they become more prominent.

We also studied the off plane slices and found no peculiar features in the residual velocity maps. This is expected as the contamination from high-mass RC stars does not extend far away from the mid-plane.



[h]

Figure 9. Age distribution of GALAXIA RC stars with $(1.8 < \log g < 3.0)$ and $4300 < T_{\rm eff}/{\rm K} < 5200$. The RC stars are defined as stars having mass greater than the RGB tipping mass. The star formation is as in GALAXIA, which is almost a constant star formation rate. (a) The RC sample is selected based on simple cuts in spectroscopic parameters ($\log g$,$T_{\rm eff}$) as defined above. The overall age distribution (blue curve) has a peak around 1.5 Gyr and this is dominated by stars with $M_K > -2.0$, while the brighter stars (orange curve) with $M_K < -2.0$ peak around 0.5 Gyr i.e. are much younger. (b) The RC selection scheme as described in this paper is based on (Bovy et al. 2014) (B14) and shown in equations (A1)–(A4) is now applied to sample in (a). This removes contamination from secondary clump stars (SRC) as well as the RGB bump and in the remaining RC sample, the younger population stands out even more clearly with the majority being Age < 1 Gyr old.

4.2 Velocity fluctuations in the mid-plane for observed data

We now discuss the results of our kinematic modelling on the observed datasets and will compare this with selection function matched mock data generated with GALAXIA as described in Section 3.1. Using Red Clump stars from *APOGEE* -DR12, B15 showed that after subtracting an axisymmetric model there remains a high residual in the $V_{\rm los}$ field in the mid-plane ($|z| < 0.25$ kpc). Their kinematical model assumed a flat rotation curve with $V_{\rm circ} = 220\,{\rm km\,s^{-1}}$ and $V_\odot = 22.5\,{\rm km\,s^{-1}}$ and the asymmetric drift was based on the Dehnen distribution (Dehnen 1999). In Fig. 10 we consider again the B15 result and explore effects that can lead to enhanced residuals. In Figs 10(a) and (b) we have reproduced their result by using the same model and data (*APOGEE* -DR12 RC sample) as them. A sharp peak of $10.4\,{\rm km\,s^{-1}}$ is obtained at a physical scale of about 2.5 kpc similar to B15.

The location and the height of the peak is essentially unchanged when we include *APOGEE* -DR14 RC sample, the peak only becomes sharper (Figs 10c and d).

Now, B15 used median statistics to compute the residual maps and power spectra. If the distribution of the residual velocity is a Gaussian then employing either mean or median statistics should not make much of a difference in the residual maps. However, if the distribution is asymmetric then it will. In the context of the Galaxy, we know that the V_ϕ distribution is asymmetric (Sharma et al. 2014). Typically one defines a kinematic model and then computes the model parameters that maximize the likelihood of the model given the data. For such a best-fitting model, it is not clear as to which statistics (mean or median) will give lower values in velocity residual maps. In Figs 10(e) and (f) we find that choosing mean statistics lowers the power by $1.0\,{\rm km\,s^{-1}}$ for the B15 model. We have checked and found that for our best-fitting *globalRz* model the results remain unchanged for either choice of statistic. So from now on for the rest of our analysis we adopt to use the mean statistics for computing the velocity residual maps. Next, we consider the volume completeness of the data sample. Fig. 11 shows the magnitude distribution of the *GALAH* and *APOGEE* Red Clump stars (*GADR14RC* dataset) in V and H passbands. In the mid-plane region most of the data are from *APOGEE* and there is a sharp fall around $H = 12$. Similarly, *GALAH* contributes significantly to the off-plane slices and the distribution falls off around $V = 14$, reflecting the survey selection function. This fall-off limit ($m_{\lambda,\rm max}$) is the faintest magnitude to which stars are observed completely (strictly speaking we mean pseudo-random-complete or unbiased in distance selection) and so we can also estimate the maximum distance this would correspond to by modifying equation (1) as

$$d_{\rm mod,max} \leq m_{\lambda,\rm max} - M_\lambda - \sigma_{M_\lambda} - A_\lambda. \quad (16)$$

Using magnitude limits for each slice, extinction factor A_λ, absolute magnitude M_λ, and its dispersion σ_{M_λ} from Table A2, we find $d_{\rm max} = 4$ kpc for the mid-plane and $d_{\rm max} = 3.25$ kpc for the off-plane regions. These distance limits are also visible in the scatter plots of Fig. 3. In Fig. 10(g) and (h) we apply the $d < 4$ kpc distance cut, which removes the high-residual pixels (beyond $x_{\rm hc} > 5$ kpc) however, there is no noticeable change in the power spectrum compared to Figs 10(g) and (h) as the amplitude is still at $9.3\,{\rm km\,s^{-1}}$. However, as a precaution, we will continue with the distance limits for the rest of the figures.

Finally, we replace the B15 model with our flexible axisymmetric model from Section 3.3 and this has the effect of further reducing the power to $7.1\,{\rm km\,s^{-1}}$ in Fig. 10(i) and (j). In Fig. 10(k) and (l) we consider the residuals for the combined dataset *GADR14RC* to increase the sample size and get essentially the same power spectrum as in Fig. 10(i) and (j) with lower amplitude of $6.3\,{\rm km\,s^{-1}}$. A characteristic pattern of blue in first quadrant, red in second, and yellow in third as seen in previous cases is also visible here. To conclude, we find that in the mid-plane after accounting for various systematics and a more flexible model the power amplitude can be reduced significantly, though interestingly it cannot be reduced to zero or to the level expected purely due to noise (pink region).

4.3 Off-plane slices and comparison with GALAXIA

We now also consider the off-plane ($|z| > 0.25$ kpc) slices of data and also compare directly with mock realizations using GALAXIA. Once again, we use the *GADR14RC* dataset and the flexible *globalRz* model. In Fig. 12, we show the residual velocity maps, power spectra as well as the $\overline{V}_\phi$ profile for each slice. To take the volume completeness of the sample into account, for the mid-plane slice we have restricted the data to $d < 4.0$ kpc and for the off-plane slices to $d < 3.25$ kpc.

As mentioned already in Section 4.2, the peak power in the mid-plane is around $6\,{\rm km\,s^{-1}}$ but moving away from the plane, the power drops (blue solid lines) and is only slightly higher than that expected from noise (dashed lines and the pink zone). Interestingly, the mock GALAXIA samples also predict this trend of high power in the mid-plane but power that is lower and only slightly higher than noise elsewhere. Note, the predicted power spectrum has intrinsic stochasticity due to Poisson noise. So we generate 100 random realizations of the GALAXIA samples and show the predicted 68 per cent confidence zone as the green shaded region. From these zones it is clear that, for GALAXIA samples, the maximum power achieved in the mid-plane is $5.2^{+1.2}_{-1.4}\,{\rm km\,s^{-1}}$. For other slices, for GALAXIA samples, the green and pink zones are almost on top of each other.

