Mon. Not. R. Astron. Soc. **379,** 755–772 (2007) doi:10.1111/j.1365-2966.2007.11964.x

The RAVE survey: constraining the local Galactic escape speed

Martin C. Smith,[1⋆] Gregory R. Ruchti,[2] Amina Helmi,[1] Rosemary F. G. Wyse,[2] J. P. Fulbright,[2] K. C. Freeman,[3] J. F. Navarro,[4] G. M. Seabroke,[5] M. Steinmetz,[6] M. Williams,[3] O. Bienaymé,[7] J. Binney,[8] J. Bland-Hawthorn,[9] W. Dehnen,[10] B. K. Gibson,[11] G. Gilmore,[5] E. K. Grebel,[12] U. Munari,[13] Q. A. Parker,[9,14] R.-D. Scholz,[6] A. Siebert,[6] F. G. Watson[9] and T. Zwitter[15]

[1]*Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, the Netherlands*
[2]*Johns Hopkins University, 366 Bloomberg Center, 3400 North Charles Street, Baltimore, MD 21218, USA*
[3]*Research School of Astronomy and Astrophysics, Mount Stromlo Observatory, Cotter Road, Weston Creek, Canberra, ACT 72611, Australia*
[4]*University of Victoria, PO Box 3055, Station CSC, Victoria, BC, Canada V8W 3P6*
[5]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*
[6]*Astrophysikalisches Institut Potsdam, An der Sterwarte 16, D-14482 Potsdam, Germany*
[7]*Observatoire de Strasbourg, 11 Rue de L'Université, 67000 Strasbourg, France*
[8]*Rudolf Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP*
[9]*Anglo-Australian Observatory, PO Box 296, Epping, NSW 1710, Australia*
[10]*University of Leicester, University Road, Leicester LE1 7RH*
[11]*University of Central Lancashire, Preston PR1 2HE*
[12]*Astronomisches Institut, Universität Basel, Venusstrasse 7, Binningen CH-4102, Switzerland*
[13]*INAF Osservatorio Astronomico di Padova, Via dell'Osservatorio 8, Asiago I-36012, Italy*
[14]*Macquarie University, Sydney, NSW 2109, Australia*
[15]*Department of Physics, University of Ljubljana, Jadranska 19, Ljubljana, Slovenia*

Accepted 2007 May 9. Received 2007 April 17; in original form 2006 November 21

ABSTRACT

We report new constraints on the local escape speed of our Galaxy. Our analysis is based on a sample of high-velocity stars from the RAVE survey and two previously published data sets. We use cosmological simulations of disc galaxy formation to motivate our assumptions on the shape of the velocity distribution, allowing for a significantly more precise measurement of the escape velocity compared to previous studies. We find that the escape velocity lies within the range $498 < v_{\rm esc} < 608\,{\rm km\,s^{-1}}$ (90 per cent confidence), with a median likelihood of $544\,{\rm km\,s^{-1}}$. The fact that $v_{\rm esc}^2$ is significantly greater than $2v_{\rm circ}^2$ (where $v_{\rm circ} = 220\,{\rm km\,s^{-1}}$ is the local circular velocity) implies that there must be a significant amount of mass exterior to the solar circle, that is, this convincingly demonstrates the presence of a dark halo in the Galaxy. We use our constraints on $v_{\rm esc}$ to determine the mass of the Milky Way halo for three halo profiles. For example, an adiabatically contracted NFW halo model results in a virial mass of $1.42^{+1.14}_{-0.54} \times 10^{12}\,{\rm M}_\odot$ and virial radius of 305^{+66}_{-45} kpc (90 per cent confidence). For this model the circular velocity at the virial radius is $142^{+31}_{-21}\,{\rm km\,s^{-1}}$. Although our halo masses are model dependent, we find that they are in good agreement with each other.

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

1 INTRODUCTION

The existence of a dark halo around the Milky Way has been known for many years, although its nature is still uncertain. The mass and extent of this halo is a significant issue in astronomy. One vital tool which can be used to tackle this problem is also one of the simplest – the escape speed. If we are able to determine the escape speed at the solar neighbourhood, that is, the velocity that a star requires to escape the local gravitational field of the Milky Way, then this can provide important constraints on the extent of the dark halo. The reason why this quantity is so important is because it is the only local dynamical measurement that can be used to probe the extent of the mass distribution outside the solar circle. Unlike the circular velocity, which depends primarily on the mass interior to the solar circle, the escape velocity contains information about the

⋆E-mail: msmith@astro.rug.nl

© 2007 The Authors. Journal compilation © 2007 RAS

mass exterior to the solar circle. Although one needs a model for this mass distribution, the escape velocity (i.e. the local gravitational potential) can be used as a constraint from which it is possible to determine the total mass.

It is possible to use more distant measurements to investigate the extent of the Galactic halo. Unfortunately, gas rotation curves cannot be traced beyond the extent of gas in circular orbits, ∼20 kpc for the Milky Way. The task of tracing the rotation curve is also complicated by the fact that velocities have to be accompanied by distances (Binney & Dehnen 1997) and, in any case, our Galaxy does not appear to have an extended H I disc. As a consequence, most methods of probing the halo rely on satellites and globular clusters, whose velocities can be measured out to significantly greater Galactocentric distances. Many authors have used the velocities of the Milky Way's satellite galaxies and globular clusters in an attempt to constrain the total halo mass. Although numerous papers have dealt with this subject (Little & Tremaine 1987; Zaritsky et al. 1989; Kulessa & Lynden-Bell 1992; Kochanek 1996), two of the more recent ones to exploit the motions of satellite galaxies and globular clusters have concluded the total mass of the halo to be around $2 \times 10^{12}\,\mathrm{M}_{\odot}$: Wilkinson & Evans (1999) found a halo mass of $\sim 1.9^{+3.6}_{-1.7} \times 10^{12}\,\mathrm{M}_{\odot}$ by adopting a halo model which produces a flat rotation curve that is truncated beyond an outer edge; whereas Sakamoto, Chiba & Beers (2003), using a halo potential that gives a flat rotation curve, also included the velocities of field horizontal-branch stars to find a total halo mass of $2.5^{+0.5}_{-1.0} \times 10^{12}\,\mathrm{M}_{\odot}$ or $1.8^{+0.4}_{-0.7} \times 10^{12}\,\mathrm{M}_{\odot}$, depending on whether or not the analysis includes Leo I. Another complementary approach that can be adopted is to analyse the radial velocity dispersion profile of halo objects; Battaglia et al. (2005, 2006) used this method to determine a total mass of $0.5\text{–}1.5 \times 10^{12}\,\mathrm{M}_{\odot}$ depending on the chosen model for the halo profile (see also Dehnen, McLaughlin & Sachania 2006, for a reanalysis of this data set). The future for this field looks promising with space missions such as Gaia (due for launch 2011; Perryman et al. 2001; Wilkinson et al. 2005) and Space Interferometery Mission (due for launch ∼2015; Allen, Shao & Peterson 1998), since such missions will be able to provide accurate proper motion information to complement the existing radial velocity measurements; with such high-quality data it should be possible to determine the mass of the Milky Way to ∼20 per cent (Wilkinson & Evans 1999).

One can see that the current results mentioned above still produce a factor of ∼2 uncertainty in the mass of the Milky Way, due to the fact that the results are still model dependent and are hindered by small-number statistics concerning the relevant data sets. Therefore it would be very valuable if one could provide tight constraints on the local escape velocity in order to pin down the gravitational potential at this point. As far back as the 1920s samples of high-velocity stars were being used to estimate the local escape velocity (e.g. Oort 1926, 1928). As the 20th century progressed, many papers refined the estimate of $v_{\rm esc}$ (see Fich & Tremaine 1991, for a review), culminating in the final decade with the seminal work of Leonard & Tremaine (1990, hereafter LT90) and the subsequent refinement by Kochanek (1996, hereafter K96). These two papers concluded that, to 90 per cent confidence, the escape velocity lies in the range $450 < v_{\rm esc} < 650\,\mathrm{km\,s^{-1}}$ and $489 < v_{\rm esc} < 730\,\mathrm{km\,s^{-1}}$, respectively. Their conclusions are hampered by several problems: first, the paucity of high-velocity stars from which to estimate $v_{\rm esc}$; secondly, the fact that biases were introduced by selecting high-velocity stars from proper motion surveys and thirdly, the uncertainty in the assumptions regarding the underlying form of the tail of the velocity distribution. In this new century the difficulties posed by the first two issues are to some extent diminishing due to the large kinematically unbiased surveys that are now underway or planned, such as RAdial Velocity Experiment (RAVE; Steinmetz et al. 2006; see also Section 4.1), Sloan Extension for Galactic Understanding and Exploration (Beers et al. 2004) and Gaia (Perryman et al. 2001). The latter problem can be tackled through various methods; one such approach could be to use predictions from cosmological simulations to estimate the form of the velocity distributions. In this paper we shall make use of the advancement afforded to us by the RAVE survey, combined with the analysis of cosmological simulations, to refine the determination of $v_{\rm esc}$.

The outline of this paper is as follows. In Section 2 we review the analytical techniques that have been developed to constrain the escape velocity from a sample of velocity measurements. Then in the following section we assess various aspects of these techniques using cosmological simulations. In particular we use the simulations to estimate the expected shape of the tail of the velocity distribution, which is a crucial ingredient in the escape velocity analysis. In Section 4 we present the data that we will use to constrain the escape velocity and undertake some tests to ensure that these data are reliable. Our new data come from the RAVE project (Steinmetz et al. 2006), but are augmented with archival data from published surveys. In Section 5 we present our results and in Section 6 we consider some of the issues arising from or concerning these results; in particular, this latter section discusses the nature of our high-velocity stars (Section 6.1), the effect of the sample volume on the recovery of the escape velocity (Section 6.2), the possible contamination from unbound stars (Section 6.3) and also uses our new constraints on $v_{\rm esc}$ to investigate the total mass of the Galactic halo (Section 6.4). In Section 7 we conclude our paper with a brief summary.

2 ANALYSIS TECHNIQUES

2.1 Likelihood

The techniques that we apply in order to constrain the escape velocity ($v_{\rm esc}$) are based on those established by LT90. They parametrize the distribution of stellar velocities around $v_{\rm esc}$ according to the following formula:

$$f(|\boldsymbol{v}| \mid v_{\rm esc}, k) \propto (v_{\rm esc} - |\boldsymbol{v}|)^k, \quad |\boldsymbol{v}| < v_{\rm esc}, \tag{1}$$

$$f(|\boldsymbol{v}| \mid v_{\rm esc}, k) = 0, \quad |\boldsymbol{v}| \geqslant v_{\rm esc}, \tag{2}$$

where $|\boldsymbol{v}|$ is the speed of the star and k describes the shape of the velocity distribution near $v_{\rm esc}$. Note that this approach is only valid if the stellar velocities do indeed extend all the way to $v_{\rm esc}$. If there is any truncation in the velocities then this approach will underestimate the true $v_{\rm esc}$.

Under the assumption that the Jeans theorem can be applied to the system, equation (1) can be understood by considering the distribution function for the energies of the stars, ϵ. Assuming there is no anisotropy in the velocities, we can write the asymptotic form of the distribution function as a power law (K96),

$$f(\epsilon) \propto \epsilon^k, \quad \text{where } \epsilon = -(\Phi + |\boldsymbol{v}|^2/2), \tag{3}$$

where Φ corresponds to the potential energy and $|\boldsymbol{v}|^2/2$ to the kinetic energy. Again k describes the shape of the velocity distribution near $v_{\rm esc}$. Clearly $\Phi = -v_{\rm esc}^2/2$, which results in the following simple form for the asymptotic behaviour of the velocity distribution,

$$f(|\boldsymbol{v}| \mid v_{\rm esc}, k) \propto \left(v_{\rm esc}^2 - |\boldsymbol{v}|^2\right)^k = [(v_{\rm esc} - |\boldsymbol{v}|)(v_{\rm esc} + |\boldsymbol{v}|)]^k. \tag{4}$$

It could be argued that the velocity distribution near $v_{\rm esc}$ will not follow the form given in equation (4) since the orbital periods of

such high-velocity stars can be comparable to or larger than the age of the system (hence invalidating Jeans theorem). In this case the velocity distribution near $v_{\rm esc}$ would be a superposition of a small number of streams. This is likely to be an important issue only in the tail of the velocity distribution (i.e. for stars with $v \gtrsim 3\sigma$, where σ is the velocity dispersion of the system, see Helmi, White & Springel 2002). With this important caveat, in this paper we shall follow the traditional assumptions reflected in the above equations, and highlight possible limitations where appropriate.

Intuitively we expect that $k > 0$, in which case $f(|\boldsymbol{v}|) \rightarrow 0$ as $|\boldsymbol{v}| \rightarrow v_{\rm esc}$. However, this is not a necessary condition provided one accepts the presence of a discontinuity at $f(|\boldsymbol{v}| = v_{\rm esc})$. The likelihood of such distributions can be observationally constrained (as will be shown in Section 5, we find that values of $k < 0$ are strongly disfavoured).

Equation (4) can be further simplified in the limit of ($|\boldsymbol{v}| \rightarrow v_{\rm esc}$) to (1) by neglecting the $(v_{\rm esc} + |\boldsymbol{v}|)^k$ term, which has a strong systematic variation with k. However, throughout this paper, unless explicitly stated otherwise, we adopt (4).

In LT90, it was argued that radial velocities alone provided the most reliable constraints on $v_{\rm esc}$, since tangential velocities are much more uncertain due to inaccuracies in measuring proper motions and estimating distances. For example, a proper motion survey with typical errors of $\sim$10 per cent in both proper motion and distance would result in an error of $\sim$60 km s^{-1} for a star with velocity of 400 km s^{-1}, whereas radial velocities can be measured with an accuracy of typically less than a few km s^{-1} (see Section 4). In addition, since our work is motivated by the current advances in radial velocity surveys, we shall not incorporate the tangential velocities into our analysis.

To apply equation (4) to a sample of radial velocities, we must average over the unknown tangential velocities,

$$f_{\rm r}(v_{\rm r} \mid v_{\rm esc}, k) = \int f(|\boldsymbol{v}| \mid v_{\rm esc}, k)\delta(v_{\rm r} - \boldsymbol{v} \cdot \hat{\boldsymbol{n}})\, {\rm d}\boldsymbol{v}, \tag{5}$$

where $\hat{\boldsymbol{n}}$ is the unit vector along the line of sight. Clearly, unless the lines of sight are isotropically distributed, this equation is only valid for an isotropic distribution function. For our data set (the RAVE survey) we do not have all-sky coverage as RAVE only monitors the Southern hemisphere. However, even if the distribution function is anisotropic, this equation is still valid for a half-sky survey provided the mean motion is small; as we shall see in Section 3.2, we are dealing almost exclusively with halo stars whose mean motion is indeed small. Although the RAVE survey does not cover the entire Southern sky, any corresponding bias should be negligible compared to the relatively large statistical uncertainties in our final result. In any case, if the distribution function is assumed to be isotropic then the issue of sky coverage is immaterial since equation (5) is then valid for any sky coverage.

