The Kinematic Properties of Milky Way Stellar Halo Populations

James M. M. Lane1*, Jo Bovy1, & J. Ted Mackereth1,2,3
1Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto ON, M5S 3H4, Canada
2Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto ON, M5S 3H8, Canada
3Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto ON, M5S 3H4, Canada

Accepted XXX. Received YYY; in original form ZZZ.

ABSTRACT

In the Gaia era stellar kinematics are extensively used to study Galactic halo stellar populations, to search for halo structures, and to characterize the interface between the halo and hot disc populations. We use distribution function-based models of modern datasets with 6D phase space data to qualitatively describe a variety of kinematic spaces commonly used in the study of the Galactic halo. Furthermore, we quantitatively assess how well each kinematic space can separate radially anisotropic from isotropic halo populations. We find that scaled action space (the “action diamond”) is superior to other commonly used kinematic spaces at this task. We present a new, easy to implement selection criterion for members of the radially-anisotropic Gaia-Enceladus merger remnant. Assuming a 1:1 ratio of Gaia-Enceladus stars to more isotropic halo, we find our selection achieves a sample purity of 82 per cent in our models with respect to contamination from the more isotropic halo. We compare this criterion to literature criteria, finding that it produces the highest purity in the resulting samples, at the expense of a modest reduction in completeness. We also show that selection biases that underlie nearly all contemporary spectroscopic datasets can noticeably impact the reduction in completeness. We conclude by providing recommendations for how authors should use stellar kinematics in the future to study the Galactic stellar halo.

Key words: Galaxy: halo – Galaxy: kinematics and dynamics – Galaxy: solar neighbourhood – Galaxy: stellar content

1 INTRODUCTION

It is an auspicious time to study the dynamics and structure of our Milky Way galaxy, in no small part thanks to a recent wealth of data obtained through large astronomical surveys. Chief among these is the astrometric space mission Gaia (Gaia Collaboration et al. 2016) which, as part of its most recent data release, has provided 5-d kinematic phase space information for nearly 1.5 billion stars, and full 6-d phase space information for a further ~ 7 million stars (Gaia Collaboration 2020). In parallel to this project, massive spectroscopic campaigns such as APOGEE (Majewski et al. 2017), LAMOST (Cui et al. 2012), Gaia-ESO (Gilmore et al. 2012), GALAH (Martell et al. 2017), SEGUE (Yanny et al. 2009), RAVE (Steinmetz et al. 2020, for the most recent data release), and H3 (Conroy et al. 2019) have provided accurate radial velocities, chemical abundances, and age estimates for hundreds of thousands of these stars. Combined, these two efforts have yielded a rich, multidimensional data set which has not only enabled novel discoveries, but in concert with powerful statistical techniques will allow us to discern among competing models of the structure, history, and composition of each part of our Milky Way galaxy with exquisite detail.

The nature of the stellar halo, in particular, has come into much sharper focus in the Gaia era. Since Gaia’s second data release, a significant amount of evidence has been amassed that the Milky Way was subject to a handful of distinct, significant mergers early in its life. The most well-studied example to date is the colourfully named “Sausage”, or “Gaia-Enceladus” (GE), which is a population of metal-rich halo stars with uncharacteristically large radial motions (Helmi et al. 2018; Belokurov et al. 2018; Haywood et al. 2018). This overdensity stands out in phase space from the remainder of halo stars, which exhibit more isotropically-distributed motions. Estimates of the mass of this population are in the range of \(10^8\) to \(10^{10}\) \(M_\odot\) (Belokurov et al. 2018; Helmi et al. 2018; Fattahi et al. 2019; Deason et al. 2019; Mackereth et al. 2019a; Mackereth & Bovy 2020; Kruĳssen et al. 2020; Naidu et al. 2021) which would place its stellar mass among the contemporary big satellites of the Milky Way, such as Sagittarius and the Large Magellanic Cloud (e.g. Niederste-Ostholt et al. 2010; van der Marel 2006).

GE is only one among an assortment of spatially incoherent halo structures which have either been discovered or extensively studied thanks to Gaia kinematics and supporting spectroscopy. These include Thamnos (Koppelman et al. 2019b), the Sequoia (Myeong et al. 2019), the Helmi streams (Helmi et al. 1999; Koppelman et al. 2018, 2019a), the Inner Galaxy Structure (IGS, also known as Hercules; Horta et al. 2021), LMS-1 (Yuan et al. 2020), as well as numerous structures identified by Naidu et al. (2020). Paralleling these results driven by stellar data, examination of the Galactic globular clusters population points to a merger known as “Kraken” (Kruĳssen et al. 2020) or “Koala” (Forbes 2020) which predates the above-mentioned events. While many of these structures require further study to confirm their nature and origin, it is nonetheless clear that significant
merger events play a major role in shaping the stellar halo we see today.

