University of Groningen

Determination of the escape velocity of the Milky Way using a halo sample selected based on proper motion

Koppelman, Helmer H.; Helmi, Amina

Published in:
Astronomy and astrophysics

DOI:
10.1051/0004-6361/202038777

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Koppelman, H. H., & Helmi, A. (2021). Determination of the escape velocity of the Milky Way using a halo sample selected based on proper motion. *Astronomy and astrophysics*, *649*, [A136]. https://doi.org/10.1051/0004-6361/202038777

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

A&A 649, A136 (2021)
https://doi.org/10.1051/0004-6361/202038777
© ESO 2021

Astronomy & Astrophysics

Determination of the escape velocity of the Milky Way using a halo sample selected based on proper motion

Helmer H. Koppelman[1,2] and Amina Helmi[1]

[1] Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands
[2] School of Natural Sciences, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA
e-mail: koppelman@ias.edu

Received 29 June 2020 / Accepted 5 March 2021

ABSTRACT

Context. The *Gaia* mission has provided the largest catalogue ever of sources with tangential velocity information. However, it is difficult to use this catalogue for dynamical studies because most of the stars lack line-of-sight velocity measurements. Recently, we presented a selection of $\sim 10^7$ halo stars with accurate distances that were selected based on their photometry and proper motions.
Aims. Using this sample, we model the tail of the velocity distribution in the stellar halo locally and as a function of distance. Our goal is to measure the escape velocity, and based on this, to constrain the mass of our Galaxy.
Methods. We fitted the tail of the velocity distribution with a power-law distribution, a commonly used approach that has long been established. For the first time, we used tangential velocities that were accurately measured for an unprecedented number of halo stars to estimate the escape velocity.
Results. In the solar neighbourhood, we obtain a very precise estimate of the escape velocity, which is $497^{+8}_{-8}\,\mathrm{km\,s^{-1}}$. This estimate is most likely biased low, our best guess is by 10%. As a result, the true escape velocity is most likely closer to $550\,\mathrm{km\,s^{-1}}$. The escape velocity directly constrains the total mass of the Milky Way. To find the best-fitting halo mass and concentration parameter, we adjusted the dark (spherical Navarro-Frenk-White) halo of a realistic Milky Way potential while keeping the circular velocity at the solar radius fixed at $v_c(R_\odot) = 232.8\,\mathrm{km\,s^{-1}}$. The resulting halo parameters are $M_{200}^{+10\%} = 1.11^{+0.08}_{-0.07} \cdot 10^{12}\,M_\odot$, and the concentration parameter is $c^{+10\%} = 11.8^{+0.3}_{-0.3}$, where we use the explicit notation to indicate that they are corrected for the 10% bias. The slope of the escape velocity with galactocentric distance is as expected in the inner Galaxy based on Milky Way models. Curiously, we find a disagreement beyond the solar radius where the estimated escape velocity is higher than at the solar radius. This result is likely an effect of a change in the shape of the velocity distribution and could be related to the presence of velocity clumps. A tentative analysis of the escape velocity as a function of (R, z) shows that the slope is shallower than expected for a spherical halo when standard values are used for the characteristic parameters describing the galactic disc.

Key words. Galaxy: kinematics and dynamics – Galaxy: structure – Galaxy: fundamental parameters

1. Introduction

Numerous studies have attempted to measure the mass of the Milky Way, but it has been notoriously difficult to obtain precise and model-independent constraints. Most works now agree that the mass of the Milky Way dark matter halo is $10^{12}\,M_\odot$ within a factor of 2 (see Fig. 7 of Callingham et al. 2019, for a recent compilation). The kinematics of globular clusters, dwarf galaxies, and halo stars have often been used in these studies (Kochanek 1996; Xue et al. 2008; Watkins et al. 2010; Deason et al. 2012; Fragione & Loeb 2017; Posti & Helmi 2019; Callingham et al. 2019; Fritz et al. 2020). The timing argument and the properties of debris streams such as those from the Sagittarius dwarf galaxy (e.g., Dierickx & Loeb 2017; Zaritsky et al. 2020) have provided additional but still similar constraints. In this work, we aim to derive a very precise estimate of the escape velocity near the Sun and hence, under further assumptions, of the mass of the Milky Way.

The escape velocity is the maximum velocity that stars can have while still being bound to the Galaxy. In principle, the single fastest moving bound star places a lower limit on the escape velocity. However, in practice, individual stars might be affected by large measurement uncertainties or they might be outliers (such as escapees). A more robust approach is to fit the velocity distribution as a whole, as was put forward by Leonard & Tremaine (1990, hereafter LT90), who described the tail of the velocity distribution with a power law.

Several works have used the LT90 method in the past. For example, Smith et al. (2007) and Piffl et al. (2014b), hereafter S07 and P14, estimated the escape velocity locally to lie in the range of $[500-600]\,\mathrm{km\,s^{-1}}$, using only radial velocity information from RAVE (Steinmetz et al. 2006). The analysis of 2017 (2017, hereafter W17) supports these values and these authors also show that the escape velocity drops to $\sim 300\,\mathrm{km\,s^{-1}}$ at a distance of 50 kpc. The advent of full phase-space information with *Gaia* DR2 has not led to a reduction in the estimated range for the escape velocity in the solar neighbourhood: It is still $[500-640]\,\mathrm{km\,s^{-1}}$ (Monari et al. 2018; Deason et al. 2019, hereafter M18 and D19), a result that can largely be attributed to the different underlying assumptions used by the authors.

In this paper, we use a sample of halo stars with only tangential velocities from *Gaia* DR2 to infer the escape velocity applying also the LT90 method. This sample comprises orders of magnitude more halo stars than any other sample used before. Samples making use of only tangential velocities have not been popular for this type of studies in the past because of the large

uncertainties in the velocities, particularly induced by the distance uncertainties. The lack of (accurate) proper motion measurements for large numbers of stars was even more dramatic. However, *Gaia* DR2, containing about ~200× more stars with proper motions than radial velocities, makes this type of study feasible now. We proceed in this work as follows. We describe the data we used and their properties in Sect. 2 and the methods in Sect. 3. In Sect. 4 we test the method for determining the escape velocity using mock data and cosmological simulations. In Sects. 5 and 6 we present our results in the solar neighbourhood and as a function of galactocentric distance, respectively. In Sect. 7 we use the local escape velocity to derive an estimate of the mass of the Milky Way dark halo and to identify likely unbound stars. In Sect. 8 we present our conclusions.

2. Data

The determination of the escape velocity is contingent upon having a sample of halo stars with high-quality measurements and large velocity amplitudes. Most of the data used in this work are provided by the *Gaia* mission (Gaia Collaboration 2016, 2018). We mainly use the sample of halo stars that was selected and analysed in Koppelman & Helmi (2021, KH21 hereafter). This sample comprises $\sim 10^7$ main-sequence (MS) halo stars, and we refer to it as the reduced proper motion or the 5D sample hereafter. Additionally, we make use of a set of nearby halo stars with full phase-space information.

2.1. Velocity information

To transform the observed motions (proper motions and radial velocities when available) into space velocities, we proceeded as follows. We computed the tangential velocity of a star by combining the proper motion and its distance as

$$v_j = 4.74057\,\mathrm{km\,s^{-1}} \left(\frac{\mu_j}{\mathrm{mas\,yr^{-1}}}\right)\left(\frac{d}{\mathrm{kpc}}\right), \tag{1}$$

where $j = (\ell, b)$. These velocities were then corrected for the solar motion using the values for the motion of the Sun with respect to the local standard of rest (LSR) given by Schönrich et al. (2010) and the motion of the LSR given by McMillan (2017); they are $(U_\odot, V_\odot, W_\odot) = (11.1, 12.24, 7.25)\,\mathrm{km\,s^{-1}}$ and $v_{\rm LSR} = 232.8\,\mathrm{km\,s^{-1}}$, respectively. The transformation to correct the tangential velocities is

$$v_j^* = v_j + v_{j,\odot}, \tag{2}$$

where $v_{\ell,\odot}$ and $v_{b,\odot}$ are defined as

$$v_{\ell,\odot} = -U_\odot \sin\ell + (V_\odot + v_{\rm LSR})\cos\ell, \tag{3a}$$

$$v_{b,\odot} = W_\odot \cos b - \sin b \cdot (U_\odot \cos\ell + (V_\odot + v_{\rm LSR})\sin\ell). \tag{3b}$$

Finally, the tangential velocity in the Galactic frame of rest as observed from the Sun is calculated as

$$v_t = \sqrt{(v_\ell + v_{\ell\odot})^2 + (v_b + v_{b\odot})^2}. \tag{4}$$

Similarly, the line-of-sight velocity can be corrected for the solar reflex motion using $v_{\rm los}^* = v_{\rm los} + v_{\rm los,\odot}$, where

$$v_{\rm los,\odot} = W_\odot \sin b + \cos b \cdot (U_\odot \cos\ell + (V_\odot + v_{\rm LSR})\sin\ell). \tag{5}$$

To derive space velocities, we used the following expressions:

$$v_x = v_{\rm los}^* \cos\ell\cos b - v_\ell^* \sin\ell - v_b^* \cos\ell\sin b, \tag{6a}$$

$$v_y = v_{\rm los}^* \sin\ell\cos b + v_\ell^* \cos\ell - v_b^* \sin\ell\sin b, \tag{6b}$$

$$v_z = v_{\rm los}^* \sin b + v_b^* \cos b. \tag{6c}$$

To transform the coordinates into a galactocentric frame, we placed the Sun at $X = -8.2$ kpc (McMillan 2017). We used this value for the distance to the Galactic centre because it is consistent with the McMillan (2017) potential that we employ later, and the same is true for the LSR velocity. We note, however, that the McMillan values agree well with the more recent determination of the distance to the Galactic centre by the GRAVITY Collaboration (2018) and with the circular velocity at the position of the Sun by Eilers et al. (2019).

To isolate a halo sample using the *Gaia* DR2 data, we considered stars with velocity vectors that deviate more than $250\,\mathrm{km\,s^{-1}}$ from the velocity vector of the LSR (i.e. the velocity vector of a typical disc star), namely $|\boldsymbol{v} - \boldsymbol{v}_{\rm LSR}| \geq 250\,\mathrm{km\,s^{-1}}$. This type of selection is known as a 'Toomre' selection.

When no line-of-sight velocity information was available, we used Eq. (6) and set $v_{\rm los}$ to zero. In that case, we refer to the velocity vector as $(\tilde{v}_x, \tilde{v}_y, \tilde{v}_z)$ to stress that these are not the true Cartesian velocities. For this set of stars, which constitute the majority of our sample, we used an adapted Toomre selection to isolate a halo sample, namely $|\tilde{\boldsymbol{v}} - \boldsymbol{v}_{\rm LSR}| \geq 250\,\mathrm{km\,s^{-1}}$.

2.2. Sample with full phase-space information

In the solar neighbourhood, we used a sample of stars with full phase-space information from *Gaia* that is known as the 6D or the radial velocity spectrometer (RVS) sample (Katz et al. 2019). We extended this dataset by adding sources with radial velocities observed by APOGEE (Wilson et al. 2010; Abolfathi et al. 2018), LAMOST (Cui et al. 2012), and RAVE (Kunder et al. 2017), see Sect. 2 of Koppelman et al. (2019) for a full description of this catalogue. The cross-matches with APOGEE and RAVE were obtained from the *Gaia* archive (Marrese et al. 2019).

For this sample and in line with M18 and D19, we used the quality cuts described in Marchetti et al. (2019), namely

- `astrometric_gof_al` < 3,
- `astrometric_excess_noise_sig` ≤ 2,
- $-0.23 \leq$ `mean_varpi_factor_al` ≤ 0.32,
- `visibility_periods_used` > 8,
- `rv_nb_transits` > 5,

and also imposed the following quality criteria:

- `ruwe` < 1.4,
- `parallax_over_error` > 5.

For the additional spectroscopic data we used the same quality cuts, with the exception of the criterion on `rv_nb_transits`. Additionally, we used survey-specific quality constraints. For APOGEE we used

- `SNR` > 20,
- `STARFLAG` $== 0$,
- abs(`SYNTHVHELIO_AVG` − `OBSVHELIO_AVG`) < 50,
- `NVISITS` > 2,

for RAVE,

- `eHRV` < 10,
- `Algo_Conv_K!` $= 1$,
- `SNR_K` > 20,

and for LAMOST,

- `snri` > 20,
- `snrg` > 20.