Evaluating equation (5) for the LT90 formalism (equation 1), we obtain

$$f_{\rm r}(v_{\rm r} \mid v_{\rm esc}, k) \propto (v_{\rm esc} - v_{\rm r})^{k+1}. \tag{6}$$

For the formalism given in equation (4), we integrate using cylindrical polar coordinates to obtain

$$f_{\rm r}(v_{\rm r} \mid v_{\rm esc}, k) \propto \left(v_{\rm esc}^2 - v_{\rm r}^2\right)^{k+1}. \tag{7}$$

In order to constrain $v_{\rm esc}$ and k for a sample of n stars, we employ the maximum likelihood method. The likelihood function $l(v_{\rm esc}, k)$ for the unknown quantities to be estimated can be defined as

$$l(v_{\rm esc}, k) = \prod_{i=1}^{n} f_{\rm r}(v_{{\rm r},i} \mid v_{\rm esc}, k), \tag{8}$$

where $f_{\rm r}(v_{{\rm r},i} \mid v_{\rm esc}, k)$ is given by either equation (6) or (7) and $v_{{\rm r},i}$ are the radial velocities of the individual n stars.

It can also be useful and sometimes more robust (especially for small samples) to apply prior knowledge about $v_{\rm esc}$ and k by way of Bayes' theorem. Given $v_{{\rm r},i}$ radial velocity observations, the probability of finding $v_{\rm esc}$ and k in the ranges $v_{\rm esc}$ to $v_{\rm esc} + {\rm d}v_{\rm esc}$ and k to $k + {\rm d}k$, respectively, is given as

$$P(v_{\rm esc}, k \mid v_{{\rm r},i=1,\ldots,n}) = \frac{P(v_{\rm esc})P(k)\Pi_{i=1}^{n} P(v_{{\rm r},i} \mid v_{\rm esc}, k)}{\int\int P\left(v'_{\rm esc}\right)P(k')\Pi_{i=1}^{n} P\left(v_{{\rm r},i} \mid v'_{\rm esc}, k'\right){\rm d}v'_{\rm esc}\,{\rm d}k'}. \tag{9}$$

$P(v_{{\rm r},i} \mid v_{\rm esc}, k)$ is proportional to $f_{\rm r}(v_{{\rm r},i} \mid v_{\rm esc}, k)$. $P(v_{\rm esc})$ and $P(k)$ are the a priori probabilities (i.e. prior knowledge) of $v_{\rm esc}$ and k, respectively. It is this equation, known as the posterior distribution, with chosen reference priors $P(v_{\rm esc})$ and $P(k)$, which will be maximized to determine $v_{\rm esc}$ and k.

In general the form of the distribution function given above (equation 3) is only true asymptotically as $v \rightarrow v_{\rm esc}$, and so to evaluate this probability we must impose a lower limit ($v_{\rm min}$) on the magnitude of the radial velocities that we will consider in our analysis; the choice of $v_{\rm min}$ is investigated further in Section 3.2. As will be shown later in Section 4, this form provides a good fit to the data over our chosen range of velocities.

Once the distributions given in equations (6) and (7) have been normalized such that $\int_{v_{\rm min}}^{v_{\rm esc}} P(v_{\rm r} \mid v_{\rm esc}, k)\, {\rm d}v_{\rm r} = 1$ we can then evaluate equation (9). One important factor in the evaluation of equation (9) is the choice of a priori probability; this is discussed in Section 2.3.

2.2 The bootstrap

A shortcoming of the method described above is that it assumes that the distribution of stellar velocities is in equilibrium and steady state. However, there are many potential limitations to this assumption, for example, velocity substructure, binary systems for which the centre-of-mass velocity has not been measured accurately or non-equilibrium stars such as those ejected from binary systems (including hypervelocity stars, which are plausibly ejected from the central regions of the Galaxy after interaction with the supermassive black hole at the Galactic Centre; e.g. Brown et al. 2006). All of these mechanisms for potential contamination can statistically perturb our underlying distribution function. They can change how the overall observed distribution function is populated (e.g. fluctuations due to the addition of some orbital velocity with amplitude dependent on phase and inclination of the binary orbit). However, these mechanisms are not well modelled, which makes it difficult to accommodate these effects. It is crucial therefore to employ a bootstrap method, which performs 'resampling' on our original data set to assess the sensitivity to possible non-equilibrium stellar velocities.

In brief, the bootstrap method involves randomly resampling the original data set (with replacement) to create artificial 'bootstrap' samples. Each bootstrap (re)sample provides one value each for k and $v_{\rm esc}$ that has maximum likelihood for that (re)sample of the data. This is repeated a large number of times (typically 5000), and the distribution of these maximum likelihood pairs is defined as the 'bootstrap distribution'. The bootstrap distribution can be used to estimate the sampling distribution of the maximum likelihood values for k and $v_{\rm esc}$, for example, as obtained in equation (9). This is extremely useful, because the fact we have a small sample means we cannot necessarily rely on maximum likelihood estimators having

converged to normality. Therefore, we can use the bootstrap distribution as a means to model the estimator sampling distributions.

Typically the bootstrap distribution is approximately normal, which allows us to rely less on hope that the original sample size is large enough for the central limit theorem to apply. The bootstrap distribution has bias if its mean values for k and $v_{\rm esc}$ are not the same as those found for the original sample. Bias and skewness in the bootstrap distribution are statistical signatures that can give us a general understanding of possible fluctuations due to any 'contaminating' velocities, as explained above. These signatures can be studied using standard techniques (Davison & Hinkley 1997). However, if the bootstrap distribution significantly deviates from a normal distribution, steps must be taken to model the distribution. The bootstrap s.e. is the s.d. of the bootstrap distribution and the 90th percentile confidence intervals correspond to the regions containing 90 per cent of the maximum likelihood pairs $(k, v_{\rm esc})$. This is useful, because now the bootstrap estimates of bias, s.e. and confidence interval endpoints are random variables. Their variances can be reduced by increasing the number of bootstrap samples. However, the quality of the bootstrap approximation still depends on the original sample size.

2.3 Considering the a priori probability for k and $v_{\rm esc}$

To evaluate equation (9) we must first choose the form of the priors for k and $v_{\rm esc}$. In previous work, the prior for $v_{\rm esc}$ was chosen to be $P(v_{\rm esc}) \propto 1/v_{\rm esc}$, since this is appropriate for a variable that varies continuously from 0 to infinity (Kendall & Stuart 1977). However, the choice of the prior for k requires more thought. In Section 2.1 we expressed the velocity distribution near the escape velocity assuming that the asymptotic behaviour follows a power law (equations 3 and 4). Unfortunately, the range of feasible values that one would expect for this exponent k is uncertain. Previous work has shown that to obtain meaningful limits on $v_{\rm esc}$ from a single maximum likelihood analysis of a solar neighbourhood sample of stars one must assume some prior for k, since only very large samples of stars would allow one to constrain both k and $v_{\rm esc}$ simultaneously. LT90 used Monte Carlo simulations to estimate that samples of $\sim$200 stars with accurate radial velocities will be required to constrain simultaneously k and $v_{\rm esc}$ for $v_{\rm min} = 250\,{\rm km\,s^{-1}}$ using the formalism given in equation (6).

For a given distribution function it is possible to predict the behaviour of k. For example, a Plummer model (1911) has the distribution function, $f(\epsilon) \propto \epsilon^{3.5}$, that is, $k = 3.5$. Similarly, to first-order approximation a Hernquist (1990) model has a distribution function with $k = 2.5$. However, LT90 argued that for a sample of stars free from selection effects, k should lie between 1 and 2. They noted that an isolated system that has undergone violent relaxation should have $k = 1.5$ (Aguilar & White 1986; Jaffe 1987; Tremaine 1987) and a collisionally relaxed system, such as a globular cluster, should have $k = 1$ (Spitzer & Shapiro 1972). K96 chose to assume a uniform prior on $k \in [0.5, 2.5]$, which brackets this violent relaxation value. Another approach one can employ to understand the nature of k is to use cosmological simulations; we implement this approach in Section 3.1.

Unlike previous work, we will also apply the bootstrap technique (Section 2.2) to ascertain whether our results are sensitive to possible problems with the data and allow confidence estimation for the maximum likelihood values of $v_{\rm esc}$ and k. Since we adopt bootstrap resampling we cannot simply adopt the reference priors chosen by LT90, without further investigation. In Appendix A we investigate the choice of prior by applying the bootstrap analysis to simulated data sets. From this work we conclude that a prior $P(k) \propto \sqrt{k}/(k+2)$, $P(v_{\rm esc}) \propto 1/(v_{\rm esc} - v_{\rm min})$ is the optimal choice, although for comparison we also evaluate the bootstrap constraints using the classical LT90 choice, that is, $P(k) \propto 1, P(v_{\rm esc}) \propto 1/v_{\rm esc}$.

3 EXPLOITING COSMOLOGICAL SIMULATIONS TO ASSESS ANALYSIS TECHNIQUES

In this section we assess the reliability of the above techniques in recovering the escape velocity. We do this by analysing a suite of four cosmological simulations. The first subsection deals with placing a priori limits on the possible values of the exponent k, which parametrizes the shape of the velocity distribution, while the second subsection deals with the effect of thin- and thick-disc contamination on the recovery of $v_{\rm esc}$ and the choice of $v_{\rm min}$.

3.1 The shape of the velocity distribution and its relation to the parameter k

In Section 2.3 we discussed the nature of the a priori information that we can incorporate into the evaluation of the likelihood function (equation 9). We explained the motivation behind the choice of prior in K96, that is, a uniform prior on $k \in [0.5, 2.5]$. However, with cosmological simulations it is possible to re-assess this choice. It was also noted in LT90 that if we have a system where the stellar velocities do not extend up to $v_{\rm esc}$, the statistical arguments of Section 2.1 will always underestimate the true $v_{\rm esc}$. So, in order to evaluate the ability of the above statistical method to recover $v_{\rm esc}$, it is vital that we understand the distribution of stellar velocities.

To test these issues we will analyse a set of four galaxies from cosmological simulations, labelled KIA1–KIA4 (Abadi, Navarro & Steinmetz 2006). These galaxies have been introduced in earlier papers and we advise interested readers to consult the following references for details regarding the code and the numerical set-up: Steinmetz & Navarro (2002), Abadi et al. (2003a,b), Meza et al. (2003, 2005). These simulations, which are carried out in a Lambda cold dark matter (ΛCDM) universe, include the gravitational effects of dark matter, gas and stars, and follow the hydrodynamical evolution of the gaseous component using the smooth particle hydrodynamics technique (Steinmetz 1996). The following cosmological parameters were adopted for the ΛCDM scenario: $\rm H_0 = 65\ km\ s^{-1}\ Mpc^{-1}$, $\sigma_8 = 0.9, \Omega_\Lambda = 0.7, \Omega_{\rm CDM} = 0.255, \Omega_{\rm bar} = 0.045$, with no tilt in the primordial power spectrum. All simulations started at redshift $z_{\rm init} = 50$ have force resolution of the order of 1 kpc, and mass resolution so that each galaxy is represented, at $z = 0$, with $\sim$125 000 star particles. The range of masses of the stellar particles in these four simulations at $z = 0$ is 10^5–$10^7\,{\rm M_\odot}$. Readers who wish to find a general overview of these simulations are recommended to consult section 2 of Abadi et al. (2006), while a more detailed description can be found in the references given above.

There are various evidences that demonstrate the reliability of these simulations. Section 3.2 of Abadi et al. (2003a) discusses the photometric properties. Here they note the good agreement between their surface brightness profile and a superposition of an $R^{1/4}$ spheroid plus exponential disc. The issue of the outer parts of the haloes of these simulated galaxies are discussed in Abadi et al. (2006). The mass distribution in these outer parts is seen to be consistent with the distributions of Galactic and M31 globular clusters. In addition, the surface brightness in these outer regions is well

Table 1. The properties of our simulated galaxies used in Section 3.1. The second column shows the rescaling factor that has been applied, which is chosen so as to enforce the circular velocity at the virial radius to be 140 km s^{-1}. The second and third columns refer to the virial mass and radius of the rescaled galaxies, respectively. The final column gives the escape velocity at the solar radius (8.5 kpc), as calculated according to the prescription described in Section 3.1.

Name	Rescaling factor	$M_{\rm vir}$ (10^{10} M$_\odot$)	$r_{\rm vir}$ (kpc)	$v_{\rm esc}$ (km s^{-1})
KIA1	1.30	120.59	230.73	581.06
KIA2	0.88	127.64	235.15	640.17
KIA3	0.87	129.87	236.50	652.62
KIA4	1.15	126.87	234.67	582.92

fitted by a Sersic law and is consistent with observations of both the haloes of M31 and the Milky Way. Support that the kinematics of these haloes are reliable can be drawn from the fact that they are consistent with the most recent observations of halo tracers (e.g. Battaglia et al. 2005, 2006).

In order to allow for a fair comparison between these four galaxies and the Milky Way, we rescaled each of them so that their circular velocity at their virial radius[1] is equal to 140 km s^{-1}. This value of 140 km s^{-1} is a typical value one could expect from using an NFW (Navarro et al. 1996,) profile to extrapolate the circular velocity at the solar neighbourhood, $\sim$220 km s^{-1}, to the virial radius. This results in a small rescaling factor (γ) of 0.9–1.3, where distances, velocities and masses are rescaled by $1/\gamma$, $1/\gamma$ and $1/\gamma^3$, respectively. The properties of our rescaled galaxies are given in Table 1.

For these simulated galaxies we need to estimate the escape velocity at the solar radius (which, for these simulations, we take to be 8.5 kpc from the centre of the model galaxy). To do this we define the escape velocity as the velocity required to get to 3 r$_{\rm vir}$, that is, we define the zero of the gravitational potential to be at this radius. Although this choice of 3r$_{\rm vir}$ is somewhat arbitrary, it is comparable to the separation of the Milky Way and its nearest massive neighbour M31. We then determine the potential at the solar radius to deduce the escape velocity. The resulting values for $v_{\rm esc}$ are between 581 and 653 km s^{-1} (see Table 1). The stellar particles contribute a significant amount of mass, namely, between 60 and 70 per cent of the total mass interior to the solar circle.

Given these four simulated galaxies, the first question which needs to be addressed is whether the velocity distribution of the stars differs from that of the dark matter particles, and in particular whether the stellar velocity distribution extends all the way to $v_{\rm esc}$. The stellar distribution must differ from that of the dark matter particles since the two populations have different density distributions.

