On the Run: mapping the escape speed across the Galaxy using high-speed stars in SDSS

Angus A. Williams[1], Vasily Belokurov[1], Andrew R. Casey[1] & N. Wyn Evans[1]

[1] *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK*

ABSTRACT

We measure the variation of the escape speed of the Galaxy across a range of $\sim$ 40 kpc in Galactocentric radius. We measure the local escape speed to be $\sim 530\,\mathrm{km s}^{-1}$, in good agreement with other studies, and find that this has already fallen to $\sim 400\,\mathrm{km s}^{-1}$ at a radius of 30 kpc. Through measuring the escape speed and its variation, we obtain constraints on the Galactic potential as a whole. In particular, the gradient in the escape speed with radius suggests that the total mass contained within 50 kpc is $\sim 3 \times 10^{11}\,\mathrm{M}_\odot$, implying a relatively light dark halo for the Milky Way. 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 dwarf we consider, we predict small pericenter 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 as they passed through pericenter.

Key words:

1 INTRODUCTION

2 METHOD

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

$$p(v) \propto \begin{cases} (v_{\mathrm{esc}} - v)^k & \text{if } v_{\min} \leqslant v < v_{\mathrm{esc}}, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

v_{esc} is the escape speed and $v_{\min}$ 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 distribution function is implicitly isotropic. Since radial velocities are measured far more precisely than transverse velocities, it is useful to marginalise the above model over proper motion. This gives

$$p(v) \propto \begin{cases} (v_{\mathrm{esc}} - |v_{||}|)^{k+1} & \text{if } v_{\min} \leqslant v_{||} < v_{\mathrm{esc}}, \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where $v_{||}$ is Galactocentric radial velocity. Since we are only interested in the absolute value of $v_{||}$ in this work, we shall now refer to $|v_{||}|$ as $v_{||}$ for brevity.

This model has been applied to data several times, most notably by Smith et al. (2007) and Piffl et al. (2014) (hereafter 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 $v_{\mathrm{esc}} \sim 530\,\mathrm{km s}^{-1}$. We seek to extend their work by constraining the escape speed of the Galaxy at a variety of locations. We do this by parameterising the escape speed as a function of position $\boldsymbol{x}$, so that

$$p(v_{||}\,|\,\boldsymbol{x}) = \begin{cases} \mathcal{N}(v_{\mathrm{esc}}(\boldsymbol{x}) - |v_{||}|)^{k+1} & \text{if } v_{\min} \leqslant v_{||} < v_{\mathrm{esc}}(\boldsymbol{x}), \\ 0 & \text{otherwise,} \end{cases}$$

(3)

where $\mathcal{N}$ is a location-dependent normalisation factor, given by

$$\mathcal{N} = \frac{k+2}{(v_{\mathrm{esc}}(\boldsymbol{x}) - v_{\min})^{k+2}}. \quad (4)$$

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

$$p(v_{||}\,|\,\ell, b, s) = \int p(v_{||}\,|\,\ell, b, s')p(s'\,|\,s)\,\mathrm{d}s'. \quad (5)$$

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

$$p_{\rm out}(v_{||}) = \frac{A}{\sqrt{2\,\pi\,\sigma^2}} \exp\frac{-v_{||}^2}{2\sigma^2}, \tag{6}$$

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

$$p_{\rm tot}(v_{||} \,|\, \ell, b, s) = (1 - f)\, p(v_{||} \,|\, \ell, b, s) + f\, p_{\rm out}(v_{||}). \tag{7}$$

This Equation represents the likelihood that we will use for the remainder of the paper, while making specific choices for $v_{\rm esc}(\boldsymbol{x})$. In practice, we compute the RHS of Equation (5) using Monte-Carlo integration (e.g. Evans et al. 2016; Bowden et al. 2016), so that

$$p(v_{||} \,|\, \ell, b, s) \simeq \frac{1}{N} \sum_{n=1}^{N} p(v_{||} \,|\, \ell, b, s_n), \tag{8}$$

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 radial velocities and distances. To obtain such a sample, we use the SDSS 9th data release (Ahn et al. 2012). We choose to utilise three distinct sets of tracers: main sequence turnoff (MSTO), K-giant and blue horizontal branch (BHB) stars. MSTO stars are numerous, and are mostly observed at distances $\sim 3\,{\rm kpc}$. These stars will 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 $\sim 50\,{\rm kpc}$ from the sun, pushing the spatial extent of our sample to a range of $\sim 40\,{\rm kpc}$. Before selecting our high-speed sample of 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 $|b| < 20°$ in order to remove possible disc contaminants, and remove stars with radii $r > 50\,{\rm kpc}$. We compute Galactocentric radii by assuming a solar radius $R_\odot = 8.5\,{\rm kpc}$. 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

$$\begin{aligned}
0.2 &< (g - r)_0 < 0.6, \\
14.5 &< r/{\rm mag} < 20,
\end{aligned} \tag{9}$$

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

$$\begin{aligned}
4800 &< T_{\rm eff}/{\rm K} < 8000, \\
3.5 &< \log g < 4, \\
-4 &< [{\rm Fe/H}] < -0.9.
\end{aligned} \tag{10}$$

This gives us a sample of metal-poor MSTO stars. We also make quality cuts to ensure high-quality photometry and radial velocity measurements for the sample, resulting in a mother sample of 22070 MSTO stars. 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 have

$$\begin{aligned}
\delta M_r &= 4.5 - 1.11\,[{\rm Fe/H}] - 0.18\,[{\rm Fe/H}]^2, \\
M_{r0} &= -5.06 + 14.32\,x - 12.97\,x^2 + 6.127\,x^3 \\
&\qquad - 1.267\,x^4 + 0.0967\,x^5, \\
M_r &= M_{r0} + \delta M_r.
\end{aligned} \tag{11}$$

In order to draw samples from $p(s' \,|\, s)$ (Equation 5), we take the quoted uncertainties on g, r, i and [Fe/H] from the SDSS pipelines and assume that they are Gaussian 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 dereddened colour-colour box

$$\begin{aligned}
-0.25 &< (g - r)_0 < 0, \\
0.9 &< (u - g)_0 < 1.4,
\end{aligned} \tag{12}$$

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

$$\begin{aligned}
-2 &< [{\rm Fe/H}] < -1, \\
3 &< \log g < 3.5, \\
8300 &< T_{\rm eff}/{\rm K} < 9300,
\end{aligned} \tag{13}$$