Several studies have reported that the sources in the RVS sample, and bright sources in general, contain a parallax offset

of ~0.05 mas (see Schönrich et al. 2019; Leung & Bovy 2019; Zinn et al. 2019; Chan & Bovy 2020). Therefore we corrected the parallaxes in the 6D sample for an offset of −0.054 mas as estimated by Schönrich et al. (2019). Following these authors, we increased the parallax uncertainties by 0.006 mas to account for the uncertainties in the offset and by 0.043 mas to account for the RMS in the offset reported by Lindegren et al. (2018), both of which were added in quadrature.

Nonetheless, to mitigate the effects of the parallax offset, we only considered sources within 2 kpc. As explained earlier, we selected halo stars as those with $|\boldsymbol{v} - \boldsymbol{v}_{\rm LSR}| > 250\,{\rm km\,s^{-1}}$. Finally, we removed the star with *Gaia* DR2 `source_id` 5932173855446728064 because its radial velocity reported in *Gaia* DR2 is known to be incorrect (Boubert et al. 2019). The final sample comprises 2067 high-quality stars, of which 495 are from the *Gaia* RVS sample, 10 from APOGEE, one from RAVE, and 1561 from LAMOST.

Since the spectroscopic surveys add a considerable number of stars, mostly from LAMOST, we verified that they do not bias our results. They are fully consistent with using only *Gaia* RVS sources. The stars from the spectroscopic surveys do not dominate the determination of $v_{\rm esc}$ because they have larger uncertainties in general. However, they do help in closing the confidence contours, as we show in Sect. 5.

2.3. Reduced proper motion sample

For the complete description of the reduced proper motion (RPM) sample, we refer to the KH21 paper. Here we summarise the details that are relevant for this paper. By virtue of the selection method, the RPM sample comprises only MS stars. The relatively linear colour-magnitude relation for these types of stars can be used to calculate photometric distances with typical uncertainties of 7%.

In KH21 we have introduced several quality cuts. Here we prune the sample even further. Firstly, we targeted the purest set of halo stars by imposing the following cut on velocities $|\tilde{\boldsymbol{v}} - \boldsymbol{v}_{\rm LSR}| > 250\,{\rm km\,s^{-1}}$ (see Sect. 2.1). Then we selected stars with tangential velocities higher than $v_t > 250\,{\rm km\,s^{-1}}$. Furthermore, we isolated stars that are least affected by extinction, that is, we considered only those with $A_G < 0.2$. Next, we selected stars in the colour range where the photometric distances have the smallest uncertainty: $0.50 < G - G_{\rm RP} < 0.715$. The blue limit here is stricter than in KH21 to remove any possible contamination from the MS turn-off. Finally, we selected stars at high latitudes to remove contamination from the disc: $|b| > 20$.

Some stars in the RPM sample have less precise photometric distance than trigonometric distance (i.e. parallax from *Gaia*). Furthermore, some stars may have been excluded because they did not satisfy the last three quality cuts described above, even though their trigonometric parallaxes are accurate. Therefore we returned these stars to the sample. We also replaced the photometric distances with trigonometric distances for stars with `parallax_over_error` > 10 if the latter has a smaller uncertainty than the first, and we only considered stars with parallaxes >0.5 mas.

As mentioned above, the trigonometric parallaxes from *Gaia* DR2 are known to contain a zero-point offset that has a complex dependence on other observational parameters (e.g., the colour and magnitude of the stars). Because most of the stars in the 5D sample (without radial velocities) are fainter than those in the 6D sample, we corrected their parallaxes with a different offset. Following Lindegren et al. (2018), we used a value of −0.029 mas for the parallax offset and increased the parallax uncertainties by 0.043 mas (the uncertainties were added in quadrature) to account for variations in the offset.

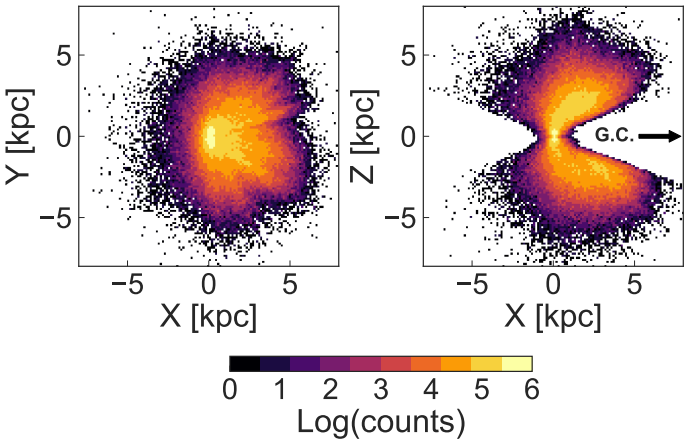

Fig. 1. Spatial distribution of the RPM sample used in this work, in heliocentric coordinates and for stars with $v_t > 250\,{\rm km\,s^{-1}}$. The Galactic centre is located at $X = 8.2$ kpc as indicated. The concentration of stars near the origin is caused by a small subset of stars with very accurate trigonometric parallaxes.

Within 1 kpc, about 90% of the distances stem from the *Gaia* parallaxes, and at 2 kpc, this percentage drops to ~50%. The final pruned sample comprises 197 449 sources, 18 236 of which have *Gaia* parallaxes.

2.4. Inspection of the RPM sample

Figure 1 shows the spatial distribution of the stars in the sample. The maps are coloured by the logarithm of the counts per bin. The quality cuts described in the previous section affect the spatial distribution of the stars, most notably, by removing low-latitude stars. The overdensity in the solar neighbourhood (centre of the figure) is caused by the addition of sources with accurate parallaxes. The median heliocentric distance of the sample is 3.6 kpc.

In Fig. 2 we show the tail of the tangential velocity distribution as a function of galactocentric distance by slicing the sample into uniformly spaced overlapping bins, ranging from 4−12 kpc, of 1 kpc width, which is larger than the typical uncertainty in the distances. A visual inspection reveals only small variations in the distributions. These clearly resemble a power law (as anticipated), but with a slight tendency to become more exponential with distance from the Galactic centre.

We propagated the uncertainties in the tangential velocities, which we denote as σ_t, using the standard uncertainty propagation approximation. This approximation uses a Taylor expansion to linearise the coordinate transformations, which is a common practice in the literature. We started the propagation from the measurement uncertainties of *Gaia* DR2, where the uncertainty of the distance is derived from the parallax uncertainty for the trigonometric sample. For the RPM sample, we derived distance uncertainties separately because this sample contains photometric distances (see Sect. 3.3. of KH21 for more information). Typically, the photometric distance uncertainties are at the 7% level.

The distribution of the relative uncertainties in the tangential velocity v_t is shown in Fig. 3 for the photometric and trigonometric distances separately. On average, the tangential

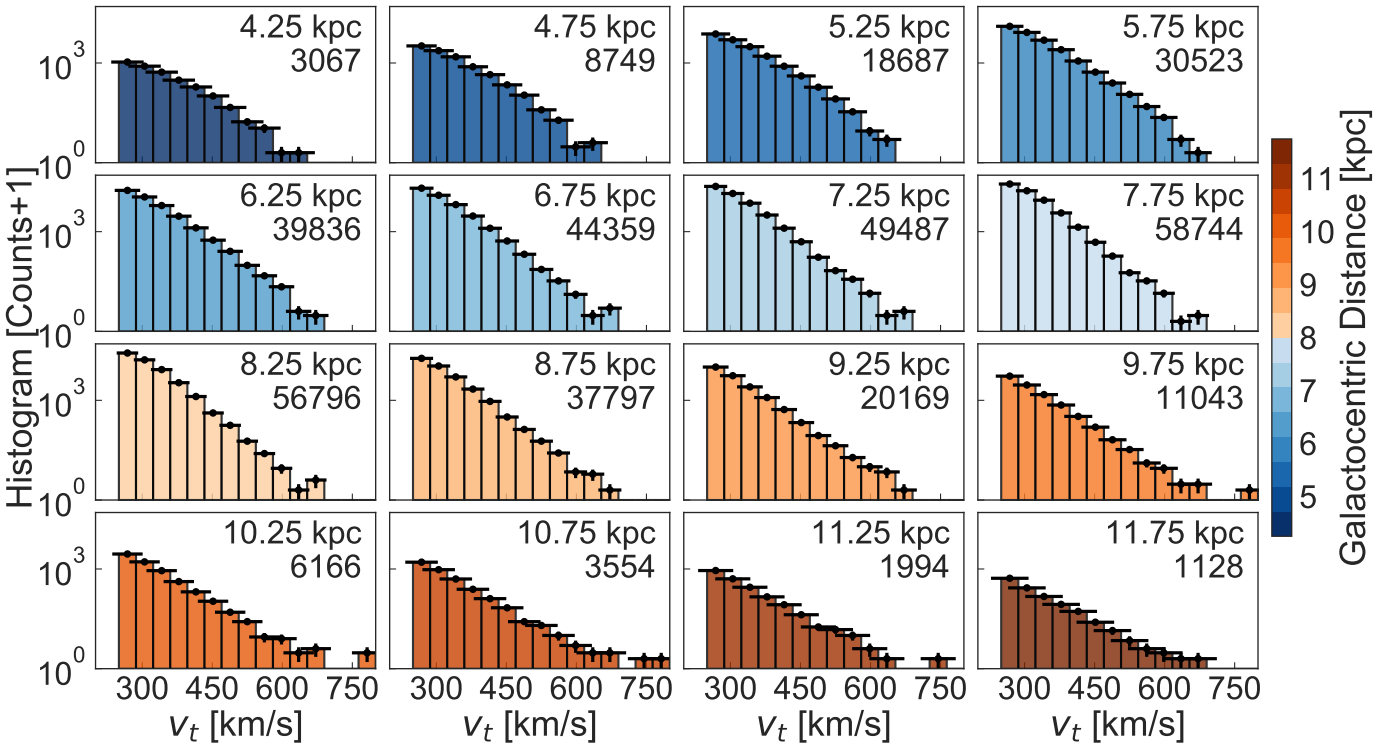

Fig. 2. Tail of the tangential velocity distributions for different galactocentric distances. The annotations in the panels indicate the central distance and number of stars per bin. The black marks give the uncertainty in the counts and the mean uncertainty in v_t for each bin.

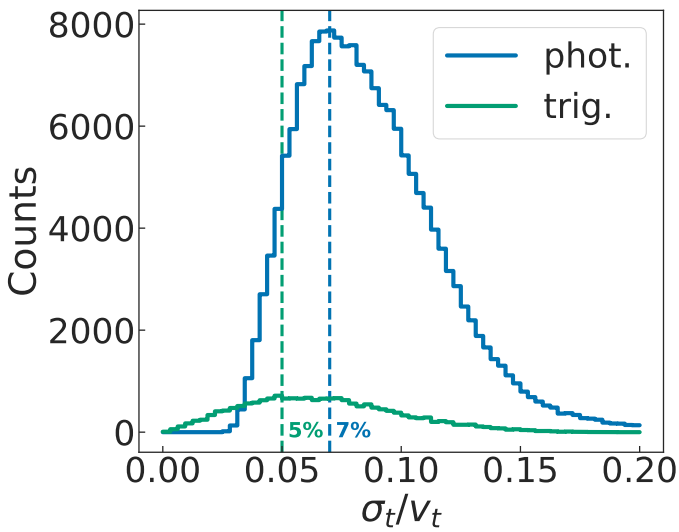

Fig. 3. Distribution of velocity uncertainties for sources in the RPM sample shown separately for sources with photometric (in blue) and trigonometric (in green) distances.

velocities derived from the trigonometric parallaxes are slightly more accurate than those based on the photometric distances. This is a selection effect because only sources with very accurate parallaxes are included in our sample. When the uncertainties are propagated in the velocities, we find that the uncertainty distribution for sources with photometric distances peaks at 7% and has a median of 8%. For the trigonometric distances, the distribution in the velocity uncertainties peaks at ~5%. The distribution of σ_t has a tail towards higher uncertainties because of proper motions uncertainties, and there is only is a small dependence on magnitude at the faint end (i.e. for $G \gtrsim 20$).