We investigate this issue by plotting the velocity distributions for the stellar and dark matter components in our four galaxies. Since the disc component of these simulations do not always provide a wholly accurate representation of the Milky Way thin and thick discs, we only plot the distribution of stellar particles defined as belonging to the 'spheroid' population, according to the decomposition of Abadi et al. (2003b). This decomposition is based on the particles' orbital parameters and results in three populations: a halo

[1] We define the virial radius as the radius at which the mean density is 200 times the critical density for closure ($\rho_{\rm cr} = 3H^2/8\pi G \approx 7.9 \times 10^{-27}$ kg m^{-3} for $H = 65$ km s^{-1} Mpc^{-1}).

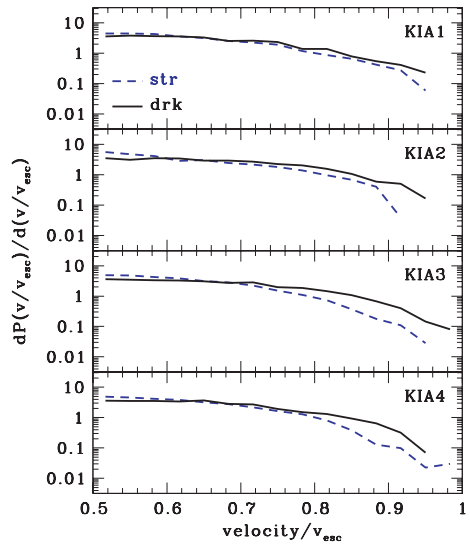

Figure 1. The velocity distribution of the stars (dashed) and dark matter particles (solid) from our four simulated galaxies. These distributions are for particles whose distance from the centre of the galaxy is between 3 and 14 kpc, and the stellar particles are those belonging to the 'spheroid' population.

with little net rotation, a thin disc with stars on nearly circular orbits and a thick disc with properties intermediate between the other two components. The halo component comprises between 10 and 40 per cent of the total number of stellar particles inside the virial radius.

Our exclusion of the disc populations from these simulations can be further justified when one considers that we expect our high-velocity sample of stars to contain only limited contamination from the Milky Way thick disc and negligible contamination from the thin disc (see Section 3.2). These velocity distributions are calculated for particles whose distance from the centre of the galaxy is between 3 and 14 kpc. Since the escape velocity varies over this large range of radii, we rescale each velocity by the mean escape velocity at that radius. The resulting distributions are shown in Fig. 1. It can easily be seen from this figure that the velocity distributions differ between the stellar and dark matter components, with the stellar particles' velocities concentrated towards smaller values. This is due to the difference between the density profiles of the stellar halo and dark components, since the stellar particles are more concentrated and hence have smaller velocities. At the solar radius, the respective density distributions have power-law slopes of ~ -2 for the dark matter and ~ -3 for the stellar halo, which is consistent with what is found for our Galaxy (e.g. Morrison et al. 2000). Despite the fact that the stellar velocities are smaller, it is encouraging to see that these velocities do indeed extend almost all the way to $v_{\rm esc}$. This is vital if we are to use this statistical analysis described in Section 2 to determine $v_{\rm esc}$ for the Milky Way. Although the decline in the velocity distribution for values close to $v_{\rm esc}$ appears to be sharper for the stellar particles, all of our galaxies have stellar particles with velocities greater than $0.9v_{\rm esc}$. As a consequence, even if there is a truncation beyond $0.9v_{\rm esc}$, the effect on $v_{\rm esc}$ cannot be more than 10 per cent.

Since these simulations indicate that the velocity distribution of stars clearly differs from that of the dark matter particles, one needs to ask how this affects the exponent k. To investigate this we evaluate the likelihood function (equation 9) for samples of stellar and dark

particles, but keeping $v_{\rm esc}$ fixed at the true value calculated above, that is,

$$P(k \mid v_{{\rm r},i=1,\ldots,n}) = \frac{P(k)\Pi_{i=1}^{n} P(v_{{\rm r},i} \mid k)}{\int P(k')\Pi_{i=1}^{n} P(v_{{\rm r},i} \mid k')\,{\rm d}k'}, \quad (10)$$

where we have assumed a uniform prior on k.

To evaluate equation (10) we need to construct a sample of particles equivalent to our observed solar neighbourhood stars. This is done by calculating the radial velocities of particles contained within a series (in azimuth around the galaxy) of non-overlapping spheres of radius 2 kpc located at a distance of 8.5 kpc from the galactic centre. These spheres are all chosen to lie in the plane of the galaxy. Since the presence of significant mean orbital rotation in our sample of stars will invalidate the assumptions made in Section 2, we only use the stellar particles defined as belonging to the non-rotating 'spheroid' population, according to the decomposition of Abadi et al. (2003b). In order to increase our statistics we also include spheres of radii 1 and 3 kpc located at radii 4.25 and 12.75 kpc; we ensured that there is no trend between the exponent k and the radial location of the spheres within our range of values. To combine spheres located at different radii, the velocities of particles within each sphere are rescaled by the escape velocity at the centre of the sphere.

By evaluating equation (10) we are able to deduce the likelihood estimates for the value of k for the stellar and dark matter populations for each of our four simulated galaxies. The resulting likelihood distributions of k are shown in Fig. 2.

As can be seen from this figure, the value of k significantly differs for the two populations. For these simulated galaxies we can see that the dark matter particles appear to match the predictions from the literature, that is, $k \in [0.5, 2.5]$ (see above). However, the stellar sample does not appear to match this expectation; for these k is shifted towards significantly larger values. Since all attempts to constrain $v_{\rm esc}$ must be based on stellar samples, it means that a uniform prior on $k \in [0.5, 2.5]$ will not provide an accurate measure of the escape velocity. To estimate a better choice of prior from these stellar likelihood intervals, we take the same range of k but centre it on the mean likelihood of 3.7, that is, $k \in [2.7, 4.7]$.

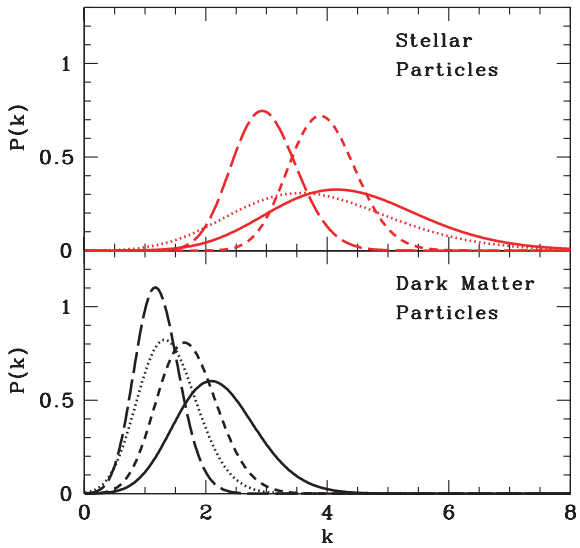

Figure 2. The likelihood estimates for the exponent k, which denotes the shape of the velocity distribution. This is shown for both the stellar (thick) and dark matter (thin) high-velocity 'solar neighbourhood' samples. The solid, dotted, short-dashed and long-dashed lines represent simulated galaxies KIA1, KIA2, KIA3 and KIA4, respectively.

At first sight it may appear from Fig. 2 that the constraints on k for galaxies KIA1 and KIA2 differ slightly from the other two galaxies. However, this can be understood when one considers the fact that KIA1 and KIA2 have fewer total numbers of stellar particles than KIA3 or KIA4. As a consequence the constraints on k appear to be shifted towards higher values for these galaxies. This is simply due to the fact that when one has fewer particles, although the regime $k \rightarrow 0$ is still well constrained (since negative k is very strongly disfavoured), it is harder to constrain large values of k. Given this fact, an alternative approach to determine a prior could be to take the range covered by the joint 90 per cent confidence interval from KIA3 and KIA4 alone. If we do this we obtain a range $k \in [2.3, 4.7]$, which is very similar to that found above. Given the fact that there is very little difference, we choose to adopt the former range (i.e. $k \in [2.7, 4.7]$) for the subsequent analysis.[2]

It is important to check whether the results presented in Fig. 2 are sensitive to our choice of rescaling, since there are a variety of different rescalings that one can adopt. To test this we adopt an alternative rescaling that enforces a circular velocity of 220 km s^{-1} at 8.5 kpc. This results in a rescaling factor of 1.11, 0.78, 0.81, 0.83 for galaxies KIA1–KIA4, respectively, which is not far removed from our original values (see Table 1). Reassuringly, when we repeat the above analysis the resulting likelihood estimates for k are wholly consistent for each of our simulated galaxies.

3.2 The choice of $v_{\rm min}$

3.2.1 Estimating the level of contamination from the disc

We wish to address the choice of $v_{\rm min}$, that is, the minimum velocity that is included in a sample of 'high-velocity' stars. For their analysis of radial velocities, LT90 chose a value of 250 km s^{-1}, while K96 chose 300 km s^{-1}, preferring to reduce any systematic errors at the expense of increased statistical errors. When considering the choice of $v_{\rm min}$ it is crucial to note that the analysis discussed in Section 2 relies on the assumption that the velocities are isotropic and there are no mean motions. This is especially important because RAVE fields are not isotropically distributed on the sky; if there is any net rotation then the averaging over the unknown tangential velocities performed in equation (5) is in error. While we might expect the mean motion of the halo stars to be negligible compared to the relatively large statistical errors that will be present in any current analysis, the possible presence of the rotating thin- and thick-disc population in any sample cannot be ignored.

In the absence of any additional information (such as metallicities, etc.), the best discriminator which can be applied to radial velocity samples is one based on the radial velocity itself. Given the expected small dispersion in velocities of disc stars, one would hope that a sufficiently large value of $v_{\rm min}$ will be able to filter out the disc population. We will investigate this issue by first estimating the fraction of disc stars that we may expect in our high-velocity sample. To do this we construct a simple toy model containing each of the three components, that is, the thin disc, thick disc and halo.

[2] If we repeat the subsequent analysis using this prior, we obtain the following constraints on $v_{\rm esc}$, that is, $494 < v_{\rm esc} < 605$ km s^{-1} (90 per cent confidence), with a median likelihood of 540 km s^{-1}. These constraints are practically indistinguishable from those found by adopting the prior using all four galaxies (see Section 5).

Before one can estimate the fraction of disc stars that may be present in our high-velocity sample, one first must ask what the total fraction of disc stars will be over all velocities. From the literature we find that in the solar neighbourhood the mass density of the three components can be estimated as 3.8×10^{-2}, 1.3×10^{-3} and $1.5 \times 10^{-4}\,\mathrm{M}_{\odot}\,\mathrm{pc}^{-3}$ for the thin disc, thick disc and halo, respectively (Jahreiß & Wielen 1997; Fuchs & Jahreiß 1998; Ojha 2001). However, our sample of stars will be drawn from a finite volume, and so we would expect to find significantly more halo stars than this value of 0.4 per cent. To obtain the final fraction we need to know how the mass density of the components varies spatially and also the approximate boundary of our sample volume. We take a value of 2 kpc for this latter quantity and estimate the former by assuming that the mass density has negligible radial dependence in the plane and is solely determined by the vertical scaleheight (for which we adopt the values of 260 and 860 pc for the thin and thick discs, respectively, consistent with the local normalizations; Ojha 2001). We assume that the halo density is constant. These assumptions should be valid provided our RAVE sample volume is not too large. The final property that we include is that our RAVE sample excludes stars within 25° of the plane. If we fold all this information into a toy model, then the resulting value for the fractional contribution of the three components before considerations of kinematics becomes 73.5, 19.5 and 7.0 per cent for the thin disc, thick disc and halo, respectively. Clearly, we cannot calculate these fractions with any high degree of certainty since we have not accounted for the luminosity function of the three components and also the local mass densities are subject to much discussion. However, we believe that these values provide a reasonable estimate and should be suitable for our purposes.

The final ingredient necessary for calculating the expected level of contamination from disc stars is to include the kinematics of the three components. This is done by simply assuming Gaussian velocity distributions with mean orbital rotation and dispersions as given in Table 2. The halo parameters are taken from Chiba & Beers (2000). Although their value of $\langle v_\phi \rangle = 40\,\mathrm{km\,s^{-1}}$ is in conflict with our assertion that the halo must have no net rotation, this value is clearly small compared to the dispersions. The thin- and thick-disc dispersions are deduced from the data of Nordström et al. (2004) following the prescription of Binney & Merrifield (1998, section 10.4; see also Reddy, Lambert & Allende Prieto 2006); namely, that thin-disc stars satisfy age < 10 Gyr, [Fe/H] > -0.4, while thick-disc stars satisfy age > 10 Gyr, [Fe/H] < -0.4.

Given this toy model we can now estimate the contribution of each component in a sample of stars with radial velocity greater than some cut-off v_{min}. To do this we take each observed star in the RAVE catalogue and generate 100 mock stars along this line of sight according to our toy model. We can then calculate the fractional contribution of each component as a function of v_{min}, which is shown in Fig. 3. From this figure it can be seen that for $v_{\mathrm{min}} > 250\,\mathrm{km\,s^{-1}}$ the contamination from the thin disc is negligible. However, thick-disc stars will be present for velocities as high as $400\,\mathrm{km\,s^{-1}}$. As stated above, K96 chose a value of $v_{\mathrm{min}} = 300\,\mathrm{km\,s^{-1}}$, for which our toy model predicts a contamination of approximately 10 per cent. The contamination rises to ∼20 per cent if $v_{\mathrm{min}} = 280\,\mathrm{km\,s^{-1}}$ and ∼40 per cent if $v_{\mathrm{min}} = 250\,\mathrm{km\,s^{-1}}$. Although the optimal value of v_{min} is dependent on the line of sight (e.g. as a function of longitude), we follow previous studies and do not adopt a spatially varying v_{min}.

Table 2. Velocity dispersions of the three components in the toy model of Section 3.2. The halo dispersions are taken from Chiba & Beers (2000). The thin- and thick-disc dispersions are deduced from the data of Nordström et al. (2004) following the prescription of Binney & Merrifield (1998; section 10.4); namely, that thin-disc stars satisfy age < 10 Gyr, [Fe/H] > -0.4, while thick-disc stars satisfy age > 10 Gyr, [Fe/H] < -0.4. We adopt a value of $220\,\mathrm{km\,s^{-1}}$ for the local standard of rest and take the solar motion from Dehnen & Binney (1998).

Component	$\langle v_\phi \rangle$ ($\mathrm{km\,s^{-1}}$)	σ_R ($\mathrm{km\,s^{-1}}$)	σ_ϕ ($\mathrm{km\,s^{-1}}$)	σ_z ($\mathrm{km\,s^{-1}}$)
Thin disc	209	29	18	14
Thick disc	174	68	55	38
Halo	40	141	106	94

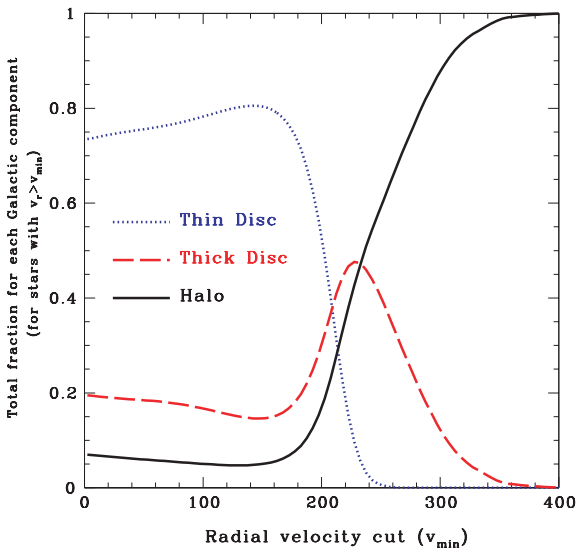

Figure 3. The fractional contribution of each Galactic component as a function of v_{min}, that is, the minimum radial velocity of the sample. These results come from the toy model described in Section 3.2.