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)

$$\begin{aligned}
M_g &= 0.434 - 0.169\,(g - r)_0 + 2.319\,(g - r)_0^2 \\
&\qquad + 20.449\,(g - r)_0^3 + 94.617\,(g - r)_0^4.
\end{aligned} \tag{14}$$

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 catalog constructed by Xue et al. (2014), taken from 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

$$\begin{aligned}
0.5 &< (g - r)_0 < 1.3, \\
0.5 &< (u - g)_0 < 2.5,
\end{aligned} \tag{15}$$

and have measured proper motions $< 11\,{\rm mas\,yr^{-1}}$. Each target category then populates distinct regions of the $(u - g)-$
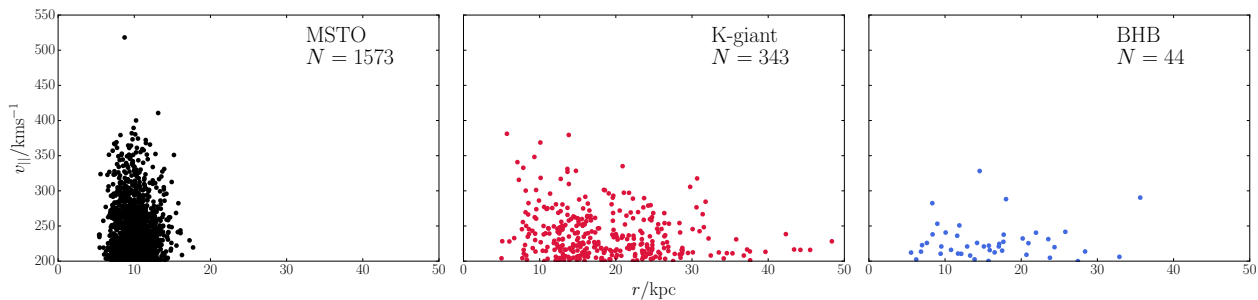
Figure 1. The distribution in Galactocentric radius and $v_{||}$ of each of our tracer samples. The MSTO stars probe a smaller range in radius, but are numerous. The K-giants probe a much larger distance range, but are fewer in number. The BHBs probe a similar range to the K-giants, but form a far smaller sample.

$-(g - r)$ plane. The precise details of the targetting criteria are given in Yanny et al. (2009). On top of the SEGUE selection, Xue et al. then make further cuts such that

$$\log g < 3.5,$$
$$E(B - V) < 0.25 \, \text{mag}.$$

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

$$[\text{Fe/H}] < -0.9. \tag{16}$$

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, $\mathcal{DM}$, of the K-giants. Their table provides the maximum a posteriori value of $\mathcal{DM}$ for each star, percentiles of the posterior samples, and a summary of the uncertainty, $\delta\mathcal{DM}$. 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 $\delta\mathcal{DM}$.

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 (3). S07 did not have access to measurements of stellar parameters for their sample, and adopted a very high threshold of $v_{\min} = 300 \, \text{kms}^{-1}$ 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 Figure 9), and then introduced a cut of $v_{\min} = 200 \, \text{kms}^{-1}$. Our cuts in latitude and metallicity serve to remove disc contaminants, which is much less of a concern in 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 (3) 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 distribution function of $v_{||}$ in the simulations does not significantly deviate

from a power law when $v_{||} > 150 \, \text{kms}^{-1}$. However, this analysis only considered a RAVE–like volume in the simulations, where $v_{\text{esc}} \sim 500 \, \text{kms}^{-1}$. 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 center than the Sun, and physical reasoning suggests that the escape speed cannot increase as a function of radius, then a cut of $200 \, \text{kms}^{-1}$ should guarantee that the power law model of Equation (3) is appropriate at the locations of all the stars in our study. Consequently, we set $v_{\min} = 200 \, \text{kms}^{-1}$. Note that this cut is applied to Galactocentric radial velocities, and so the motion of the sun must be removed beforehand. For this, we assume a local standard of rest $v_{\text{LSR}} = 240 \, \text{kms}^{-1}$ and a solar peculiar motion $(U_\odot, V_\odot, W_\odot) = (11.1, 12.24, 7.25) \, \text{kms}^{-1}$ (Schönrich et al. 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 Figure 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 $v_{\text{esc}}(\boldsymbol{x})$

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

$$v_{\text{esc}}(r) = v_{\text{esc}}(R_\odot) \left(\frac{r}{R_\odot} \right)^{-\alpha/2}, \tag{17}$$

where r is Galactocentric radius and $0 \leqslant \alpha \leqslant 1$, on physical grounds. We normalise the model to the escape speed at the solar radius. The gravitational potential Φ and escape speed are simply related by

$$v_{\text{esc}} = \sqrt{-2\Phi}, \tag{18}$$

meaning that the SPL model corresponds to a gravitational potential of the form

$$\Phi(r) = -\frac{v_{\text{circ}}(R_\odot)^2}{\alpha} \left(\frac{r}{R_\odot} \right)^{-\alpha}. \tag{19}$$

Thus, unlike other studies that measure only the local escape speed, we can translate our measurements of $v_{\text{esc}}(R_\odot)$ and α into inference of the circular speed of the potential, and the mass enclosed within spherical shells.

We will also report on results when two other models

are used. The first is a simple generalisation of the SPL, a power law in elliptical radius (EPL)

$$v_{\rm esc}(r_q) = v_{\rm esc}(R_\odot) \left(\frac{r_q}{R_\odot} \right)^{-\beta/2},$$ (20)

with $r_q = \sqrt{R^2 + z^2/q^2}$, where R and z are the usual cylindrical coordinates, and $0 \leqslant q \leqslant \infty$. 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). This model corresponds to a gravitational potential of the same form as Equation (19), but with elliptical radius r_q replacing spherical radius r.

The final model we consider is the truncated flat rotation curve (TF) model (see Gibbons et al. 2014; Wilkinson & Evans 1999). We primarily investigate this model in order to facilitate a direct comparison with the mass inference of Gibbons et al. (2014). The model has a flat rotation curve of amplitude v_0 in the inner parts, which then starts to decline as a power law $\gamma/2$ around a scale radius r_s

$$v_0^2(r) = \frac{v_0^2 \, r_s^\gamma}{(r_s^2 + r^2)^{\gamma/2}}.$$ (21)

The form of the escape speed for this model cannot be written in terms of elementary functions, so we omit it here.

