On the run: mapping the escape speed across the Galaxy with SDSS

Angus A. Williams,¹* Vasily Belokurov,¹ Andrew R. Casey^{1,2} and N. Wyn Evans¹

¹Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK ²Monash University, Wellington Rd, Clayton, VIC 3800, Australia

Accepted 2017 February 24. Received 2017 February 24; in original form 2016 December 21

ABSTRACT

We measure the variation of the escape speed of the Milky Way across a range of ~40 kpc in Galactocentric radius. The local escape speed is found to be 521_{-30}^{+46} km s⁻¹, in good agreement with other studies. We find that this has already fallen to 379_{-28}^{+34} km s⁻¹ at a radius of 50 kpc. Through measuring the escape speed and its variation, we obtain constraints on the Galactic mass profile and rotation curve. The gradient in the escape speed suggests that the total mass contained within 50 kpc is $30_{-5}^{+7} \times 10^{10}$ M_☉, implying a relatively light dark halo for the Milky Way. The local circular speed is found to be $v_c(R_0) = 223_{-34}^{+40}$ km s⁻¹ and falls with radius as a power law with index -0.19 ± 0.05 . Our method represents a novel way of estimating the mass of the Galaxy, and has very different systematics to more commonly used models of tracers, which are more sensitive to the central parts of the halo velocity distributions. Using our inference on the escape speed, we then investigate the orbits of high-speed Milky Way dwarf galaxies. For each considered dwarf, we predict small pericentre radii and large orbital eccentricities. This naturally explains the large observed ellipticities of two of the dwarfs, which are likely to have been heavily disrupted at pericentre.

Key words: Galaxy: halo-Galaxy: kinematics and dynamics.

1 INTRODUCTION

The fastest moving stars have long been a subject of fascination and speculation. Vibert Douglas (1956) recounts a discussion between Kopff and Eddington on horse-racing, an enthusiasm of the latter. Kopff stated he was not interested because of course one horse will always run faster than another. 'But why', retorted Eddington. 'When one star moves faster, you are very interested!' The fastest moving stars are intriguing both because of the processes that accelerated them and because they are probes of their environment.

The local escape speed v_{esc} is a measure of the depth of the potential well at the solar position $\Phi(R_0)$. Therefore, the velocities of the speediest stars passing through the solar neighbourhood can in principle provide information on the enclosed mass. There were early investigations on the local escape speed by Caldwell & Ostriker (1981) and Alexander (1982), who suggested values ~450 km s⁻¹. The analysis methods were extended and systematized by Leonard & Tremaine (1990), who emphasized the importance of modelling the shape of the tail of the velocity distribution. They provided a number of physically inspired models and concluded that v_{esc} lay between 450 and 650 km s⁻¹ with 90 per cent confidence. The recent spectroscopic surveys of the Galaxy have stimulated a surge of activity, notably by Smith et al. (2007, hereafter S07) and Piffl et al. (2014, hereafter P14). These authors used data from the RAdial Velocity Experiment spectroscopic survey,

which provided an abundance of information on the kinematics of local stars. P14, who defined the escape speed as the minimum speed needed to reach three virial radii, concluded that it was $533^{+54}_{-41} \text{ km s}^{-1}$ (90 per cent confidence), albeit on the basis of a small sample.

There are some obvious drawbacks. First, there is no guarantee that the high-velocity tail of the distribution function (DF) is actually occupied all the way up to the escape speed. This may mean that the velocity of the fastest moving star is an underestimate of the true escape speed. Second, the method is sensitive to interlopers or contaminants, which may be unrepresentative of a smooth, relaxed stellar population. This includes stars in the process of leaving the Galaxy, such as hypervelocity stars ejected by interaction of binaries with black holes (e.g. Brown 2015; Boubert & Evans 2016). Stars may also be unbound from the Milky Way but none the less bound to the Local Group. Although no such stars are known, the phenomenon is familiar to us through the intergalactic stars identified in nearby clusters like Fornax (Theuns & Warren 1997). Third, and perhaps most awkwardly, the spatial distribution of the highest energy stars is set by the stochastic patterns of cosmic accretion. Therefore, the concept of a smooth velocity distribution may be a fiction at the highest energies, even at locations in the inner Galaxy.

These disadvantages are offset by a number of assets. First, only the high-velocity tail of the DF need be modelled, as opposed to its entirety. This is clearly a massive simplification that sidesteps much of the complexity in building DFs. Second, the method provides a nice counterpoint to other ways of measuring the potential (or

^{*} E-mail: aamw3@ast.cam.ac.uk

equivalently the mass) of the Galaxy. For example, methods based on the Jeans equations use the first and second moments of the distribution, and so are controlled by the main bulk rather than the tail. And, third, future prospects are bright, with huge new data sets of line-of-sight velocities from spectroscopic surveys and of proper motions from the *Gaia* astrometric satellite becoming available.

All previous work has focused on measuring the escape speed locally. This is because samples of high-velocity stars have been small (sometimes minute, for example S07 used just 16 stars, whilst P14 relied on 86 stars) and concentrated in the solar neighbourhood. In this paper, we present the first measurements of the escape speed throughout the Galaxy using a variety of tracers – main-sequence turn-off stars (MSTOs), blue horizontal branch stars (BHBs) and K-giants – extracted from the Sloan Digital Sky Survey (SDSS). Although the MSTOs are located at heliocentric distances within \sim 3 kpc, the BHBs and K-giants in our sample extend out to Galactocentric radii of \sim 50 kpc, providing much greater reach.

Section 2 describes the likelihood function and models for the tail of the velocity distribution, primarily following the formalism established by Leonard & Tremaine (1990). We discuss our tracers and distance estimators in Section 3. This includes the cuts used to build the samples of MSTOs, BHBs and K-giants, as well as to extract the high-velocity stars. Our chosen parametrizations of v_{esc} are given in Section 4, and our analysis methods are described in Section 5. Section 6 presents our results and their implications for the mass profile and rotation curve of the Galaxy, and discusses the possible orbits of three fast moving Galactic satellite galaxies (Bootes III, Triangulum II and Hercules).

2 METHOD

A star travelling at the Galactic escape speed $v_{\rm esc}(\mathbf{x})$ possesses just enough energy to free itself from the Milky Way's potential well. Consequently, a measurement of the escape speed is linked to a measurement of the potential, $\Phi(\mathbf{x})$. In a spherical galaxy, the two quantities are related through the energy equation

$$\frac{1}{2}v_{\rm esc}^{2}(r) + \Phi(r) = \Phi(r_{\rm max}),$$
(1)

so that a star travelling at speed $v_{esc}(r)$ at radius r will be stationary at r_{max} , the radius beyond which stars are no longer bound to the Milky Way. If the Galaxy existed in isolation then $r_{max} = \infty$, but in practice r_{max} quantifies the distance a star must reach from the Galactic Centre before it falls into another potential well.¹ A measurement of $v_{esc}(r)$ therefore amounts to quantifying $\sqrt{-2 [\Phi(r) - \Phi(r_{max})]}$.

Importantly, if one *maps* the escape speed across some range in radius, then the mass enclosed within this range may be inferred directly from the escape speed and its derivative. The mass of a spherically symmetric system is given by

$$M(r) = \frac{r^2}{G} \frac{\mathrm{d}\Phi}{\mathrm{d}r}.$$
 (2)

Eliminating $\Phi(r)$ using equation (1) gives

$$M(r) = \frac{-r^2 v_{\rm esc}}{G} \frac{\mathrm{d}v_{\rm esc}}{\mathrm{d}r},\tag{3}$$

and so the mass within radius *r* is directly related to v_{esc} and dv_{esc}/dr . A direct parametrization of the escape speed is preferable because it reduces possible systematic biases that could arise from instead

¹ Andromeda is \sim 800 kpc from the Milky Way, so a rough estimate for r_{max} is 400 kpc.

modelling the potential. First, if the potential is modelled, then a value for r_{max} must be assumed or inferred. Second, the mass profile of the Galaxy must be assumed out to r_{max} . Since r_{max} is large, and lies far beyond where data are available, this extrapolation can introduce further systematic bias. Indeed, P14 showed that setting $r_{\text{max}} = 3R_{\text{vir}}$ instead of $r_{\text{max}} = \infty$ (as was assumed by S07) introduced a 20 per cent difference in the inferred virial mass. In this work, we will therefore work with direct parametrizations of the escape speed itself, rather than models of the gravitational potential.