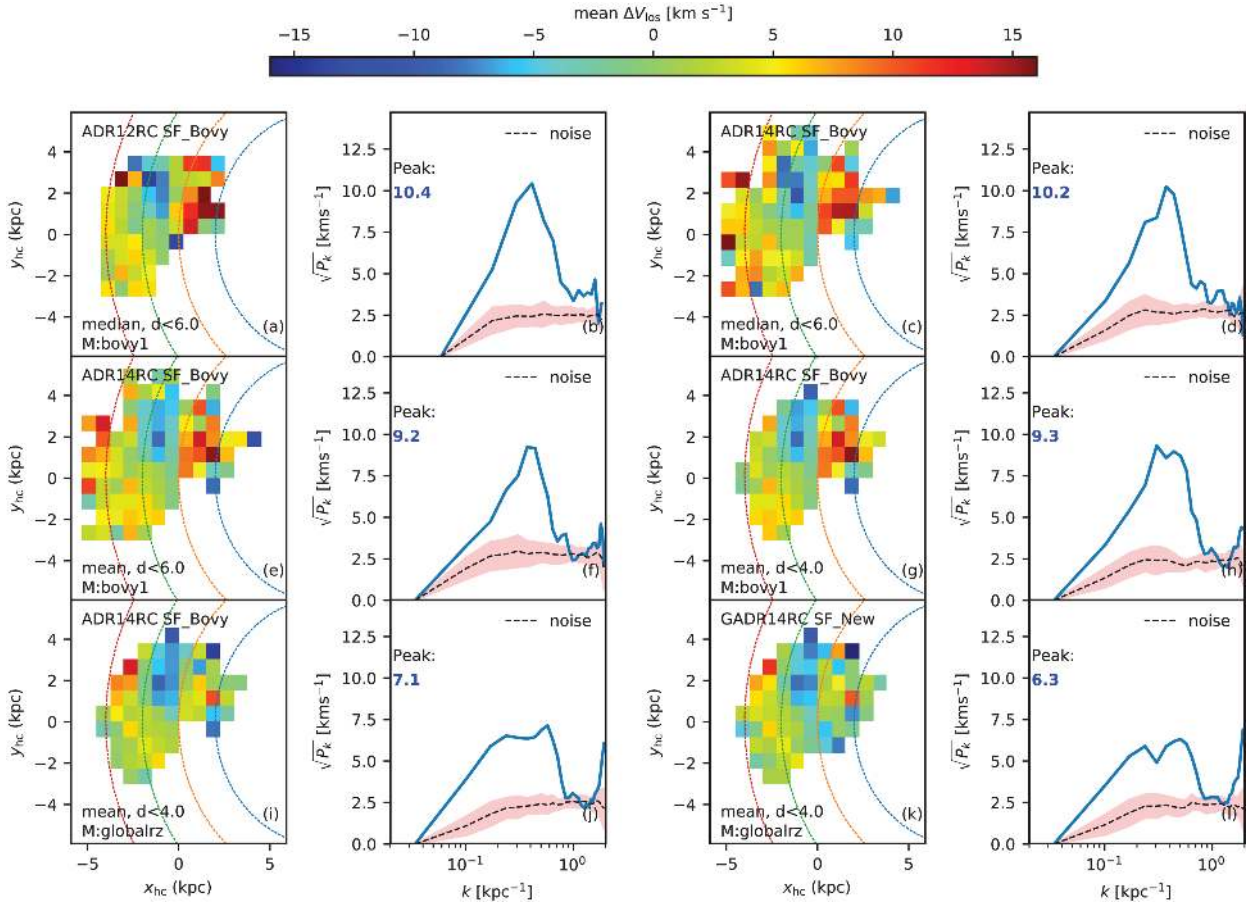


Figure 10. Residual velocity maps and power spectrum for the observed data in the $|z| < 0.25$ kpc slice. Shown are cases for different data sets, with different radial cuts and kinematic models to illustrate the effect of systematics. (a and b) Data used are *APOGEE*_RC_DR12 with stars restricted to radial distance $d < 6$ kpc, using median statistics to compute the residual and using Bovy's analytical model for the kinematics. (c and d) Same as a and b but with *APOGEE*_RC_DR14. (e and f) Same as c and d but now residual is computed using mean statistics. (g and h) Same as e and f but data restricted to $d < 4$ kpc to satisfy volume completeness. (i and j) Same as g and h but we now apply the flexible 2d polynomial kinematic model named as *globalRz* to show it reduces power. (k and l) We apply *globalRz* model to the combined *APOGEE*_DR14 and *GALAH* dataset that uses the new Red Clump selection criteria and distance estimation scheme described in the paper.

However, the maximum power in observed data sets is higher by about $2\,\mathrm{km\,s^{-1}}$ as compared to GALAXIA samples.

We note that for the observed data and the $0.75 < |z| < 1.25$ slice the $\overline{V}_\phi$ profile obtained using only line-of-sight motion traces well the $\overline{V}_\phi$ profile obtained using both line-of-sight and proper motions. This suggests that, for this slice, there is minimal systematic error associated with distance, proper motion or line-of-sight velocities. However, for the other two slices which are closer to the plane we do see differences. The slice closest to the plane shows most pronounced deviations. The mock GALAXIA samples also show similar behaviour. This is most likely due to systematic errors in distances as discussed in Section 4.1. If there are systematic errors with distances then its effect on the inferred $\overline{V}_\phi$ profile will be different depending upon if we infer the profile based on line-of-sight velocities or both line-of-sight velocities and proper motions.

The shape of the rotation profiles for the mock and observed data sets also show differences. For the mock data, the $\overline{V}_\phi$ profile is predominantly flat across all the slices. In contrast, for the observed data a clear variation with R is visible, and the variation becomes more pronounced as we move further away from the mid-plane. While our model is flexible enough to account for simple radial trends in rotation curves, this flexibility can over fit the data if the spatial coverage is not uniform. This is particularly a concern in the mid-plane where the coverage in the (x, y) plane is not uniform, as there is a dearth of stars in the fourth quadrant. This is because both APOGEEE and GALAH have not observed enough stars in the mid-plane and in the Southern Sky.

Basically the constraints on $\overline{V}_\phi$ for $R < 8$ kpc come from data in the first and the fourth quadrant. As evidenced by the red and blue patches in Fig. 6(b), the systematics in distances lead to incorrect values for the mean $V_{\rm los}$ in the first and the fourth quadrant. If data from only one quadrant is available the model can adjust the value of V_ϕ for $R < 8$ kpc to fit the $V_{\rm los}$ in that quadrant perfectly, however this will not match the mean $V_{\rm los}$ in the other quadrant. If the data from the other quadrant was also available the model would not have the freedom to do this, but in the absence of it the model over fits the data.

The GALAXIA samples are generated from a simulation for which the kinematics are known by design, so we can avoid over fitting a model which is similar to the input model. The input model has kinematics as a function of age, but since we do not have ages in the observed data, we approximate the kinematics by *Strom_z* model which is based on the Strömberg equation and described in Section 3.3. In Fig. 13, we employ this new fitting model *Strom_z*

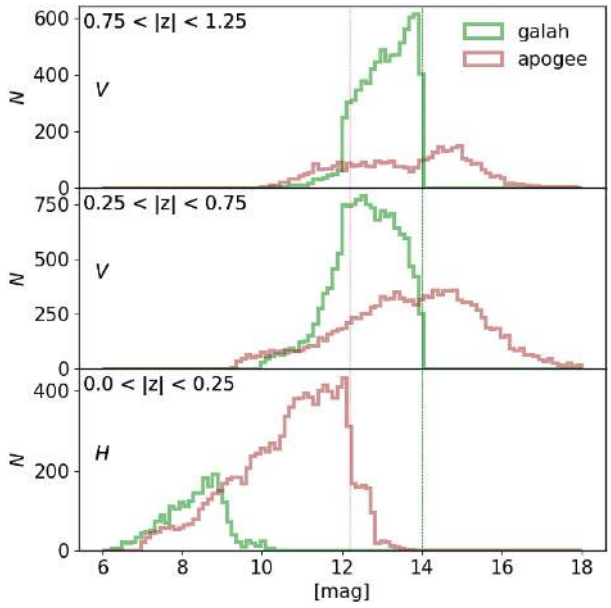


Figure 11. Magnitude distribution for *GALAH* and *APOGEE* Red Clump datasets shown for different slices in z (kpc). The position where the magnitude distribution falls sharply sets the maximum distance up to which the stars are unbiased in distance (pseudo volume complete). The magnitude limits are shown by dotted vertical lines. For the off-plane slices the completeness limit is $V = 14$ and set by the *GALAH* survey that dominates the number counts. For the mid-plane slice the limit is $H = 12$ and set by *APOGEE* survey that dominates the number counts.

for GALAXIA and compare its results with that of *globalRz* model fitted to the observed data. Overall the trends in velocity maps and the power spectrum for the different slices are the same as in Fig. 12, i.e. high power in the mid-plane and negligible power away from the plane. The characteristic pattern of red in first quadrant and yellow in third as seen in observed data for the mid-plane slice is also reproduced in the mid-plane slice of the simulated data. For the GALAXIA samples there is a slight increase in the power by about 3 km s^{-1} for the *Strom_z* as compared to *globalRz*. This is not surprising, as the *globalRz* model is more flexible and has more degrees of freedom than the *Strom_z* model. Moreover, in the plane *globalRz* model can overfit the data due to incomplete coverage of the (x, y) plane.