3. Methods

3.1. Determining v_{esc}

As described earlier, we used the LT90 method to determine the escape velocity, denoted hereafter as $v_{\rm esc}$. The motivation of this method is that the tail of the velocity distribution can be described by a power law, and $v_{\rm esc}$ is the velocity at which the probability of finding a star reaches zero. Although we closely followed Sections IIa and IIc from LT90 and adopted their notation, the formalism we used reveals some differences.

As just stated, the probability of finding a star in a local volume with a velocity in the range $(v, v + {\rm d}v)$ is described close to the escape velocity as a power law,

$$f(v|v_{\rm esc}, k) = \begin{cases} A(v_{\rm esc} - v)^k, & v_{\rm cut} < v < v_{\rm esc}, \\ 0, & v \geq v_{\rm esc}, \end{cases} \tag{7}$$

where k is the exponent, $v_{\rm esc}$ is the escape velocity, and $v_{\rm cut}$ is a threshold velocity below which the distribution is poorly represented by a power law. It is important to set $v_{\rm cut}$ accordingly such that only the tail of the distribution is fit. The normalisation constant is defined as $A = \frac{k+1}{(v_{\rm esc}-v_{\rm cut})^{k+1}}$, which is obtained from $A \int_{v_{\rm cut}}^{v_{\rm esc}} f(v|v_{\rm esc}, k)\,{\rm d}v = 1$.

The expression in Eq. (7) describes the distribution of velocities corrected for the solar motion (including peculiar and LSR), for example at the location of the Sun. If $f{\rm d}v$ is the probability of finding a star with velocity v in the range $(v, v + {\rm d}v)$, this implies that some distribution $g(\mathbf{v})$ exists such that $\int_\Omega g(\mathbf{v}){\rm d}\mathbf{v} = 4\pi v^2 g(v){\rm d}v = f(v){\rm d}v$ under the assumption that the velocity distribution is isotropic.

We now wish to obtain the probability distribution for the tangential velocity (i.e. $f_t(v_t|v_e, k)$). This can be derived from the joint distribution $f_{r,t}(v_r, v_t|v_e, k)$, which gives the probability of

finding a star with a given line-of-sight velocity and tangential velocity as $f_{r,t}(v_r, v_t|v_e, k)\mathrm{d}v_r\mathrm{d}^2v_t$. By performing a transformation of variables

$$f_{r,t}(v_r, v_t|v_e, k) = \int g(\boldsymbol{v}|v_e, k)\delta(v_r - \boldsymbol{v}\cdot\hat{\boldsymbol{n}})\delta(v_t - |\boldsymbol{v}\times\hat{\boldsymbol{n}}|)\mathrm{d}\boldsymbol{v}, \tag{8}$$

where $\hat{\boldsymbol{n}}$ is a unit vector along the line of sight. To express the distribution function in terms of v_t alone, we integrated over the line-of-sight component (and over angle),

$$f_t(v_t|v_e, k) = \frac{1}{2\pi}\int g(\boldsymbol{v}|v_{\rm esc}, k)\delta(v_t - |\boldsymbol{v}\times\hat{\boldsymbol{n}}|)\mathrm{d}\boldsymbol{v}. \tag{9}$$

The distribution $f_t\mathrm{d}v_t$ gives the probability of finding a tangential velocity v_t in the range $(v_t, v_t + \mathrm{d}v_t)$.

Evaluating this integral in spherical coordinates, with $\hat{\boldsymbol{n}}$ aligned with the z-axis (implicitly assuming the stars are distributed isotropically), we obtain

$$f_t(v_t|v_e, k) = \iint f(v)\delta(v_t - v\sin\theta)\, v^2\sin\theta\,\mathrm{d}v\mathrm{d}\theta, \tag{10}$$

which, by changing the order of integration and the substitution of $u = v_t - v\sin\theta$, reduces to

$$f_t(v_t|v_e, k) = -\int_{v_t}^{v_{\rm esc}} f(v)\left[v_t^{-2} - v^{-2}\right]^{-\frac{1}{2}}\mathrm{d}v. \tag{11}$$

The resulting integral for $f(v)$ given by Eq. (7) can be solved with `Mathematica` (and depends on the regularised hypergeometric ${}_2F_1$ function). When the Taylor series expansion of $v_t \to v_e$ is evaluated for the integral, we obtain

$$f_t(v_t|v_{\rm esc}, k_t) \propto \begin{cases}(v_{\rm esc} - v_t)^{k_t+\frac{1}{2}}, & v_{\rm cut} < v_t < v_{\rm esc},\\ 0, & v_t \geq v_{\rm esc},\end{cases} \tag{12}$$

which is the expression found by LT90. It can be normalised by multiplying with the constant $A_t = \frac{k_t+1.5}{(v_{\rm esc}-v_{\rm cut})^{k_t+1.5}}$, which is derived from the requirement that $A_t\int_{v_{\rm cut}}^{v_{\rm esc}} f_t(v_t|v_{\rm esc}, k_t)\,\mathrm{d}v_t = 1$, and where we have replaced k with k_t for clarity. The reason for this is that only in the case of $v_t \to v_e$ are the two power-law indices of Eqs. (7) and (12) related, and $k_t = k$. It is thus best to think of $f_t(v_t|v_{\rm esc}, k_t)$ in Eq. (12) simply as a power-law description of the tangential velocity tail, an approximation that is supported by Fig. 2. We show in Sect. 4.1 that it is in general not quite true that $k_t = k$ for the $v_{\rm cut}$ values that are typically considered in the literature. In what follows, we therefore reserve the notation k_t for the power-law exponent of f_t, use k for the exponent using the distribution from Eq. (7), and use k_r to indicate the exponent for a sample using only line-of-sight velocities (e.g., when comparing to values in the literature).

So far, we have assumed that the velocities (v and v_t) are the true velocities. However, in reality, we are dealing with 'observed' velocities, which are a combination of the true velocity and some unknown uncertainty. In this section we use v_t' to indicate the observed tangential velocity and reserve v_t for the true value. To account for the uncertainty, we smoothed the velocities by convolving them with an error distribution $\epsilon(v_t - v_t', \sigma_t)$, where σ_t is the uncertainty in v_t'. When we assume that v_ℓ and v_b have Gaussian errors, then the distribution $\epsilon(v_t - v_t', \sigma_t)$ follows the Beckmann distribution[1] (i.e. it is non-Gaussian). However, when it is evaluated far away from the origin ($v_t'/\sigma_t \gg 0$), this distribution is well approximated by a Gaussian. This gives us another reason to choose a sufficiently large $v_{\rm cut}$. In what follows, we therefore approximate $\epsilon(v_t - v_t', \sigma_t)$ by a Gaussian $f_G(v_t - v_t', \sigma_t)$. The convolution of the power law from Eq. (12) and the Gaussian is given by

$$C(v_t', \sigma_t, \boldsymbol{\theta}) = \int_{v_{\rm cut}}^{v_{\rm esc}} f_t(v_t|v_{\rm esc}, k_t) f_G(v_t - v_t', \sigma_t)\mathrm{d}v_t, \tag{13}$$

where $\boldsymbol{\theta} = (v_{\rm esc}, k_t, v_{\rm cut})$.

We note that we took $v_{\rm cut}$ as the lower boundary and not zero, as in Eq. (17) of LT90, because the velocity distribution below $v_{\rm cut}$ is poorly described by a power law. In this range, it is better described by a different distribution function $f^\dagger(v_t)$ for which the convolution over the range $0 < v_t < v_{\rm cut}$ would take the form

$$C^\dagger(v_t', \sigma, v_{\rm cut}) = \int_0^{v_{\rm cut}} f_t^\dagger(v_t)\epsilon(v_t - v_t', \sigma)\mathrm{d}v_t, \tag{14}$$

which neither depends on $v_{\rm esc}$ nor on k. As we show below, we may thus ignore this part of the velocity distribution. This also means that we also ignored stars with an apparent v_t' below the cut, but with a finite probability of having a true v_t above it. We show in Sect. 4 that these assumptions do not affect the ability of the method to infer $v_{\rm esc}$.

By normalising Eq. (13), we find $P(v_t', \sigma_t|\boldsymbol{\theta})$, the probability of finding a star with v_t' in the range $(v_t', v_t' + \mathrm{d}v_t')$,

$$P(v_t', \sigma_t|\boldsymbol{\theta}) = \frac{C(v_t', \sigma_t, \boldsymbol{\theta})}{\int_0^\infty C(v_t', \sigma_t, \boldsymbol{\theta})\mathrm{d}v_t'}. \tag{15}$$

By definition, because both f_t and f_G are normalised, the integral in the denominator is unity because the area under a convolution is $\int(f\circledast g)\mathrm{d}t = [\int f(u)\mathrm{d}u][\int g(t)\mathrm{d}t] = 1$. The resulting likelihood function is given by

$$\mathcal{L} = \prod_{i=1}^{n} P(v_t'^i, \sigma_t^i|\boldsymbol{\theta}). \tag{16}$$

The probability distribution of the model parameters $v_{\rm esc}$ and k_t is found by using Bayes' theorem,

$$P(\boldsymbol{\theta}|\Sigma_i^n v_t^i, \sigma_t^i) \propto P(v_{\rm esc})P(k_t)\prod_{i=1}^{n} P(v_t'^i, \sigma_t^i|\boldsymbol{\theta}), \tag{17}$$

where $P(v_{\rm esc})$ and $P(k_t)$ are priors for $v_{\rm esc}$ and k_t. For numerical reasons, the logarithm of the probability is evaluated, which also allows us to ignore the normalisation because that is constant and independent of the model parameters.

The procedure that is outlined above implicitly makes the following assumptions. Firstly, the tail of the velocity distribution is populated up to the escape velocity and, secondly, the tail of the velocity distribution is smooth. Furthermore, we assume that there are no unbound stars in our sample and that there is no contamination from a rotating (disc-like) population. The latter would break the isotropy on the sky.

[1] The Beckmann distribution is the most general form of the distribution $p(r)$ of parameter $r = \sqrt{x^2 + y^2}$, where x and y drawn from a bivariate Gaussian distribution (see https://reference.wolfram.com/language/ref/BeckmannDistribution.html). The distribution can generally only be expressed in integral form (e.g., Eq. (31) of Beckmann 1962), but takes an explicit form for specific cases. For example, when x and y are independent and drawn from a standard normal distribution, $p(r)$ takes the form of the Rice distribution. The more general case of $p(r)$, in which x and y are drawn from an uncorrelated bivariate Gaussian distribution, is known as the non-central chi distribution.

Perhaps the most problematic assumption is the first. There is no guarantee that the velocity distribution locally, or at any other location in the Milky Way, extends up to the escape velocity. Most likely, it is truncated at some lower value. As a result, the LT90 method may easily underestimate the true $v_{\rm esc}$. For example, cosmological simulations show velocity distributions that are truncated at 90% of $v_{\rm esc}$ (e.g., S07). The exact location of the truncation depends on the assembly history of the galaxy, and quite possibly, also on the resolution of the simulation. We quantify the truncation of the velocity distribution using mock data in Sect. 4. We stress that because of this truncation, whatever value we derive for $v_{\rm esc}$ it most likely is a lower limit.

The second assumption has recently been tested by Grand et al. (2019), who found that clustering in the velocity distribution biases the estimation of $v_{\rm esc}$, and can result in both under- and overestimates. Nonetheless, these authors showed that the estimated $v_{\rm esc}$ is typically underestimated by 7%. To emphasise the importance of this bias: A difference of 7% in the escape velocity results in a 21% bias in the estimated mass.

It seems unlikely that our sample contains many unbound stars because hyper-velocity stars are typically young stars ejected from the Galactic centre and not old stars in the halo (e.g., Brown 2015; Boubert et al. 2018). Furthermore, the velocity distributions shown in Fig. 2 are relatively smooth, suggesting the presence of a single population dominated by main-sequence halo stars. Nonetheless, it would be interesting to spectroscopically follow up stars near the escape velocity. In Sect. 7.3 we revisit possible outliers in the solar neighbourhood.