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 than are independent of our choice for $v_{\rm esc}(\boldsymbol{x})$, and potential parameters, which concern our explicit model of the escape speed.

5.1   Global priors

Our global parameter set consists of four numbers: $(\boldsymbol{k}, f)$. We allow a different power law slope in the velocity distribution of each tracer $\boldsymbol{k} = (k_{\rm MSTO}, k_{\rm K-giant}, k_{\rm BHB})$, and we fit for the outlier fraction, f. 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, so that our prior 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 $0 \leqslant f \leqslant 1$.

5.2   Potential priors

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

$$0 \leqslant \alpha \leqslant 1.$$ (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 $v_{\rm esc}(R_\odot)$ a positive definite scale parameter, we adopt a scale invariant Jeffreys prior

$$p\left(v_{\rm esc}(R_\odot)\right) \propto 1/v_{\rm esc}(R_\odot).$$ (23)

The EPL model is identical to the SPL model other than the inclusion of the axis ratio, q, so that $\theta = (v_{\rm esc}(R_\odot), \beta, q)$. The priors we use for β and $v_{\rm esc}(R_\odot)$ are the same as in Equations (22) and (23). For q, we follow Bowden et al. (2016) and use a prior

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

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

Our final model, the TF model, has three parameters: $\theta = (v_0, r_s, \gamma)$. For the velocity normalisation v_0, we use a Jeffreys prior, limited so that $150 < v_0/{\rm kms}^{-1} < 400$. For the scale length, r_s, we use a prior that mirrors the result found by Gibbons et al. (2014), so that

$$\log p(r_s) \propto -\frac{1}{2}(r_s - \mu)^2/\sigma^2,$$ (25)

where $\mu = 15\,{\rm kpc}$ and $\sigma = 7\,{\rm kpc}$. We adopted this prior after experimentation with less informative priors, where we found that our model, in combination with these data, does not constrain r_s (i.e., the posterior distribution on r_s always resembles the prior). The power law index γ can take any value between 0 and 1, and hence we adopt the same prior as Equation (22).

5.3   Sampling method

Armed with our priors, we can now perform a fully Bayesian analysis. Our full parameter space is $\Theta = (\boldsymbol{k}, f, \theta)$, which has 6 (SPL) or 7 (EPL and TF) dimensions. Bayes' theorem for our problem can be written in the following way

$$p(\Theta \,|\, {\rm data}) =$$

$$\frac{p(\theta)\, p(\boldsymbol{k}, f) \prod\limits_{j=1}^{3} \prod\limits_{i=1}^{N_j} p_{\rm tot}(v_{||,i}^j \,|\, \ell_i^j, b_i^j, s_i^j, k_j, f, \theta)}{p({\rm data})},$$ (26)

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 (5) using 200 Monte-Carlo samples from the distance uncertainties of each star. We draw these samples
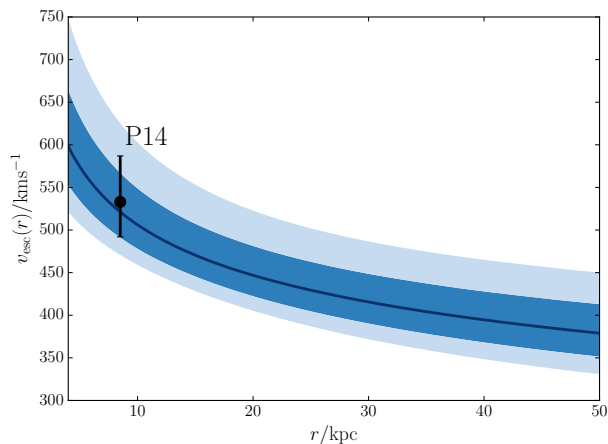
Figure 3. Our inference on the escape speed as a function of Galactocentric radius. The median posterior result is shown as a dark blue line, and the 68% (94%) credible interval is a dark (light) blue band. The result using RAVE data from Piffl et al. (2014) and the associated 90% credible interval is 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 $30\,\mathrm{kpc}$.
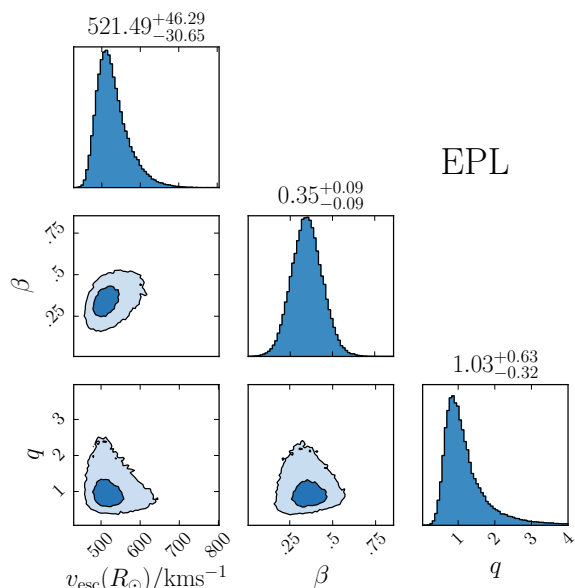

Figure 4. Posterior distribution for the EPL model. The results are very similar to the SPL model, and the extra parameter q is found to be ~ 1, suggesting that the Galactic potential is probably spherical, though our uncertainties are quite large.
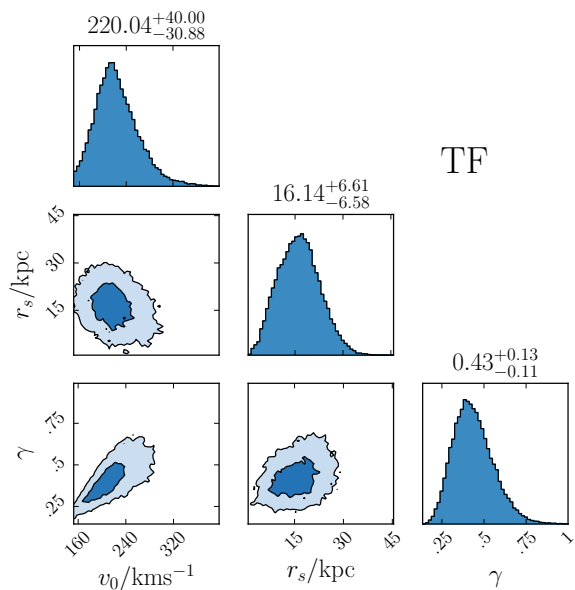