In Section 4.1, we showed that the presence of high-mass RC stars can contaminate the kinematics in the mid-plane and can give rise to high $V_{\rm los}$ residuals. Fig. 13(c) and (f) shows that if we remove this population, by restricting stars to $d_{\rm RC}/d_{\rm true} > 0.8$, the excess power disappears. This suggests that the observed excess power is spurious and is due to contamination from high mass stars whose distances are underestimated. For the off-plane slices this additional cut makes no difference as the density of high mass RC is negligible for these slices.

5 DISCUSSION AND CONCLUSIONS

Over the past few years several surveys have hinted at non-axisymmetric motion in the disc of the Milky Way. Bovy et al. (2015) used Red Clump stars from *APOGEE* to show velocity fluctuations of 11 km s^{-1} in the mid-plane region on scales of 2.5 kpc. In this paper we have made use of all the *APOGEE* Red Clump stars available up to date along with data from *GALAH*. Our results do not dispute the presence of deviation from mean axisymmetric motion in the mid-plane of the Galaxy. However, simulations using GALAXIA show that RC samples are likely to be contaminated by intrinsically brighter Red Clump stars, these stars are young and have high mass. Distance is underestimated for such stars. Being young, such stars lie preferentially closer to the mid-plane. This has the effect of contaminating the population at any given location with distant stars in that direction whose kinematics is different. This results in strange features when residual velocity maps are constructed in the (x, y) plane.

From Fig. 10, we conclude that for the mid-plane slice the peak power $p_{\rm max}$ occurs at physical scales of $k^{-1} \approx 3$ kpc for the observed data, and is either 9.3 km s^{-1}, using the original *Bovy1* model, or 6.3 km s^{-1}, using the more flexible axisymmetric model *globalRz*. On the other hand, the simulations from GALAXIA in Figs 12 and 13 show that the peak power is $8.1^{+2.0}_{-1.5}$ km s^{-1}using the *Strom_z* model or $5.2^{+1.2}_{-1.4}$ km s^{-1}with the flexible *globalRz* model. The peak in the power spectrum is also at the same physical scale of 3 kpc for both the observed sample and GALAXIA sample. We have also demonstrated that the power in GALAXIA is due to contamination from young high-mass Red Clump stars, as the sample with $d_{\rm RC}/d_{\rm True} > 0.8$ does not show excess power. So we do expect the high-mass stars to contribute to the power in the observed data, but how much is the contribution from real streaming motion is not obvious at this stage. The streaming and spurious perturbations in the velocity field could be correlated or uncorrelated. For the first case the streaming perturbations will add on to spurious perturbations and will enhance the power linearly. This would mean that the real streaming motion (observed peak power minus the average predicted peak power by GALAXIA is less than 1.2 km s^{-1}, adopting either *StromR_z* or *globalRz* as the reference model. Note, the observed fluctuations using the *Bovy1* model are best compared with GALAXIA predictions using the *Strom_z* model, as both models are inflexible models. If instead they are uncorrelated, we would expect the contributions to be added quadratically (given that power is physically a measure of dispersion), leading to an estimate of 4.6 km s^{-1} using *StromR_z* and 3.6 km s^{-1} using *globalRz*.

In the mid-plane using the flexible *globalRz* model we have been able to reduce the power from 9.3 to 6.3 km s^{-1}. The red pattern in the first quadrant and the yellow in the third are subdued. However, the blue pattern in second quadrant still exists, which could be due to a real feature in the data.

For slices away from the plane, $0.25 < |z/\text{kpc}| < 0.75$ and $0.75 < |z/\text{kpc}| < 1.25$, we find that for the observed data the power decreases with height above the plane and is no more than 5.1 km s^{-1}. This rules out large non-axisymmetric streaming motion extending beyond the $|z| > 0.25$ kpc. The GALAXIA samples also predict very little power (3 km s^{-1}) for slices away from the plane. However, the power in the observed data is higher than that predicted by GALAXIA by about 2 km s^{-1}. So, small streaming motion is not ruled out. Assuming streaming motion to be uncorrelated with other effects, we estimate the power to be less than 4.4 km s^{-1} for $0.25 < |z/\text{kpc}| < 0.75$ and less than 2.9 km s^{-1} for $0.75 < |z/\text{kpc}| < 1.25$.

If the excess power in the observed data is real and not an artefact of high mass clump stars, then it is interesting to consider the cause behind the decrease of power with height. This could be indicative of the fact that it is much easier to excite streaming motion in young dynamically cold populations than old dynamically hot populations.

We note that the analysis presented here has limitations when applied to data away from the mid-plane. The average age of stars increases with height above the plane due to the age scale height relation in the Galaxy. The mean azimuthal motion depends upon age and hence is also a function of $|z|$. Now, if a slice in $|z|$ is