3.2. Adopting a prior on v_{esc} and k_t

In line with the literature, we assumed a simple prior on $v_{\rm esc}$ of the form $P(v_{\rm esc}) \propto 1/v_{\rm esc}$. For k (we use the notation k here, understanding that it only compares to k_t and k_r in the limiting case) there is some debate in the literature on what to assume, and because $v_{\rm esc}$ and k are highly degenerate (see next section), the prior assumed might bias the resulting $v_{\rm esc}$. For example, the M18 and D19 estimates of $v_{\rm esc}$ differ by $\sim$50 km s^{-1} mainly because different ranges were considered for k. Attempting to estimate $v_{\rm esc}$ and k simultaneously is only possible with a large sample with very small uncertainties. For example, LT90 estimated that a sample of >200 stars with high-quality radial velocities above $v_{\rm cut}$ is necessary to estimate both values simultaneously.

LT90 argued that k values should be in the range [0.5−2.5] because this brackets $k = 1.5$, which is the value expected for a system that has undergone violent relaxation (Aguilar & White 1986; Jaffe 1987; Tremaine 1987). S07 have compared stellar halos in cosmological simulations of Milky Way-like galaxies and found a range of [2.7−4.7] to be more appropriate. P14, building on more recent simulations like this, reduced this range to [2.3−3.7], which is also the range used by M18. D19 updated the criteria for finding Milky Way analogues based on recent discoveries regarding the merger history of the Milky Way (Belokurov et al. 2018; Helmi et al. 2018). When using cosmological simulations, the range [1.0−2.5] was found to be more favoured. Using a sample of blue horizontal branch (BHB) stars, K giants, and main-sequence turnoff stars from the Sloan Digital Sky Survey (SDSS) with only line-of-sight velocities, W17 determine both k and $v_{\rm esc}$ simultaneously. They report a value for k_r of 4 ± 1 for the local stars.

The above paragraph shows that no consensus has been reached on the value of k for the Milky Way. To complicate matters, the ranges mentioned above were determined for the stellar halo at the position of the Sun (in the simulations). It is not clear whether k remains constant as a function of distance to the Galactic centre. In this work, we mainly rely on the estimate of k at the location of the Sun. This is where our sample contains many stars with reliable parallax information that are approximately isotropically distributed on the sky. For this local sample of stars, we calculated the marginalised posterior distribution for k_t. We applied this posterior as a prior to other distance bins in which we estimate $v_{\rm esc}$. In doing so, we assumed that k_t does not vary (much) over the distance range that we probe, which is also justified by the analysis we carry out in Sect. 4.2.

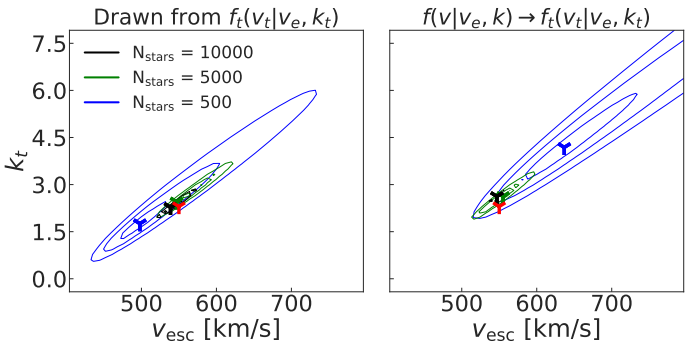

Fig. 4. Probability distributions of $v_{\rm esc}$ and k_t derived using tangential velocities for a mock data sample drawn from a power-law distribution in the velocity modulus, and convolved with realistic uncertainties. The *left panel* shows the results based on mock data drawn directly from Eq. (12), and on the *right* we draw the data from Eq. (7) and then transform the velocities into tangential velocities. The closed contours mark the probability levels at which the probability has dropped to 61%, 14%, and 1% of the maximum a posteriori value (they correspond to the 1, 2, and 3σ levels if the distribution were Gaussian). The coloured markers indicate the highest probability parameter combinations, with the red marker showing the input parameters.

4. Validating the method

Before applying the method to the data, we attempted to establish the accuracy of the method, the sample size required to estimate both $v_{\rm esc}$ and k_t at the same time, and the effect of the cut-off parameter $v_{\rm cut}$. We first consider mock data and then apply the method to cosmological simulations.

4.1. Tests with mock data

The mock data were drawn from an idealised power-law distribution. We sampled velocities according to the distribution given by Eq. (12) assuming $v_{\rm esc} = 550$ km s^{-1} and $k_t = 2.3$. We convolved the resulting distribution with realistic uncertainties drawn from the distribution of uncertainties (i.e. that shown in Fig. 3) for the sample of photometric distances.

The left panel of Fig. 4 shows the results of applying the formalism described in Sect. 3 to this dataset for three different sample sizes (see annotation) above $v_{\rm cut} = 250$ km s^{-1}. The contours mark the 1, 2, and 3σ levels estimated by the level where the probability has dropped to 61%, 14%, and 1% of the probability of the most likely parameter combination. The true $v_{\rm esc}$ and k_t of the parent distribution are indicated with a red marker. Decreasing the number of stars (from 10 000 to 500) results in higher uncertainties in the estimates of the $v_{\rm esc}$ and k_t parameters. A sample with $\sim 10^4$ stars is sufficiently large to determine both $v_{\rm esc}$ and k_t at the same time, given the amplitude of the velocity uncertainties.

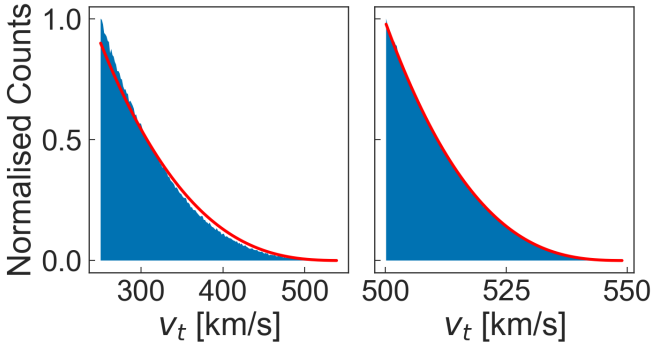

Fig. 5. Mock tangential velocity distributions drawn using Eq. (7) for two cut-off velocities (250, 500) km s^{-1} (*left and right panels*, respectively). The red line indicates the distribution that is expected when $v_t \rightarrow v_{\rm esc}$ (i.e. Eq. (12)).

We also tested a procedure drawing the mock distribution of v_t starting from a parent distribution of full 3D velocities (e.g., starting from Eq. (7) rather than from Eq. (12)). The 3D velocities were then transformed into v_t velocities by artificially setting one component to zero, assuming that the velocities are distributed isotropically. Arguably, the resulting distribution of mock v_t more closely describes the observed distribution of v_t than the one drawn directly from Eq. (12). The right panel of Fig. 4 shows that the resulting values for $v_{\rm esc}$ are close to the input value. However, the values for k_t are slightly overestimated. This overestimate arises from the difference between k and k_t for low $v_{\rm cut}$ and was already anticipated in Sect. 3.1.

To emphasise this behaviour, we show in Fig. 5 the behaviour of the artificial v_t distribution (drawn without uncertainties) compared to the expected distribution of $k_t \rightarrow k$ if $v_t \rightarrow v_{\rm esc}$ (i.e. Eq. (12)). The two distributions are only equivalent for cut-off velocities $v_{\rm cut}$ very close to the escape velocity.

Although this does not invalidate our approach at all because $v_{\rm esc}$ is robustly determined without any biases, we nonetheless have to be cautious when we compare the value of k_t obtained using tangential velocities. Similar considerations are in order when the LT90 method is applied to a sample of line-of-sight velocities only.

4.2. Tests on Aurigaia Milky Way-like halos

We now test the method on two halos from the Aurigaia suite of mock *Gaia* catalogues (Grand et al. 2019). We explore here whether the tail of the velocity distribution is well described by a power law, the effect of velocity clumps, the behaviour of k_t as a function of distance, and the power of the method given the typical uncertainties in the tangential velocity in our sample.

The Aurigaia catalogues have been generated from the Auriga suite of Milky Way-like galaxies (Grand et al. 2017), which is a suite of high-resolution, zoomed-in re-simulations based on galaxies extracted from the EAGLE simulations (Schaye et al. 2014). The mock catalogues that we analysed correspond to halos 6 and 27, have the bar at 30 degrees orientation, and were generated with the `SNAPDRAGONS` code (Hunt et al. 2015). We refer to these simulations as Au-06 and Au-27. These specific halos are chosen somewhat at random, although Au-06 is the most similar to the Milky Way (based on halo spin, Grand et al. 2018). The halo of Au-06 has a similar mass as the Milky Way (i.e. $M_{200} \approx 10^{12}\,M_\odot$), whereas that of Au-27 is slightly more massive: $M_{200} \approx 1.7 \cdot 10^{12}\,M_\odot$ (Grand et al. 2017). Both halos are mildly prograde (~30–70 km s^{-1}), as measured by the mean rotational velocity of accreted stars with heliocentric distances smaller than 1 kpc.

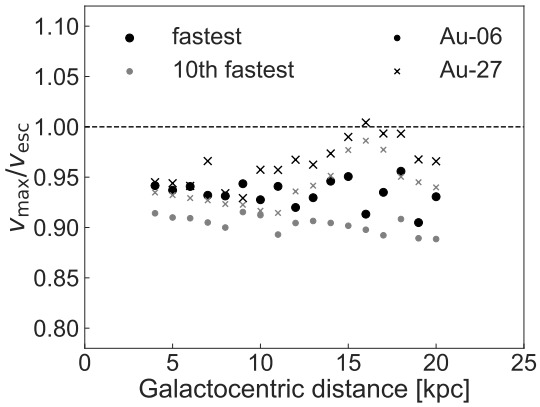

Fig. 6. Truncation of the velocities in the Aurigaia halos as a function of galactocentric distance. The velocity distribution in the halos is truncated at ~95%, except for a few bins in the outer regions of Au-27. The black markers show the stars that are the closest to the escape velocity. To indicate how densely populated the high-velocity tail is, the tenth fastest star is also shown (grey markers).

Because the original Auriga simulations do not have the resolution of *Gaia* DR2 (~10^9 stars), the `SNAPDRAGONS` code was used to artificially increase the number of objects, whereby simulated stellar particles are split into multiple 'stars'. This leads to artificial enhancement of the clustering of stars in phase-space, which can lead to biases in the determination of the escape velocity. Therefore we only used unique stellar particles by filtering all duplicates using the true `HCoordinates` and `HVelocities` parameters in the Aurigaia catalogue.

4.2.1. High-velocity tail in Aurigaia

Figure 6 shows the velocity of the fastest moving stars relative to the escape velocity and as a function of distance. We see that the fastest star typically moves at 90–100% of the true $v_{\rm esc}$ throughout the range of galactocentric distances probed. The escape velocity was calculated as the velocity needed to reach $r \rightarrow \infty$ for the potential given in the Aurigaia catalogue (parameter: `GravPotential`), and using

$$v_{\rm esc}(r \rightarrow \infty) \equiv \sqrt{2|\Phi(r) - \Phi(\infty)|} = \sqrt{2|\Phi(r)|}. \quad (18)$$

In Fig. 6 the black markers correspond to the fastest star, and the grey markers indicate the location of the tenth fastest star and provide an idea of the steepness of the velocity tail.

Interestingly, for the halo of Au-27, the velocity distribution is truncated close to the escape velocity around a galactocentric radius of ~15 kpc. In the inner regions, particularly for Au-6 but for Au-27 to some extent as well, the difference between the fastest and the tenth fastest moving star typically shows less scatter, indicating that there are many stars near the truncation of the velocity distribution.