Figure 5. Posterior for the TF model when the prior based on the inference of Gibbons et al. (2014) is used for r_s. We infer a rotation curve amplitude of $220\,\mathrm{km\,s^{-1}}$ and a power law decline of $\gamma \simeq 0.4$, both of which are entirely consistent with the results of Gibbons et al. (2014).

once, and then use the same set for each posterior 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 the remaining $N - 1$ walkers. Here, we choose $N = 80$. We initialise 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. We then assess convergence using the Gelman & Rubin (1992) criterion, where the within–chain variances are compared to the between-chain variance in order to compute the Potential Scale Reduction Factor (PSRF). If the PSRF is large ($\gg 1$), then the chains are not well-mixed and the N walkers have not converged to the target posterior distribution. We adjudge a given run to be converged if the PSRF is less than 1.1 in all parameters. For each of the runs presented in this paper, the PSRF is $\simeq 1.01$ or smaller in all parameters.

6 RESULTS

Figure 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 $\pm 34\%$ credible intervals. All of the model parameters are well constrained, save for k_{BHB}, which is unsurprising given that BHBs are by far the least numerous tracer in our sample.

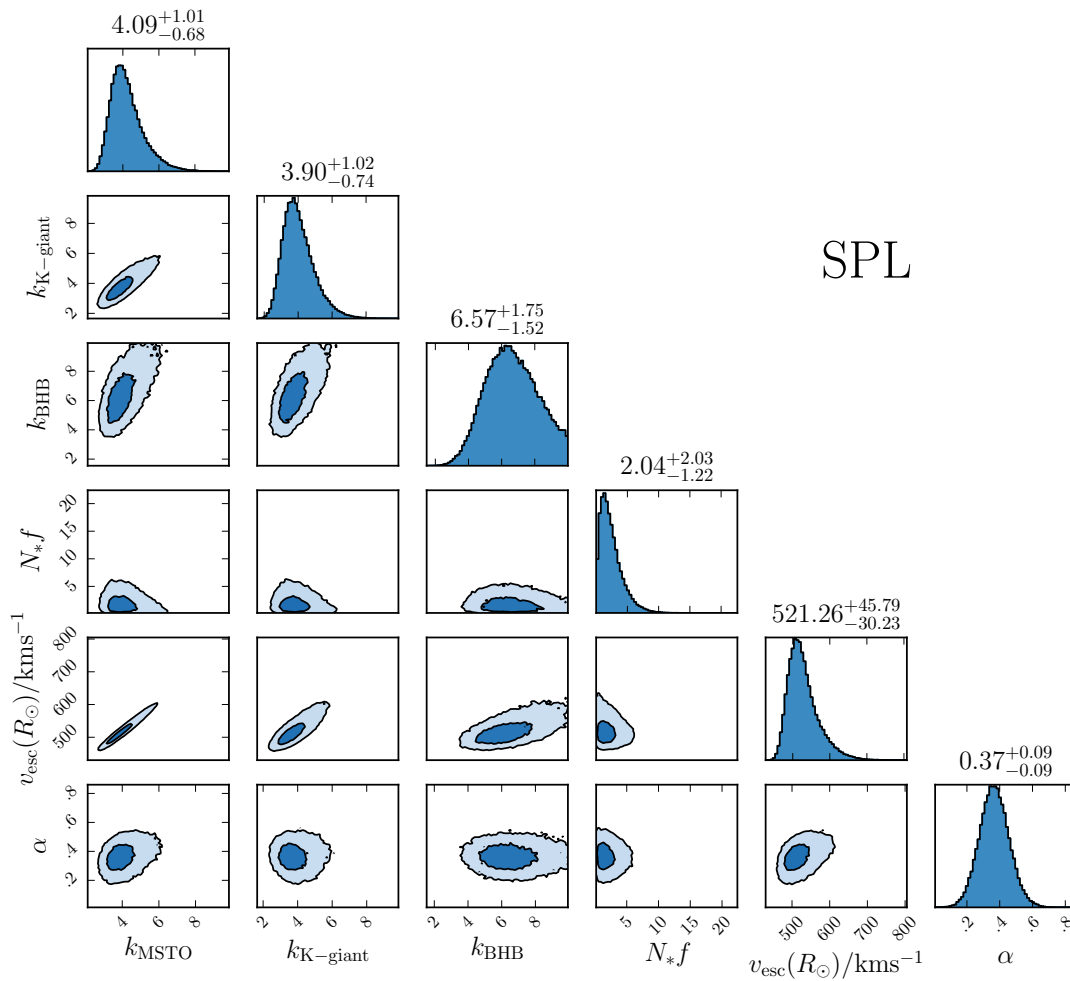Our results imply a local escape speed of $521^{+46}_{-30}\,\mathrm{km\,s^{-1}}$,

Figure 2. One and two–dimensional projections of our MCMC samples for the SPL fit. The 68% and 94% credible intervals are shown in the 2D projections, and the median parameter values and uncertainties computed using the ±34% credible intervals of the 1D projections are shown above the 1D 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 $k_{\mathrm{K-giant}}$, which are constrained to be close to equal, and between $v_{\mathrm{esc}}(R_\odot)$ and k_{MSTO} and $k_{\mathrm{K-giant}}$. Larger values of k suggest larger values of $v_{\mathrm{esc}}(R_\odot)$, as expected. See text for discussion.

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 Figure 1 is prophetic of this, because the edge of the K-giant distribution in the $r - v_{||}$ plane is steep. Figure 3 shows the run of v_{esc} with radius implied by our inference, with associated 68% and 94% credible intervals, and the steep drop in the escape speed is clear. The Milky Way loosens its grip on its inhabitants significantly: our model predicts that the escape speed at 30 kpc from the Galactic centre is $415^{+33}_{-25}\,\mathrm{kms}^{-1}$, and by 50 kpc this has dropped to $379^{+34}_{-28}\,\mathrm{kms}^{-1}$.

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, then, 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. Nonetheless, the inference on k_{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.

Figure 2 shows a strong degeneracy between k_{MSTO} and $v_{\mathrm{esc}}(R_{\odot})$. This is to be expected, and is the reason why a narrow prior on k was necessary in previous work. Figure 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 $k_{\mathrm{K-giant}}$ and the local escape speed, though it is broader.