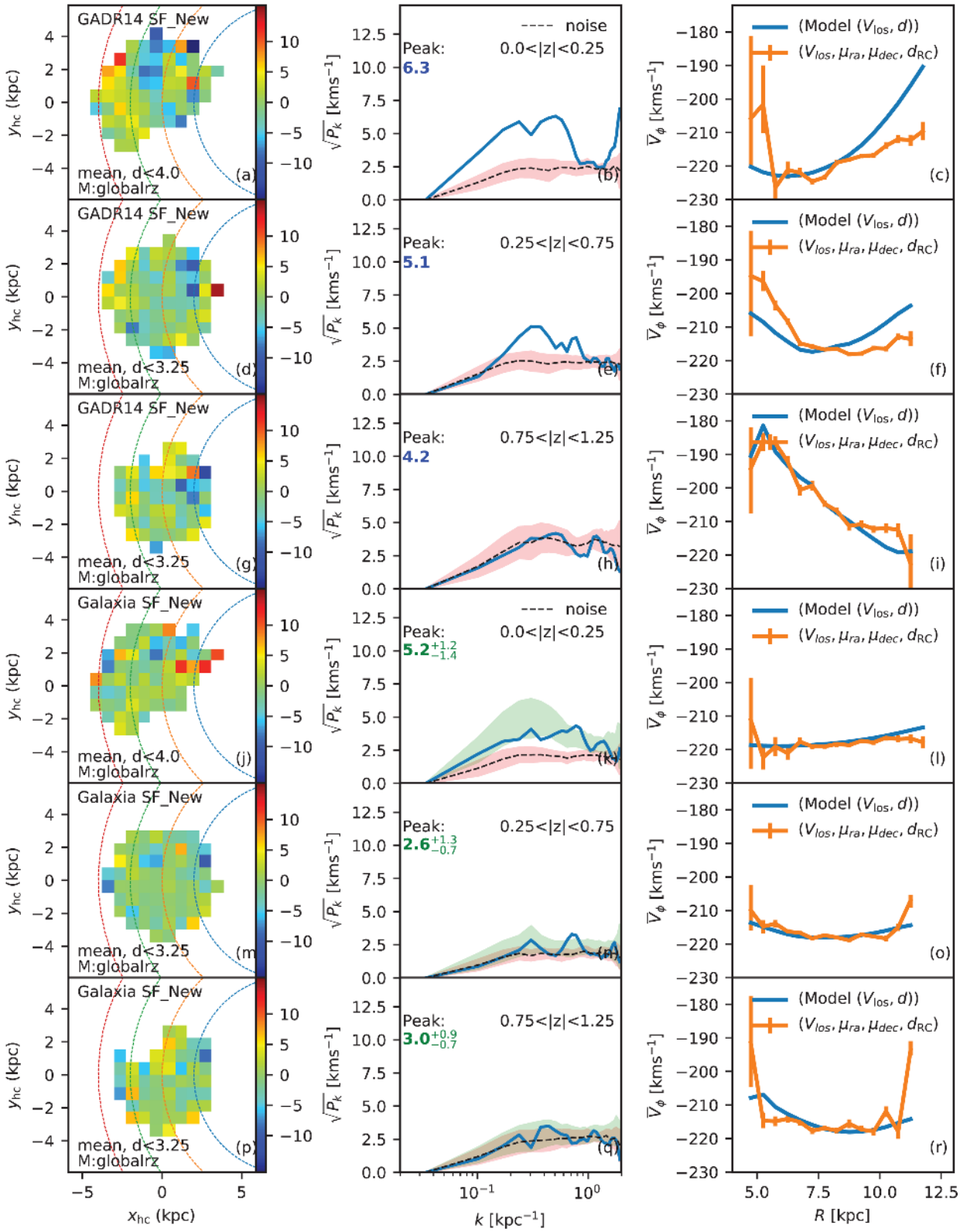


Figure 12. Residual V_{los} velocity maps, power spectrum, and $\overline{V}_\phi$ profiles for observed and simulated data for different slices in z. The top three panels correspond to *GADR14RC* while the bottom three panels correspond to mock GALAXIA. In each case, the power spectrum corresponding to the velocity map is shown in blue, the 1σ noise spread based on 20 random realizations in pink, and median noise in dotted black. For GALAXIA the green region represents the stochastic spread over 100 realizations, in power spectrum with the same selection function as data. We find that except for the case of mid-plane the $\overline{V}_\phi$ profiles are captured well by the *globalRz* models and the power spectrum of residual V_{los} velocity approaches noise with amplitude $\approx 2\,\mathrm{km\,s^{-1}}$.

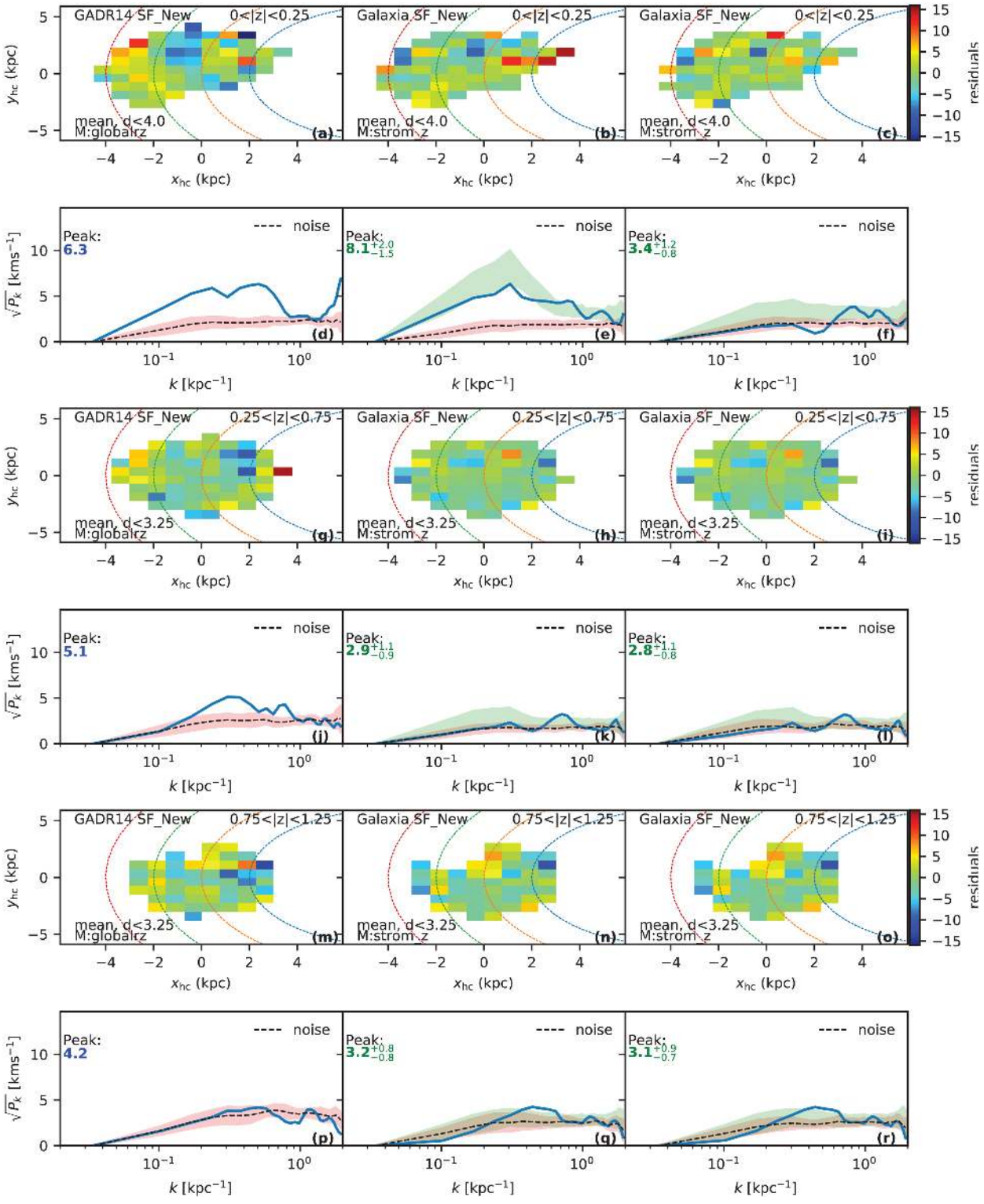


Figure 13. Residual velocity maps and power spectrum for observed and simulated data. Left column shows results of the observed data. Middle column shows results for data simulated with GALAXIA. Right column also shows results with GALAXIA but when high mass stars with systematically underestimated distances are removed. First and second rows are for $|z/\mathrm{kpc}| < 0.25$, third and fourth are for $0.25 < |z/\mathrm{kpc}| < 0.75$ and fifth and sixth are for $0.75 < |z/\mathrm{kpc}| < 1.25$. For the observed data a 2d polynomial of degree 3 (nine coefficients) is employed to create the residual velocity map. For simulated data the kinematic model is based on the Strömberg equation and is known a priori.

not sampled uniformly in the (x, y, z) space, the mean residual motion will show large variance just due to incomplete sampling. It is quite common for spectroscopic surveys to have such incomplete sampling at high $|z|$, as they observe in small patches across the sky. In such cases, one should always compare the power spectrum of observed data with selection function matched mock data which will correctly capture the power due to incomplete sampling.