4.2.2. Determining the escape velocity in Aurigaia

We followed a similar procedure as for the data to select stars with high tangential velocities from the Aurigaia halos. Firstly, the tangential velocities were convolved with uncertainties drawn from the 'observed' uncertainty distribution shown in Fig. 3 (and as in Sect. 4.1). We then selected stars that have

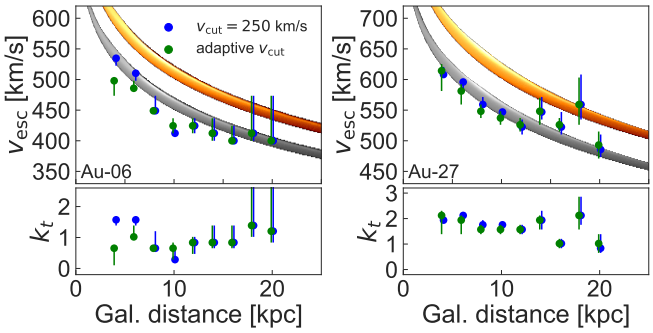

Fig. 7. Determination of v_{esc} and k_t as a function of galactocentric distance. The results for both a fixed (blue) and adaptive (green) cut-off velocity are shown. The yellow contour in the background shows the true escape velocity, calculated as $\sqrt{2|\texttt{GravPotential}|}$. The grey contour has been shifted downwards by 10%. The error bars indicate the 3σ confidence levels.

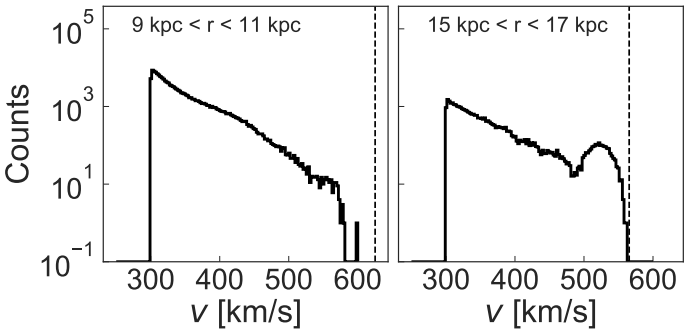

Fig. 8. Velocity distribution in Au-27 in two distance ranges. *Left*: smooth distribution of stars in the range 9−11 kpc that is truncated shortly before the escape velocity indicated by the dashed vertical line. *Right*: clumpy velocity distribution for stars in the range 15−17 kpc that reaches up to the escape velocity. This figure shows the full velocities (and not v_t) to emphasise the clumpiness.

$v_t > 200\,\mathrm{km\,s^{-1}}$. Next, we artificially set the line-of-sight velocities to zero and selected stars with $|\tilde{\boldsymbol{v}} - \boldsymbol{v}_{\rm LSR}| > 250\,\mathrm{km\,s^{-1}}$ (as in Sect. 2.3). Although the Aurigaia catalogues do not exactly represent the Milky Way, these velocity cuts serve to remove the thin disc and (most of) the thick disc present in the simulations.

We then determined v_{esc} and k_t in concentric shells of 1 kpc in width centred on the galaxy centre, with radii ranging from 2−21 kpc. For both Auriga halos, the cut-off velocity was set at $v_{cut} \approx 250\,\mathrm{km\,s^{-1}}$. This is well below the escape velocity in all the distance bins we probed. We also tested a heuristic procedure to determine v_{cut} by taking the maximum of $250\,\mathrm{km\,s^{-1}}$ and the velocity of the 10 000th fastest moving star (20 000th for Au-27). For bins with a large number of stars, this pushes the cut-off to higher values.

We note that for the Au-27 halo, the top 20 000 stars work better to determine v_{cut} than the top 10 000. Because this halo is more massive than that of Au-06, its escape velocity is higher and there are more stars with extreme velocities. However, because the Au-06 halo is more similar to the Milky Way, we expect that 10 000 is a realistic number of stars for the Milky Way.

Figure 7 shows the results of fitting the tangential velocity tail in the halos of Au-06 (left) and Au-27 (right). In yellow we show the mean v_{esc} that is calculated from Eq. (18) by using the pre-computed potential energies of every particle (i.e. the `GravPotential` parameter). The width of the yellow region indicates that there is a range of escape velocities at a fixed radius. This range exists because the potential is not spherically symmetric. Stars close to the disc experience a stronger potential than those slightly farther away.

The results obtained from a fixed v_{cut} are indicated with blue markers, and green markers are for the adaptive v_{cut}. The top panels of Fig. 7 show that the estimates are systematically too low compared to the expected v_{esc} for both halos. However, they match the grey curve well that has been obtained by lowering the yellow region by 10%. This is a reflection of a truncation in the tangential velocity distribution in that it does not extend all the way to v_{esc}. In both halos, the features in the v_{esc} curve are closely matched by the velocity features apparent for the tenth fastest stars shown in Fig. 6. An interesting result is that k_t varies only weakly over distance, as is shown in the bottom panels of Fig. 7.

The above results mean that the determination of v_{esc} with the method described in Sect. 3 is sensitive to the behaviour of the tail of the velocity distribution. This is particularly clear for Au-27, which shows a bump in v_{esc} at $d \gtrsim 15$ kpc. An excess of stars (a clump) is visible in the halo of Au-27 that moves at a velocity close to v_{esc}, as can be seen by comparing the panels of Fig. 8, which plot the velocity distributions for the distance ranges 9−11 kpc and 15−17 kpc.

In summary, the analysis of the Aurigaia experiments analysis shows that v_{cut} can be determined from the top 10 000 stars. Secondly, we may assume that k_t varies only weakly over the distance range probed by the RPM sample. Furthermore, on average, the method underestimates v_{esc} by ∼10%. This is slightly more than the 7% estimated by Grand et al. (2019), which might be related to differences in the method (e.g., the convolution with an uncertainty distribution and the typically large uncertainties on v_t). Finally, we note that by determining v_{esc} over a range of galactocentric distances, we can determine local 'biases'.

5. Results: Solar neighbourhood

We determined the escape velocity at the solar position using the two samples of stars described in Sect. 2, one with full 6D information and the other with only tangential velocities (5D). We considered only stars with a heliocentric distance of 2 kpc or less. We evaluated the probability (Eq. (17)) on a grid of 100 × 100 points ranging from $400\,\mathrm{km\,s^{-1}} < v_{esc} < 800\,\mathrm{km\,s^{-1}}$ and $1 < k_t < 6$ (both for the 5D and 6D cases). These ranges bracket the values that are presented in the literature. For the 5D sample, the cut-off velocity is based on the 10 000th fastest star and set to $317\,\mathrm{km\,s^{-1}}$, and for the 6D sample, it is $250\,\mathrm{km\,s^{-1}}$. Although the results for the 6D sample are consistent when v_{cut} is set to $317\,\mathrm{km\,s^{-1}}$, in this case, the inference on k is weaker.

The confidence contours for the two samples are presented in Fig. 9. For the sample with full phase-space information, we plot the results for the *Gaia*-only data (6D) and also including the additional data from ground-based spectroscopic surveys (6D+). The arrows in the figure indicate the maximum probability values for each sample. The contours correspond to estimates of the 1, 2, and 3σ levels (see Sect. 4.1). The side panels show the marginalised distributions $P(v_{esc})$ and $P(k)$ ($P(k_t)$ for the 5D sample). These distributions are best constrained for the 5D sample (in blue) because of its large number of stars.

The marginal distributions of v_{esc} agree very well with each other for all samples. The slight difference in the 5D and 6D curves (the contours are consistent within the 2σ level, however), is driven by the anticipated differences that are the result of using the full velocity modulus or tangential velocity information only

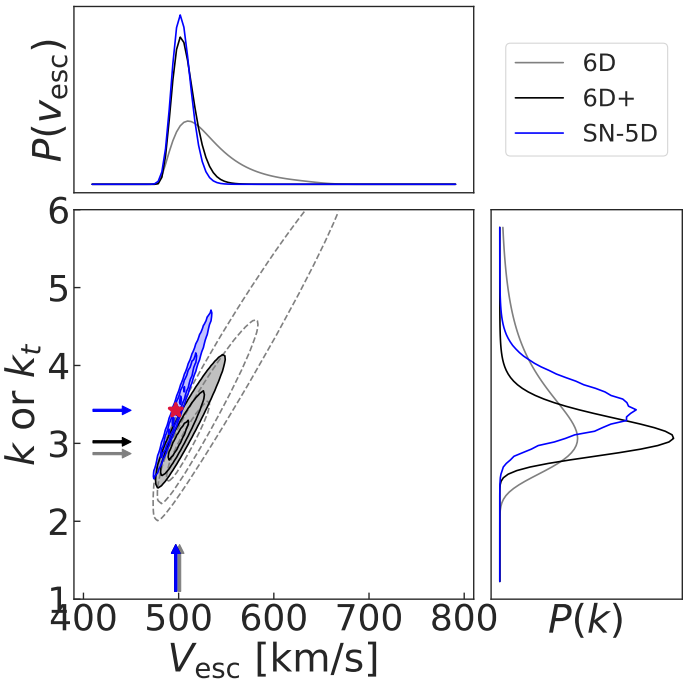

Fig. 9. Confidence levels obtained by applying the LT90 method to the 5D and 6D samples in the solar neighbourhood. For each curve the 1, 2, and 3σ levels are shown, and the arrows indicate the maximum probability values. The *side panels* show the marginalised posterior distributions for $P(v_{esc})$ and $P(k)$. For the 6D sample we show the results for both the augmented dataset and when using *Gaia* data alone. For the 5D sample, we recall that the method determines k_t, and this is what is shown on the y-axis of the *main panel*, while the blue curve in the *right panel* represents $P(k_t)$.

Table 1. Escape velocity (v_{esc}), power-law exponent (k_t, for the 5D and k for the 6D samples), and the number of stars (N_{stars}) for different distance estimates in the solar neighbourhood.

Sample	v_{esc} (in km s^{-1})	k_t/k	N_{stars}
5D	497^{+8}_{-8}	$3.4^{+0.4}_{-0.3}$	10 000
6D+	497^{+12}_{-8}	$3.0^{+0.3}_{-0.2}$	2067
6D (*Gaia* only)	505^{+32}_{-16}	$3.0^{+0.7}_{-0.4}$	495

Notes. The uncertainties in v_{esc} and k_t (or k) are given by the marginalised 1σ confidence levels.

(i.e. $k \neq k_t$ for v_{cut} far from v_{esc}, cf. the contours and red marker in Fig. 4).

Marginalising over k_t, we find a maximum probability value $v_{esc} = 497^{+8}_{-8}$ km s^{-1} for the 5D sample, which we stress is most likely biased low compared to the actual v_{esc}. The quoted uncertainties correspond to the marginalised 68% confidence levels (i.e. the ~1σ level). Table 1 presents v_{esc} and k_t (or k) derived for all the samples considered and for the curves shown in Fig. 9.

6. Results: Beyond the solar neighbourhood

6.1. Determining v_{esc}

We now proceed to determine v_{esc} as a function of galactocentric distance. As we saw in the Aurigaia halos, the behaviour of v_{esc} as a function of distance can help in identifying local 'biases' or issues. We assumed as prior for k_t the marginalised distribution obtained for the solar neighbourhood $P(k_t)_{SN}$ that is shown in the right panel of Fig. 9 with the blue curve. Therefore we implicitly assumed that k_t remains constant over the distance range probed. This assumption is justified by the Aurigaia simulations, as shown in Fig. 7.

We sliced the data into 16 concentric shells of 1 kpc width, with $4 < r < 12$ kpc and centred on the Galactic centre (as in Fig. 2). The number of stars per bin, with velocities higher than v_{cut}, varies from 58 744 to 1128. In each shell, v_{cut} is determined adaptively by selecting the top 10 000 fastest stars. We note that the results do not change significantly when the cut-off is fixed to $v_{cut} = 250$ km s^{-1}.

Figure 10 shows the trend of our estimate of v_{esc} with galactocentric distance. In each bin, the probability map is marginalised over the range $2.6 < k_t < 4.8$ after applying $P(k_t)_{SN}$. This range in k_t corresponds to the 3σ interval of the posterior distribution of $P(k_t)_{SN}$. We note that this is a very similar range to that assumed in S07. The use of the $P(k_t)_{SN}$ prior beyond the solar neighbourhood has helped in determining v_{esc} for all the distance bins considered, despite the sometimes relatively small number of stars used. With the size of the samples that are currently available, we could not have constrained both k_t and v_{esc} simultaneously for all radial bins.