Leonard & Tremaine (1990) proposed that the velocity distribution of high-speed stars is a power law of the form

$$p(v) \propto \begin{cases} (v_{\rm esc} - v)^k & \text{if } v_{\rm min} \le v < v_{\rm esc}, \\ 0 & \text{otherwise,} \end{cases}$$
(4)

where v_{min} is a cut-off such that $p(v < v_{min})$ begins to deviate from a power law. Since the model depends only on the speed v, the full velocity DF is implicitly isotropic. An obvious yet important assumption of this model is that stellar velocity distribution really does extend to v_{esc} , instead of truncating at some smaller value. Cosmological simulations provide some reassurance that this is likely the case: S07 showed that star particles are seen with speeds $\gtrsim 0.9 v_{esc}$, so that premature truncations in the stellar velocity distribution may only cause small biases. Reasoning that line-of-sight velocities are measured far more precisely than transverse velocities, it is useful to marginalize the above model over proper motion. This gives

$$p(v_{||}) \propto \begin{cases} (v_{esc} - |v_{||}|)^{k+1} & \text{if } v_{min} \le v_{||} < v_{esc}, \\ 0 & \text{otherwise,} \end{cases}$$
(5)

where $v_{||}$ is the line-of-sight velocity from the perspective of an observer who is stationary with respect to the Galactic Centre at the location of the Sun (hereafter 'Galactocentric line-of-sight velocity' or simply 'line-of-sight velocity'). Since we are only interested in the absolute value of $v_{||}$ in this work, we shall refer to $|v_{||}|$ as $v_{||}$ for brevity.

This model has been applied to data several times, most notably by S07 and P14 using data from the RAVE survey (Kordopatis et al. 2013). Both studies used small samples of stars (<100) close to the sun and found $v_{esc} \sim 530 \,\mathrm{km \, s^{-1}}$. We seek to extend their work by constraining the escape speed of the Galaxy at a variety of locations. We do this by parametrizing the escape speed as a function of position x, so that

$$p(v_{||} | \mathbf{x}) = \begin{cases} C \left(v_{\text{esc}}(\mathbf{x}) - |v_{||}| \right)^{k+1} & \text{if } v_{\min} \le v_{||} < v_{\text{esc}}(\mathbf{x}), \\ 0 & \text{otherwise}, \end{cases}$$
(6)

where C is a location-dependent normalization factor, given by

$$C = \frac{k+2}{(v_{\rm esc}(\mathbf{x}) - v_{\rm min})^{k+2}}.$$
(7)

By far the largest source of uncertainty in our analysis is in the distance to each star. Consequently, we consider the uncertainty in the line-of-sight velocity, longitude and latitude to be negligible. Our likelihood function should therefore be the probability of a line-of-sight velocity, given Galactic coordinates and an imperfect inference on the distance to the star. Writing $\mathbf{x} = (\ell, b, s)$, where (ℓ, b) are Galactic longitude and latitude, respectively, and *s* is the measured line-of-sight distance to the star, we then have

$$p(v_{||} | \ell, b, s) = \int p(v_{||} | \ell, b, s') p(s' | s) \,\mathrm{d}s'.$$
(8)

have

$$\delta M_r = 4.5 - 1.11 \,[\text{Fe/H}] - 0.18 \,[\text{Fe/H}]^2,$$

$$M_{r0} = -5.06 + 14.32 \,x - 12.97 \,x^2 + 6.127 \,x^3$$

$$- 1.267 \,x^4 + 0.0967 \,x^5,$$

$$M_r = M_{r0} + \delta M_r.$$
(14)

Using these formulae, we compute a point estimate of the distance to each star and remove objects with implied heliocentric distances >15 kpc in order to remove objects with spurious absolute magnitude estimates. This finally leaves us with a mother sample of 22071 MSTO tracers. In order to draw samples from p(s' | s) (equation 8), we take the quoted uncertainties in g, r, i and [Fe/H] from the SDSS pipelines and assume that they are normally distributed and uncorrelated. We then draw Monte Carlo samples from each of these distributions and compute the distance for each sample.

3.2 BHB sample

To obtain a clean sample of BHBs, we first select in the de-reddened colour–colour box

$$-0.25 < (g - r)_0 < 0,$$

$$0.9 < (u - g)_0 < 1.4,$$
(15)

which was used by Deason, Belokurov & Evans (2011). There will still be significant contamination from high surface gravity blue stragglers within this box, so we then make the spectroscopic cuts

$$-2 < [Fe/H] < -1,$$

$$3 < \log g < 3.5,$$

$$8300 < T_{eff}/K < 9300,$$
 (16)

giving us 1039 BHBs in total. We estimate the absolute g-band magnitudes, and hence distances, of the BHBs using the relation derived by Deason et al. (2011)

$$M_g = 0.434 - 0.169 (g - r)_0 + 2.319 (g - r)_0^2 + 20.449 (g - r)_0^3 + 94.617 (g - r)_0^4.$$
(17)

We then use the same Monte Carlo strategy as for the MSTO stars to sample from the distance uncertainties.

3.3 K-giant sample

Rather than gathering a sample of K-giants directly from SDSS, as we did for our other two tracer samples, we use the catalogue constructed by Xue et al. (2014), taken from Sloan Extension for Galactic Understanding and Exploration (SEGUE). The SEGUE survey (Yanny et al. 2009) explicitly targeted K-giants in three different categories: 'I-colour K-giants', 'red K-giants' and 'proper-motion K-giants'. All three categories satisfy the photometric constraints

$$0.5 < (g - r)_0 < 1.3,$$

$$0.5 < (u - g)_0 < 2.5,$$
(18)

and have measured proper motions $<11 \text{ mas yr}^{-1}$. Each target category then populates distinct regions of the (u - g) - (g - r) plane. The precise details of the targeting criteria are given in Yanny et al. (2009). On top of the SEGUE selection, Xue et al. then make further cuts such that

$$\log g < 3.5,$$

 $E(B - V) < 0.25 \,\mathrm{mag}.$

p(s' | s) is the probability that s' is the true distance to the star given our imperfect inference s. Finally, we also include a Gaussian outlier model for possibly unbound stars

$$p_{\text{out}}(v_{||}) = \frac{A}{\sqrt{2\pi\sigma^2}} \exp\left(-v_{||}^2/2\sigma^2\right),$$
 (9)

where we fix $\sigma = 1000 \,\mathrm{km \, s^{-1}}$, and *A* is the normalization of the Gaussian over the interval $[\nu_{\min}, \infty]$. We then introduce a free parameter *f* for the fraction of outliers, so that the overall likelihood function is

$$p_{\text{tot}}(v_{||} \mid \ell, b, s) = (1 - f) p(v_{||} \mid \ell, b, s) + f p_{\text{out}}(v_{||}).$$
(10)

This equation represents the likelihood that we will use for the remainder of the paper, while making specific choices for $v_{esc}(x)$. In practice, we compute the RHS of equation (8) using Monte Carlo integration so that

$$p(v_{||} \mid \ell, b, s) \simeq \frac{1}{N} \sum_{n=1}^{N} p(v_{||} \mid \ell, b, s_n),$$
(11)

where each of the s_n is drawn from p(s' | s).