Finally, Bovy et al. (2015), using their axisymmetric model, obtained a power excess in the mid-plane region, of $\approx$12 km $^{-1}$ and strongly suggested that the LSR itself is streaming at this velocity. They add this excess to the Schönrich et al. (2010) value for the Sun's peculiar motion to give the new $V_{\odot} \approx 12.1 + 12.0 = 24.1\,\mathrm{km\,s^{-1}}$. Following our analysis, we suggest that the adjustment to $V_{\odot}$ should be no more than 4.2 km s^{-1}, provided the excess power in the residual velocity field is not due to high-mass Red Clump stars. Interestingly, Kawata et al. (2018) using *Gaia* DR1 Cepheids also obtain $V_{\odot} = 12.5 \pm 0.8\,\mathrm{km\,s^{-1}}$, i.e. consistent with Schönrich et al. (2010), although they do not assert it to be conclusive given the small size of their sample.

We find that the spectro-photometric RC selection criterion given by Bovy et al. (2014) is quite efficient at isolating the RC stars. Based on GALAXIA simulations, the criterion can isolate RC stars with a purity of 98 per cent. We further refined the criteria and made it purely based on spectroscopic parameters. However, we find that such selection criteria have a strong age bias, Red Clump stars below 2 Gyr are significantly underrepresented.

Looking further to the future, *Gaia* can resolve some of the questions raised by our analysis. First, with accurate parallaxes from *Gaia*, we can confirm if the *APOGEE* Red Clump catalog contains high-mass stars with underestimated distances. If so, then does removing this population get rid of the excess power in the residual velocity map? Moreover, with proper motion we can construct and study velocity maps of V_{ϕ}, V_R, and V_z separately instead of just $V_{\rm los}$. We can also make use of all type of stars and not just the Red Clump.

ACKNOWLEDGEMENTS

The GALAH survey is based on observations made at the Australian Astronomical Observatory, under programmes A/2013B/13, A/2014A/25, A/2015A/19, A/2017A/18. We acknowledge the traditional owners of the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present. Parts of this research were conducted by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. SS is funded by University of Sydney Senior Fellowship made possible by the office of the Deputy Vice Chancellor of Research, and partial funding from Bland-Hawthorn's Laureate Fellowship from the Australian Research Council. DMN was supported by the Allan C. and Dorothy H. Davis Fellowship.

This research has made use of Astropy13, a community-developed core PYTHON package for Astronomy (Astropy Collaboration, 2013). This research has made use of NUMPY (Walt et al., 2011), SCIPY, and MATPLOTLIB (Hunter, 2007).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

REFERENCES

Abolfathi B. et al., 2018, ApJS, 235, 42
Altmann M., Roeser S., Demleitner M., Bastian U., Schilbach E., 2017, A&A, 600, L4
Anguiano B., Rebassa-Mansergas A., García-Berro E., Torres S., Freeman K. C., Zwitter T., 2017, MNRAS, 469, 2102
Antoja T., de Bruijne J., Figueras F., Mor R., Prusti T., Roca-Fàbrega S., 2017, A&A, 602, L13
Bertelli G., Bressan A., Chiosi C., Fagotto F., Nasi E., 1994, A&AS, 106, 275
Binney J., Tremaine S., 2008, Galactic Dynamics. 2nd edn. Princeton Univ. Press, Princeton, NJ
Binney J. et al., 2014, MNRAS, 437, 351
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Bovy J., 2010, ApJ, 725, 1676
Bovy J. et al., 2012, ApJ, 759, 131 (B12)
Bovy J. et al., 2014, ApJ, 790, 127
Bovy J., Bird J. C., García Pérez A. E., Majewski S. R., Nidever D. L., Zasowski G., 2015, ApJ, 800, 83
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Cannon R. D., 1970, MNRAS, 150, 111
Casagrande L., Ramírez I., Meléndez J., Bessell M., Asplund M., 2010, A&A, 512, A54
Cassisi S., Salaris M., 1997, MNRAS, 285, 593
Chen Y., Girardi L., Bressan A., Marigo P., Barbieri M., Kong X., 2014, MNRAS, 444, 2525
Chen Y., Bressan A., Girardi L., Marigo P., Kong X., Lanza A., 2015, MNRAS, 452, 1068
Dehnen W., 1998, AJ, 115, 2384
Dehnen W., 1999, AJ, 118, 1201
Duong L. et al., 2018, MNRAS, 476, 5216
Gaia Collaboration, 2016, A&A, 595, A2
Girardi L., 1999, MNRAS, 308, 818
Girardi L., 2016, ARA&A, 54, 95
Hawkins K., Leistedt B., Bovy J., Hogg D. W., 2017, MNRAS, 471, 722
Holtzman J. A. et al., 2015, AJ, 150, 148
Hunt J. A. S. et al., 2018, MNRAS, 474, 95
Kawata D., Bovy J., Matsunaga N., Baba J., 2018, MNRAS, 482, 40
Kos J. et al., 2017, MNRAS, 464, 1259
Kunder A. et al., 2017, AJ, 153, 75

Majewski S. R., Skrutskie M. F., Weinberg M. D., Ostheimer J. C., 2003, ApJ, 599, 1082
Majewski S. R., Zasowski G., Nidever D. L., 2011, ApJ, 739, 25
Majewski S. R., APOGEE Team, APOGEE-2 Team, 2016, Astron. Nachr., 337, 863
Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, A&A, 482, 883
Martell S. L. et al., 2017, MNRAS, 465, 3203
Miglio A. et al., 2017, Astron. Nachr., 338, 644
Nataf D. M., Gould A. P., Pinsonneault M. H., Udalski A., 2013, ApJ, 766, 77
Nataf D. M. et al., 2016, MNRAS, 456, 2692
Nordström B. et al., 2004, A&A, 418, 989
Reid M. J., Brunthaler A., 2004, ApJ, 616, 872
Robin A. C., Reylé C., Derrière S., Picaud S., 2003, A&A, 409, 523
Robin A. C., Bienaymé O., Fernández-Trincado J. G., Reylé C., 2017, A&A, 605, A1
Roeser S., Demleitner M., Schilbach E., 2010, AJ, 139, 2440
Salaris M., Girardi L., 2002, MNRAS, 337, 332
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Sharma S., 2017, ARA&A, 55, 213
Sharma S., Bland-Hawthorn J., 2013, ApJ, 773, 183
Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, ApJ, 730, 3
Sharma S. et al., 2014, ApJ, 793, 51 (S14)
Sharma S., Stello D., Bland-Hawthorn J., 2016, Astron. Nachr., 337, 875
Sharma S. et al., 2018, MNRAS, 473, 2004
Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., Zentner A. R., 2008, ApJ, 683, 597
Strömberg G., 1946, ApJ, 104, 12
Tang J., Bressan A., Rosenfield P., Slemer A., Marigo P., Girardi L., Bianchi L., 2014, MNRAS, 445, 4287
Wegg C., Gerhard O., Portail M., 2015, MNRAS, 450, 4050
Williams M. E. K. et al., 2013, MNRAS, 436, 101
Wittenmyer R. A. et al., 2018, AJ, 155, 84
Zacharias N., Finch C. T., Girard T. M., Henden A., Bartlett J. L., Monet D. G., Zacharias M. I., 2013, AJ, 145, 44
Zacharias N., Finch C., Frouard J., 2017, AJ, 153, 166
Zasowski G. et al., 2013, AJ, 146, 81
Zhao G., Zhao Y., Chu Y., Jing Y., Deng L., 2012, Research in Astronomy and Astrophysics, 12, 723

APPENDIX A: RED CLUMP CALIBRATION AND SELECTION

Following from Section 2.1, here we describe details of our Red Clump selection and calibration. A crude sample of RC stars can be selected based on cuts in surface gravity $\log g$ and dereddened colour $(J-K)_0$, for example, Williams et al. (2013) used the simple cuts of $1.8 \leq \log g \leq 3.0$ and $0.55 < (J-K)_0 < 0.8$ on *RAVE* data. However, this was estimated to be contaminated by about 30–60 per cent of non-RC stars, including the secondary Red Clump (Girardi 1999), and the red giant branch bump, which is a metallicity-dependent localized excess in the luminosity function of first-ascent red giant branch stars (Cassisi & Salaris 1997; Nataf et al. 2013).