3.2.2 *Quantifying the effect of disc contamination*

We now need to test whether this level of contamination will affect our results. To do this we return to the simulations of Abadi et al. (2006) that were introduced in Section 3.1. Since the disc components in these galaxies do not always provide a wholly accurate representation of the Milky Way disc, we instead include the contamination from the thick disc according to the toy model described above. We investigate two fiducial cases, namely, thick-disc contamination of 10 and 20 per cent (note that for velocity cuts of $v_{\mathrm{min}} \geqslant 250\,\mathrm{km\,s^{-1}}$ we predict that the contamination from the thin disc will be negligible). For the halo component we take the 'spheroid' star particles as defined in Section 3.1. In addition, we restrict ourselves to the analysis of only two of our simulated galaxies (KIA3 and KIA4) since they have significantly more particles than the other two. In order to make a fair comparison between the two galaxies we rescale the velocities of the halo component so that $v_{\mathrm{esc}} = 600\,\mathrm{km\,s^{-1}}$ and ensure that each sample contains 100 stars. For the minimum velocity cut we take a value of $v_{\mathrm{min}} = 300\,\mathrm{km\,s^{-1}}$.

Given these two samples of simulated stars with 10 and 20 per cent thick-disc contamination, we evaluate equation (9). The results from this analysis indicate that the presence of a thick-disc population has a noticeable, but not hugely significant, effect. The peak of the likelihood distribution is shifted towards larger k and v_{esc}, although for a given value of k the predicted value of v_{esc} is lower. Therefore, when one applies the prior calculated in Section 3.1 (i.e. $k \in [2.7, 4.7]$) the predicted escape velocity is reduced. For a

thick-disc contamination of 20 per cent, this shift is approximately $-30\,\mathrm{km\,s^{-1}}$ for both KIA3 and KIA4; for a contamination of 10 per cent this shift is reduced to $-15\,\mathrm{km\,s^{-1}}$ for KIA3 and $-10\,\mathrm{km\,s^{-1}}$ for KIA4.

In conclusion, we concur with K96 and propose that a value of $v_{\mathrm{min}} = 300\,\mathrm{km\,s^{-1}}$ is a safe choice which should allow us to obtain a reliable and robust determination of the escape velocity.

4 DATA

The following sections describe the data that we will use to constrain the escape velocity.

4.1 The RAVE project

The RAVE is an ambitious survey to measure radial velocities and stellar atmosphere parameters (temperature, metallicity, surface gravity) of up to one million stars using the 6dF multi-object spectrograph on the 1.2-m United Kingdom Schmidt Telescope of the Anglo-Australian Observatory. The RAVE survey is ideal for constraining v_{esc} because it is a magnitude-limited survey ($9 < I < 12$), which means that it avoids any kinematical biases, unlike catalogues constructed using high proper motion stars. Given the RAVE magnitude range, the catalogue will consist of giants up to a distance of a few kpc and local dwarfs. The project is described in detail in the paper accompanying the first data release (Steinmetz et al. 2006). It is foreseen that the RAVE project will run until 2010. This survey will represent a giant leap forward in our understanding of our own Milky Way galaxy, providing a stellar kinematic data base larger than any other survey proposed for this coming decade. For our analysis we use an internal data release (Summer 2006) containing radial velocities for over 50 000 stars covering an area of nearly 8000 $\mathrm{deg^2}$.

4.2 Our high radial velocity RAVE sample

To construct a catalogue of high radial velocity objects we must first convert our heliocentric velocities into Galactic standard of rest-frame velocities (also known as Galactocentric radial velocities; see e.g. equation 10-8 of Binney & Tremaine 1987). This is done by taking the local standard of rest to be $220\,\mathrm{km\,s^{-1}}$ and the solar motion to be $(10.00, 5.25, 7.17)\,\mathrm{km\,s^{-1}}$ (Dehnen & Binney 1998). Note that unless stated otherwise, all velocities will be quoted in the Galactic standard of rest frame. We then take all stars greater than $v_{\mathrm{min}} = 300\,\mathrm{km\,s^{-1}}$.

It is vital to ensure that our sample is free from contamination from spurious measurements. To this end we enforce a high threshold for the value of the Tonry–Davis (Tonry & Davis 1979) correlation coefficient between the observed spectrum and the template spectrum ($R > 10$). We perform additional checks to verify the reliability of the data in the following section. It should also be noted that currently the metallicities and gravities have not been accurately determined for the RAVE data, so we will only use the radial velocities in our analysis of the escape velocity.

A total of 16 stars pass these criteria. Two of these have $10 < R < 15$, but the velocities for both stars have been confirmed by follow-up observations (see the following section).

4.3 Data validation

This subsection discusses the methods that have been adopted to verify the reliability of the RAVE radial velocity measurements of our sample of 16 high-velocity stars. For a comprehensive discussion of the quality of the RAVE data in general, we refer readers to the data release paper (Steinmetz et al. 2006).

We have obtained follow-up observations of a selection of RAVE stars using the 2.3-m Advanced Technology Telescope at Siding Springs Observatory (Steinmetz et al. 2006). The data were taken using the double beam spectrograph, giving slit spectra between 8000 and 8900 Å at a resolution similar to the RAVE observations ($0.55\,\text{Å pix}^{-1}$). In general, the radial velocity measurements show reasonable agreement with the RAVE values to within $\sim 3\,\mathrm{km\,s^{-1}}$. Given the fact that we wish to use only the highest radial velocity stars for our v_{esc} analysis, this level of discrepancy is negligible. Among this selection of re-observed stars there were 12 of our high radial velocity RAVE stars, all of which show good agreement. A smaller subsample of RAVE stars were also observed with the echelle spectrograph on the 3.5-m Telescope at Apache Point Observatory (APO), New Mexico. The single-slit echelle has a spectral resolution of 37 000 and covers the entire optical wavelength range (3500 Å–1 μm). Nine of the observed stars were high-velocity stars. Similarly to the 2.3-m data, the uncertainties are dominated by RAVE errors, and the radial velocity measurements agreed with the RAVE velocities within 3σ. In Fig. 4 we show the results of the follow-up observations for these 11 stars. All show good agreement, with none having a discrepancy of more than $7\,\mathrm{km\,s^{-1}}$.

Although we found good agreement for our follow-up velocities, there were two exceptions not mentioned above. Two of our stars showed behaviour consistent with being binary stars. One of these stars (C2129509−080418) has four observations taken over a period of 18 months, while another (C1437570−280154) has two observations taken over a period of 26 months. From our measured velocities we can place a lower limit on the semi-amplitude of the velocity variations and we find values of 10.6 and $15.4\,\mathrm{km\,s^{-1}}$ for C1437570−280154 and C2129509−080418, respectively. These values are typical of single-lined spectroscopic binaries in Latham et al. (2002), for which 68 per cent have semi-amplitudes within the range 3–$15\,\mathrm{km\,s^{-1}}$. Therefore if we assume that the actual

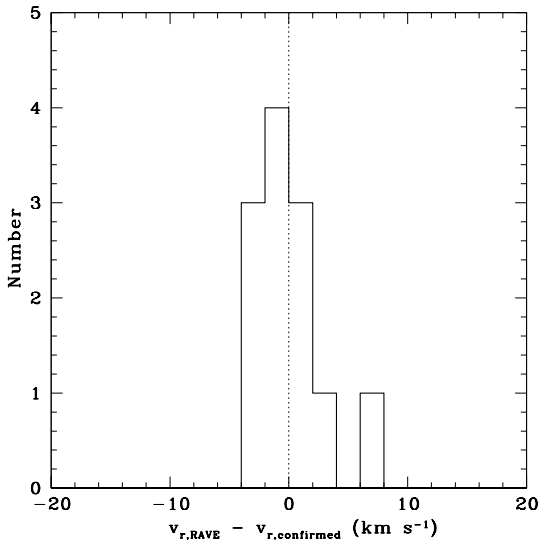

Figure 4. Results from the follow-up work described in Section 4.3 for 12 of our high radial velocity RAVE stars. The horizontal axis shows the difference between the velocity as reported in the RAVE catalogue and the weighted mean of all velocities taken during the follow-up campaign. Note the good agreement between the two measurements. Typical errors are $\sim 2\,\mathrm{km\,s^{-1}}$ for the RAVE catalogue and $< 1\,\mathrm{km\,s^{-1}}$ for the follow-up velocities.

Table 3. The high-velocity sample of RAVE stars. See Section 4 for details about these stars. The identifiers are taken from the RAVE input catalogue, while the *I*-band magnitudes are from the DENIS catalogue (Epchtein et al. 1997). Note that the magnitudes of three stars, C1437570−280154, C2129509−080418 and T7524_00065_1, are not available in the DENIS catalogue and so the values for these stars are taken from USNO-B (Monet et al. 2003). Column 1 shows the Galactocentric radial velocity and column 2 shows the error in the velocity; where we have more than one observation for a star we adopt the weighted mean. Two stars in our sample are believed to be binaries (C1437570−280154 and C2129509−080418) and for these we give an estimate of the centre of binary mass motion (see Section 4.3).

Name	v_r (km s^{-1})	δv_r (km s^{-1})	RA (J2000) (°)	Dec. (J2000) (°)	ℓ (°)	b (°)	I (mag)	Total number of observations
C1012254−203007	314.7	0.5	153.106 25	−20.501 94	259.83	28.76	12.2	3
C1022369−140345	391.8	0.5	155.654 08	−14.062 67	257.04	35.23	12.0	2
C1032508−244851	315.3	1.4	158.211 99	−24.814 39	267.08	28.24	11.9	1
C1100242−024226	344.0	0.7	165.101 14	−2.707 22	256.67	49.92	11.7	4
C1250398−030748	309.8	1.1	192.666 12	−3.130 19	302.55	59.74	12.1	2
C1437570−280154	−317.6	–	219.487 79	−28.031 69	329.89	29.21	11.3	2
C1508217−085010	−304.3	1.3	227.090 82	−8.836 22	350.44	41.07	11.3	2
C1519196−191359	−448.8	0.7	229.831 92	−19.233 09	344.64	31.41	12.1	4
C1536201−144228	−345.9	1.0	234.083 97	−14.707 92	351.75	32.12	12.4	3
C2041305−113156	−345.8	0.4	310.377 35	−11.532 44	34.81	−29.61	11.8	3
C2129509−080418	305.91	–	322.4620	−8.0717	44.9807	−38.7363	12.5	4
C2214430−480306	−332.9	1.8	333.6800	−48.0519	347.71	−53.16	11.7	1
T4931_00266_1	390.0	0.4	175.273 88	−1.545 36	269.56	56.70	10.6	4
T7524_00065_1	322.7	0.8	4.618 00	−39.009 44	330.93	−76.27	9.8	3
T7535_00160_1	311.3	0.8	10.376 17	−40.981 89	310.70	−76.00	10.2	3
T8395_01513_1	−320.5	0.9	300.801 70	−48.107 83	351.10	−31.50	10.0	3

semi-amplitude for our two RAVE stars is close to our observed lower limit, we can estimate the centre-of-mass velocity by averaging the highest and lowest velocities for each star. Given the statistics of the Latham et al. (2002) sample, in particular that less than 10 per cent of their binaries have semi-amplitude greater than 20 km s^{-1}, we can predict that the probable error for the estimated velocity of our RAVE stars is less than 10 km s^{-1}. Both of these centre-of-mass velocities are above $v_{\rm min}$ and hence we retain these for our analysis. Our fraction of binary stars (2 out of 16) is comparable to the fraction of single-lined spectroscopic binaries found in the Latham et al. (2002) sample (171 out of 1359).

In summary, out of our final sample of 16 RAVE stars with Galactocentric velocity greater than 300 km s^{-1}, 14 have repeat observations (including the two binary stars). The total number of observations for each star varies from one to four (see Table 3).

4.4 The final high-velocity RAVE sample

Given this high-quality RAVE data, we are now able to construct a final catalogue of high radial velocity stars. Since many of our stars now have repeat observations, we choose to adopt the weighted mean of all measurements as our definitive velocity, with the exception of the two binary stars for which we give our estimate of the centre of binary mass motion. These are given in Table 3 and the velocity distribution is shown in the inset of Fig. 6. In Fig. 5 we show how the radial velocities vary as a function of Galactic longitude. This plot clearly shows the signature of the Galactic disc and from this one can obtain an understanding of why a value of $v_{\rm min} \approx 250$ km s^{-1} results in significant contamination from the disc; if the mean rotational velocity of our sample is close to zero (as we would like for a halo population), then there should be an equal number of stars with positive and negative radial velocity for a given longitude. However, for $l \approx 270$ there is clearly a greater number of stars with radial velocities in the range $v_r \in (-300, -250)$ compared to $v_r \in (250, 300)$, indicating contamination by a rotating component. Note

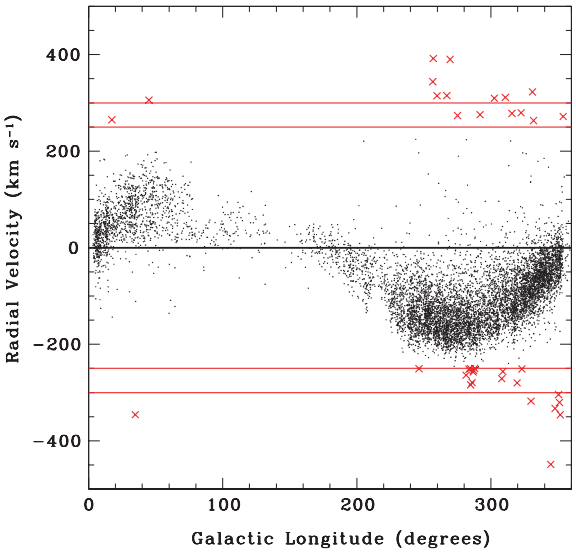

Figure 5. The relation between radial velocity (corrected for solar motion) and longitude for stars in the RAVE catalogue. Note that the signature of the disc is clearly visible. The horizontal lines correspond to $v_r = -300$, $-250, 0, +250, +300$ km s^{-1}. The crosses simply denote stars with $|v_r| > 250$ km s^{-1}.

that this asymmetry is not evident for stars with $|v_r| > 300$ km s^{-1}, which supports our choice of $v_{\rm min} = 300$ km s^{-1}.