The stellar halo is not solely composed of ex-situ accretion remnants, but also in-situ stars formed within the Milky Way. Mounting kinematic and chemical evidence suggests that this in-situ component consists mostly of stars from the thicker components of the disc (Helmi et al. 2018; Haywood et al. 2018; Di Matteo et al. 2019). This idea, while not new (e.g. Nissen & Schuster 2010), is put into new perspective given the emerging importance of mergers in the early evolution of the Milky Way. Additionally, an emerging class of halo substructures known as “Splash” populations, which may be remnants of a proto-disc disrupted by an infalling satellite, may be evidence of the dynamical link between halo and thick disc (Belokurov et al. 2020; Naidu et al. 2020; Lian et al. 2020).

While each halo population is often discussed as being “distinct” in terms of its spatial, kinematic, and/or chemical properties, there is significant overlap in the properties of many, if not all, of these structures. This presents an issue when attempting to dissect the stellar halo on a granular level, as is the intent with current and upcoming surveys. We recognize a gap in the established literature, whereby there does not exist simple kinematic models to contextualize these large range of data. N-body models are commonly used to interpret results, but tailoring the simulations to individual datasets is challenging.

In this work we examine mock samples of stars generated from distribution functions which represent the disc and halo populations commonly studied in the Milky Way. Our goals are twofold: first we seek to provide context for observed kinematics from the current and next generation of surveys. Second we use these models to numerically evaluate the best way to separate radially biased halo populations from more isotropic halo populations, an effective proxy for isolating GE remnant stars.

2 MODEL POTENTIAL AND DISTRIBUTION FUNCTIONS

Here we introduce the distribution functions (DFs) and accompanying gravitational potentials that we use to model the Milky Way disc and halo. First, we present the potential and then we describe the DFs and the techniques used to draw kinematic samples from them. Throughout we assume that: the Sun is located at \((R,z,\phi) = (8.178 \text{ kpc}, 0.028 \text{ kpc}, 0)\) (Gravity Collaboration et al. 2019; Bennett & Bovy 2019), and the solar motion is \((U,V,W) = (11.1, 12.24, 7.25) \text{ km s}^{-1}\) (Schönrich et al. 2010). We use a left-handed Galactocentric cylindrical coordinate system where angular momentum is positive in the direction of Galactic rotation. By default we express quantities in cylindrical coordinates and are explicit when we do otherwise. The code we use for this project is available on github.

### 2.1 Milky Way Potential

To represent the Milky Way we use the \texttt{MWPotential2014} model from Bovy (2015). The potential has three components. The disc is represented by a Miyamoto & Nagai (1975) profile with mass \(M_d = 6.8 \times 10^{10} \text{ M}_\odot\), a scale length of 3 kpc, and a scale height of 280 pc. The bulge is an exponentially-truncated power law density profile with mass \(M_b = 0.5 \times 10^{10} \text{ M}_\odot\), power law index \(-1.8\), and exponential scale radius of 1.9 kpc. The halo is an NFW (Navarro et al. 1997) profile with a virial mass of \(8 \times 10^{11} \text{ M}_\odot\) and concentration of 15.3. For a discussion of the derivation of the model parameters which describe \texttt{MWPotential2014}, see section 3.5 of Bovy (2015).

### 2.2 Disc Distribution Functions

We approximate the disc’s intricate structure (e.g., Bovy et al. 2012) as a simple “thin” and “thick” superposition and to represent both disc components we use the quasi-isothermal DF of Binney & McMillan (2011) (see also Binney 2010). This DF takes the form

\[
f(J_r, L_z, J_z) = f_{\sigma r}(J_r, L_z) f_{\sigma z}(J_z),
\]

where \(f_{\sigma r}\), largely controls the dynamics within the plane of the disc, and \(f_{\sigma z}\) largely controls the vertical dynamics perpendicular to the disc. Each of these are functions of the actions \(J_r, L_z, \) and \(J_z, f_{\sigma r}\) is given by

\[
f_{\sigma r} = \frac{\Omega \Sigma}{\pi \sigma_r^2} \left[ 1 + \tanh(L_z/L_0) \right] e^{-\kappa J_r/\sigma_r^2},
\]

where \(\Sigma \) and \(\Omega\) are the local surface density and orbital frequency respectively, \(R_c \) and \(\kappa\) the guiding center radius and epicycle frequency of a given orbit. \(L_0\) is a cutoff angular momentum. \(\sigma_r\) is the radial velocity dispersion profile. \(\Sigma\) has the exponential form

\[
\Sigma(L_z) = \Sigma_0 e^{(R_0-R_c)/R_d}
\]