In the APOGEE-RC catalog Bovy et al. (2014) use PARSEC isochrones (Bressan et al. 2012) and asteroseismic constraints to improve the sample purity, resulting in the following comprehensive selection scheme

$$1.8 \leq \log g \leq 0.0018\,\mathrm{dex\,K^{-1}}\ (T_{\rm eff} - T_{\rm eff}^{\rm ref}([\mathrm{Fe/H}])) + 2.5 \qquad \text{(A1)}$$

$$Z > 1.21[(J-K)_0 - 0.05]^9 + 0.0011, \qquad \text{(A2)}$$

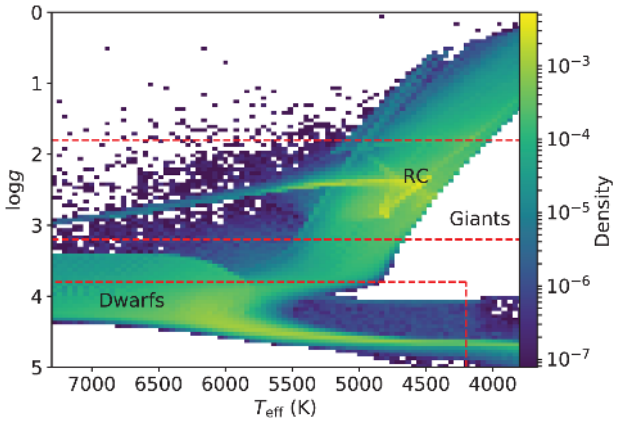


Figure A1. Spectroscopic HR diagram of the GALAXIA $J < 15$ all-sky sample used to derive $(J-K)_0$ as a function of $T_{\rm eff}$ and $\log g$. Red dashed lines mark the approximate boundary between Dwarfs and Giants and the typical location of Red Clump (RC) is indicated.

Table A1. Best-fitting coefficients for equation (A8) used to derive $(J-K)_0$ for the three populations: Dwarfs, Giants, and Red Clump. The fitting is carried out over the temperature range $4200 < T_{\rm eff} < 8000$.

Population	a_0	a_1	a_2	a_3	a_4	a_5
Dwarfs	−0.637	−0.107	−0.007	0.093	0.915	0.251
Giants	−0.957	0.000	−0.006	−0.020	1.489	0.002
Red Clump	−0.800	0.046	0.008	−0.060	1.199	0.132

Table A2. Median absolute magnitude $M_{\rm RC}$, and dispersion in absolute magnitude $\sigma_{M_{\rm RC}}$ for Red Clump stars selected from GALAXIA using the scheme in Appendix A. We have tabulated the values for a few common passbands only for a comparison with literature. Also listed are the extinction factors (f_λ) for the four passbands and these are taken from (Schlegel et al. 1998).

Passband (λ)	$M_{\rm RC}$	$\sigma_{M_{\rm RC}}$	$f_\lambda = \frac{A_\lambda}{E(B-V)}$
J	−0.98	0.11	0.902
H	−1.52	0.12	0.576
K	−1.60	0.13	0.367
V_{JK}	+0.75	0.15	3.240

$$Z < \mathrm{Min}(2.58[(J-K)_0 - 0.40]^3 + 0.0034, 0.06), \qquad \text{(A3)}$$

$$0.5 < (J-K)_0 < 0.8, \qquad \text{(A4)}$$

where

$$T_{\rm eff}^{\rm ref}([\mathrm{Fe/H}]) = -382.5\,\mathrm{K\,dex^{-1}}\,[\mathrm{Fe/H}] + 4607\,\mathrm{K}\,. \qquad \text{(A5)}$$

and Z is the PARSEC isochrone metallicity. However, this requires de-reddened $(J-K)_0$ colour and to get them extinction is required. In the *APOGEE* Red Clump catalog by Bovy et al. (2014) extinction was estimated using the Rayleigh Jeans Colour Excess method (RJCE; Majewski et al. 2011) which requires photometry in $[4.5\mu]$ band. Extinction estimates based purely on photometry are useful but have inaccuracies associated with them. To overcome this, we use pure Red Clump stars from GALAXIA to derive empirical relations expressing $(J-K)_0$ in terms of spectroscopic parameters [Fe/H] and $T_{\rm eff}$. Such relations have previously been derived for

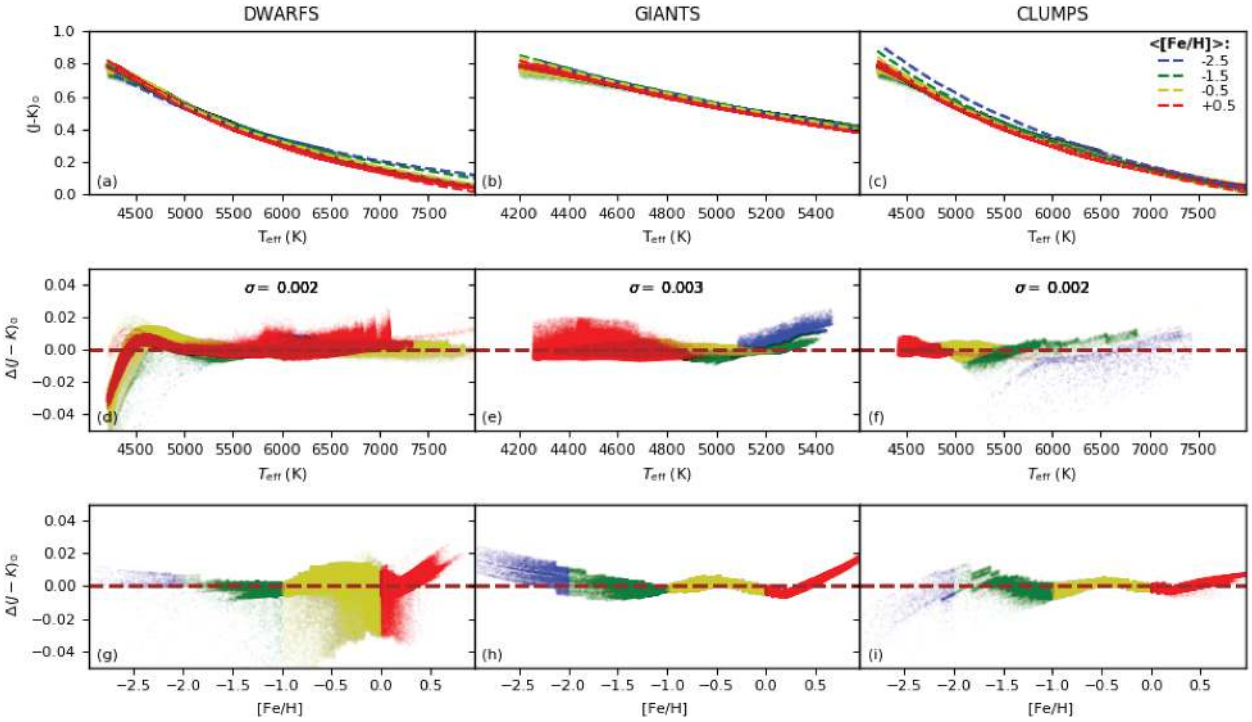


Figure A2. Empirical $((J-K)_0 \,|T_{\rm eff}, \log g)$ calibration using GALAXIA all-sky sample based on equation (A8): Panels (a, b, c) show the best-fitting (dotted curves) for Dwarfs, Giants, and Red Clump stars and the colour-coding is the mean [Fe/H]. Panels (d, e, f) and (g, h, i) show the residuals against temperature $T_{\rm eff}$ and [Fe/H], respectively. While for Dwarfs the derived relations fit well at low metallicity and high temperatures, for the Giants and Red Clump, residuals are low everywhere except for at [Fe/H]<-2.