4.5 Augmenting our high-velocity sample with stars from archival surveys

Since we would like our sample of stars to be as large as possible, we incorporate additional stars from the Beers et al. (2000) catalogue of metal-poor stars. It is important to note that this sample is

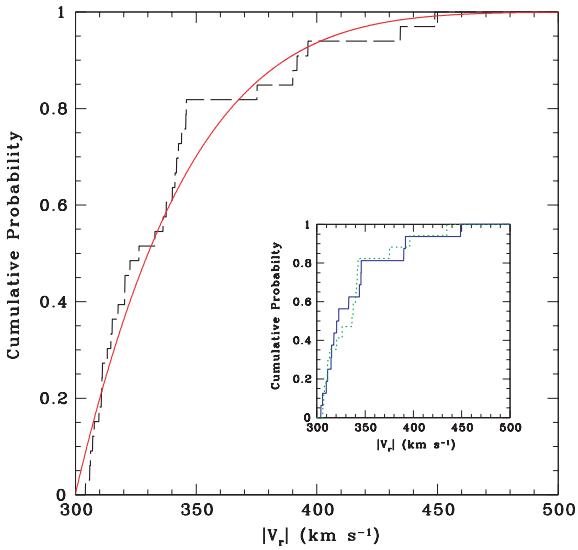

Figure 6. The cumulative distribution for the magnitude of the radial velocities. The dashed line shows the velocities for our sample of high radial velocity stars, which has been constructed by augmenting a sample of RAVE stars with a supplementary archival survey (Beers et al. 2000; see Section 4). The solid line corresponds to the maximum likelihood evaluated in Section 5 ($v_{\rm esc} = 616\,{\rm km\,s^{-1}}, k = 7.1$). The inset shows the comparison between the 16 RAVE stars (solid) and the 17 archival stars (dotted), showing no clear discrepancy between these two samples.

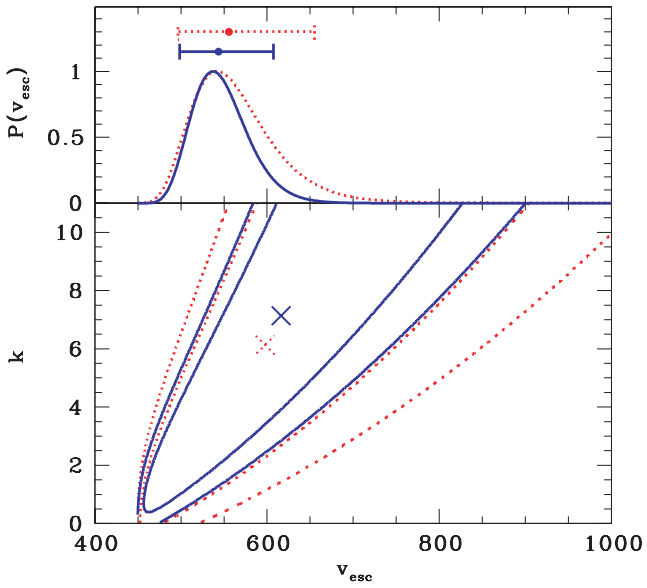

Figure 7. The lower panel shows the 2D likelihood contours that can be placed on the local escape velocity ($v_{\rm esc}$) and the shape of the velocity distribution (k; see Section 2) from our combined high-velocity sample. The cross corresponds to the peak likelihood, while the contours denote 10 and 1 per cent of this peak likelihood value. The upper panel shows the likelihood distribution for $v_{\rm esc}$ obtained by assuming a uniform prior on $k \in [2.7, 4.7]$ (see Section 3.1); the corresponding error bar shows the 90 per cent confidence interval. The dotted quantities show the results from a sample containing only the high-velocity RAVE stars, that is, a smaller sample of 16 stars.

kinematically unbiased, which is important if we are to combine data sets in this way. This sample is ideal since it contains metal-poor stars, which are preferentially halo stars. The Beers et al. (2000) sample provides a total of 17 stars faster than $300\,{\rm km\,s^{-1}}$, once we have removed three stars for which the distance estimate indicates that they are farther than 5 kpc away (all of the retained stars have distances of less than 2.5 kpc). These archival stars are given in Table B1.

This brings the total number of stars in our full augmented sample to 33, which is a significant improvement on the number of stars used in LT90 (15 stars with $v_{\rm r} > 250\,{\rm km\,s^{-1}}$) and K96 (10 stars with $v_{\rm r} > 300\,{\rm km\,s^{-1}}$).

The velocity distribution for this larger augmented sample is shown in Fig. 6. Note that the inset of this figure compares the distribution of RAVE stars with the distribution of our archival stars from Beers et al. (2000). The Kolmogorov–Smirnov test indicates no significant discrepancy between these two distributions (14 per cent probability that they come from different distributions), so there is no inconsistency in combining the two data sets. In addition, similar to the RAVE sample (as was shown in Fig. 5), we reassuringly find no correlation between radial velocity and Galactic longitude. In Section 5 we check that the process of combining data sets does not introduce any obvious bias by carrying out the likelihood analysis on both the combined sample as well as a sample consisting solely of our 16 RAVE stars.

5 RESULTS

We now wish to use the sample of high-velocity stars described above to constrain the local escape velocity. To do this we apply the techniques outlined in Section 2.

5.1 Maximum likelihood analysis of the sample

Evaluating equation (9) results in 2D likelihood contours for k and $v_{\rm esc}$, shown in the lower panel of Fig. 7. The peak of the likelihood lies at $v_{\rm esc} = 616\,{\rm km\,s^{-1}}$ and $k = 7.1$. Although this value of k is slightly greater than that predicted from our cosmological simulations in Section 3.1, we find that our peak likelihood is relatively broad. In addition, as was discussed in Section 3.2.2, although our predicted level of thick-disc contamination should have little effect on our $v_{\rm esc}$ constraints, any contamination will shift the peak likelihood to higher values of k. In Fig. 6 we show the probability distribution for the radial velocities (given by equation 7) corresponding to this peak likelihood.

As can be seen from Fig. 7, given the size of our sample of high radial velocity stars it is not possible to constrain simultaneously both $v_{\rm esc}$ and k using this method.

From this figure we can see that there is a degeneracy between k and $v_{\rm esc}$. This can be understood as follows: to retain the same shape of the velocity distribution over the observed range of velocities, an increase in $v_{\rm esc}$ must be accompanied by an increase in the steepness of the slope of the distribution (i.e. k). Given this degeneracy, we need to apply a suitable prior on k to obtain constraints on $v_{\rm esc}$. The results from Section 3.1 indicate that a suitable choice is to adopt a uniform prior in the range $k \in [2.7, 4.7]$. We have done this, and the results are shown in the top panel of Fig. 7. The resulting 90 per cent confidence interval is

$$498 < v_{\rm esc} < 608\,{\rm km\,s^{-1}}, \quad (11)$$

with a median likelihood of $v_{\rm esc} = 544\,{\rm km\,s^{-1}}$. Note that despite the degeneracy between k and $v_{\rm esc}$, it is immediately evident that negative values of k are strongly disfavoured.

Table 4. Bootstrap results for combined data with $v_{\rm min} = 300\,{\rm km\,s^{-1}}$. The third column gives the maximum likelihood values of $v_{\rm esc}$ and k for the original (non-bootstrapped) sample, while the fourth column gives the mean of the bootstrap distribution of maximum likelihood values (note that the means quoted here for Prior 1 differ from those given for the maximum likelihood method in Section 5.1 since the former employs the LT90 approximation, i.e. equation 6). The SE column gives the s.e. values from the s.d. of the bootstrap distribution. The last column gives the 90 per cent confidence intervals for $v_{\rm esc}$ and k. The $v_{\rm esc}$ confidence intervals are computed using the bootstrap normal distribution approximation, while the k intervals are calculated from a χ^2 distribution with the number of degrees of freedom equal to the bootstrap mean value.

Label	Prior form	Initial MLE values	Bootstrap mean values	SE	90 per cent confidence
1	$P_{\rm LT}(v_{\rm esc}, k) \propto 1/v_{\rm esc}$	$v_{\rm esc} = 588.7\,{\rm km\,s^{-1}}$	569.9	98.3	(445.0 769.4)
		$k = 4.2$	4.1	3.0	(0.3,9.0)
2	$P_{\rm J}(v_{\rm esc},\ k) \propto \frac{\sqrt{k}}{(v_{\rm esc}-v_{\rm min})\sqrt{k+2}}$	$v_{\rm esc} = 535.2\,{\rm km\,s^{-1}}$	521.5	52.5	(462.0, 640.0)
		$k = 2.9$	2.7	1.6	(0.1, 5.4)

It is important to check whether any kinematical bias is introduced by combining data from separate surveys. To do this we repeat the above analysis using only the 16 RAVE stars from Section 4.4. This is shown by the dotted quantities in Fig. 7. Reassuringly there is no noticeable offset between the two sets of contours; the only difference is a general broadening of the contours. When we apply the prior $k \in [2.7, 4.7]$ we find that the 90 per cent confidence interval becomes $496 < v_{\rm esc} < 655\,{\rm km\,s^{-1}}$, with a median likelihood of $v_{\rm esc} = 556\,{\rm km\,s^{-1}}$.

5.2 Bootstrap analysis

To further assess the likelihood constraints presented in the previous section, we also apply the bootstrap technique (see Section 2.2) to our data.

We apply the bootstrap approach to the combined data set of 33 stars, but unlike Section 5.1 we apply the LT90 approximation (equation 6) when calculating the maximum likelihood. The bootstrap computed the values of $v_{\rm esc}$ and k that maximized equation (9) using 5000 resamples of the original RAVE sample. Table 4 shows the resulting values of $v_{\rm esc}$ and k for the two chosen priors (see Appendix A).

When we compare the bootstrap interval with the likelihood interval obtained in Section 5.1, we find that the interval is shifted towards smaller $v_{\rm esc}$. This is consistent with what one would expect for the bootstrap method, since (unlike the method described in Section 5.1) the process of bootstrapping can result in values of $v_{\rm esc}$ that are smaller than the highest velocity star in the sample. This is a consequence of the fact that the bootstrap approach accounts for possible unreliable or inconsistent data. However, it is also important to note that the bootstrap mean values of k and $v_{\rm esc}$ found with both priors are identical, within s.e. values, to those found in the previous section using the non-bootstrap technique with the LT90 prior. Fig. 8 shows the bootstrap distributions and corresponding confidence intervals calculated for k and $v_{\rm esc}$ when each prior is applied to equation (9). The dashed curves correspond to Prior 1, while the solid curves correspond to Prior 2. The confidence intervals obtained using Prior 2 are clearly smaller than that from Prior 1, owing to the fact that Prior 2 contains more information about our expectations of k.

As a result of our analyses with a simulated data set (see Appendix A), we adopted the confidence regions from Prior 2, obtaining the bootstrap 90 per cent confidence intervals $462 < v_{\rm esc} < 640\,{\rm km\,s^{-1}}$ and $0.1 < k < 5.4$.

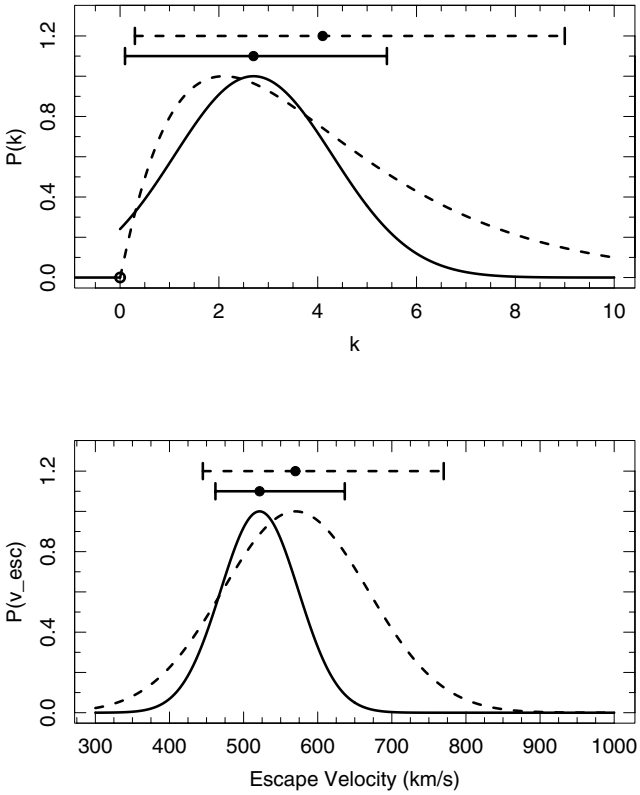

Figure 8. The bootstrap distributions for k (top panel) and $v_{\rm esc}$ (lower panel). The dashed and solid curves correspond to the bootstrap distributions calculated when solving equation (9) with Priors 1 and 2, respectively. The corresponding error bars show the 90 per cent confidence intervals for the parameters for each prior. Note that for both parameters, Prior 2 gives the smallest confidence interval. Furthermore, note that the bootstrap distribution for k follows a χ^2 distribution.

6 DISCUSSION

6.1 The nature of the fastest RAVE stars

One might wonder about the nature of the fastest stars in the RAVE catalogue. Although we do not currently possess any additional information from the RAVE spectra, such as metallicities, we do have estimates for each star's proper motion from various sources (Steinmetz et al. 2006). When combined with accurate photometry

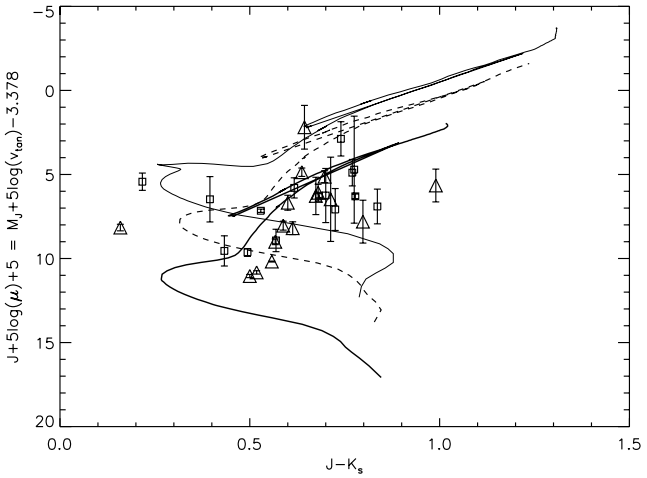

Figure 9. *J*-band RPM diagram of RAVE stars selected to have Galactocentric radial velocity greater than 250 km s^{-1}. The triangle symbols represent stars with Galactocentric radial velocity greater than 300 km s^{-1}. Note that the errors in the photometry are smaller than the symbols. Isochrones from Bonatto et al. (2004) are overplotted, with different tangential velocities adopted, to represent the three main stellar components of the Galaxy. The thin solid line is appropriate for the thin disc ($Z = 0.019$, age = 2.5 Gyr, $v_{\rm tangential} = 20$ km s^{-1}), the dashed line describes the thick disc ($Z = 0.004$, age = 10 Gyr, $v_{\rm tangential} = 42$ km s^{-1}) and the thick solid line describes the halo ($Z = 0.001$, age = 10 Gyr, $v_{\rm tangential} = 200$ km s^{-1}). The majority of the stars have positions in this diagram consistent with being halo giants. However, the star with $J - K$ corresponding to $\sim$0.15 is most likely a BHB star.