The inferred outlier fraction is very small, $f \simeq 0.001$, but non-zero. This suggests that there are one or two outliers in our sample. Inspection of Figure 1 suggests one clear candidate: there is an MSTO star at $r \simeq 10\,\mathrm{kpc}$ with a measured line of sight velocity of $518.2\,\mathrm{kms}^{-1}$, which is more than $100\,\mathrm{kms}^{-1}$ 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

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

using the model parameters obtained by taking the median values of each of the one dimensional marginalised posterior distributions, $\bar{\Theta}$. The largest outlier probability is > 0.999, and belongs to the star identified visually in Figure 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 verified this status, it is now interesting to ask whether this star is genuinely unbound. Unfortunately, this conclusion seems unlikely after further investigation. The outlier has a quoted radial velocity uncertainty of $17\,\mathrm{kms}^{-1}$, which is $7\,\mathrm{kms}^{-1}$ larger than the second largest uncertainty in the MSTO sample, and $14\,\mathrm{kms}^{-1}$ larger than the median uncertainty. Furthermore, its measured effective temperature is the largest of all of the MSTO sample, with $T_{\mathrm{eff}} = 6877\,\mathrm{K}$. Given its outlier status in other spectroscopic parameters, we conclude that it is likely that this star has a spurious radial velocity measurement.

Figures 4 and 5 show the results of the EPL and TF analyses. We removed the global parameters from the figures because the results were essentially identical to those found for the SPL model, without any interesting additional degeneracies. The addition of a halo axis 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.

Our inference on the TF model suggests an inner rotation curve with amplitude $v_0 = 220^{+40}_{-31}\,\mathrm{kms}^{-1}$, which starts to fall as $\gamma/2 = 0.22^{+0.07}_{-0.06}$ over a scale $r_s = 16^{+7}_{-7}\,\mathrm{kms}^{-1}$. The posterior on the scale radius is no different from our prior, and so is uninteresting. However, having assumed similar inference on the scale radius as Gibbons et al. (2014), the amplitude and power law index are entirely consistent with their result, which we will discuss further in Section 6.1. When translated into an escape speed profile, the inference appears virtually identical to Figure 3. This suggests that our inference is relatively robust against different parameterisations of the escape speed.

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 curve $v_c(r)$ for the Galaxy. For a spherically symmetric escape speed profile, we have

$$M(r) = \frac{-r^2\,v_{\mathrm{esc}}}{G}\frac{\mathrm{d}v_{\mathrm{esc}}}{\mathrm{d}r}, \tag{28}$$

$$v_c(r) = \sqrt{-r\,v_{\mathrm{esc}}\frac{\mathrm{d}v_{\mathrm{esc}}}{\mathrm{d}r}}. \tag{29}$$

Given our model of $v_{\mathrm{esc}}(r)$, and our inference on its parameters, we can compute posterior distributions on $M(r)$ and $v_c(r)$ using these formulae. For example, the local circular speed that we measure is

$$v_c(R_{\odot}) = 223^{+40}_{-34}\,\mathrm{kms}^{-1}. \tag{30}$$

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 distribution function 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 moments of the velocity distributions. Distribution functions satisfy the full collisionless Boltzmann equation, and therefore the entire infinite hierarchy of Jeans 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.

Figure 6 shows the mass and circular speed profiles implied by the SPL model, along with associated 68% and 94% credible regions. Our model predicts $M(50\,\mathrm{kpc}) = 29.8^{+6.9}_{-5.2} \times 10^{10}\mathrm{M}_{\odot}$. 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 distribution function models to the data, and infer systematically higher masses than we do here, with $M(50\,\mathrm{kpc}) \simeq 45 \times 10^{10}\mathrm{M}_{\odot}$. Both are consistent with the 94% credible interval of our inference, but there is a hint that there is a discrepancy between distribution function 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.
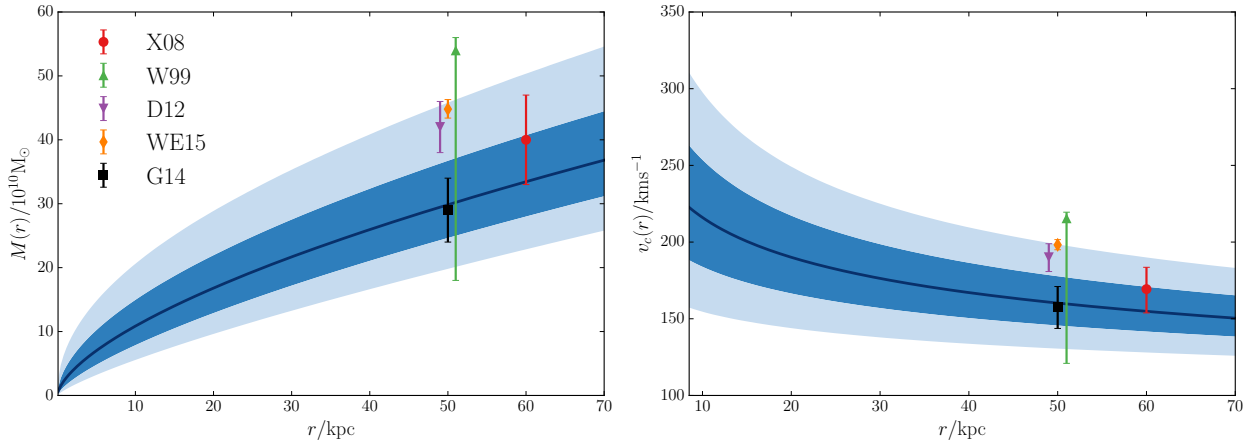
Figure 6. Left: cumulative mass distribution within spherical shells, predicted by the SPL model. The 68% (94%) credible interval is shown as a dark (light) blue band around our median result. We predict a relatively light Milky Way within 50 kpc, with $M(50\,\mathrm{kpc}) = 29.8^{+6.9}_{-5.2} \times 10^{10}\mathrm{M}_\odot$. In both panels are the results from various other studies, see text for discussion.

W99, like D12 and WE15, used a distribution function 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 $M(50\,\mathrm{kpc}) = 54^{+2}_{-36} \times 10^{10}\mathrm{M}_\odot$, like the other two distribution function 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 et al. (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 $M(50\,\mathrm{kpc}) = 29 \pm 5 \times 10^{10}\mathrm{M}_\odot$. We can make this comparison even more explicit because we have also estimated the parameters of the model they used in their analysis (TF). When we use the TF model, the mass enclosed is $M(50\,\mathrm{kpc}) = 33^{+8}_{-6} \times 10^{10}\mathrm{M}_\odot$. The two analyses, though very different in detail, produce near identical results.