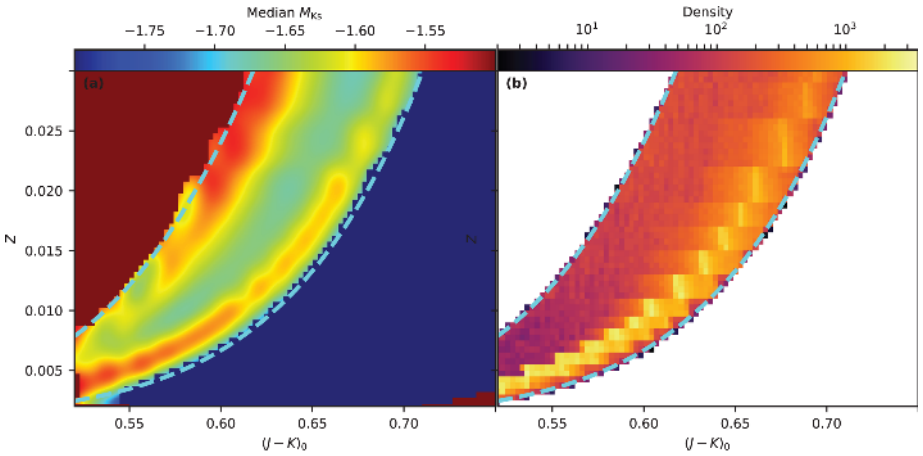


Figure A3. The all-sky GALAXIA Red Clump sample selected using $T_{\rm eff}$ and log g cuts based on equations (A1) and (A4), and with additional cuts (cyan dotted line) based on equations (A2) and (A3) using colour-temperature–metallicity calibration necessary to remove contamination from non RC stars. In the final selected sample, the median $M_{\rm k}$ lies in a narrow band around −1.60 (Panel a) and most stars are concentrated around this value (Panel b).

K-type dwarf stars by Casagrande et al. (2010), where one fits for a function of the form

$$5040\,{\rm K}/T_{\rm eff} = a_0 + a_1 X + a_2 X^2 + a_3 XY + a_4 Y + a_5 Y^2, \qquad \text{(A6)}$$

where $X = (J-K)_0$, $Y =$ [Fe/H] and $(a_0 ... a_5)$ are the fit coefficients. While this is a valid function to use, it is not analytically invertible to derive $(J-K)_0$, unless the dependence on [Fe/H] can be neglected in which case equation (A6) can be easily inverted to give[8]

$$(J-K)_0 \sim \frac{1}{2a_2}\left[-a1 + \sqrt{a_1^2 - 4a_2\left(a_0 - \frac{5040\,{\rm K}}{T_{\rm eff}}\right)}\right]. \qquad \text{(A7)}$$

[8]For completion we also perform the fitting using equation (A6) with and without the [Fe/H] term and found that the derived temperature had residuals below 20 K for Red Clump and Giants, though Dwarfs had higher (≈50 K) residuals without the [Fe/H] term. We provide these results in Table A4 but do not use it for our analysis in this paper.

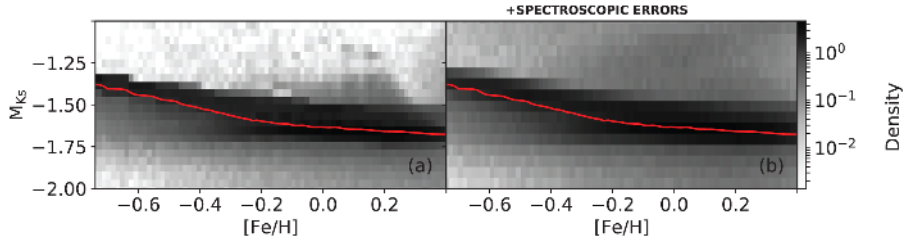


Figure A4. M_K - [Fe/H] distribution for the GALAXIA Red Clump sample. (a) This is the case without any errors in spectroscopic parameters. The sample has a tight distribution with dispersion in estimated distance modulus of $\sigma_{\rm dmod} = 0.1$; the running average (solid red curve) can thus be used to derive absolute magnitude M_K from spectroscopy. (b) This is the case with spectroscopic errors of $(\sigma_{\log T_{\rm eff}}, \sigma_{\rm [Fe/H]}, \sigma_{\log g}) = (0.011, 0.05, 0.1)$ dex. Here dispersion in distance modulus increases to $\sigma_{M_k} = 0.18$. The red curve is same as in (a).

Table A3. Tabulated values for mean absolute magnitude M_{Ks} as function of [Fe/H] as derived with GALAXIA (red line in Fig. A4a). Distances to Red Clump stars are estimated by linear interpolating over these values.

Fe/H (dex)	−0.8	−0.7	−0.6	−0.5	−0.4	−0.3	−0.2	−0.1	0.0	0.1	0.2	0.3	0.4
$M_{\rm Ks}$	−1.390	−1.405	−1.442	−1.475	−1.520	−1.564	−1.598	−1.622	−1.633	−1.646	−1.659	−1.666	−1.676

Table A4. Best-fitting coefficients for equation (A6) for the three populations: Dwarfs, Giants, and Red Clump stars. Equation (A6) derives 5040 K/$T_{\rm eff}$ which can be analytically be inverted to derive $(J-K)_0$ if we neglect the [Fe/H] term. This alters the coefficients slightly and so we also list that case. Note: This table is only provided for completion and we do not use it for our analysis in this paper.

Population	[Fe/H] used?	a_0	a_1	a_2	a_3	a_4	a_5	No. of stars
Dwarfs	Yes	0.5985	0.8148	−0.104	−0.053	0.0382	0.0045	212120
Dwarfs	No	0.6045	0.7700	−0.051	–	–	–	212120
Red Clump	Yes	0.6511	0.6410	0.0298	−7e-05	0.0116	−0.002	141666
Red Clump	No	0.5701	0.8421	−0.083	–	–	–	141666
Giants	Yes	0.6447	0.6651	0.0010	0.0044	0.0113	0.0042	135891
Giants	No	0.5458	0.9260	−0.171	–	–	–	135891

We show below that the dependence on [Fe/H] is weak but not negligible. So we alter equation (A6) to fit directly for $(J-K)_0$ as

$$(J-K)_0 = a_0 + a_1 X + a_2 X^2 + a_3 XY + a_4 Y + a_5 Y^2, \qquad \text{(A8)}$$

where $X = $ [Fe/H] and $Y = 5040\,{\rm K}/T_{\rm eff}$.

To derive the coefficients, we use data simulated by the code GALAXIA, which allows us to obtain relations valid for majority of the stars that we observe. More specifically, we generate an all-sky catalogue with $J < 15$, where the stellar parameters are generated using PARSEC[9] isochrones (Bressan et al. 2012; Chen et al. 2014, 2015; Tang et al. 2014), and choose the 2MASSWISE photometric system. From this we select three populations using boundaries in $\log g$, namely Dwarfs ($\log g \geq 3.8$), Giants ($\log g \geq 3.2$), and Red Clump stars ($1.8 \leq \log g \leq 3.0$). Fig. A1 marks the approximate boundaries between the three populations in the spectroscopic HR diagram. For a given age and metallicity of a star, stellar models can predict the initial mass required to reach the tip of the giant branch, and so for Red Clump stars the initial mass must exceed this threshold tipping mass (i.e. $>M_{\rm RGB,\,tip}$). We make this additional cut to identify the real Red Clump stars in GALAXIA. We also exclude M-dwarfs from our analysis by applying a temperature cut of $4200 < T_{\rm eff}$ (K) < 8000, as the $(J-K)$ colour is not a good indicator of temperature for them.