3 SAMPLE SELECTION

In order to measure the variation in the escape speed with position in the Galaxy, we require a set of halo tracers that spans a sufficiently large volume, and has measured line-of-sight velocities and distances. To obtain such a sample, we use the SDSS 9th data release (Ahn et al. 2012). We choose to utilize three distinct sets of tracers: MSTO, K-giant and BHB stars. MSTO stars are numerous, and are mostly observed at distances \sim 3 kpc. These stars allow us to constrain the local escape speed, and its variation relatively close to the sun. K-giants and BHBs are fewer in number, but are bright and have been observed at distances \sim 50 kpc from the sun, pushing the spatial extent of our sample to a range of ~ 40 kpc in Galactocentric radius. Before selecting our sample of high-speed stars, we first constructed a 'mother sample' for each class of tracer using a series of cuts. In addition to the cuts described below, we also removed stars with latitudes $|b| < 20^{\circ}$ in order to remove possible disc contaminants, and stars with radii r > 50 kpc. We compute Galactocentric radii by assuming a solar radius $R_0 = 8.5$ kpc. The full SQL queries used are given in Appendix A.

3.1 MSTO sample

To extract the mother sample of MSTO stars, we start by selecting in the de-reddened colour-magnitude box

$$0.2 < (g - r)_0 < 0.6,$$

14.5 < $r/\text{mag} < 20,$ (12)

. .

where the *r*-band extinction $A_r < 0.3$. We then make cuts on spectroscopic parameters, so that

$$4500 < T_{\rm eff}/{\rm K} < 8000,$$

$$3.5 < \log g < 4,$$

$$-4 < [{\rm Fe}/{\rm H}] < -0.9.$$
(13)

This gives us a sample of metal-poor MSTO stars. We also make cuts to ensure high-quality photometry and line-of-sight velocity measurements for the sample. To compute distances to these stars, we estimate the absolute magnitude in the *r* band using the prescription derived by Ivezić et al. (2008). With $x = (g - i)_0$, we



Figure 1. The distribution in Galactocentric radius and v_{\parallel} of each of our tracer samples. The MSTO stars probe a smaller range in radius, but are numerous, whereas the K-giants probe a much larger distance range, but are fewer in number. The BHBs probe a similar range to the K-giants, but form a far smaller sample. The black points in the MSTO panel highlight the 11 objects that are discussed in Section 6.

Besides our cuts on latitude and Galactocentric radius, we make one further cut on this sample to ensure a clean halo population, namely

$$[Fe/H] < -0.9.$$
 (19)

After applying our latitude and metallicity cuts, there remain 5160 K-giants in our mother sample. Xue et al. compute posterior samples for the distance moduli, μ , of the K-giants. Their table provides the maximum a posteriori value of μ for each star, percentiles of the posterior samples and a summary of the uncertainty, $\delta\mu$. In order to compute distance samples for the K-giants, we simply assume that the uncertainty in the distance modulus is Gaussian, with a standard deviation of $\delta\mu$.

3.4 High-speed sample

Given our three mother samples, we now seek to select members of the high-speed tail of the velocity distribution. In order to do this, we must make a decision for the value of v_{min} in equation (6). S07 did not have access to measurements of stellar parameters for their sample, and adopted a very high threshold of $v_{min} = 300 \text{ km s}^{-1}$ in order to avoid contamination from the Galactic disc. P14 used proper motion measurements to remove stars with disc-like signatures of rotation (although in practice this cut operates very similarly to a metallicity threshold, see their fig. 9), and then introduced a cut of $v_{min} = 200 \text{ km s}^{-1}$. Our cuts in latitude and metallicity serve to remove disc contaminants, which is much less of a concern when using more distant SDSS tracers than it is for samples of RAVE stars. Since we expect very little disc contamination, our choice of v_{min} now depends only on the range of velocities for which we believe the power-law model in equation (6) is valid.

P14 performed a detailed analysis of the simulation suite of Scannapieco et al. (2009) in order to make an educated choice for v_{\min} , and found that the DF of v_{\parallel} in the simulations does not significantly deviate from a power law when $v_{\parallel} > 150 \,\mathrm{km \, s^{-1}}$. In our work, we must consider the fact that our cut must be appropriate across a range of locations. Since the vast majority of our sample are further from the Galactic centre than the sun, and physical reasoning suggests that the escape speed cannot increase as a function of radius, then a cut of 200 km s⁻¹ should guarantee that the power-law model of equation (6) is appropriate at the locations of all the stars in our study. Consequently, we set $v_{\min} = 200 \text{ km s}^{-1}$. Note that this cut is applied to Galactocentric line-of-sight velocities, and so the motion of the sun must be removed beforehand. For this, we assume a local standard of rest $v_{LSR} = 240 \,\mathrm{km \, s^{-1}}$ and a solar peculiar motion $(U_{\odot}, V_{\odot}, W_{\odot}) = (11.1, 12.24, 7.25) \text{ km s}^{-1}$ (Schönrich, Binney & Dehnen 2010). Once this cut is applied, there remain 1573 MSTO

stars, 343 K-giants and 44 BHBs. The distributions of each of our final tracer samples in the $r - v_{||}$ plane is shown in Fig. 1. We note that our sample of 1960 stars represents a factor of >100 increase from the 16 stars used by S07 and a factor >20 compared to the 86 used by P14.

4 CHOICE OF $v_{esc}(x)$

Our primary model for the escape speed is a spherically symmetric power-law model (SPL)

$$v_{\rm esc}(r) = v_{\rm esc}(R_0) \left(\frac{r}{R_0}\right)^{-\alpha/2},$$
(20)

where *r* is Galactocentric radius and $0 \le \alpha \le 1$, on physical grounds. We normalize the model to the escape speed at the solar radius.

We will also report on results when a simple generalization of the SPL, a power law in elliptical radius (EPL), is used. The EPL is parametrized as

$$v_{\rm esc}(r_q) = v_{\rm esc}(R_0) \left(\frac{r_q}{R_0}\right)^{-\beta/2},\tag{21}$$

with $r_q = \sqrt{R^2 + z^2/q^2}$, where *R* and *z* are the usual cylindrical coordinates, and $0 \le q \le \infty$. If the escape speed falls off more rapidly with height above the Galactic plane than it does in the radial direction, then q < 1 (oblate), whereas larger escape speeds high above the plane suggest q > 1 (prolate).

5 PRIORS AND NUMERICAL IMPLEMENTATION

Having constructed our likelihood function and gathered our data, we now need to choose explicit priors on each of our model parameters. We split these parameters into two groups: global parameters, which are the model parameters that are independent of our choice for $v_{esc}(x)$, and $v_{esc}(x)$ parameters, which concern our explicit model of the escape speed.

5.1 Global priors

We allow a different power-law slope in the velocity distribution of each tracer $\mathbf{k} = (k_{\text{MSTO}}, k_{\text{K-giant}}, k_{\text{BHB}})$, and we fit for the outlier fraction, *f*. These four parameters constitute our global model parameters. We anticipate that the three values of *k* should be similar, since all three sets of tracers belong to the halo. On the other hand, the MSTO sample spans a different radial range to the BHB and K-giant samples, and so any differences in our inferred *k* values could suggest a radial variation in *k*. S07 and P14 both used cosmological simulations in order to place informative priors on the value of k. S07 used a flat prior over the range 2.7 < k < 4.7 and P14 adjusted this range to 2.3 < k < 3.7. These priors were necessary for both analyses because of the limited number of stars in both the S07 and P14 samples. Given our much larger sample of stars, we opt for a much less informative prior on each k, which is flat in the range 0 < k < 10. This loose prior will allow us to critically assess the similarity between cosmological simulations and the Galaxy. Our prior on f is flat in the permissible range $0 \le f \le 1$.