(in our case, *J*- and *K*-band magnitudes from 2MASS Skrutskie et al. 1997), proper motions can be used to place the stars on a reduced proper motion (RPM) diagram. An RPM diagram can be used to differentiate types of stars such as dwarfs and giants of different stellar populations because it is closely related to a standard colour–magnitude diagram, modified by kinematics. The *J*-band RPM is given by

$$J_{\rm RPM} \equiv J + 5\log\mu + 5 = M_J + 5\log\frac{v_{\rm tangential}}{47.4\,{\rm km\,s^{-1}}} + 5, \qquad (12)$$

where μ is the proper motion in arcsec yr^{-1} and $v_{\rm tangential}$ is the tangential velocity corresponding to this proper motion. Thus one obtains a quantity that is related to the absolute magnitude, but with stellar populations of different mean motions offset from one another, with scatter induced by the velocity dispersion.

We show the RPM for our fastest RAVE stars in Fig. 9. Three Bonatto, Bica & Girardi (2004) isochrones, adjusted for different mean $v_{\rm tangential}$, are also plotted. The three isochrones have ages and metallicities corresponding to different assumptions regarding the kinematics: namely, thin-disc, thick-disc and halo populations. The thin solid isochrone has been plotted assuming thin-disc kinematics, with tangential velocity $v_{\rm tangential} = 20$ km s^{-1}, $Z = 0.019$ and age = 2.5 Gyr. The dashed isochrone represents the thick-disc component, $v_{\rm tangential} = 42$ km s^{-1}, $Z = 0.004$ age = 10 Gyr. Finally, the thick solid isochrone depicts the halo with $v_{\rm tangential} = 200$ km s^{-1}, $Z = 0.001$ and age = 10 Gyr. In a probabilistic assignment of any one star to a given population, one would take account of the expected $\sigma \sim 150$ km s^{-1} spread around the halo isochrone.

The RAVE high-velocity stars clearly scatter around the red giant branch of the halo isochrone; this association is even stronger for the highest velocity stars, those with $v_{\rm r} > 300$ km s^{-1}, denoted by the triangle symbols. This is reassuring because it indicates that our sample should not suffer from significant contamination from the thick disc, thus justifying our choice of $v_{\rm min}$ in Section 3.2. Furthermore, using the echelle data from APO (section 4.3), we were able to calculate gravities and temperatures for a few of the high-velocity stars. These gravities were low, confirming giants. The bluest star appears from its spectrum to be a halo blue horizontal branch (BHB) star.[3]

If these stars are indeed mostly halo giants, then they are not in the immediate solar neighbourhood. This could have implications for our determination of $v_{\rm esc}$ and so we investigate this issue in the following section.

It is also worth noting that our sample contains three relatively high proper motion stars with $\mu > 50$ mas yr^{-1} (C1022369−140345, C1250398−030748 and T8395_01513_1). These all appear to be halo dwarfs or relatively nearby halo giants. By calculating distances from either our follow-up spectra or from the crude distance calibration of Scholz, Meusinger & Jahreiß (2005) we find that their total 3D velocities are all well within our value of $v_{\rm esc}$.

6.2 The effect of sample volume on the escape velocity

In Section 3.1 we used a sample of stellar particles from cosmological simulations to analyse our ability to recover the escape velocity. In that section our solar neighbourhood volumes were constructed from spheres of radius 2 kpc centred at 8.5 kpc (and also equivalent spheres of radii 1 and 3 kpc centred at 4.25 and 12.75 kpc, respectively, thus preserving the ratio of sample radius to distance from the Galactic Centre). However, the catalogue of radial velocity measurements used to try to constrain the Milky Way escape velocity could contain stars at significantly larger distances. The magnitude limit of the RAVE survey is $I = 12$ mag, which implies that very bright giants may be a significant distance away; for example, a K2 giant with $M_I \approx -0.9$ mag could be at a distance of around 4 kpc. Indeed, Section 6.1 suggested that some of these fastest RAVE stars may well be halo giants.

We test whether our results are sensitive to the sample volume by analysing the KIA3 and KIA4 simulations introduced in Section 3.1 (we restrict ourselves to these two galaxies as they have the largest number of stellar particles). We generate samples of stellar particles by taking spheres of varying radius centred on 8.5 kpc. From these samples we can estimate the escape velocity by evaluating equation (9), adopting the prior determined in Section 3.1 (i.e. $k \in [2.7, 4.7]$). This analysis shows how the recovered escape velocity varies as a function of the radius of the 'solar neighbourhood' spheres. We find that for larger spheres the escape velocity is slightly overestimated: the error in $v_{\rm esc}$ is $\lesssim$15 per cent for all radii up to 5 kpc, with an error of $\lesssim$10 per cent for radii of less than 4 kpc. This overestimation is probably due to the fact that larger volumes include stars that probe regions of higher $v_{\rm esc}$ and hence result in larger estimates for $v_{\rm esc}$. It should also be noted that in reality this is an upper limit since only a small proportion of our sample of observed stars will probe such large distances, whereas these simulated samples do not incorporate any such distance effects (i.e. the probability of selecting a simulated star is independent of its distance).

6.3 Unbound stars

One obvious deficiency of the analysis presented in Section 2 is that it requires all stars in our sample to be bound to the Galaxy. It is

[3] The BHB star has identifier, C1100242−024226.

clear that the presence of any unbound stars would invalidate our results, although the bootstrap technique should reduce our sensitivity to any spurious data points. There are currently seven known 'hypervelocity', probably unbound, stars in our Galaxy (Edelmann et al. 2005; Hirsch et al. 2005; Brown et al. 2006). The Galactic standard of rest-frame radial velocities computed for these stars range from $\sim$550 to $\sim$720 km s^{-1}. The first few were found serendipitously while the most recent were found by a targeted survey of faint early-type stars (Brown et al. 2006); indeed all of the known hypervelocity stars are blue and therefore almost certainly young (Kollmeier & Gould 2007).

It is suggested that such high velocities may originate from encounters with the massive black hole at the Galactic Centre. However, Yu & Tremaine (2003) found that the numbers of hypervelocity stars produced would be about 10^{-5} yr^{-1} (see also Perets, Hopman & Alexander 2006). That would give about 10^5 stars in total, assuming a long-lived black hole, which suggests that hypervelocity stars are very rare. Furthermore, they predicted that the number of lower velocity 'hyper-velocity' stars is even smaller. Although such objects should be rare, and hence the chance of one appearing in our catalogue is small, this can nevertheless not be ruled out entirely. There is one simple test one can apply to investigate this issue; if one can obtain the full 3D velocities for the fastest stars, then it can be ascertained whether their direction of motion is consistent with that of an object ejected from the centre of the Galaxy. However, this method could fail if the unbound star(s) originate from different circumstances, such as stars which are bound to the local group but unbound with respect to our Galaxy.

High-velocity stars can also originate from stellar binary systems in which one of the stars undergoes a supernova and 'kicks' out the companion star. This would have larger effects on main-sequence stars as they can acquire an extra velocity of $\sim$100 km s^{-1}. However, since our sample contains mostly halo giants (see Section 6.1) this should be a small effect.

One final point to note is that given the consistency between our constraints on $v_{\rm esc}$ with other independent methods for estimating the mass of the Galaxy, the presence of any unbound stars in our sample seems to be an unlikely scenario.

6.4 The mass and extent of the Galactic halo

The escape velocity can be used as a powerful tool to probe the mass distribution in the Galaxy. Our constraints on $v_{\rm esc}$ allow us to constrain the halo potential at the solar radius,

$$\Phi_{\rm total}({\rm r}_\odot) = -\frac{v_{\rm esc}^2}{2}. \tag{13}$$

Halo masses derived from our constraints on $v_{\rm esc}$ rely on the assumption that our fastest stars probe velocities all the way to the local escape velocity, that is, there is no truncation in the stellar velocity distribution. If this assumption is invalid then the mass constraints quoted in this section would be lower limits to the true mass. However, as has been shown for the simulated galaxies in Section 3.1, the level of truncation in the stellar component cannot be more than 10 per cent. When the likelihood analysis (equation 9) is applied to these simulations we do not find any systematic underestimation in the recovered $v_{\rm esc}$, which gives us confidence that such weak levels of truncation (if present) will not pose significant problems for our analysis.

In this section we will use $v_{\rm esc}$ to constrain the total mass of the halo. However, before undertaking any detailed calculations, one obtains a qualitative understanding of the situation from the simple relation (see e.g. equation 2-22 of Binney & Tremaine 1987)

$$v_{\rm esc}^2 = 2\,v_{\rm circ}^2 + 8\pi G \int_{r_\odot}^{\infty} \rho(r)\, r\, {\rm d}r, \tag{14}$$

where $v_{\rm circ}$ is the circular velocity at the solar radius. The fact that our measured lower limit of $v_{\rm esc} > 498$ km s^{-1} is significantly greater than $\sqrt{2}v_{\rm circ} \sim 311$ km s^{-1} shows that the second term in equation (14) cannot be small, that is, there must be a significant amount of mass exterior to the solar circle. This simple argument convincingly demonstrates the presence of a dark halo in the Galaxy.

Another straightforward calculation that can be made is the extent of an isothermal halo that has a constant circular velocity out to a truncation radius $r_{\rm cut}$. For this elementary model we find that

$$r_{\rm cut} = {\rm r}_\odot \exp\left(\frac{v_{\rm esc}^2}{2v_{\rm circ}^2} - 1\right). \tag{15}$$

Our measured value of $v_{\rm esc}$ ($v_{\rm esc} \approx 544$ km s^{-1}) corresponds to a truncation radius of $r_{\rm cut} \approx 58$ kpc.

The above simple arguments demonstrate the existence of a dark halo, but we now undertake a more detailed calculation in order to constrain the total mass of the halo. To do this we first need a model for the baryonic contribution to the total potential from the bulge and disc of the Galaxy. We adopt a spherical Hernquist (1990) bulge with mass $1.5 \times 10^{10}\,{\rm M}_\odot$ and scale radius 0.6 kpc, and a Miyamoto & Nagai (1975) disc with mass $5 \times 10^{10}\,{\rm M}_\odot$, scalelength 4 kpc and scaleheight 0.3 kpc. The halo potential is then simply given by

$$\Phi_{\rm halo} = \Phi_{\rm total} - \Phi_{\rm bulge} - \Phi_{\rm disc}. \tag{16}$$

When $\Phi_{\rm halo}$ is combined with a halo profile it is possible to probe the total halo mass. One popular profile is the NFW halo (Navarro, Frenk & White 1996). It can be shown that the radial potential for an NFW density profile can be expressed as

$$\Phi_{\rm NFW}(r) = -\frac{4\pi G \rho_s r_{\rm vir}^3}{c^3 r} \ln\left(1 + \frac{cr}{r_{\rm vir}}\right), \tag{17}$$

where c is a concentration parameter equal to the ratio of the virial radius to the scale radius, and $\rho_{\rm s}$ is a characteristic density given by

$$\rho_s = \frac{\rho_{\rm cr}\Omega_0\delta_{\rm th}}{3} \frac{c^3}{\ln(1+c) - c/(1+c)}, \tag{18}$$

where $\rho_{\rm cr} = 3H^2/8\pi G$ is the critical density of the Universe, Ω_0 is the contribution of matter to the critical density and $\delta_{\rm th}$ is the critical overdensity at virilization. The virial mass can then be determined from the virial radius,

$$M_{\rm vir} = \frac{4\pi}{3}\rho_{\rm cr}\Omega_0\delta_{\rm th} r_{\rm vir}^3. \tag{19}$$

If we take the solar radius to be 7.5 kpc, $\Omega_0 = 0.3$, $\delta_{\rm th} = 340$, $H = 65$ km s^{-1} Mpc^{-1}, and enforce the circular velocity at the solar radius to be $v_{\rm circ} = 220$ km s^{-1}, our 90 per cent confidence constraint on the escape velocity (equation 11) allow us to obtain the following constraints on the virial radius, mass and concentration parameter for the Milky Way:

$$M_{\rm vir} = 0.85^{+0.55}_{-0.29} \times 10^{12}\,{\rm M}_\odot,$$

$$r_{\rm vir} = 257^{+47}_{-33}\,{\rm kpc},$$

$$c = 24.3^{+6.5}_{-5.1}.$$

It is worth noting that these constraints have little dependence on the adopted value for the solar radius; taking a value of 8 kpc will change the limits on $M_{\rm vir}$ and $r_{\rm vir}$ by less than 1 per cent.

One common variation of this classical NFW model is to account for the adiabatic contraction of the dark matter halo due to the presence of the baryons in the disc and bulge (Mo, Mao & White 1998). In this instance we simplify the calculation by adopting an exponential disc, retaining the same mass and scalelength as above. The resulting constraints on the virial mass, radius and concentration are

$$M_{\rm vir} = 1.42^{+1.14}_{-0.54} \times 10^{12}\,{\rm M}_{\odot},$$

$$r_{\rm vir} = 305^{+66}_{-45}\,{\rm kpc},$$

$$c = 7.7^{+2.8}_{-2.0}.$$

Both of these models predict low values for the virial velocity of the halo. The contracted NFW model has $v_{\rm vir} = 142^{+31}_{-21}\,{\rm km\,s^{-1}}$, while the uncontracted one has $v_{\rm vir} = 124^{+20}_{-14}\,{\rm km\,s^{-1}}$, that is, around half $v_{\rm circ}$. These low values of $v_{\rm vir}$ pose problems for semi-analytic models of galaxy formation that require the peak circular velocity of the disc to be $\sim v_{\rm vir}$ in order to simultaneously explain the normalization of the Tully–Fisher relation and the galaxy luminosity function in a ΛCDM universe (e.g. Cole et al. 2000). More recent semi-analytic models (Croton et al. 2006a,b) have been able to relax this requirement slightly, so that the peak circular velocity of the disc is only required to be comparable to the maximum circular velocity of the dark matter halo. However, it can be seen that our models are only marginally consistent with this requirement, for example, for our contracted model, the peak circular velocity of the halo is only $167\,{\rm km\,s^{-1}}$, while the peak circular velocity of the disc is $220\,{\rm km\,s^{-1}}$.