The behaviour of the escape velocity in the inner halo ($r <$ 8 kpc) matches the expectation from several Milky Way models well. This can be seen by comparison to the predicted escape velocity plotted in the background of Fig. 10 for the Piffl14, McMillan17, BT08 (model I), and MW14 potentials (Piffl et al. 2014a; McMillan 2017; Binney & Tremaine 2008; Bovy 2015, all computed using the implementation from `AGAMA`, Vasiliev 2019). The behaviour for the estimated v_{esc} shows small variations: a slight elevation at ~6 kpc and a dip at ~4.5 kpc, although it is fully consistent with a smooth increase towards the inner Galaxy. Furthermore, the amplitude of these variations is of a similar level as we observed in the Aurigaia simulations. Curiously, our estimate of v_{esc} is higher outside of the solar radius (i.e. distance >8 kpc). This cannot be driven by the mass profile of the Milky Way and can only mean that something biases the determination of v_{esc}, as we discuss in detail next.

6.2. High v_{esc} outside of the solar radius

Several effects could lead to a higher v_{esc} outside the solar radius, namely (*i*) biases in the data (e.g., in the distance estimate), (*ii*) biases in the method (e.g., sample size), and (*iii*) variations in the dynamical properties of the stars with distance. We already explored the biases introduced by the first two categories in Sect. 2 (see also KH21) and Sect. 4.1. Nevertheless, we also tested that when the sample is downsized to a random subset of 5000 stars and bins with fewer stars are excluded, the results do not change. We therefore now focus on the third possibility, that the velocity distribution might be different outside of the solar radius.

Careful inspection of Fig. 2 shows that the velocity distribution is not contaminated by single outliers, even though in a relative sense (to the absolute number of objects), there seem to be more extreme velocity values in the outer radial bins. However, as mentioned earlier, the figure does show that the distributions seem to become more exponential with distance.

Curiously, we have seen a similar behaviour for Au-27 of the v_{esc} profile as observed for the 5D sample, see Fig. 7. In this case, the increase in v_{esc} was tentatively attributed to the presence of tidal debris (or at least lumpiness) that moved with speeds close to the true escape velocity.

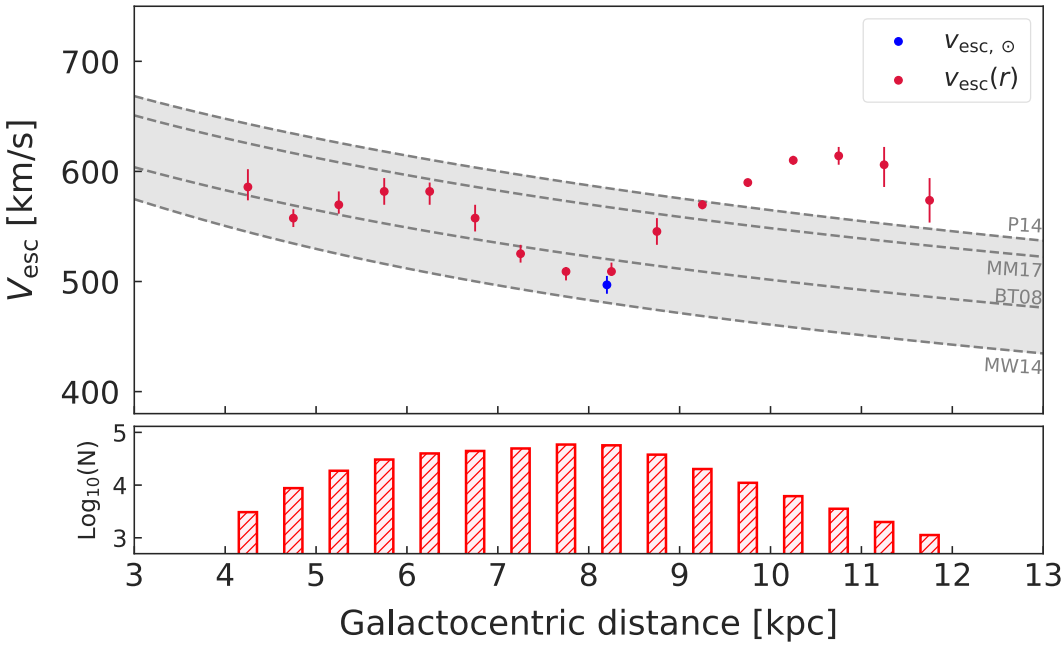

Fig. 10. Escape velocity as a function of galactocentric distance (*top*). We also shown in grey the expected behaviour of the escape velocity for four often-used Milky Way models. The *bottom panel* shows the logarithm of the number of stars for each distance bin. The blue marker indicates the v_{esc} that we determined using a local sample of stars, see Sect. 5. The error bars indicate the 1σ confidence levels. We note that the local sample is not the same as the data sample at the bin of ~8.2 kpc, therefore the two markers there do not overlap exactly.

With a two-point velocity correlation function, we tested the statistical clustering of the stars in the tail of the velocity distribution of our 5D sample. An excess of pairs implies that the velocity distribution is not smooth. The two-point velocity correlation function is given by

$$\xi(\Delta v) = \frac{DD(\Delta v)}{\langle RR(\Delta v)\rangle}, \tag{19}$$

where $DD(\Delta v)$ is the number of data-data pairs with a velocity separation of Δv, and similarly, $\langle RR(\Delta v)\rangle$ is the mean number of random-random pairs obtained by randomly shuffling the velocities 100 times. To this end, v_ℓ and v_b were shuffled and the pseudo-Cartesian velocities were re-calculated from Eq. (6) assuming $v_{los} = 0$. Both the data and re-shuffled samples were cut off at the velocity of the 10 000th star, or 250 km s^{-1} if there were not enough stars per bin. Because of this re-sampling, most of the bins have an equal number of stars, except for those at large radii.

Figure 11 shows the results of the correlation function ξ for the 5D sample for the same distance bins as used throughout this paper. The curves in Fig. 11 are coloured by the mean distance of the bin, and the error bars show the uncertainty in ξ estimated by the Poisson error in the number of counts per bin. A value $\xi = 1$ indicates no excess correlation. Inside 8 kpc, ξ decreases with distance. Meanwhile, the bins just outside of this radius (light red) show the highest level of correlation over the full velocity range probed. The inner and outermost bins (dark colours) show the least correlation, although the uncertainties are large because of the low number of stars in these bins. There might also be an effect associated with the area of the shells increasing with distance squared, which results in the stars in the outer shells being physically more separated than those in the inner shells. This might give rise to gradients in the trajectories of the stars and hence to lower correlation amplitudes.

The analysis of the velocity correlation function confirms that the properties of the velocity distribution change with distance. An indication of velocity clustering at $r \sim 10$ kpc in our 5D sample similar (although of lower amplitude) to that seen for Au-27 could thus be responsible for this change.

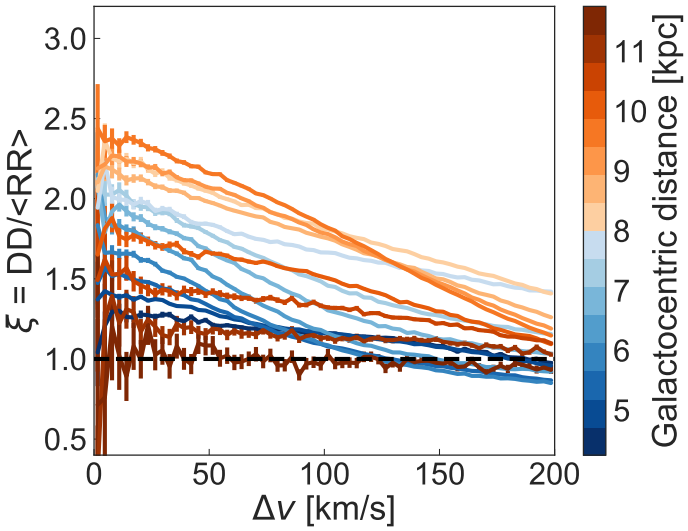

Fig. 11. Two-point correlation function ξ of the pseudo-Cartesian velocities of the stars, binned by galactocentric distance. A correlation of $\xi > 1$ indicates an excess of pairs compared to a random sample. The random sample $\langle RR\rangle$ is obtained by randomly shuffling the velocities in the Galactic rest frame.

7. Discussion

7.1. Relating v_{esc} to the Milky Way potential

In Eq. (18) we defined the escape velocity as the velocity required to reach $r = \infty$. A more realistic definition is obtained by taking a different zero-point. No matter how 'escaping from

the Milky Way' is defined, stars do not have to travel to infinity to be considered as escapees. For example, stars escaping to M31 make a much shorter journey ($r \approx 800$ kpc).

Therefore we used the definition of P14, who took the escape velocity to be the velocity required to reach $3r_{340}$,

$$v_{\rm esc}(r \to 3r_{340}) \equiv \sqrt{2|\Phi(r) - \Phi(3r_{340})|}, \tag{20}$$

where r_{340} is the radius within which the average halo density is $340 \times \rho_{\rm crit}$ (which is equal to $3H^2/8\pi G$, and where we assume $H = 73\,{\rm km\,s^{-1}\,Mpc^{-1}}$). We note that this zero-point is set somewhat arbitrarily, the 'true' value depends on the direction and might be a few km s^{-1} higher or lower. D19 used a different definition, which is for the star to escape to $2r_{200}(\approx 2.5r_{340})$. At the solar position, these two definitions result in a difference of 5 km s^{-1}.

Because the potential is axisymmetric, $v_{\rm esc}$ varies as a function of cylindrical R and z for a fixed spherical r. In the plane of the disc, where the potential is the steepest, the escape velocity is highest. Using the McMillan (2017) potential, we estimate that $v_{\rm esc}$ decreases by $\sim$20 km s^{-1} at a distance of 5 kpc away from the plane of the disc, whereas at 10 kpc, the difference is about 50 km s^{-1}.

To develop some intuition of how properties such as the mass of the Milky Way are related to $v_{\rm esc}$, we used the following equations: For a spherical potential, the gradient $dv_{\rm esc}/dr$ is related to the mass, circular velocity, and potential as

$$\frac{d\Phi(r)}{dr} = -v_{\rm esc}(r)\frac{dv_{\rm esc}(r)}{dr} = \frac{v_{\rm circ}^2(r)}{r} = \frac{GM(r)}{r^2}. \tag{21}$$

Another insightful equation, given by Eqs. (2)–(22) of Binney & Tremaine (1987), is

$$v_{\rm esc}(r_\odot)^2 = 2v_{\rm circ}(r_\odot)^2 + 8\pi G \int_{r_\odot}^{\infty} r\rho(r)\,dr \tag{22}$$

(see S07). The circular velocity $v_{\rm circ}$ at the solar position is a direct measure of the mass inside of the solar radius. On the other hand, the escape velocity $v_{\rm esc}$ is a measure of the total gravitational potential. The two are related through a factor of $\sqrt{2}$ only if there is no mass outside of the radius where both are measured. In other words, the difference $v_{\rm esc}^2 - 2v_{\rm circ}^2$ in the solar neighbourhood probes the potential, and with it, the mass distribution beyond the solar radius.

7.2. Estimating the mass of the Milky Way halo

We used our estimate of $v_{\rm esc}$ at the position of the Sun to constrain the mass of the halo of the Milky Way. The escape velocity and the gravitational potential of the Milky Way are related through Eq. (20). A straightforward procedure to derive the mass of the Milky Way is to take an existing model and adjust the parameters of the halo such that it matches the $v_{\rm esc}$ measured for the solar neighbourhood. We closely followed the procedure outlined in Sect. 5 of D19, but we used the McMillan (2017) potential and only varied the parameters of its dark halo, which is represented by a Navarro-Frenk-White (NFW) profile (Navarro et al. 1997).

The only issue with this procedure is that $v_{\rm esc}(r_\odot)$ is mostly sensitive to the mass outside of the solar radius. As a result, fitting $v_{\rm esc}$ constrains the concentration of mass inside the solar radius only weakly. A solution is to use the circular velocity ($v_{\rm circ}$), which is sensitive to the mass inside the solar radius, as an additional constraint. That is, when fitting $v_{\rm esc}(r_\odot)$, we forced the model to have a certain $v_{\rm circ}(r_\odot)$.