5.2 $v_{\rm esc}(x)$ priors

We now detail the priors used on the parameters for our two forms for $v_{esc}(\mathbf{x})$. The SPL model has two parameters: $\boldsymbol{\theta} = (v_{esc}(R_0), \alpha)$. If the Galaxy is to possess a flat or declining rotation curve, as we would expect from physical reasoning, then

$$0 \le \alpha \le 1. \tag{22}$$

Besides this basic physical constraint, we do not include any more information about the value of α , so our prior is flat over the above range. Since $v_{esc}(R_0)$ is a positive definite scale parameter, we adopt a scale invariant Jeffreys prior

$$p(v_{\rm esc}(R_0)) \propto 1/v_{\rm esc}(R_0).$$
 (23)

The EPL model is identical to the SPL model other than the inclusion of the axial ratio, q, so that $\theta = (v_{esc}(R_0), \beta, q)$. The priors we use for β and $v_{esc}(R_0)$ are the same as in equations (22) and (23). For q, we follow Bowden, Evans & Williams (2016) and use a prior

$$p(q) \propto \frac{1}{1+q^2},\tag{24}$$

which places equal weight on oblate $(0 \le q < 1)$ and prolate (q > 1) axial ratios.

5.3 Sampling method

Armed with our priors, we can now perform a Bayesian analysis. Our full parameter space is $\Theta = (k, f, \theta)$, which has six (SPL) or seven (EPL) dimensions. Bayes' theorem for our problem can be written in the following way:

$$p(\boldsymbol{\Theta} \mid \text{data}) = \frac{p(\boldsymbol{\theta}) \, p(\boldsymbol{k}, f) \prod_{j=1}^{3} \prod_{i=1}^{N_j} p_{\text{tot}}(v_{\mid\mid,i}^j \mid \ell_i^j, b_i^j, s_i^j, k_j, f, \boldsymbol{\theta})}{p(\text{data})},$$
(25)

where the index *j* refers to the tracer type (MSTO, K-giant or BHB). Since we do not seek to compare the evidence for our three models, we do not explicitly compute the denominator of Bayes' theorem. In order to constrain the model parameters, we use an MCMC approach.

In order to compute the likelihood, we evaluate the integral in equation (8) using 200 Monte Carlo samples from the distance uncertainties of each star. We draw these samples once, and then use the same set for each likelihood evaluation, thus avoiding random noise in the posterior (McMillan & Binney 2013). We then use the EMCEE code (Foreman-Mackey et al. 2013) to Monte Carlo sample the posterior distribution. EMCEE is a PYTHON implementation of the affine-invariant ensemble sampling approach suggested by Goodman & Weare (2010), where an ensemble of N 'walkers' is used, and the proposal distribution for a given walker is based on the

current positions of N/2 of the remaining walkers. Here, we choose N = 80. We initialize the ensemble by randomly sampling from our prior distributions, and then evolve each walker for 5000 steps. We then inspect the trace plots in each dimension in order to prune our samples for burn-in, which typically requires ~ 200 steps. To assess the convergence of our chains, we first compute the acceptance fraction a_f . For all the analyses presented in this paper, $0.3 < a_f < 0.5$ for the entire ensemble of walkers. In addition, we compute the integrated autocorrelation time τ_f for each of our chains, which is the number of posterior evaluations required to produce independent samples. We find $\tau_f \simeq 50$ in each dimension, so that our chains are run for ~ 100 autocorrelation times, which provides us with ~ 8000 independent samples from the posterior.

6 RESULTS

Fig. 2 shows the one- and two-dimensional projections of the posterior samples of the six SPL model parameters, as well as the inferred median posterior values and uncertainties based on the ± 34 per cent credible intervals. All of the model parameters are well constrained, save for $k_{\rm BHB}$, which is unsurprising given that BHBs are by far the least numerous tracer in our sample.

Our results imply a local escape speed of 521^{+46}_{-30} km s⁻¹, which is in good agreement with S07 and P14. We infer a power-law index $\alpha = 0.37 \pm 0.09$, suggesting that the escape speed is falling rapidly as a function of radius. The middle panel of Fig. 1 is prophetic of this because the edge of the K-giant distribution in the $r - v_{||}$ plane is steep. Fig. 3 shows the run of v_{esc} with radius implied by our inference, with 68 per cent and 94 per cent credible intervals, and the steep drop in the escape speed is clear. For perspective, we also show the distribution of the mother samples of each tracer group in the $r - v_{||}$ plane in Fig. 4. The Milky Way loosens its grip on its inhabitants significantly: our model predicts that the local escape speed is 521^{+46}_{-30} km s⁻¹, and by 50 kpc this has dropped to 379^{+34}_{-28} km s⁻¹.

A priori, it is unclear what the value of k should be. Leonard & Tremaine (1990) point out that violent relaxation would lead to k = 3/2, whereas collisional relaxation gives k = 1 (Spitzer & Shapiro 1972). S07 further showed that the Plummer and Hernquist spheres (Binney & Tremaine 2008) have k = 2.5 and k = 3.5, respectively. The simulations analysed by S07 and P14 both suggest $k \simeq 3$. Clearly, there is a relatively large range of possible values. Due to small sample sizes in previous studies, k has never been measured from data on the Milky Way. Given our significantly larger sample of stars, we are able to do this for the first time. The two tracer samples containing the most stars, MSTO and K-giants, both favour $k \simeq 4 \pm 1$, which is in comfortable agreement with simulations. These results also suggest that k is not a strong function of position, given the rather different radial ranges probed by the MSTO and K-giant samples. The inference on k for the BHB sample is much weaker, and favours a slightly higher value. S07 points out that this is to be expected for small sample sizes. None the less, the inference on $k_{\rm BHB}$ is not in significant tension with the hypothesis that k is constant. Our results vindicate the choice of prior by S07, while the range used by P14 is a touch on the low side.

Fig. 2 shows a strong degeneracy between k_{MSTO} and $v_{\text{esc}}(R_0)$, which can be encoded by the empirical covariance matrix of the samples

$$\operatorname{Cov}\left(k_{\mathrm{MSTO}}, v_{\mathrm{esc}}(R_{0})\right) = \begin{bmatrix} 0.84 & 37\,\mathrm{km\,s^{-1}}\\ 37\,\mathrm{km\,s^{-1}} & 1713\,\mathrm{km^{2}\,s^{-2}} \end{bmatrix}, \quad (26)$$



Figure 2. One- and two-dimensional projections of our MCMC samples for the SPL fit. The 68 per cent and 94 per cent credible intervals are shown in the two-dimensional projections, and the median parameter values and uncertainties computed using the ± 34 per cent credible intervals of the one-dimensional projections are shown above the one-dimensional histograms. We multiply the outlier fraction *f* by the number of stars in our sample, $N_* = 1960$ in order to make the inferred value easier to interpret. All of our parameters are well constrained, except k_{BHB} , which is unsurprising given that this is the tracer sample with the smallest number of stars by an order of magnitude. Notable degeneracies are between k_{MSTO} and $k_{K-giant}$, which are constrained to be close to equal, and between $v_{esc}(R_0)$ and $k_{K-giant}$. Larger values of k suggest larger values of $v_{esc}(R_0)$, as expected. See text for discussion.

i.e. if k_{MSTO} increases by unity, then $v_{\text{esc}}(R_0)$ increases by 37 km s⁻¹. This is to be expected, and is the reason why a narrow prior on k was necessary in previous work. Fig. 1 of P14 nicely demonstrates the appearance of this degeneracy for varying sample sizes. Fortunately, our sample is large enough to locate the maximum along the degeneracy. The same degeneracy is seen between $k_{\text{K-giant}}$ and the local escape speed, though it is broader. Note that this explains why our statistical uncertainty on the local escape speed is larger than that of P14, who found $533^{+54}_{-41} \text{ km s}^{-1}$ at 90 per cent confidence, compared to our 90 per cent credible interval of $521^{+88}_{-45} \text{ km s}^{-1}$. Our larger 95th percentile of 690 km s⁻¹ is a consequence of the degeneracy between k and $v_{\text{esc}}(R_0)$: the 95th percentile of the posterior on k_{MSTO} is 6, which is considerably larger than the upper end of P14's prior.