6.2 The orbits of Milky Way dwarf galaxies

The left panel of Figure 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 radial 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 (right hand panel). However, some of the dwarfs seem likely unbound based on our estimates of the escape speed. On the other hand, ΛCDM simulations predict that 99.9% of subhalos should be bound to their hosts (Boylan-Kolchin et al. 2013). A reconciliation of these two statements is to conclude that the $\sqrt{3}v_{||}$ 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 Figure 7 are Bootes III (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 radial 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 et al. 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 characterise their orbits through their pericenter radii r_{peri}, apocenter radii r_{apo} and orbital eccentricities ϵ.

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

$$p(X \,|\, v_r, r) = \int p(X \,|\, v_r, r, \Theta)\, p(\Theta \,|\, \mathrm{data})\, \mathrm{d}\Theta, \qquad (31)$$

where $X = (r_{\mathrm{peri}}, r_{\mathrm{apo}}, \epsilon)$ and $p(\Theta \,|\, \mathrm{data})$ 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

$$p(X \,|\, v_r, r, \Theta) = \int p(X \,|\, v_r, v_T, r, \Theta)\, p(v_T \,|\, v_r, r, \Theta)\, \mathrm{d}v_T,$$
$$= \int \delta\left[X - f_X(v_r, v_T, r, \Theta)\right] p(v_T \,|\, v_r, r, \Theta) \mathrm{d}v_T, \qquad (32)$$

where v_T is the transverse velocity. In a spherical potential, the two velocities (v_T, v_r) and radius r completely specify the orbit, and hence r_{peri}, r_{apo} and ϵ. This gives rise to the delta–function in Equation (32), where f_X represents the relation between (v_r, v_T, r) and X. Finally, we need to specify $p(v_T \,|\, v_r, r, \Theta)$. To do this, we assume that the high speed dwarfs follow the same distribution function as the high speed stars, given by Equation (1). Some measure of justification of this assumption is possible through inspection of Figure 7, where both distributions have a similar 'zone of avoidance' away from the escape speed. This gives

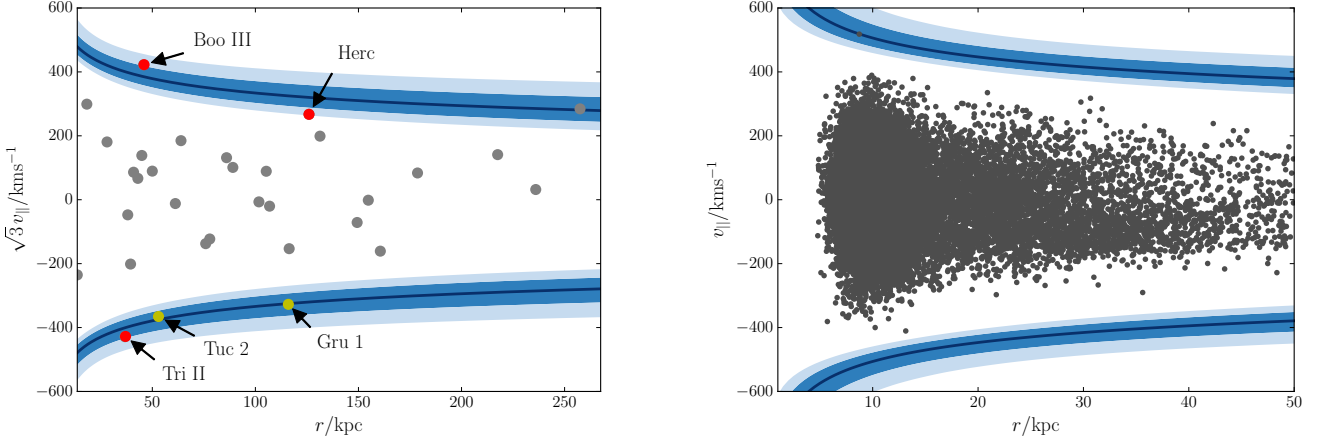$$p(v_T \,|\, v_r, r, \Theta) \propto \left(v_{\mathrm{esc}}(r) - \sqrt{v_T^2 + v_r^2}\right)^k. \qquad (33)$$

Figure 7. Left: 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. Radial 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. Right: the distribution of our mother samples of stars in the $r - v_{||}$ plane. The distribution, though significantly better sampled, broadly resembles the distribution of dwarfs, suggesting that the high–speed dwarfs might follow a similar power law tail in speed.

Since f_X is not analytic, we must solve for r_{peri}, r_{apo} and ϵ numerically. We do this by solving the equation

$$E - \frac{L^2}{2r^2} - \Phi(r) = 0, \qquad (34)$$

where $\Phi(r)$ is the gravitational potential implied by our model of the escape speed, E is the orbital energy and L is the total angular momentum. This equation has two roots: r_{peri} and r_{apo}. ϵ is then given by

$$\epsilon = \frac{r_{\mathrm{apo}} - r_{\mathrm{peri}}}{r_{\mathrm{apo}} + r_{\mathrm{peri}}}. \qquad (35)$$

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 (33). For k, we use k_{MSTO}, and we normalise the speed distribution between v_r and $v_{\mathrm{esc}}(r)$. Then, we solve Equation (34) for r_{peri}, r_{apo} and ϵ and store the result. This process is repeated for every set of model parameters in our posterior samples. The resulting histograms in r_{peri}, r_{apo} and ϵ are then faithful representations of $p(X \,|\, v_r, r)$.

The results of this procedure are shown in Figure 8. All three dwarfs are expected to be on very eccentric orbits. Bootes III and Triangulum II have $\epsilon \simeq 0.95$, while Hercules has $\epsilon \simeq 0.8$, though with a somewhat broader distribution. This is aligned with our intuition: if the radial 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 apocenters, $r_{\mathrm{apo}} \sim 100 - 300$ kpc and considerably smaller pericenters $r_{\mathrm{peri}} \sim 10$ kpc (although the posterior on the pericenter of Hercules is significantly less peaked than for Bootes III and Triangulum II).