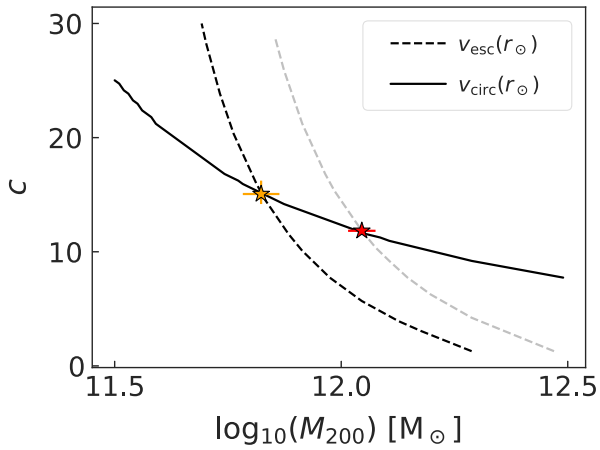

Fig. 12. Best-fit combinations of the halo mass and concentration parameter. The orange marker indicates the best-fitting M_{200} and c parameters, and the red marker shows the best-fitting model after correcting $v_{\rm esc}$ for a 10% offset. The error bars indicate the uncertainty in the halo parameters that is related to 1σ variations in the estimate of $v_{\rm esc}$.

The best-fitting potential is defined as the one that minimises

$$\eta = (v_{\rm esc}(r_\odot) - v_{\rm esc}^{\rm est})^2 + (v_{\rm circ}(r_\odot) - 232.8\,{\rm km\,s^{-1}})^2, \tag{23}$$

where we take $v_{\rm esc}^{\rm est}$ to be the maximum probability value found in the solar neighbourhood for the 5D sample (see Table 1). The value for the circular velocity that we assumed, $v_{\rm circ}(r_\odot) = 232.8$ km s^{-1}, is the value that was used in the original McMillan (2017) potential. We note that there is no freedom in choosing $v_{\rm circ}(r_\odot)$ because the data are only consistent with the value above because it is used in the correction for the solar motion.

Figure 12 shows the values for Eq. (23) for the ranges of M_{200} and c that we explored, namely $\log_{10}(M_{200})[M_\odot] \in [11.5, 12.5]$ and $c \in [1, 30]$. The solid line marks all models that have a correct $v_{\rm circ}(r_\odot)$, and the dashed lines mark all models that have the correct $v_{\rm esc}^{\rm est}$. The best-fitting potential lies at the intersection of the two lines. The curves illustrate the benefit of including the $v_{\rm circ}$ in the fit. As we expected, the dashed curve is only weakly sensitive to c. The orange marker highlights the combination of M_{200} and c that best fits $v_{\rm esc}$ and $v_{\rm circ}$. Therefore the best-fitting estimate of the mass is $M_{200} = 0.67^{+0.06}_{-0.06} \cdot 10^{12}\,M_\odot$ and the corresponding concentration parameter is $c = 15.0^{+1.2}_{-0.9}$. The uncertainties are derived by calculating the best-fitting M_{200} and c for the extreme cases of $v_{\rm esc}^{\rm est} + 8$ km s^{-1} and $v_{\rm esc}^{\rm est} - 8$ km s^{-1}, which are the limits given by the 1σ level (e.g., Table 1).

As we mentioned above, the LT90 method is likely to underestimate the $v_{\rm esc}$. Therefore the mass and concentration parameters quoted above should be seen as a lower limit. It is accordingly consistent with the original potential of McMillan (2017) in the sense that it is smaller and more concentrated. Moreover, this lower limit is also lower than most recent mass estimates (cf. Fig. 7 Callingham et al. 2019, for a recent compilation). If we now use the results from the analysis of the Auriga simulations and adjust for the 10% underestimation of $v_{\rm esc}$ (grey dashed curve in Fig. 12), we find that the best-fit mass and concentration parameter are $M_{200} = 1.11^{+0.08}_{-0.07} \cdot 10^{12}\,M_\odot$ and $c = 11.8^{+0.3}_{-0.3}$.

7.3. Stars that might be unbound

A possibly interesting follow-up project is to measure the radial velocities of the stars in the 5D sample that lie near the truncation of the best-fit power law. Using the maximum probability fit of the velocity distribution, we can calculate which stars have a high probability of being unbound. Given the apparent v'_t and its uncertainty, we can calculate the probability of these stars having a true v_t higher than $v_{\rm esc}$. We note that strictly speaking, the uncertainties are non-Gaussian, see also Sect. 3. However, we assume that they are small enough such that they may be approximated to be Gaussian.

For the set of stars that have apparent tangential velocities higher than the estimated $v_{\rm esc}$, we calculated the probability of the star being bound as

$$P_{\rm bound} = \sum_{v_{e,i}} P_{\rm SN}(v_{e,i}) \int_0^{v_{e,i}} f_G(v'_t, v_t, \sigma_t) {\rm d}v_t, \tag{24}$$

where $P_{\rm SN}(v_e)$ is the posterior of $v_{\rm esc}$ marginalised over k_t (i.e. the blue curve in the right panel of Fig. 9). The uncertainty distribution $f_G(v'_t, v_t, \sigma_t)$ is defined such that it gives the probability of finding the star with a true velocity v_t and uncertainty σ_t with an apparent velocity in the range $(v'_t, v'_t + {\rm d}v'_t)$. The probability of the star being unbound is simply $P_{\rm unbound} = 1 - P_{\rm bound}$.

The list of sources that fall outside of the maximum probability value of $v_{\rm esc}$ is given in Table 2. We stress that very likely, the actual $v_{\rm esc}$ is higher than our best estimate. The values for $P_{\rm unbound}$ given here should therefore be considered as upper limits. To emphasise this, we also calculated the probability of these stars being unbound after correcting $v_{\rm esc}$ for a 10% offset, based on our analysis in Sect. 4.2. Only two sources remain unbound in this case, and one of these just barely. The source with the highest probability of being unbound, with *Gaia* DR2 `source_id` `2655054950237153664`, has been flagged in the `faststars`[2] database (Guillochon et al. 2017) as a potential hyper-velocity star. The source was first identified by Du et al. (2019) based on its high tangential velocity.

About half of the sources in Table 2 have an inward-pointing velocity vector based on the pseudo-velocities in the galactocentric frame. This makes it likely that the majority of these stars are bound to the Milky Way. A possibility remains, however, that the velocity vectors of the stars point radially outwards when line-of-sight velocities are measured. However, for some stars the vectors will always point inwards, even in the extreme case of $v_{\rm los} = \pm 500\,{\rm km\,s^{-1}}$. One of these stars has the highest probability of being unbound (`source_id` `2655054950237153664`), which has an outward-pointing velocity vector even for adopted line-of-sight velocities of $\pm 500\,{\rm km\,s^{-1}}$, and therefore it might truly be unbound.

7.4. $v_{\rm esc}$ as tracer of the mass distribution

The luminous components of the Milky Way are most definitely not spherically symmetric. Because the escape velocity traces the potential, we should ultimately measure it in axisymmetric coordinates rather than as a function of spherical radius. By estimating $v_{\rm esc}$ as a function of z, we can perhaps constrain the flattening of the halo, although with the current sample we are more sensitive to the contribution of the disc to the total potential of the Milky Way. This means that such an analysis would benefit from a large sample of stars that probe the Milky Way halo at a greater depth, such as what may become available with *Gaia* (e)DR3.

Table 2. High-velocity sources that are close to the escape velocity.

`source_id`	v'_t (ϵ_v)	$P_{\rm unbound}$	$P^{+10\%}_{\rm unbound}$
5456509319663300096	499 (24)	0.50	0.02
4966291540726119936	500 (48)	0.50	0.13
1301068812277635968	500 (47)	0.51	0.13
3176805236597893248	502 (39)	0.53	0.10
600589157020469120	503 (50)	0.53	0.15
1300879558838744832	506 (41)	0.57	0.13
6669606511542374656	507 (61)	0.55	0.22
1268796702891972864	513 (46)	0.61	0.19
3289306720892701056	513 (64)	0.58	0.26
6843814817473042176	514 (35)	0.65	0.14
6587991790636824960	519 (36)	0.70	0.17
1142600839930233216	521 (65)	0.63	0.30
1948677828145591296	523 (57)	0.66	0.29
1831456179092459264	537 (43)	0.80	0.35
6085387089802067968	538 (63)	0.73	0.40
5845412900328041856	543 (57)	0.77	0.42
1981230244289202176	546 (48)	0.83	0.43
3495222399548253440	553 (68)	0.78	0.49
5175122643183339392	554 (30)	0.96	0.49
6270738976140076928	560 (49)	0.88	0.54
2655054950237153664	614 (62)	0.96	0.82

Because of the large number of sources in our 5D sample, it is possible for the first time to explore the escape velocity as a function of cylindrical R and z. We sliced our 5D sample into overlapping bins of 8×11 volumes of $|R_c| < 1$ kpc and $|z_c| < 1$ kpc, where R_c and z_c are the centres of the volumes. Assuming that the Milky Way is perfectly axisymmetric, we included sources independent of their azimuthal angles. Bins with fewer than 500 stars were discarded. We used the same method to determine the escape velocity as we used in Sect. 6 and presented in Fig. 10. That is, we again assumed the posterior distribution of $P_{\rm SN}(k_t)$ from the solar neighbourhood as prior on k_t. For computational reasons, we decreased the size of the grid on which the probability distribution was evaluated to 50×22 points ranging from $400\,{\rm km\,s^{-1}} < v_{\rm esc} < 800\,{\rm km\,s^{-1}}$ and $2.6 < k_t < 4.8$ (which corresponds to the 3σ levels in the solar neighbourhood).

Figure 13 shows the escape velocity in these volumes (large coloured markers) with a colour map corresponding to the escape velocity predicted by the McMillan17 model, with the updated lower limit of the halo mass computed in Sect. 7.2. Therefore this model is based on the estimate for $v_{\rm esc}$ that is biased low, and we used it to predict this estimate at other locations for a spherical NFW halo. The large plus markers in Fig. 13 indicate volumes in which the 3σ levels of the $v_{\rm esc}$ include the expected value. The large crosses indicate volumes where the $v_{\rm esc}$ expected from the updated McMillan17 potential lies outside of the 3σ confidence level of the maximum probability determination of $v_{\rm esc}$. Interestingly, the distribution is not fully symmetric in z. The fact that $v_{\rm esc}$ does not match the expected value in many locations might indicate a bias in the estimated $v_{\rm esc}$ in the solar neighbourhood. Another possibility is that the decrease in the strength of the potential with z is less steep than expected for a spherical halo (e.g., pointing to a prolate halo or a weaker effect from the disc).

[2] https://faststars.space/

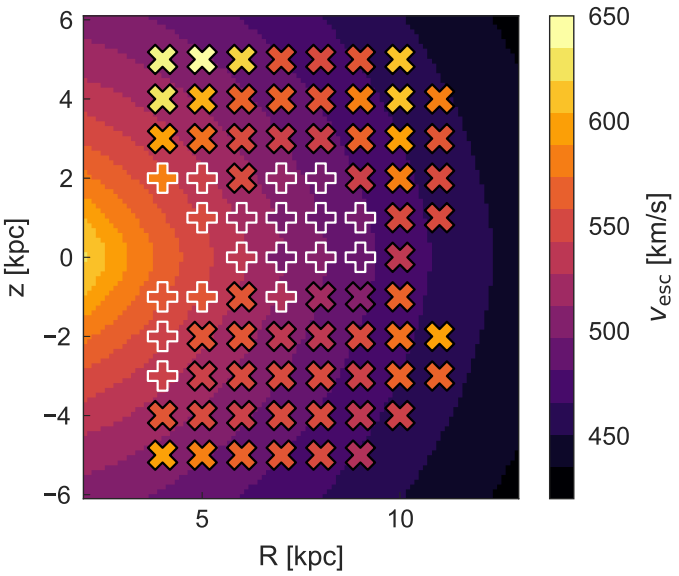

Fig. 13. Escape velocity determined using the LT90 method in rings of constant cylindrical R and z. The colour of the markers indicates the maximum probability value for v_{esc}. The colour map in the background shows expected isocontours for the escape velocity according to the McMillan17 potential, which assumes a spherical halo, but whose parameters we have updated with the values from Sect. 7.2. Volumes in which the 3σ levels enclose the background value are indicated with plus markers, the crosses indicate volumes in which the determined v_{esc} is higher than expected.