The inferred outlier fraction is very small, $f \simeq 0.001$, but nonzero. This suggests that there are one or two outliers in our sample. Inspection of Fig. 1 suggests one clear candidate: there is an MSTO star at $r \simeq 10$ kpc, shown as a black point, with a measured lineof-sight velocity of 518.2 km s⁻¹, which is more than 100 km s⁻¹ larger than any other star at a comparable radius in our sample. Otherwise, there are no obvious outliers through visual inspection. As a check of this intuition, we calculated the outlier probability of each star in our sample as

$$p(\text{outlier} | v_{||}, \ell, b, s) = \frac{\bar{f} p_{\text{out}}(v_{||})}{\bar{f} p_{\text{out}}(v_{||}) + (1 - \bar{f}) p(v_{||}|\ell, b, s, \bar{\theta}, \bar{k})},$$
(27)

using the model parameters obtained by taking the median values of each of the one-dimensional marginalized posterior distributions, $\bar{\Theta}$. The largest outlier probability is >0.999, and belongs to the object identified visually in Fig. 1. Otherwise, the largest outlier probability is <0.01, and so we conclude that this object is the only probable outlier in the sample. Having identified this outlier, we visually inspected its spectrum and image data from SDSS. From



Figure 3. Our inference on the escape speed as a function of Galactocentric radius. The median posterior result is shown as a dark blue line, and the 68 per cent (94 per cent) credible interval is a dark (light) blue band. The result using RAVE data from P14 and the associated 90 per cent credible interval is also shown, and is in good agreement with our inference. We measure a significant gradient in the escape speed, such that it has already fallen by $\sim 100 \,\mathrm{km \, s^{-1}}$ by a radius of 30 kpc.



Figure 4. The distribution of our mother samples of stars in the $r - v_{||}$ plane, with horizontal dashed lines at $v_{||} = \pm 200 \text{ km s}^{-1}$, our cut in line-of-sight velocity. The coloured bands are our inference on the escape speed as a function of radius. The 'spur' at negative line-of-sight velocities is from K-giants belonging to the Sagittarius stream. Note that the contamination in our high-speed sample from these stars is negligible, since the maximum velocity that the stream centroid reaches is ~150 km s^{-1} (Belokurov et al. 2014) with a dispersion of ~20 km s^{-1}.

the image data it is clear that this object is a galaxy, and has been misclassified by the spectroscopic pipeline of SDSS. Having found a galaxy contaminant in our sample, we added a further constraint to our sqL query that all of the MSTO targets should be morphologically classified as stars (as well as spectroscopically, as in our original query) in order to check for any other similar contaminants. The high-speed sample is reduced in size by 11 objects when this stricter constraint is applied: all 11 of these targets, including the galaxy outlier previously discussed, are shown as black points in Fig. 1. We visually inspected the spectra and image data for all of these objects, and concluded that, other than the original outlier that we identified, they are likely partially blended stars. Although this



Figure 5. Posterior distribution for the EPL model. The results are very similar to the SPL model, and the extra parameter q is found to be ~ 1 , suggesting that the Galactic potential is probably spherical, though our uncertainties are quite large.

might mean that the photometry for some of these objects may be unrepresentative to some degree, visual inspection of their spectra suggests that they likely have reliable stellar parameters and lineof-sight velocities. For this reason, we opted not to remove them from our sample, but we did re-run our analysis with these objects removed to check for inconsistencies. We found that none of our conclusions change when these targets are removed from the sample: the comparison of the posterior distributions with and without these objects is shown in Appendix B. We also experimented with a stricter colour cut of $0.2 \le (g - r)_0 \le 0.35$, to further remove objects with potentially spurious stellar parameters, but again found that this made no quantitative difference to our conclusions. Later, when we compare our model to the data, we remove the bona-fide galaxy contaminant from the sample, but retain the other 10 objects.

Fig. 5 shows the results of the EPL analysis. We removed the global parameters from the figure because the results were essentially identical to those found for the SPL model, without any interesting additional degeneracies. The addition of a halo axial ratio, q, leads to the conclusion that the escape speed falls at the same rate in all directions from the Galactic Centre, with $q = 1.03^{+0.63}_{-0.32}$. This in turn implies that the Galactic potential is likely spherical, although our uncertainties are large: the data are compatible with q = 0.7–1.6. Since q corresponds to the flattening of the potential, this corresponds to a relatively wide range of flattenings in the dark matter halo. Hopefully, with more data, the method will provide a more useful constraint on the symmetries of the Galactic potential. The other two parameters of the EPL model are the same as those in the SPL model, and the inferred values and uncertainties are indistinguishable.

6.1 The mass and rotation curve of the Milky Way

Having mapped the escape speed across the Galaxy, we are now able to convert this measurement into a mass profile M(r) and rotation



Figure 6. Left: cumulative mass distribution within spherical shells, predicted by the SPL model. The 68 per cent (94 per cent) credible interval is shown as a dark (light) blue band around our median result. We predict a relatively light Milky Way within 50 kpc, with $M(50 \text{ kpc}) = 29.8^{+6.9}_{-5.2} \times 10^{10} \text{ M}_{\odot}$. Right: the rotation curve, also from the SPL model. The local circular speed is $v_c(R_0) = 223^{+40}_{-34} \text{ km s}^{-1}$, and the rotation curve falls with radius as a power law with index -0.19 ± 0.05 . In both panels are the results from various other studies, see text for discussion.

curve $v_c(r)$ for the Galaxy using equation (3) and $v_c^2(r) = G M(r)/r$. The mass and circular speed profiles for the Galaxy can be inferred over the range for which we have mapped $v_{esc}(r)$: $6 \leq r/\text{kpc} \leq$ 50. For example, the local circular speed implied by our method is $v_c(R_0) = 223^{+40}_{-34} \text{ km s}^{-1}$, which is pleasingly consistent with a multitude of other methods.

Fig. 6 shows the mass and circular speed profiles implied by the SPL model, along with associated 68 per cent and 94 per cent credible regions. Our model predicts $M(50 \text{ kpc}) = 29.8^{+6.9}_{-5.2} \times 10^{10} \text{ M}_{\odot}$. For reference, we have also plotted the results from a selection of other studies. Xue et al. (2008, X08), Deason et al. (2012, D12) and Williams & Evans (2015, WE15) all used samples of halo BHBs taken from SDSS. D12 and WE15 applied DF models to the data, and infer systematically higher masses than we do here, with $M(50 \text{ kpc}) \simeq 45 \times 10^{10} \text{ M}_{\odot}$. Both are consistent with the 94 per cent credible interval of our inference, but there is a hint that there is a discrepancy between DF methods and the present approach. X08, on the other hand, compared SDSS BHBs to cosmological simulations, and their result is comfortably in agreement with ours.

W99, like D12 and WE15, used a DF approach, but applied their method to globular clusters and dwarf galaxies. Their sample was small, with only 27 objects, and so their uncertainty is large. Their preferred mass of $M(50 \text{ kpc}) = 54^{+2}_{-36} \times 10^{10} \text{ M}_{\odot}$, like the other two DF approaches, is significantly larger than our result, although the large asymmetric uncertainty removes any possible tension.

The final study to which we compare is that of Gibbons, Belokurov & Evans (2014, G14), who modelled the disruption of the Sagittarius stream. They exploited the fact that the apocentric precession of the stream should be sensitive to the details of the gravitational potential. Their inference produces very similar results to our work, with $M(50 \text{ kpc}) = 29 \pm 5 \times 10^{10} \text{ M}_{\odot}$. The two analyses, though very different in detail, produce near identical results.