Küpper et al. (2016) argued that the observed ellipticity of Hercules ($e \simeq 0.7$) and its large radial velocity are suggestive that it has 'exploded' as a consequence of its last pericenter passage. Using N-body simulations, they arrived at an estimate of the orbital eccentricity $\epsilon \simeq 0.95$, which is

larger than our value but not in significant tension with it. Bootes III has a similar morphology ($e \simeq 0.5$) in keeping with the picture that these satellites are on orbits that will cause them to disrupt into streams very after comparatively few orbital periods. On the other hand, Triangulum II has a relatively small ellipticity ($e \simeq 0.2$). However, both Bootes II and Hercules have positive radial velocities: they are travelling away from the Galactic centre, suggesting that they have undergone at least one pericenter passage. Triangulum II, on the other hand, has a negative radial velocity. It is plausible, therefore, that it is on first infall and is about to undergo a large amount of disruption on its first pericenter passage.

6.3 Model performance and systematics

We now seek to assess how well our model fits the data. We do this by calculating the posterior predictive distribution, which is the probability density of new observations (i.e. the radial velocities of stars yet to be observed), conditioned on the data we have used to constrain the model:

$$p(v_{||} \,|\, \mathrm{data}) = \int p(v_{||} \,|\, \Theta)\, p(\Theta \,|\, \mathrm{data})\, \mathrm{d}\Theta. \qquad (36)$$

This distribution is a faithful representation of our model predictions, because it takes into account the uncertainty in the model parameters by marginalising over the posterior distribution. It is preferable, for example, to using the likelihood function at the median posterior values, which would not properly account for uncertainty. It is possible to calculate this distribution by using the N samples drawn from the posterior distribution during our MCMC sampling, so that

$$\int p(v_{||} \,|\, \Theta)\, p(\Theta \,|\, \mathrm{data})\, \mathrm{d}\Theta \simeq \frac{1}{N} \sum_{n=1}^{N} p(v_{||} \,|\, \Theta_n). \qquad (37)$$

We want to compare the distribution in $v_{||}$ of each tracer sample to our model using this technique. There is a subtlety
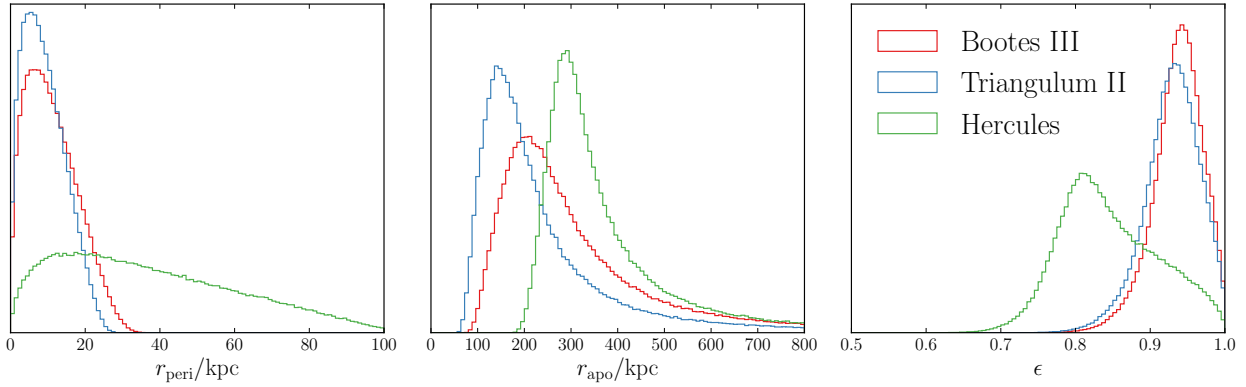
Figure 8. Posterior distributions on orbital pericenter r_{peri}, apocenter r_{apo} and eccentricity ϵ for Bootes III, Triangulum II and Hercules. If the dwarfs are bound, they must be on very eccentric orbits.

in how we must compare our model to the data, however. The velocity distribution of high-speed stars is position dependent in our 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

$$
\begin{aligned}
p(v_{||}|\Theta) &= \frac{\int p(v_{||}|r,\Theta)\, p(r)\, \mathrm{d}r}{\int p(v_{||}|r,\Theta)\, p(r)\, \mathrm{d}r\, \mathrm{d}v} \\
&= \frac{\int p(v_{||}|r)\, p(r)\, \mathrm{d}r}{\int (k+2)^{-1}(v_{\text{esc}}(r) - v_{\text{min}})^{k+2}\, p(r)\, \mathrm{d}r}, \quad (38)
\end{aligned}
$$

where $p(r)$ is the probability of observing a star at radius r. We approximate $p(r)$ by constructing a histogram of the mother sample of each tracer, and then evaluating $p(r)$ as a cubic spline between the bin centres. The two integrals in Equation (38) are then computed using 12 point Gauss-Legendre quadrature. Our posterior predictive distribution is then calculated using Equations (37) and (38).

Figure 9 shows the comparison of the posterior predictive distributions with the data for our MSTO, K-giant and BHB samples. The model generally seems to well-represent the data. The MSTO sample, which is our largest tracer group, is in particularly excellent agreement with the model. The K-giant velocity distribution is also in generally good agreement with the model, although there is one bin close to our minimum radial velocity that is not following the overall power law trend. On the other hand, the uncertainties suggest that this is not a major concern.

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 $v_{\text{LSR}} = 240\,\text{kms}^{-1}$. 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 $v_{\text{LSR}} = 220\,\text{kms}^{-1}$. The local escape speed is then inferred