The resulting best-fitting coefficients $(a_0 ... a_5)$ for each population are listed in Table A1, using which we derive $(J-K)_0$. Fig. A2 shows the predicted $(J-K)_0$ and residuals as a function of $T_{\rm eff}$ and [Fe/H]. The best-fitting curves trace the colour well and the residuals for all three populations are below 0.003 mag. As mentioned earlier, weak metallicity dependence is visible. For the Red Clump and giants, the residuals show very little variation with temperature (Fig. A2e and f), but with metallicity (Fig. A2h and i) a systematic effect can be seen for [Fe/H] < -2. In comparison for dwarfs higher metallicities and lower temperatures have high residuals (Fig. A2d and g).

Finally using these derived colours we can now use equations (A1)–(A4) to produce a sample of Red Clump stars from our mock $J < 15$ GALAXIA catalogue. Here and throughout the paper for the purpose of selection function we make use of the $(J-K)_0$ relation corresponding to the Red Clump stars. Fig. A3 shows the Red Clump selection in metallicity-colour space and illustrates the effect of applying additional cuts from equations (A2) and (A3) (using colour–temperature–metallicity selection) in order to remove contamination from non-RC stars. It is clear that the final selection has a very narrow range in the median M_k and lies around −1.60.

With the RC sample selected, in Fig. A4 we plot the M_K against [Fe/H] and a running median curve (shown in red) that can be used to approximate Red Clump magnitude from metallicity. The dispersion in estimated distance modulus ($\sigma_{\rm dmod}$) increases from 0.1 to 0.16 by adding spectroscopic errors, however, if the uncertainty in temperature is a factor of two lower this lower $\sigma_{\rm dmod}$ to 0.12. For the GALAH data we can get such precision for good signal to noise data (Duong et al. 2018; Sharma et al. 2018).

For some simple calculations it is useful to know the typical absolute magnitude of Red Clump stars in different photometric

[9]The isochrones were downloaded from http://stev.oapd.inaf.it/cmd

Table B1. Coordinate transformation matrices.

lzR2xyz	*VlzR2xyz*	*Vxyz2lbr*
$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} Rcos(l) \\ Rsin(l) \\ z \end{bmatrix}$;	$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} = \begin{bmatrix} V_\phi & V_z & V_R \end{bmatrix} \begin{bmatrix} -sin(l) & cos(l) & 0 \\ 0 & 0 & 1 \\ cos(l) & sin(l) & 0 \end{bmatrix}$;	$\begin{bmatrix} V_l \\ V_b \\ V_r \end{bmatrix} = \begin{bmatrix} V_x & V_y & V_z \end{bmatrix} \begin{bmatrix} -y/rc & -zx/rc & x \\ -x/rc & -zy/rc & y \\ 0 & rc & z \end{bmatrix}$ $rc = \sqrt{x^2+y^2}$, $V_r = V_{\rm los}$

bands, e.g. to estimate the volume completeness of various surveys. Hence, in Table A2 we list the median absolute magnitude and dispersion based on 68 per cent confidence region for the J, H, K, and V_{JK} pass bands. Here,

$$V_{JK} = Ks + 2.0(J - Ks + 0.14) + 0.382\exp[(J - Ks - 0.2)/0.50] \quad \text{(A9)}$$

is the Johnson V-band magnitude computed using 2MASS magnitudes (Sharma et al. 2018). Our derived values are in good agreement with literature (Girardi 2016).

APPENDIX B: PHASE-SPACE TRANSFORMATION EQUATIONS

For our main analysis we fit a model for the mean $V_{\phi,\rm GC}$ to the $V_{\rm los}$ data. For this we require the following transformation from Galactocentric to heliocentric coordinates:

$$(l, z, R, V_\phi, V_z, V_R)_{\rm GC} \rightarrow (l, b, d, V_l, V_b, V_{\rm los})_{\rm HC}. \quad \text{(B1)}$$

This is achieved in the sequence,

(i) $(x, y, z)_{\rm GC} = \boldsymbol{lzR2xyz}(l, z, R)_{\rm GC}$,
(ii) $(V_x, V_y, V_z)_{\rm GC} = \boldsymbol{VlzR2xyz}(V_\phi, V_z, V_R)_{\rm GC}$,
(iii) $(V_x, V_y, V_z)_{\rm HC} = (V_x, V_y, V_z)_{\rm GC} - (U, \Theta, W)_\odot$,
(iv) $(x, y, z)_{\rm HC} = (x, y, z)_{\rm GC} - (x, y, z)_\odot$,
(v) $(V_l, V_b, V_{\rm los})_{\rm HC} = \boldsymbol{Vxyz2lbr}(x, y, z, V_x, V_y, V_z)_{\rm HC}$,

where following Schönrich et al. (2010) we adopt $(U, V)_\odot = (11.1, 7.25)$ km s^{-1}, and the azimuthal component $\Theta_\odot = 242.0$ km s^{-1} for data (239.08 for GALAXIA). The $\Theta_\odot$ for data is estimated as $\Omega_\odot R_\odot$, with $R_\odot = 8$ kpc and $\Omega_\odot = 30.24$ km s^{-1}kpc^{-1} as set by the proper motion of Sgr A* (Reid & Brunthaler 2004). The transformation matrices (in bold) are defined in Table B1.

On the other hand, to obtain the 'true' rotation profiles, we first convert the longitudinal and latitudinal proper motions to heliocentric velocities:

$$V_l = \mu_l \times d \times 4.74 \times 10^3 \quad \text{(B2)}$$

$$V_b = \mu_b \times d \times 4.74 \times 10^3 \quad \text{(B3)}$$

and then combined with $V_{\rm los}$ use the sequence above in reverse order to obtain V_ϕ.

[1]*Sydney Institute for Astronomy, School of Physics, A28, The University of Sydney, NSW 2006, Australia*
[2]*Centre of Excellence for Astrophysics in Three Dimensions (ASTRO-3D), Australia*
[3]*Miller Professor, Miller Institute, UC Berkeley, Berkeley, CA 94720, USA*
[4]*Center for Astrophysical Sciences and Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, MD 21218, USA*
[5]*Institute for Advanced Study, Princeton, NJ 08540, USA*
[6]*Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA*
[7]*Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA*
[8]*School of Physics, University of New South Wales, Sydney, NSW 2052, Australia*
[9]*Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia*
[10]*Australian Astronomical Observatory, 105 Delhi Rd, North Ryde, NSW 2113, Australia*
[11]*Research School of Astronomy & Astrophysics, Australian National University, ACT 2611, Australia*
[12]*Max Planck Institute for Astronomy (MPIA), Koenigstuhl 17, D-69117 Heidelberg, Germany*
[13]*Department of Astronomy, University of Virginia, PO Box 400325, Charlottesville, VA 22904-4325, USA*
[14]*University of Southern Queensland, Center for Astrophysics, Toowoomba, Qld 4350, Australia*
[15]*ICRAR, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*
[16]*Computational Engineering and Science Research Centre, University of Southern Queensland, Toowoomba, Qld 4350, Australia*
[17]*Australian Centre for Astrobiology, University of New South Wales, Sydney, NSW 2052, Australia*

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