It is worth noting that our method clearly possesses very different systematic uncertainties when compared to more common approaches in the literature. Most dynamical models of halo tracers, like DF and Jeans analyses, are most sensitive to the central parts of the velocity distributions. This is particularly true of Jeans analyses, which generally only model the first and second moments of the velocity distributions. DFs satisfy the full collisionless Boltzmann equation, and therefore the entire infinite hierarchy of Jeans



Figure 7. The distribution of dwarf galaxies around the Milky Way in the $r - \sqrt{3} v_{||}$ plane. The blue bands are our inference on the escape speed as a function of radius. Line-of-sight velocities have been multiplied by $\sqrt{3}$, as is done in the literature to account for unknown tangential velocities. If the true speeds of Triangulum II, Tucana II, Grus 1, Bootes III and Hercules are close to $\sqrt{3} v_{||}$, then they are likely to be unbound.

equations, but this generally comes at the cost of large systematic uncertainties that arise from the chosen form of the model (Wang et al. 2015). Our approach moves the focus to the tail of the velocity distribution, and is therefore complimentary to other approaches.

6.2 The orbits of Milky Way dwarf galaxies

Fig. 7 shows the distribution of known Milky Way dwarf galaxies in the $r - \sqrt{3} v_{||}$ plane. It is typical in the literature to multiply the line-of-sight velocity by $\sqrt{3}$ as a crude way of accounting for unknown tangential velocities. We see that most of the dwarfs are enveloped by the escape speed curves, with a similar shape to the $r - v_{||}$ distribution of stars (Fig. 4). However, some of the dwarfs seem likely unbound based on our estimate of the escape speed. On the other hand, Λ cold dark matter simulations predict that 99.9 per cent of subhaloes should be bound to their hosts (Boylan-Kolchin et al. 2013). A reconciliation of these two statements is to conclude that the $\sqrt{3}v_{\parallel}$ approximation for the total speed of these dwarfs is likely unrealistic in these cases. Given our inference on the escape speed, the assumption that these objects are bound allows us to place constraints on their orbits. The red points in Fig. 7 are Bootes III (Carlin et al. 2009; Grillmair 2009), Triangulum II (Laevens et al. 2015) and Hercules (Belokurov et al. 2007), all of which are likely associated with the Milky Way and have large lineof-sight velocities. They are also at radii where our inference on the escape speed is trustworthy (although Hercules is a somewhat marginal case). The two yellow points are Tucana 2 and Grus 1 (Koposov et al. 2015; Bechtol et al. 2015), which are probably dwarfs of the Magellanic clouds (Jethwa, Erkal & Belokurov 2016). We do not seek to constrain the orbital properties of these dwarfs because of their more complex orbital histories. For the three Milky Way dwarfs, we would like to characterize their orbits through their pericentre radii r_{peri} , apocentre radii r_{apo} and orbital eccentricities

$$e = \frac{r_{\rm apo} - r_{\rm peri}}{r_{\rm apo} + r_{\rm peri}}.$$
(28)

In order to compute posterior distributions on r_{peri} , r_{apo} and e, we first write

$$p(X \mid v_r, r) = \int p(X \mid v_r, r, \mathbf{\Theta}) \, p(\mathbf{\Theta} \mid \text{data}) \, d\mathbf{\Theta},$$
(29)

where $X = (r_{\text{peri}}, r_{\text{apo}})$ and $p(\Theta | \text{data})$ is the posterior distribution on our model of the escape speed. We have written $v_{||} \simeq v_r$, where v_r is the velocity away from the centre of the Galaxy, because these objects are distant. We then have

$$p(X | v_r, r, \boldsymbol{\Theta}) = \int p(X | v_r, v_T, r, \boldsymbol{\Theta}) p(v_T | v_r, r, \boldsymbol{\Theta}) dv_T,$$

=
$$\int \delta [X - f_X(v_r, v_T, r, \boldsymbol{\Theta})] p(v_T | v_r, r, \boldsymbol{\Theta}) dv_T,$$
(30)

where $v_{\rm T}$ is the transverse speed. In a spherical potential, $v_{\rm T}$, v_r and *r* completely specify the orbit, and hence $r_{\rm peri}$, $r_{\rm apo}$ and *e*. This gives rise to the delta-function in equation (30), where f_X represents the relation between $(v_r, v_{\rm T}, r)$ and *X*. Finally, we need to specify $p(v_{\rm T} | v_r, r, \Theta)$. To do this, we assume that the high-speed dwarfs follow the same DF as the high-speed stars, given by equation (4)

$$p(v_{\rm T} | v_r, r, \boldsymbol{\Theta}) \propto \left(v_{\rm esc}(r) - \sqrt{v_{\rm T}^2 + v_r^2} \right)^k.$$
 (31)

We make this assumption for simplicity, although Erkal et al. (2016) note that the velocity distributions of subhaloes in the VLII sim-

ulations (Diemand et al. 2008) are not the same as those of the dark matter particles, and have velocity dispersions of size 160–200 km s⁻¹. Measured velocity dispersions for Milky Way halo stars tend to be somewhat smaller than this, at \sim 100–150 km s⁻¹ (Evans et al. 2016). Hence, for a given radial velocity, larger tangential velocities could be more probable than the estimate of equation (31), which in turn means that the orbit of the dwarf may be more eccentric than our simple calculation suggests.

Since f_X is not analytic, we must solve for r_{peri} , r_{apo} and e numerically by solving the energy equation in the potential $\Phi(r)$ implied by our model of the escape speed. In practice, we sample $p(X | v_r, r)$ in the following way. Given a set of parameters in our posterior samples for the SPL model, we first draw a tangential velocity from equation (31). For k, we use k_{MSTO} , and we normalize the speed distribution between v_r and $v_{\text{esc}}(r)$. Then, we solve for r_{peri} , r_{apo} and e and store the result. This process is repeated for every set of model parameters in our posterior samples, and the resulting histograms in r_{peri} , r_{apo} and e are then faithful representations of $p(X | v_r, r)$.

The results of this procedure are shown in Fig. 8. All three dwarfs are expected to be on very eccentric orbits. Bootes III and Triangulum II have $e \simeq 0.95$, while Hercules has $e \simeq 0.8$, though with a somewhat broader distribution. This is aligned with our intuition: if the line-of-sight velocity alone is relatively close to the escape speed at the radius of the dwarf, then the tangential velocity cannot be large and hence the orbit must be eccentric. As a consequence, the dwarfs have large apocentres, $r_{\rm apo} \sim 100-300$ kpc and considerably smaller pericentres $r_{\rm peri} \sim 10$ kpc (although the posterior on the pericentre of Hercules is significantly less peaked than for Bootes III and Triangulum II).

Küpper et al. (2017) argued that the observed ellipticity of Hercules ($\epsilon \simeq 0.7$) and its large line-of-sight velocity are suggestive that it has 'exploded' as a consequence of its last pericentre passage. Using N-body simulations, they arrived at an estimate of the orbital eccentricity $e \simeq 0.95$, which is larger than our value but not in significant tension with it. Bootes III has a similar morphology $(\epsilon \simeq 0.5)$ in keeping with the picture that these satellites are on orbits that will cause them to disrupt into streams after comparatively few orbital periods. Both Bootes II and Hercules have positive line-of-sight velocities: they are travelling away from the Galactic Centre, suggesting that they have undergone at least one pericentre passage. Triangulum II, on the other hand, has a negative line-of-sight velocity and a relatively small ellipticity ($\epsilon \simeq 0.2$). Therefore, it is plausible that Triangulum II is on first infall and is about to undergo a large amount of disruption on its first pericentre passage.



Figure 8. Posterior distributions on orbital pericentre r_{peri} , apocentre r_{apo} and eccentricity *e* for Bootes III, Triangulum II and Hercules. If the dwarfs are bound, they must be on very eccentric orbits.