to be $542^{+56}_{-37}\,\text{kms}^{-1}$, 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 $v_{\text{min}} = 200\,\text{kms}^{-1}$ 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 $v_{\text{min}} = 250\,\text{kms}^{-1}$. This results in a significantly smaller sample of 644 stars (539 MSTO, 99 K-giants, 6 BHBs). We thus expect much larger uncertainties on 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 by comparison to our full analysis, leading to a worse determination of the local escape speed. This is because the degeneracy between k and $v_{\text{esc}}(R_\odot)$ is no longer broken by the data, which pushes up our estimates of k and $v_{\text{esc}}(R_\odot)$ as predicted by S07. Specifically, we find $v_{\text{esc}}(R_\odot) = 617^{+77}_{-84}\,\text{kms}^{-1}$. This value is nonetheless consistent with our full analysis, due to the inflated uncertainties. Correspondingly, the inferred values of k are ~ 6, with uncertainties ~ 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 $> 200\,\text{kms}^{-1}$, thus vindicating our choice of $v_{\text{min}} = 200\,\text{kms}^{-1}$. 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 $v_{\text{min}} = 250\,\text{kms}^{-1}$ 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% credible regions of the other run.
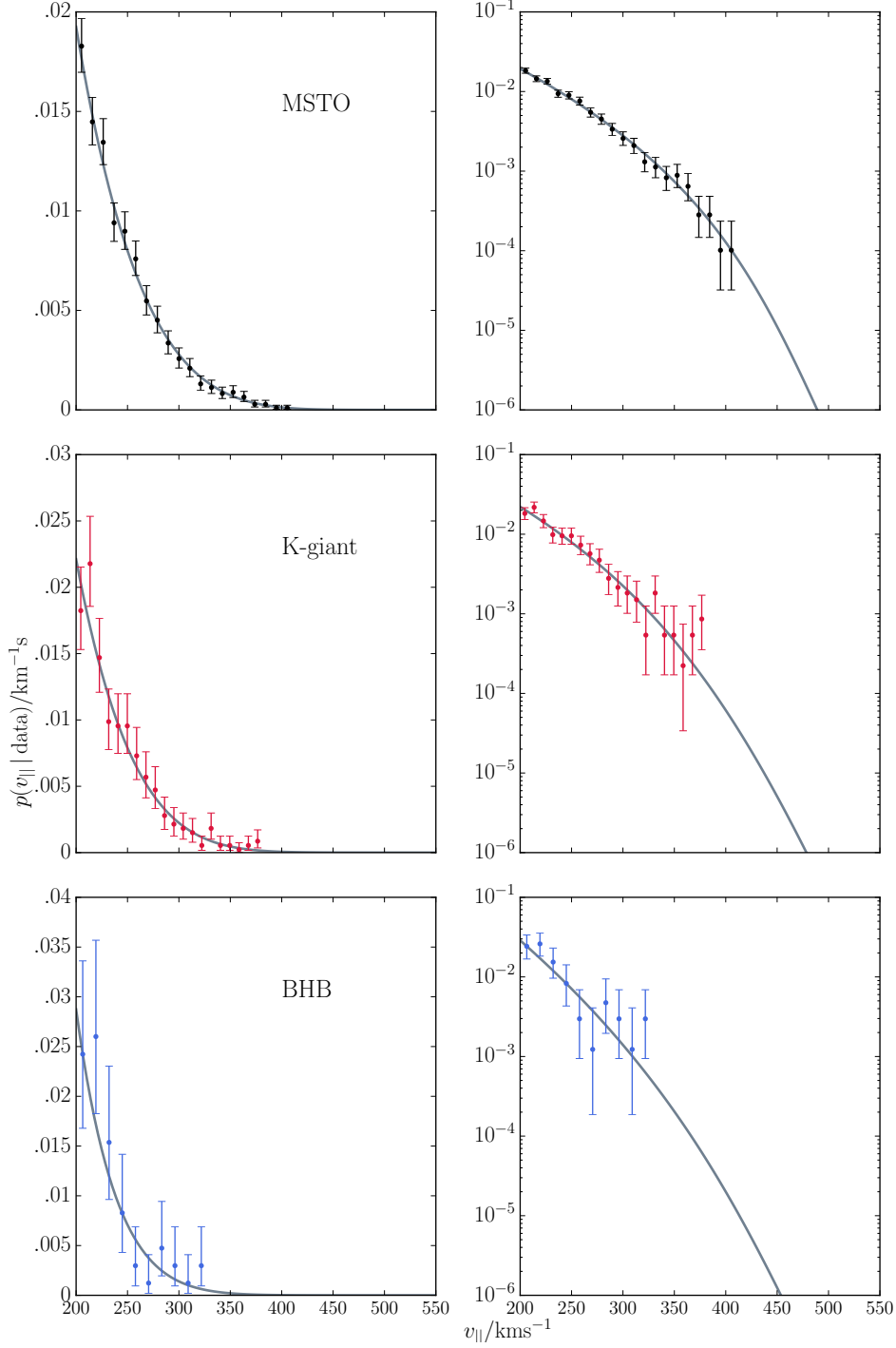
Figure 9. Posterior predictive checks of the SPL model. The coloured circles and uncertainties are the binned radial velocities of our three tracer groups, with 90% credible intervals computed using the Poisson distribution. The black line is the posterior predictive distribution of the SPL model. Our model well reproduces the data over ~ 2 orders of magnitude in the number of counts.

7 CONCLUSIONS

8 ACKNOWLEDGEMENTS

REFERENCES

Ahn C. P., Alexandroff R., Allende Prieto C., Anderson S. F., Anderton T., Andrews B. H., Aubourg É., Bailey S., Balbinot E., Barnes R., et al. 2012, ApJS, 203, 21

Bechtol K., et al., 2015, ApJ, 807, 50

Belokurov V., et al., 2007, ApJ, 654, 897

Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press

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

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

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., Rubin B., 1992, Statistical Science, 7, 457

Gibbons S. L. J., Belokurov V., Evans N. W., 2014, MN-RAS, 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., 2016, ArXiv e-prints

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

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

Spitzer Jr. L., Shapiro S. L., 1972, ApJ, 173, 529

Wang W., Han J., Cooper A. P., Cole S., Frenk C., Lowing B., 2015, MNRAS, 453, 377

Wilkinson M. I., Evans N. W., 1999, MNRAS, 310, 645

Williams A. A., Evans N. W., 2015, MNRAS, 454, 698

Xue X.-X., Ma Z., Rix H.-W., Morrison H. L., Harding P., Beers T. C., Ivans I. I., Jacobson H. R., Johnson J., Lee Y. S., Lucatello S., Rockosi C. M., Sobeck J. S., Yanny B., Zhao G., Allende Prieto C., 2014, ApJ, 784, 170

Xue X. X., Rix H. W., Zhao G., Re Fiorentin P., Naab T., Steinmetz M., van den Bosch F. C., Beers T. C., Lee Y. S., Bell E. F., Rockosi C., Yanny B., Newberg H., Wilhelm R., Kang X., Smith M. C., Schneider D. P., 2008, ApJ, 684, 1143

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 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.loggadopunc < 0.1
AND (spp.zwarning=0 OR spp.zwarning=16)
AND spp.snr > 20.
```

(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.
```