It is possible to work out similar constraints for other halo models. For example, one can take the model of Wilkinson & Evans (1999), which has a flat rotation curve in the inner regions and a sharp fall-off in density beyond an outer edge (hereafter referred to as the WE model). For this halo profile the total halo mass is given by

$$M_{\rm WE} = \frac{v_{\rm esc}^2}{2G} a_{\rm WE} \left[\ln\left(\frac{\sqrt{r^2 + a_{\rm WE}^2} + a_{\rm WE}}{{\rm r}_{\odot}} \right) \right]^{-1}, \qquad (20)$$

where $a_{\rm WE}$ is the scalelength. For this model we revert to our original Miyamoto & Nagai (1975) disc, as used for the uncontracted NFW profile. By again enforcing the circular velocity to be $220\,{\rm km\,s^{-1}}$ we can use this model to obtain the following 90 per cent confidence constraints on the total mass and scalelength,

$$M_{\rm WE} = 1.89^{+76.46}_{-1.13} \times 10^{12}\,{\rm M}_{\odot},$$

$$a_{\rm WE} = 314^{+986}_{-188}\,{\rm kpc}.$$

It is clear that our constraints on the halo mass are consistent with previous estimates; within the last 10 yr estimates tend to lie in the range 1–2 $\times\, 10^{12}{\rm M}_{\odot}$ (see Section 1), which is fully consistent with our findings. It is also important to note that our constraints are relatively model independent, with only small differences between the mass estimates for our three halo profiles.

It is also useful to quote the total mass within a certain radius of the Galactic Centre. For example, our three models predict the following constraints on the mass contained within 50 kpc: $4.04^{+1.10}_{-0.76} \times 10^{11}, 3.87^{+0.64}_{-0.56} \times 10^{11}, 3.58^{+0.04}_{-0.17} \times 10^{11}\,{\rm M}_{\odot}$, for the uncontracted NFW, contracted NFW and WE model, respectively. These masses are relatively well constrained mainly due to the fact that our models must adhere to the $v_{\rm circ} = 220\,{\rm km\,s^{-1}}$ constraint. Wilkinson & Evans (1999) obtained a value of $5.4^{+0.2}_{-3.6} \times 10^{11}\,{\rm M}_{\odot}$, which is consistent with our models.

Given these halo models, we can estimate the maximum Galactic radius that our fastest star will reach. From Table 3 we see that the fastest star has a rest-frame radial velocity of $-449\,{\rm km\,s^{-1}}$. We assume that the star's kinetic energy is dominated by its radial component, so

$$E({\rm r}_{\odot}) = \frac{v_{\rm r}^2 - v_{\rm esc}^2}{2}. \qquad (21)$$

At apocentre the kinetic energy is solely due to the angular momentum L, which we assume to be conserved. Assuming that the distance to the Sun is small, we have

$$L = \begin{pmatrix} 7.5\ {\rm kpc} \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} -v_{\rm r}\cos b \cos l \\ v_{\rm r}\cos b \sin l \\ v_{\rm r}\sin b \end{pmatrix}. \qquad (22)$$

The apocentre distance $r_{\rm apo}$ can now be found from

$$E(r_{\rm apo}) = \frac{L^2}{2r_{\rm apo}^2} + \Phi(r_{\rm apo}). \qquad (23)$$

Equations (21) and (23) can be solved using any of the models for the potential described above. For our constraints on $v_{\rm esc}$ determined in Section 5.1, we find that $r_{\rm apo} = 102^{+51}_{-22}\,{\rm kpc}$ for our contracted NFW halo model. Since this calculation assumes that there is no tangential velocity, it is actually a lower limit for $r_{\rm apo}$; for example, if the star has a tangential velocity of $100\,{\rm km\,s^{-1}}$ then $r_{\rm apo}$ increases to $\sim 122^{+81}_{-31}\,{\rm kpc}$. Although we do not know the distance to our fastest radial velocity star, by adopting a distance of 1 kpc we find that the observed proper motion ($17.2 \pm 9.4\,{\rm mas\,yr^{-1}}$; Hambly et al. 2001) corresponds to a tangential velocity of $82 \pm 46\,{\rm km\,s^{-1}}$.

The question of whether there is a truncation in the stellar halo is currently open to debate (e.g. Ivezić 2000; Dehnen et al. 2006). Therefore our large value for $r_{\rm apo}$ is important since it could disfavour Galactic models with an early truncation.

6.5 Uncertainties and the next steps

As discussed in Section 2, our results are based on the assumption that the velocity distribution near the escape velocity can be approximated by a power law $f(v) \propto (v_{\rm esc} - |\boldsymbol{v}|)^k$. In a hierarchical universe, this model is unlikely to be valid because the motions of the fastest moving stars are expected to be strongly clumped (Helmi et al. 2002). Since the velocity distribution would be dominated by very few streams, it would be the sum of a few delta functions centred on the mean velocities of those streams. It may even be possible that each stream originates in the same object, in which case, only one orbit is probed. This invalidates the use of statistical arguments to derive the mass of the Galaxy. Such a high degree of lumpiness only is important when the fastest 1 per cent of the stars are considered, that is, those with velocities $v_{\rm r} \gtrsim 3\sigma \sim 450\,{\rm km\,s^{-1}}$.

Notice that there are no stars in our sample with such large radial velocities. This is the reason why substructure appears not be an issue in our case (e.g. the results of the bootstrap analysis obtained by resampling are consistent with the maximum likelihood performed using the actual data set). However, a new approach will be necessary in the future. By the time the RAVE survey concludes, it is intended that the catalogue will contain up to one million stars. If we assume that the rate of detecting high-velocity stars is unchanged for the rest of the survey, this would result in a final sample of $\sim$200 stars with radial velocities greater than $300\,{\rm km\,s^{-1}}$. Such a sample will reveal a large amount of substructure. This should also be evident when

the RAVE metallicities are added. Gravity determinations, plus the metallicities, will allow robust distance estimates and hence the use of tangential velocities. Full phase-space information should enable one to perform more sophisticated models to determine the mass of our Galaxy.

All of this highlights the importance of understanding the expected shape of the velocity distribution near the tails (Section 3.1). The limited number of high-resolution cosmological simulations of galaxies like the Milky Way is also a source of uncertainty in our estimates of the escape velocity. Both higher resolution (a significantly larger number of stellar particles) and a larger set of simulations are necessary.

7 CONCLUSION

In this work we have reported new constraints on the local escape speed of our Galaxy. We argued that the choice of prior on k may have been incorrectly estimated in previous work and deduced a different range for this prior (see Section 3.1). We then applied this to a catalogue of radial velocities from the RAVE collaboration, augmented with additional stars from the existing survey of Beers et al. (2000). Our results provide a 90 per cent confidence interval of $498 < v_{\rm esc} < 608\,{\rm km\,s^{-1}}$, with a median likelihood of $v_{\rm esc} = 544\,{\rm km\,s^{-1}}$. We also applied a bootstrap technique, which allowed us to reduce dependencies on extreme velocities that may be unbound or from stars in binary systems, while additionally allowing us to perform confidence analysis on the distribution of maximum likelihood values, $v_{\rm esc}$ and k, simultaneously. This resulted in a 90 per cent confidence interval of $462 < v_{\rm esc} < 640\,{\rm km\,s^{-1}}$ and $0.1 < k < 5.4$.

The fact that $v_{\rm esc}^2$ is significantly greater than $2v_{\rm circ}^2$ implies that there must be a significant amount of mass exterior to the solar circle. Furthermore, from our constraints on $v_{\rm esc}$ we can infer model-dependent estimates for the mass of the halo (see Section 6.4); for example, an adiabatically contracted NFW halo profile predicts that $M_{\rm vir} = 1.42^{+1.14}_{-0.54} \times 10^{12}\,{\rm M}_\odot$, $r_{\rm vir} = 305^{+66}_{-45}\,{\rm kpc}$ and $c = 7.7^{+2.8}_{-2.0}$. Similar results for the halo mass were found for both an uncontracted NFW halo and a Wilkinson & Evans (1999) halo. Although our results are model dependent, we find that our three models are in good agreement. The mass within 50 kpc is found to be 3.6–$4.0 \times 10^{11}\,{\rm M}_\odot$. It is interesting to note that our models predict low values for the virial velocity, for example, $v_{\rm vir} = 142^{+31}_{-21}\,{\rm km\,s^{-1}}$ for the contracted NFW model.

By the time the RAVE survey reaches completion (currently predicted to be 2010) it will provide an unparalleled data base of stellar kinematics, thus allowing dramatic progress in this field.

ACKNOWLEDGMENTS

The authors wish to thank M. Abadi for advice and support regarding the cosmological simulations utilized in Section 3, A. Villalobos for assistance with Section 6.4 and P. Maiste and D. Naiman for their assistance regarding the bootstrap analysis discussion. The authors are indebted to S. White and the anonymous referee for comments that greatly improved the clarity of the paper. MCS and AH acknowledge financial support from the Netherlands Organization for Scientific Research (NWO). GRR and RFGW acknowledge financial support from the US National Science Foundation through grant AST-0508996. This work benefited from the support of the European Community's Sixth Framework Marie Curie Research Training Network Programme, Contract No. MRTN-CT-2004-505183 'ANGLES'.

Funding for RAVE has been provided by the Anglo-Australian Observatory, Astrophysical Institute Potsdam, Australian Research Council, German Research Foundation, National Institute for Astrophysics at Padova, The Johns Hopkins University, Netherlands Research School for Astronomy, Natural Sciences and Engineering Research Council of Canada, Slovenian Research Agency, Swiss National Science Foundation, National Science Foundation of the USA, Netherlands Organization for Scientific Research, Particle Physics and Astronomy Research Council of the UK, Opticon, Strasbourg Observatory and Universities of Basel, Cambridge and Groningen.

The RAVE web site is at www.rave-survey.org.

REFERENCES

Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003a, ApJ, 591, 499
Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003b, ApJ, 597, 21
Abadi M. G., Navarro J. F., Steinmetz M., 2006, MNRAS, 365, 747
Aguilar L. A., White S. D. M., 1986, ApJ, 307, 97
Allen R., Shao M., Peterson D., 1998, Proc. SPIE, 2871, 504
Battaglia G. et al., 2005, MNRAS, 364, 433
Battaglia G. et al., 2006, MNRAS, 370, 1055
Beers T. C., Chiba M., Yoshii Y., Platais I., Hanson R. B., Fuchs B., Rossi S., 2000, AJ, 119, 2866
Binney J., Dehnen W., 1997, MNRAS, 287, L5 Astronomy. Princeton Univ. Press, Princeton, NJ
Binney J., Merrifield M., 1998, Galactic Astronomy. Princeton Univ. Press, Princeton
Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton Univ. Press, Princeton, NJ
Bonatto C., Bica E., Girardi L., 2004, A&A, 415, 571
Brown W. R., Geller M. J., Kenyon S. J., Kurtz M. J., 2006, ApJ, 647, 303
Chiba M., Beers T. C., 2000, AJ, 119, 2843
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Croton D. J. et al., 2006a, MNRAS, 365, 11
Croton D. J. et al., 2006b, MNRAS, 367, 864
Davison A. C., Hinkley D. V., 1997, Bootstrap Methods and their Application. Cambridge Univ. Press, New York
Dehnen W., Binney J. J., 1998, MNRAS, 298, 387
Dehnen W., McLaughlin D. E., Sachania J., 2006, MNRAS, 369, 1688
Edelmann H., Napiwotzki R., Heber U., Christlieb N., Reimers D., 2005, A&A, 634, L181
Edvardsson B., Andersen J., Gustafsson B., Lambert D. L., Nissen P. E., Tomkin J., 1993, A&AS, 102, 603
Epchtein N. et al., 1997, The Messenger, 87, 27
Fich M., Tremaine S. D., 1991, ARA&A, 310, 409
Fuchs B., Jahreiß H., 1998, A&A, 329, 81
Hambly N. C., 2001, MNRAS, 326, 1279
Hernquist L., 1990, ApJ, 356, 359
Helmi A., White S. D. M., Springel V., 2002, Phys. Rev. D, 66, 3502
Hirsch H. A., Heber U., O'Toole S. J., Bresolin F., 2005, A&A, 444, L61
Ivezić Ž., et al., 2000, AJ, 120, 963
Jaffe W., 1987, in de Zeeuw T., ed., IAU Symp., 127, Structure and Dynamics of Elliptical Galaxies. Reidel, Dordrecht, p. 511
Jahreiß H., Wielen R., 1997, ESA SP-402: Hipparcos – Venice '97, Vol. 402. ESA Publications Division, Noordwijk p. 675
Jeffreys H., 1961, Theory of Probability. Oxford Univ. Press, Oxford
Kendall M., Stuart A., 1977, The Advanced Theory of Statistics. Griffin, London
Kochanek C. S., 1996, ApJ, 547, 228 (K96)
Kollmeier J. A., Gould A. 2007, ApJ, submitted (astro-ph/0701350)
Kulessa A. S., Lynden-Bell D., 1992, MNRAS, 310, 105
Latham D. W., Stefanik R. P., Torres G., Davis R. J., Mazeh T., Carney B. W., Laird J. B., Morse J. A., 2002, ApJ, 124, 1144
Lee P. M., 2004, Bayesian Statistics, An Introduction, 3rd edn. Arnold, Hodder Headline Group, London, England
Leonard P. J. T., Tremaine S., 1990, ApJ, 353, 486 (LT90)

Little B., Tremaine S. D., 1987, ApJ, 320, 493
Meza A., Navarro J. F., Steinmetz M., Eke V. R., 2003, ApJ, 590, 619
Meza A., Navarro J. F., Abadi M. G., Steinmetz M., 2005, MNRAS, 359, 93
Miyamoto M., Nagai T., 1975, PASJ, 27, 533
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Monet D. G. et al., 2003, AJ, 125, 984
Morrison H. L., Mateo M., Olszewski E. W., Harding P., Dohm-Palmer R. C., Freeman K. C., Norris J. E., Morita M., 2000, AJ, 119, 2254
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Nordström B. et al., 2004, A&A, 418, 989
Ojha D. K., 2001, MNRAS, 322, 426 O
Oort J. H., 1926, Groningen Publ., 40
Oort J. H., 1928, Bull. Astron. Inst. Neth., 4, 269
Perets H. B., Hopman C., Alexander T., 2006, ApJ, 656, 709
Perryman M. A. C. et al., 2001, A&A, 369, 339
Plummer H. C., 1911, MNRAS, 71, 460
Reddy B. E., Lambert D. L., Allende Prieto C., 2006, MNRAS, 367, 1329
Sakamoto T., Chiba M., Beers T. C., 2003, A&A, 397, 899
Scholz R. D., Meusinger, Jahreiß, 2005, A&A, 442, 211
Skrutskie M. F. et al., 1997, in Garzón F., ed., The Impact of Large Scale Near-IR Sky Surveys. Kluwer, Dordrecht, p. 187
Spitzer L. Jr, Shapiro S. L., 1972, ApJ, 173, 529
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Steinmetz M., 1996, MNRAS, 278, 1005
Steinmetz M., Navarro J. F., 2002, New Astron., 7, 155
Steinmetz M. et al., 2006, AJ, 132, 1645
Tonry J., Davis M., 1979, AJ, 84, 1511
Tremaine S., 1987, in de Zeeuw T., ed., IAU Symp., 127, Structure and Dynamics of Elliptical Galaxies. Reidel Dordrecht, p. 367
Wilkinson M. I., Evans N. W., 1999, MNRAS, 310, 645
Wilkinson M. I. et al., 2005, MNRAS, 359, 1306
Wong O. I., Drinkwater M. J., Jones J. B., Gregg M. D., Freeman K. C., 2004, in Ryder S. D., Pisano D. J., Walker M. A., Freeman K. C., eds, IAU Symp. 220, Dark Matter in Galaxies. Astron. Soc. Pac., San Francisco, p. 213
Yu Q., Tremaine S., 2003, ApJ, 599, 1129
Zacharias N., Urban S. E., Zacharias M. I., Wycoff G. L., Hall D. M., Monet D. G., Rafferty T. J., 2004, AJ, 127, 3043
Zaritsky D., Olszewski E. W., Schommer R. A., Peterson R. C., Aaronson M., 1989, ApJ, 345, 759

APPENDIX A: THE PRIOR PROBABILITY DISTRIBUTIONS FOR THE BOOTSTRAP METHOD

A1 The choice of prior

A1.1 Jeffreys' rules

It can be quite difficult to choose a priori probabilities for $v_{\rm esc}$ and k especially when not much is known about the quantities. Note that the posterior distribution (equation 9), with uniform prior distributions [i.e. $P(v_{\rm esc}) = P(k) = 1$], is equivalent to the normalized likelihood equation. Furthermore, even if the prior probabilities are non-uniform, the maximization of equations (8) and (9) are asymptotically equivalent. This is due to the fact that as n goes to infinity, the product in equation (9) will dominate. However, for small n, there may be large differences between the parameter values maximizing $l(v_{\rm esc},\, k)$ and those maximizing the posterior distribution. Therefore it is important, for small n, to choose 'good' reference priors for $v_{\rm esc}$ and k.