6.3 Model performance and systematics

We now seek to assess how well our model fits the data by performing posterior predictive checks. Since our model is only generative in line-of-sight velocities, and not in the positions of stars, we should compare the distribution in $v_{||}$ of the data with our model. There is a subtlety in how this must be done, however. The velocity distribution of high-speed stars is position dependent in the model, owing to the spatial variation of the escape speed. The fastest star at a given radius is likely to be travelling slower than its counterparts at smaller radii, where the escape speed is larger. This effect means that we must take into account the number of stars that have been observed at each radius in the Galaxy for our comparison between model and data to be meaningful. We therefore write

$$p(v_{||} | \mathbf{\Theta}) = \frac{\int p(v_{||} | r, \mathbf{\Theta}) p(r) dr}{\int p(v_{||} | r, \mathbf{\Theta}) p(r) dr dv}$$

=
$$\frac{\int (v_{esc}(r) - v_{||})^{k+1} p(r) dr}{\int (k+2)^{-1} (v_{esc}(r) - v_{min})^{k+2} p(r) dr},$$
(32)

where p(r) is the probability of observing a star at radius r. We approximate p(r) by binning the mother sample of each tracer in radius and fitting a cubic spline to the resulting histogram. The two integrals in equation (32) are then computed using 12 point Gauss–Legendre quadrature. We then draw samples from the posterior predictive distribution

$$p(v_{||} | \text{data}) = \int p(v_{||} | \boldsymbol{\Theta}) p(\boldsymbol{\Theta} | \text{data}) d\boldsymbol{\Theta}, \qquad (33)$$

which is the predicted distribution of $v_{||}$ given the knowledge we have gained by analysing our SDSS sample. To sample from this distribution, we draw from equation (32) a mock data set that is the same size as the SDSS data for each set of parameters in our MCMC chain (Gelman et al. 2013), producing many replicated data sets. In order to draw from equation (32), we use inverse transform sampling, where we numerically compute the CDF

$$F(v_{||} | \mathbf{\Theta}) = \int_{v_{\min}}^{v_{||}} p(v_{||}' | \mathbf{\Theta}) dv_{||}'$$
(34)

on a grid of points in v_{\parallel} , and compute the inverse function F^{-1} as a cubic spline. Finally, we draw a set of points *u*, uniformly distributed between 0 and 1, and compute the corresponding velocities via $v_{\parallel} = F^{-1}(u)$.

Fig. 9 shows the comparison of the posterior predictive distributions with the data for our MSTO, K-giant and BHB samples. For each tracer, we constructed a histogram in line-of-sight velocity, and show the number of counts as a black point at each bin centroid. The median number of counts in each bin from our replicated data sets is shown as a solid line in each panel. The 68 per cent and 95 per cent intervals are shown as bands around the median, and the final band shows the full extent of the number of counts. Our model reproduces the data very well over a range of ~ 2.5 orders of magnitude in the number of counts for the MSTO sample. Similarly good agreement is seen for our other two tracers.

Besides verifying that the model is a good representation of the data, we also seek to understand some of the possible sources of systematic uncertainty. We investigated three possibilities: our choice of the local standard of rest v_{LSR} , our choice of the cut velocity v_{min} and inconsistencies between the different tracer groups. In the analysis presented in the rest of the paper, we assumed $v_{LSR} = 240 \text{ km s}^{-1}$. In order to test the influence that this assumption has on our inference, we re-ran all of our analyses with a lower value of $v_{LSR} = 220 \text{ km s}^{-1}$. The local escape speed is then inferred

to be 542^{+56}_{-37} km s⁻¹, which is consistent with our previous analysis (note that different stars will enter our high-speed sample when a different value of v_{LSR} is used). The same is true for the rest of the model parameters.

We chose $v_{\min} = 200 \text{ km s}^{-1}$ because we were not particularly concerned about contamination from disc stars. In order to check the sensitivity of our work to this value, we re-ran the analysis with $v_{\rm min} = 250 \,\rm km \, s^{-1}$. This results in a significantly smaller sample of 644 stars (539 MSTO, 99 K-giants, 6 BHBs). We thus expect much larger uncertainties in all of our model parameters. This is indeed the case, and we find that the values of k for each of the tracers are poorly constrained compared to our full analysis, leading to a worse determination of the local escape speed. This is because the degeneracy between k and $v_{esc}(R_0)$ is not broken as effectively by these data, which pushes up our estimates of k and $v_{esc}(R_0)$, as predicted by S07. Specifically, we find $v_{\rm esc}(R_0) = 617^{+77}_{-84} \,\rm km \, s^{-1}$. This value is none the less consistent with our full analysis, due to the inflated uncertainties. Correspondingly, the inferred values of k are ~ 6 , with uncertainties \sim 2, which again are consistent with our previous estimates, but systematically higher. Given how well our full model represents the data, we would suggest that the speed distribution of halo stars does not significantly deviate from a power law at speeds >200 km s⁻¹, thus vindicating our choice of $v_{\rm min} = 200 \,\rm km \, s^{-1}$. Larger choices of v_{min} should only be necessary in circumstances where disc contamination is a more serious concern. If this is the case, then our checks with $v_{\min} = 250 \text{ km s}^{-1}$ imply that a prior on the value of k is probably necessary.

Our final check was to understand whether there is any tension between the different tracer samples. We ran our analysis again using only the MSTO sample, and then with the K-giants alone. We did not perform a run with the very small sample of BHBs. We found no tension whatsoever between the results from the MSTO only run and the K-giant only run. The K-giants favour a marginally larger value of α , but the difference is not marked: the median values of α for each run lie comfortably within the 68 per cent credible regions of the other run.

7 CONCLUSIONS

The main achievement of this paper is the measurement of the escape speed and its variation over Galactocentric radii between \sim 8 and 50 kpc. This was done using a variety of tracer populations – MSTOs, BHBs and K giants – extracted from the SDSS. We find that the local escape speed is 521^{+46}_{-30} km s⁻¹ in good agreement with the studies by Smith et al. (2007) and Piffl et al. (2014) using data from the RAVE survey.

Our sample extends out to Galactocentric radii ~50 kpc, so we can track the gradient in the escape speed with radius for the first time. At 50 kpc, the escape speed has fallen to a value of 379^{+34}_{-28} km s⁻¹, indicative of a rapid power-law decline ($\alpha = 0.37 \pm 0.09$). This suggests that the total mass contained within 50 kpc is $30^{+7}_{-5} \times 10^{10}$ M_{\odot}, implying a relatively light dark halo for the Milky Way. Our mass is pleasingly consistent with the work of Gibbons et al. (2014), who modelled the disruption of the Sagittarius dwarf galaxy, and therefore made very different assumptions in the analysis.

Perhaps the most striking thing about this method is its simplicity. We have demonstrated that modelling the tail of the velocity distribution can recover results in excellent agreement with other, much more elaborate, methods. In particular, there is no evidence of any bias as the values of the escape velocity and circular speed



Figure 9. Posterior predictive checks of the SPL model. The black points are the number of stars observed in bins of $v_{||}$. The solid lines are the median counts from our replicated data sets. The coloured bands around the median line are the 68 per cent, 95 per cent intervals and full range of the counts in the replicated data sets. Our model produces replicated data that resemble the SDSS sample.

are in good agreement with other methods. As the size of data sets becomes huge, the computational cost of any algorithm will become an important yardstick. Our method is computationally much cheaper than full DF approaches, and will be able to rapidly provide robust results when faced with the output of enormous surveys.

The advent of the second *Gaia* release in 2017 and the first results from the WEAVE spectroscopic survey in 2018 will provide proper motions and line-of-sight velocities for massive samples of halo stars. The computational challenges of modelling such data sets are severe. We suggest that analysis of the high-velocity tails may be the optimum method to quickly extract information on the potential of the Galaxy. Current data sets preclude a stringent measurement on the shape of the halo using the variation in escape speed with position, but we anticipate that this too will become possible in the very near future.

Although our model provides a counterpoint to other methods in the literature, it comes with some caveats. It is not known whether the velocity distributions of halo stars extend to v_{esc} , and it may be the case that they truncate at a smaller value. If this is the case, then our work provides a lower limit on the escape speed. Cosmological simulations analysed by S07 and P14 are indicative that stars are indeed found at such high speeds, with particles consistently found with speeds in excess of 0.9 v_{esc} , which provides some reassurance that this systematic bias is unlikely to be large. Another caveat is that the periods of orbits with speeds approaching v_{esc} are long, so that the DF may not be phase mixed at very high speeds. Although S07 and P14 showed that the model is effective in recovering the escape speeds and virial masses of simulated haloes, this may not be the case for the Galaxy.