8. Conclusions

We used a sample of halo stars with high tangential velocities to constrain the escape velocity in the vicinity of the Sun and as a function of galactocentric distance. We applied the well-known LT90 method, which fits the high-velocity tail (i.e. above some velocity v_{cut}) of the velocity distribution with a power law of the form $(v_{esc} - v)^k$. In the process of applying the method, we identified a number of shortcomings.

The study presented here constitutes the first application of the method to a sample of stars using only tangential velocities. We have found that in practice, the estimated value for the parameter k is not exactly what is predicted by LT90 (namely $k_t = k$), except in the very tail of the distribution, in the limit where v_{cut} differs by 10% from v_{esc}. Unfortunately, the value of v_{cut} typically chosen is farther away from v_{esc} because enough stars ($\sim 10^4$) with high velocity need to be present in the sample to precisely estimate v_{esc}. A similar conclusion may be reached when the method is applied to radial velocity samples. This means that care is necessary when the values of k are compared for different studies in the literature. Fortunately, v_{esc} is not affected.

In addition, and as previously discussed in the literature, the v_{esc} determined via the LT90 method is most likely a lower limit. To correct for this bias, we tested the method on two mock *Gaia* catalogues from the Aurigaia project (Grand et al. 2018). In these simulated galaxies, the estimated v_{esc} are $\sim$10% lower than the true values, close to the 7% bias found in a similar study by Grand et al. (2019). Based on this result, we also quote the value obtained by applying a 10% correction when we report our estimates of the escape velocity. However, we note that there is no guarantee that the Milky Way halo is truncated at a similar level as the Aurigaia halos. The truncation of the velocity distribution depends on the (recent) assembly history of the Galaxy, and for the simulations, it might depend on the numerical resolution.

In the solar neighbourhood, using a 5D sample, we determine a very precise estimate of the escape velocity, $v_{esc} = 497^{+8}_{-8}\,\mathrm{km\,s^{-1}}$, and power-law exponent $k_t = 3.4^{+0.4}_{-0.3}$. The quoted uncertainties are given by the level where the probability has dropped to 61% of the maximum value (i.e. the $\sim 1\sigma$ level). These values agree well with previous works, but this is the first time that we can determine (a lower limit to) the escape velocity with such high confidence. This value for v_{esc} agrees remarkably well with the value that is obtained when we use a local sample of halo stars with full phase-space information. Applying the 10% fix would mean that the true escape velocity is $v_{esc}^{+10\%} = 552\,\mathrm{km\,s^{-1}}$.

We also determined v_{esc} as a function of galactocentric distance. We find that the escape velocity is higher in the inner halo than at the solar radius. This matches the behaviour expected from smooth Milky Way models well. Unexpectedly, for radii beyond 8 kpc, v_{esc} is also higher than at the solar radius (see Fig. 2). Indications of a similar trend were picked up by M18, but at a much lower significance level because of their limited sample size.

Interestingly, we find that the behaviour of v_{esc} outside of the solar radius is paired with a change in shape of the velocity distribution. For example, the tail of the velocity distribution becomes more exponential (and less power law-like) with galactocentric distance (see Fig. 2). Moreover, the velocities in the bins outside of 8 kpc show a higher degree of correlation as measured by the velocity correlation function. Therefore we conclude that the bump in v_{esc} in the outskirts is likely driven by a change in the kinematic properties of the sample as a function of galactocentric distance. Coincidentally, we found a similar effect in one of the Aurigaia halos we analysed, where a velocity bump (presumably related to a clump or a non-phase-mixed structure in the halo) dominates the tail near the escape velocity.

The estimated v_{esc} can be used to provide a very precise estimate of the mass of the halo of the Milky Way. To this end, we adjusted the halo component of the McMillan (2017) Milky Way potential (which is a spherical NFW profile) while keeping the other components fixed. The halo parameters that best fit the estimated $v_{esc}(r_\odot)$ are $M_{200} = 0.67^{+0.06}_{-0.06} \cdot 10^{12}\,M_\odot$ and $c = 15^{+1.2}_{-0.9}$, where we used $v_{circ}(r_\odot)$ as an additional constraint. When we applied the tentative 10% fix, we find that the best-fitting halo has $M_{200}^{+10\%} = 1.11^{+0.08}_{-0.07} \cdot 10^{12}\,M_\odot$ and $c^{+10\%} = 11.8^{+0.3}_{-0.3}$.

The method for determining v_{esc} consists of fitting the tail of the velocity distribution with a parametrised model. Using the best-fitting model obtained, we can also establish if there are any unbound stars in the solar neighbourhood. That is, we may calculate which stars have a high probability of having a true velocity that is higher than the determined escape velocity. We list these stars in Table 2. Their pseudo-velocities (without the line-of-sight velocity) suggest they are not all unbound, however: their velocity vectors point both inwards and outwards. If these high-velocity stars were truly escaping, we would expect them to all be on radially outbound trajectories. Nonetheless, it might be interesting to follow these stars up. When the tentative 10% fix is taken into account, only one candidate with a high probability of being unbound remains: *Gaia* DR2 `source_id` `2655054950237153664`. This star was first flagged as being unbound by Du et al. (2019).

Finally, we discussed a tentative method to probe the mass distribution of the Milky Way by determining v_{esc} as a function of (R, z). We find that escape velocity values are weakly asymmetric with respect to the Galactic plane, and also a tentative

indication that the halo may be prolate. However, for more robust conclusions, a larger sample with more accurate distances that probes deeper into the Milky Way is necessary. We hope that such a sample will become available with *Gaia* (e)DR3.

Acknowledgements. We gratefully acknowledge financial support from a VICI grant and a Spinoza Prize from the Netherlands Organisation for Scientific Research (NWO) and HHK is grateful for the support from the Martin A. and Helen Chooljian Membership at the Institute for Advanced Study. HHK thanks Daniel Foreman-Mackey, Scott Tremaine, and Rosemary Wyse for stimulating discussions on an early version of this work, that took place during the KITP Santa Barbara long-term program 'Dynamical Models for Stars and Gas in Galaxies in the *Gaia* Era', which was supported in part by the National Science Foundation under Grant No. NSF PHY-1748958. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (`http://www.cosmos.esa.int/gaia`), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, `http://www.cosmos.esa.int/web/gaia/dpac/consortium`). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. For the analysis, the following software packages have been used: `vaex` (Breddels & Veljanoski 2018), `scipy` (Virtanen et al. 2020), `numpy` (Van Der Walt et al. 2011), `matplotlib` (Hunter 2007), `jupyter notebooks` (Kluyver et al. 2016), and `Mathematica` (Wolfram Research Inc. 2020).

References

Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, ApJS, 235, 42
Aguilar, L. A., & White, S. D. M. 1986, ApJ, 307, 97
Beckmann, P. 1962, J. Res. Nat. Bur. Stand. Sect. D: Radio Propag., 66D, 231
Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611
Binney, J., & Tremaine, S. 1987, Galactic dynamics, v1 edn. (Princeton University Press)
Binney, J., & Tremaine, S. 2008, Galactic Dynamics (Princeton University Press), 885
Boubert, D., Guillochon, J., Hawkins, K., et al. 2018, MNRAS, 479, 2789
Boubert, D., Strader, J., Aguado, D., et al. 2019, MNRAS, 486, 2618
Bovy, J. 2015, ApJS, 216, 29
Breddels, M. A., & Veljanoski, J. 2018, A&A, 618, A13
Brown, W. R. 2015, ARA&A, 53, 15
Callingham, T. M., Cautun, M., Deason, A. J., et al. 2019, MNRAS, 484, 5453
Chan, V. C., & Bovy, J. 2020, MNRAS, 493, 4367
Cui, X.-Q., Zhao, Y.-H., Chu, Y.-Q., et al. 2012, Res. Astron. Astrophys., 12, 1197
Deason, A. J., Belokurov, V., Evans, N. W., et al. 2012, MNRAS, 425, 2840
Deason, A. J., Fattahi, A., Belokurov, V., et al. 2019, MNRAS, 485, 3514
Dierickx, M. I. P., & Loeb, A. 2017, ApJ, 847, 42
Du, C., Li, H., Yan, Y., et al. 2019, ApJS, 244, 4
Eilers, A.-C., Hogg, D. W., Rix, H.-W., & Ness, M. K. 2019, ApJ, 871, 120
Fragione, G., & Loeb, A. 2017, New Ast., 55, 32
Fritz, T. K., Di Cintio, A., Battaglia, G., Brook, C., & Taibi, S. 2020, MNRAS, 494, 5178
Gaia Collaboration (Prusti, T., et al.) 2016, A&A, 595, A1
Gaia Collaboration (Brown, A. G. A., et al.) 2018, A&A, 616, A1
Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179
Grand, R. J. J., Helly, J., Fattahi, A., et al. 2018, MNRAS, 481, 1726
Grand, R. J. J., Deason, A. J., White, S. D. M., et al. 2019, MNRAS, 487, L72
GRAVITY Collaboration (Abuter, R., et al.) 2018, A&A, 615, L15
Guillochon, J., Parrent, J., Kelley, L. Z., & Margutti, R. 2017, ApJ, 835, 64
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Nature, 563, 85
Hunt, J. A. S., Kawata, D., Grand, R. J. J., et al. 2015, MNRAS, 450, 2132
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jaffe, W. 1987, Struct. Dyn. Elliptical Galaxies, 127, 511
Katz, D., Sartoretti, P., Cropper, M., et al. 2019, A&A, 622, A205
Kluyver, T., Ragan-Kelley, B., Pérez, F., et al. 2016, Jupyter Notebooks-a Publishing Format for Reproducible Computational Workflows (IOS Press)
Kochanek, C. S. 1996, ApJ, 457, 228
Koppelman, H. H., & Helmi, A. 2021, A&A, 645, A69
Koppelman, H. H., Helmi, A., Massari, D., Roelenga, S., & Bastian, U. 2019, A&A, 625, A5
Kunder, A., Kordopatis, G., Steinmetz, M., et al. 2017, AJ, 153, 75
Leonard, P. J. T., & Tremaine, S. 1990, ApJ, 4, 486
Leung, H. W., & Bovy, J. 2019, MNRAS, 489, 2079
Lindegren, L., Hernandez, J., Bombrun, A., et al. 2018, A&A, 616, A2
Marchetti, T., Rossi, E. M., & Brown, A. G. A. 2019, MNRAS, 490, 157
Marrese, P. M., Marinoni, S., Fabrizio, M., & Altavilla, G. 2019, A&A, 621, A144
McMillan, P. J. 2017, MNRAS, 465, 76
Monari, G., Famaey, B., Carrillo, I., et al. 2018, A&A, 616, L9
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Piffl, T., Scannapieco, C., Binney, J., et al. 2014a, A&A, 562, A91
Piffl, T., Binney, J., McMillan, P. J., et al. 2014b, MNRAS, 445, 3133
Posti, L., & Helmi, A. 2019, A&A, 621, A56
Schaye, J., Crain, R. A., Bower, R. G., et al. 2014, MNRAS, 446, 521
Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829
Schönrich, R., Mcmillan, P., & Eyer, L. 2019, MNRAS, 487, 3568
Smith, M. C., Ruchti, G. R., Helmi, A., et al. 2007, MNRAS, 772, 755
Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, AJ, 132, 1645
Tremaine, S. 1987, Struct. Dyn. Elliptical Galaxies, 127, 367
Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Comput. Sci. Eng., 13, 22
Vasiliev, E. 2019, MNRAS, 482, 1525
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nat. Methods, 17, 261
Watkins, L. L., Evans, N. W., & An, J. H. 2010, MNRAS, 406, 264
Williams, A. A., Belokurov, V., Casey, A. R., & Evans, N. W. 2017, MNRAS, 468, 2359
Wilson, J. C., Hearty, F., Skrutskie, M. F., et al. 2010, in Ground-based and Airborne Instrumentation for Astronomy III, eds. I. S. McLean, S. K. Ramsay, H. Takami, et al., Int. Soc. Opt. Photonics, 7735, 77351C
Wolfram Research Inc. 2020, Mathematica, Version 12.1, champaign, IL, 2020
Xue, X. X., Rix, H. W., Zhao, G., et al. 2008, ApJ, 684, 1143
Zaritsky, D., Conroy, C., Zhang, H., et al. 2020, ApJ, 888, 114
Zinn, J. C., Pinsonneault, M. H., Huber, D., & Stello, D. 2019, ApJ, 878, 136