One method that can be used to attain reference priors is that defined by Sir Harold Jeffreys (Jeffreys 1961). First, define the log of the normalized likelihood equation (equation 8) as $L(v_{\rm esc},\, k)$. Next, we define a 2×2 information matrix, known as the Fisher information matrix:

$$I(\boldsymbol{\theta} \mid \boldsymbol{v}_{\rm r})_{i,\,j} = -E\left[\frac{\partial^2 L(v_{\rm esc}, k)}{\partial\theta_i\, \partial\theta_j}\right], \tag{A1}$$

where $\boldsymbol{\theta} = (v_{\rm esc}, k)$ and the right-hand side is the negative of the expectation value of the second derivative. Note that if $\boldsymbol{v}_{\rm r}$ is a vector of n independent observations, by linearity of expectation the information is $nI(\boldsymbol{\theta} \mid v_{\rm r})_{i,j}$. However, we do not include this factor n since it is a constant and will not affect the maximum likelihood. Using this information matrix, Jeffreys argues that a good reference prior can be estimated as

$$P_{\rm J}(\theta) \propto \sqrt{|I(\boldsymbol{\theta}|v_{\rm r})|}, \tag{A2}$$

where $|I(\boldsymbol{\theta} \mid v_{\rm r})|$ is the determinant of the information matrix. This equation is a useful reference prior because it does not require the selection of any specific parametrization. However, it is often argued that this form of the prior is improved by considering that the parameters are independent of each other (Lee 2004). We follow this approach, which requires us to neglect the off-diagonal terms in the determinant. By applying this analysis to the LT90 formalism (equation 6) we obtain

$$P_{\rm J}(v_{\rm esc},\, k) \propto \frac{1}{(v_{\rm esc} - v_{\rm min})\sqrt{k(k+2)}}. \tag{A3}$$

LT90 apply another of Jeffreys' rules that says if a variable varies continuously from 0 to infinity, one must adopt an a priori probability distribution equal to the reciprocal of the variable (Kendall & Stuart 1977). LT90 adopt this form of the prior probability distribution for the escape velocity, $P_{\rm LT}(v_{\rm esc}) \propto 1/v_{\rm esc}$. However, note that $v_{\rm esc}$ does not necessarily vary over the range $(0,\, \infty)$. We have applied a minimum cut-off velocity, as defined in Section 3.2. Therefore, an arguably more appropriate approach would be to say that the escape velocity varies continuously in $(v_{\rm min},\, \infty)$, that is, $P_{\rm J'}(v_{\rm esc}) \propto 1/(v_{\rm esc} - v_{\rm min})$. This approach can also be applied to the parameter k. From equation (6) we can see that for the LT90 formalism the variable k varies over the range $(-1,\, \infty)$ and so one could also adopt the prior $P_{\rm J'}(k) \propto 1/(k+1)$.

A1.2 A priori probability

The question remains, which prior probabilities should be used? In an ideal world, the samples should be large enough so that maximization of equation (9) would be the same for any chosen prior. However, we cannot rely on this fact. It is very important to think of Jeffreys' rules as only guidelines for choosing prior distributions, especially when dealing with more than one parameter. Therefore we have chosen three priors for our analysis. The first is that used by LT90 in their maximum likelihood analysis, which will be labelled as Prior 1. The other two are those derived using Jeffreys' rules described above, that is, $P_{\rm J}$ and $P_{\rm J'}$. These priors are useful because they provide much more information about k than the LT90 prior.

However, these priors need slight modifications. From equation (1) it can be seen that $k > 0$, since negative k would imply that the probability does not go to zero as $v \rightarrow v_{\rm esc}$. Therefore for the priors derived from Jeffreys' rules we enforce that the $P_{\rm J} = P_{\rm J'} = 0$ for $k < 0$. In addition, as can be seen from equation (A3), prior $P_{\rm J}$ diverges as $k \rightarrow 0$. This is clearly unacceptable because our likelihood only goes to zero as $k \rightarrow -1$ (equation 6).[4] In order to

[4] Note that the reason why this likelihood does not go to zero as $k \rightarrow 0$ is because of the integral over the unknown tangential velocities performed in equation (5).

Table A1. Data from the optimization of 35 stars with simulation parameters: $v_{\rm esc} = 600\,{\rm km\,s^{-1}}$, $k = 2.0, v_{\rm min} = 300\,{\rm km\,s^{-1}}$ after applying our derived priors to equation (9). The third column gives the maximum likelihood values of $v_{\rm esc}$ and k for the original (non-bootstrapped) sample, while the fourth column gives the mean of the bootstrap distribution of maximum likelihood values. The SE column gives the s.e. values from the s.d. of the bootstrap distribution. The last column gives the 90 per cent confidence intervals for $v_{\rm esc}$ and k. The $v_{\rm esc}$ confidence intervals are estimated using the bootstrap normal distribution approximation, while the k intervals are calculated from a χ^2 distribution with the number of degrees of freedom equal to the bootstrap mean value. Prior 1 is the same as that used by LT90, while Priors 2 and 3 are derived from Jeffreys' rules. All priors are 0 for $k \leqslant 0$. Note that the non-bootstrap analysis employs Prior 1.

Label	Prior form	Initial MLE values	Bootstrap mean values	SE	90 per cent confidence
1	$P_{\rm LT}(v_{\rm esc}, k) \propto 1/v_{\rm esc}$	$v_{\rm esc} = 631.1\,{\rm km\,s^{-1}}$	625.6	132.8	(502.7, 832.7)
		$k = 2.5$	2.7	2.9	(0.3, 7.3)
2	$P_{\rm J}(v_{\rm esc},\ k) \propto \frac{\sqrt{k}}{(v_{\rm esc}-v_{\rm min})\,(k+2)}$	$v_{\rm esc} = 604.6\,{\rm km\,s^{-1}}$	588.0	74.1	(516.3, 730.6)
		$k = 2.1$	2.0	1.5	(0.1, 6.0)
3	$P_{J'}(v_{\rm esc},\ k) \propto \frac{1}{(v_{\rm esc}-v_{\rm min})\,(k+1)}$	$v_{\rm esc} = 563.5\,{\rm km\,s^{-1}}$	544.0	54.2	(495.9, 634.6)
		$k = 1.2$	1.0	1.1	(0.0, 3.8)

avoid this problem we introduce an additional factor of $k/\sqrt{k+2}$ into prior $P_{\rm J}$, which means that $P_{\rm J} \to 0$ as $k \to 0$, which is useful for continuity at $k = 0$. Importantly, we still retain the property that $P_{\rm J} \to 0$ as $k \to \infty$.

All priors are described in Table A1. These different forms for the a priori probability distributions apply different assumptions and weights upon the variables. To further assess these prior distributions, maximization of equation (9) using the parameters $v_{\rm esc}$ and k (refer to Section 2.2) was performed after applying both of the above priors.

A2 Using simulations to assess the choice of prior

We tested the choice of prior by applying the bootstrap technique to a simulated data set. A random sample of stars was drawn from the assumed LT90 probability density function (equation 6) for user-defined values of $v_{\rm esc} = 600\,{\rm km\,s^{-1}}$, $k = 2.0$ and $v_{\rm min} = 300\,{\rm km\,s^{-1}}$. These values are similar to what we might expect for our Galaxy. We chose to analyse a random sample size of 35 simulated stars as this is comparable to the number of stars in our observed sample (Section 4). This mock sample was used to test the bootstrap maximization methods as well as assess chosen priors in equation (9) for calculating the escape velocity and k. Note that we apply the LT90 approximation (equation 6) when calculating the maximum likelihood.

The mock 35 star sample was run through the bootstrap maximization technique for the three priors given in Table A1 applied to equation (9). The bootstrap performed maximization on 5000 resamples of the original mock data set. The bootstrap distribution of peak likelihood points for $v_{\rm esc}$ is very close to normal. Therefore, we used techniques as described in Section 2.2 to compute confidence for $v_{\rm esc}$. However, the bootstrap distribution for k is clearly not normal. We determined that the distribution of peak likelihoods follows a χ^2 distribution with the number of degrees of freedom equalling the mean of the bootstrap distribution. Therefore confidence limits for k were derived from a χ^2 distribution instead of a normal distribution.

Table A1 gives the maximization results for each prior defined in Section A1. The third column of the table represents the maximum likelihood values of $v_{\rm esc}$ and k calculated from the original mock sample before the bootstrap has been applied. The fourth column give the mean bootstrap values of $v_{\rm esc}$ and k obtained from the bootstrap distribution. The s.e. values were computed from the s.d. of the bootstrap resamples as explained in Section 2.2. The last column gives the 90 per cent confidence intervals of $v_{\rm esc}$ and k. Note that confidence endpoints for $v_{\rm esc}$ are computed using the bootstrap normal distribution approximation, while that for k were computed from a χ^2 distribution.

Comparing the results from the three different priors, there are two very noticeable attributes. First, the bootstrap distribution of $v_{\rm esc}$ from Prior 1 has slightly less bias than that from Priors 2 and 3. This suggests that Priors 2 and 3 are 'stronger' priors, introducing bias into the maximum likelihood analysis. However, the bootstrap mean from Prior 2 gives estimates for $v_{\rm esc}$ that are still identical, within s.e. values, to the chosen values for the simulation. Prior 3, however, seems to slightly underestimate the value of $v_{\rm esc}$. The other attribute to notice is that the errors and confidence regions from Prior 2 are less than that of Prior 1, while also still containing the desired maximum likelihood values for $v_{\rm esc}$ and k. Therefore it would be useful to use the confidence regions from Prior 2 when concluding on the confidence of our actual data sample, especially since it contains more information about k. Notice that Prior 3 has similar confidence regions to Prior 2, but since the bootstrap mean estimate for $v_{\rm esc}$ is beyond 1σ of the assumed value we believe that Prior 3 may not be as useful. Therefore we will compute our analysis using Priors 1 and 2 for comparison.

To further assess the reliability of the bootstrap method, we repeated our analysis on a large number of different mock samples. These mock samples were chosen to cover a wide range of values for the input parameters of our distribution ($v_{\rm esc} \in [400, 700]$, $k \in [2, 5]$). For only 4 per cent of these samples did the bootstrap method produce maximum likelihood values beyond 2σ of the actual values defined. Otherwise, the technique extracted values of $v_{\rm esc}$ and k that were consistent within s.e. values to their assumed values in the simulation for all a priori probability definitions. Furthermore, the technique was applied using several different initial guesses for $v_{\rm esc}$ and k during maximization and the conclusions from all priors remained unaffected.

APPENDIX B: THE ARCHIVAL HIGH-VELOCITY STARS

Here we list the archival stars that were used to augment the RAVE data set. The stars given in Table B1 are from the Beers et al. (2000) catalogue of metal-poor stars.

© 2007 The Authors. Journal compilation © 2007 RAS, MNRAS **379,** 755–772

Table B1. The archival high-velocity stars, taken from Beers et al. (2000). The first column shows the Galactocentric radial velocity. Note that three stars have been excluded from this sample as their estimated distances are greater than 5 kpc; all stars listed in this table have estimated distances of less than 2.5 kpc (see Section 4.5).

Name	v_r (km s^{-1})	δv_r (km s^{-1})	RA (J2000) (°)	Dec. (J2000) (°)	ℓ (°)	b (mag)	V
BPS CS 30339−0040	337.6	10	5.109 00	−36.505 53	336.03	−78.55	13.00
BPS CS 22166−0024	−341.2	10	15.979 33	−12.697 61	134.99	−75.28	13.86
BPS CS 22189−0007	−307.8	10	39.392 17	−12.939 44	188.44	−61.42	13.21
BPS CS 22173−0015	−320.5	10	61.449 04	−17.351 06	211.13	−44.26	13.22
BPS CS 22177−0009	−307.2	10	61.919 21	−25.045 22	221.74	−46.17	14.27
BPS CS 22871−0070	−311.1	10	220.318 08	−18.611 42	336.14	37.08	14.76
BD+01 3070	−306.2	3	230.667 00	1.264 69	3.87	45.48	10.06
BPS CS 30312−0006	−342.7	10	233.240 71	−1.889 31	2.76	41.49	13.65
HD 178443	336.4	10	287.653 58	−43.276 58	354.18	−21.52	9.99
BPS CS 22964−0074	−340.1	10	297.372 50	−39.710 78	0.09	−27.58	14.46
BPS CS 22943−0087	434.7	10	304.826 96	−46.457 64	353.38	−34.05	14.24
V⋆ V1645 Sgr	−326.3	14	305.185 25	−41.118 17	359.80	−33.68	11.96
V⋆ AO Peg	310.8	35	321.770 38	18.633 86	69.94	−22.58	12.83
BPS CS 22948−0093	396.3	10	327.631 37	−41.130 33	0.21	−50.54	15.18
BPS CS 22951−0055	341.9	10	328.686 83	−46.524 06	351.66	−50.36	14.78
HD 214161	−375.3	10	339.283 50	−40.510 67	358.40	−59.31	9.10
BPS CS 22949−0026	313.2	10	350.759 79	−5.214 28	75.14	−59.61	15.23

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