As an application, we have used our model of the escape speed to study the orbital properties of the extreme Galactic satellites, Hercules, Bootes III and Triangulum II. All three must have highly eccentric orbits, with Bootes III and Triangulum II possessing an eccentricity $e \simeq 0.95$, and Hercules $e \simeq 0.8$. The dwarfs must have large apocentres, $r_{\rm apo} \sim 100-300$ kpc and small pericentres $r_{\rm peri} \sim 10$ kpc. Given that Hercules and Bootes III are now moving radially outwards, this implies that they must have at least one pericentric passage already and so we expect their morphology to be elongated and disrupted, as is the case. Triangulum II is moving inwards, and so must be on first infall, which is consistent with its roundish shape and ellipticity of 0.2.

ACKNOWLEDGEMENTS

AAW acknowledges the support of STFC. We thank Sergey Koposov for plenty of helpful advice, and Prashin Jethwa, Jason Sanders, Kathryn Johnston, Jorge Peñarrubia and Gabriel Torrealba for useful conversations. This project was developed in part at the 2016 NYC *Gaia* Sprint, hosted by the Center for Computational Astrophysics at the Simons Foundation in New York City. We thank the referee, James Binney, for a helpful report that led to important clarifications in the paper.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

REFERENCES

- Ahn C. P. et al., 2012, ApJS, 203, 21
- Alexander J. B., 1982, MNRAS, 201, 579
- Bechtol K. et al., 2015, ApJ, 807, 50
- Belokurov V. et al., 2007, ApJ, 654, 897
- Belokurov V. et al., 2014, MNRAS, 437, 116
- Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton, NJ
- Boubert D., Evans N. W., 2016, ApJ, 825, L6
- Bowden A., Evans N. W., Williams A. A., 2016, MNRAS, 460, 329
- Boylan-Kolchin M., Bullock J. S., Sohn S. T., Besla G., van der Marel R. P., 2013, ApJ, 768, 140
- Brown W. R., 2015, ARA&A, 53, 15
- Caldwell J. A. R., Ostriker J. P., 1981, ApJ, 251, 61
- Carlin J. L., Grillmair C. J., Muñoz R. R., Nidever D. L., Majewski S. R., 2009, ApJ, 702, L9
- Deason A. J., Belokurov V., Evans N. W., 2011, MNRAS, 416, 2903

- Deason A. J., Belokurov V., Evans N. W., An J., 2012, MNRAS, 424, L44 Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J.,
- 2008, Nature, 454, 735
- Erkal D., Belokurov V., Bovy J., Sanders J. L., 2016, MNRAS, 463, 102
- Evans N. W., Sanders J. L., Williams A. A., An J., Lynden-Bell D., Dehnen W., 2016, MNRAS, 456, 4506
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
- Gelman A., Carlin J., Stern H., Dunson D., Vehtari A., Rubin D., 2013, Bayesian Data Analysis, 3rd edn. Chapman & Hall/CRC/Taylor & Francis
- Gibbons S. L. J., Belokurov V., Evans N. W., 2014, MNRAS, 445, 3788
- Goodman J., Weare J., 2010, Commun. Appl. Math. Comput. Sci., 5, 65
- Grillmair C. J., 2009, ApJ, 693, 1118
- Ivezić Ž. et al., 2008, ApJ, 684, 287
- Jethwa P., Erkal D., Belokurov V., 2016, MNRAS, 461, 2212
 Koposov S. E., Belokurov V., Torrealba G., Evans N. W., 2015, ApJ, 805, 130
- Kordopatis G. et al., 2013, AJ, 146, 134
- Küpper A. H. W., Johnston K. V., Mieske S., Collins M. L. M., Tollerud E. J., 2017, ApJ, 834, 112
- Laevens B. P. M. et al., 2015, ApJ, 802, L18
- Leonard P. J. T., Tremaine S., 1990, ApJ, 353, 486
- McMillan P. J., Binney J. J., 2013, MNRAS, 433, 1411
- Piffl T. et al., 2014, A&A, 562, A91 (P14)
- Scannapieco C., White S. D. M., Springel V., Tissera P. B., 2009, MNRAS, 396, 696
- Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
- Smith M. C. et al., 2007, MNRAS, 379, 755 (S07)
- Spitzer L., Jr, Shapiro S. L., 1972, ApJ, 173, 529
- Theuns T., Warren S. J., 1997, MNRAS, 284, L11
- Vibert Douglas A., 1956, Arthur Stanley Eddington. Cambridge Univ. Press, Cambridge
- Wang W., Han J., Cooper A. P., Cole S., Frenk C., Lowing B., 2015, MNRAS, 453, 377
- Williams A. A., Evans N. W., 2015, MNRAS, 454, 698
- Xue X. X. et al., 2008, ApJ, 684, 1143
- Xue X.-X. et al., 2014, ApJ, 784, 170
- Yanny B. et al., 2009, AJ, 137, 4377

APPENDIX A: SQL QUERIES

Here, we give the sQL queries that we used to obtain the MSTO and BHB samples.

(i) MSTO query: SELECT *

FROM sdssdr9.specphotoall AS spa, sdssdr9.sppparams AS spp

WHERE spp.specobjid=spa.specobjid AND spp.scienceprimary=1 AND spa.class=`STAR' AND spa.extinction_r<0.3 AND spa.dered_g-spa.dered_r BETWEEN 0.2 AND 0.6 AND spa.dered_r BETWEEN 14.5 AND 20. AND spp.fehadop BETWEEN -4. AND -0.9 AND spp.loggadop BETWEEN 3.5 AND 4. AND spp.teffadop BETWEEN 3.5 AND 4. AND spp.teffadop BETWEEN 4500. AND 8000. AND spa.psfmagerr_g BETWEEN 0. AND 0.04 AND spa.psfmagerr_r BETWEEN 0. AND 0.04 AND spa.psfmagerr_i BETWEEN 0. AND 0.04 AND spa.fehadopunc < 0.1

AND (spp.zwarning=0 OR spp.zwarning=16) AND spp.snr > 20.

When we apply our extra morphological cut, the condition

spa.TYPE = 6

is included in the above query. (ii) BHB query: SELECT *

FROM sdssdr9.specphotoall AS spa, sdssdr9.sppparams AS spp

WHERE spp.specobjid=spa.specobjid AND spp.scienceprimary=1 AND spa.class=`STAR' AND spa.psfmag_g-spa.extinction_g-spa .psfmag_r

```
+spa.extinction_r BETWEEN -0.25 AND 0.
AND spa.psfmag_u-spa.extinction_u-spa
.psfmag_g
+spa.extinction_g BETWEEN 0.9 AND 1.4
AND spp.fehadop BETWEEN -2. AND -1.
AND spp.loggadop BETWEEN 3. AND 3.5
AND spp.teffadop BETWEEN 8300. AND 9300.
AND (spp.zwarning=0 OR spp.zwarning=16)
AND spp.snr>20.
```

APPENDIX B: RESULTS WITH STRICTER MORPHOLOGICAL CUTS

Here, we show the results when objects that have not been morphologically classified as stars are removed from the MSTO catalogue. Fig. B1 shows the comparison of the posterior distributions from our original SPL run and when the stricter morphological cut is applied.



Figure B1. One- and two-dimensional posterior distributions of two runs: our original SPL run, and the run where the stricter morphological cut discussed in Section 6 is applied. The projections of the posterior distributions from the two runs are completely consistent with one another, save for the outlier fraction. This is because, when the morphological cut is applied, the galaxy contaminant with a spurious line-of-sight velocity discussed in Section 6 is removed.

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