where \(\Sigma_0\) is an arbitrary normalization, \(R_0\) is a scale radius (which we set to the location of the Sun), and \(R_d\) is the scale length of the disc. The second component, \(f_{\sigma z}\), is given by a pseudo-isothermal DF

\[
f_{\sigma z} = \frac{v_z e^{-v_z J_z/\sigma_z^2}}{2\pi \sigma_z^2},
\]

where \(v_z\) and \(J_z\) are the vertical epicycle frequency and action and \(\sigma_z\) is the vertical velocity dispersion profile. The parameters we use for the thin and thick disc DFs are summarized in Table 1. Both the radial and vertical exponential profiles have exponential forms given by

\[
\sigma_{r}(L_z) = \sigma_{r,0} e^{(R_0-R_c)/R_d}
\]

\[
\sigma_{z}(L_z) = \sigma_{z,0} e^{(R_0-R_c)/R_d}.
\]

Here \(\sigma_{r,z,0}\) are characteristic velocity dispersions at \(R_0\). We use, overall, hotter disc DF parameters than are advocated for by Binney & McMillan (2011), but that are not unlike other measurements have suggested (e.g. Bovy et al. 2012; Mackereth et al. 2019b). By choosing disc DF parameters on the hot end of what is normally used we hedge our results towards including disc samples which intrude most on the kinematic space normally assigned to the halo.

To sample from this DF, we use \texttt{galpy}\(^3\) (Bovy 2015) where it

\footnote{1 https://github.com/jamesmlane/mw-dfs}

\footnote{2 Binney & McMillan (2011) have a multiplicative factor \(q\) in these exponentials, but it is unity here.}

\footnote{3 https://github.com/jobovy/galpy}

\begin{table}[h]
\centering
\caption{Properties of quasi-isothermal thin and thick disc DFs}
\begin{tabular}{ccc}
\hline
Parameter & Thin disc & Thick disc \\
\hline
\(R_d\) & 3 kpc & 2 kpc \\
\(R_0\) & 8.178 kpc & 8.178 kpc \\
\(\sigma_{r,0}\) & 40 km s\(^{-1}\) & 60 km s\(^{-1}\) \\
\(\sigma_{z,0}\) & 30 km s\(^{-1}\) & 45 km s\(^{-1}\) \\
\(L_0\) & 10 kpc km s\(^{-1}\) & 10 kpc km s\(^{-1}\) \\
\hline
\end{tabular}
\end{table}
is implemented as the class `galpy.df.quasiisothermaldf`. The phase space samples we generate are described further on in § 3, but for our purposes we are only interested in sampling velocities from this DF given fixed positions. Velocity samples are obtained using rejection sampling of the DF at each individual radius and height above the midplane. Actions for the DF are calculated using the “Stückel fudge” method of Binney (2012). In this approach our Milky Way potential is locally approximated for each individual sample orbit as a Stückel potential using a focal length estimated with equation (9) of Sanders (2012).

2.3 The Stellar Halo Distribution Function

For the stellar halo we use spherical DFs with constant anisotropy. These types of DFs generally take the form

\[ f(E, L) = L^{-2\beta} f_1(E) . \]  (5)

Conventionally, here, \( E = \Psi - \frac{1}{2} v^2 \) is the binding energy, where \( \Psi = -\Phi + \Phi_0 \) is the negative of the gravitational potential offset to be 0 at infinity, and \( v \) is the magnitude of the velocity. \( L \) is the total 3-dimensional angular momentum. The parameter \( \beta \) is the orbital anisotropy, defined as

\[ \beta = 1 - \frac{\langle v_\theta^2 \rangle + \langle v_\phi^2 \rangle}{2\langle v_r^2 \rangle} , \]  (6)

where \( \langle v_{\theta,\phi,r} \rangle \) are the spherical velocity dispersions in the polar, azimuthal, and radial directions respectively. \( \beta \) ranges from \(-\infty\) (completely tangential orbits), over 0 (ergodic, isotropic, or unbiased orbits), and up to 1 (completely radial orbits). As shown in chapter 4.3 of Binney & Tremaine (2008) the function \( f_1 \) is related to the mass density of the DF\(^3\), \( \rho \) as

\[ \frac{2^\beta - 1}{2\pi I_\beta} r^\beta \rho = \int_0^\Psi \frac{f_1(E)}{(\Psi - E)^{\beta - 1/2}} dE . \]  (7)

\( I_\beta \) is a constant given by

\[ I_\beta = \sqrt{\pi} \frac{\Gamma(1 - \beta)}{\Gamma(3/2 - \beta)} , \]  (8)

where \( \Gamma \) is the usual Gamma function.

For \( 1/2 < \beta < 1 \), a DF of this form is an Abel integral equation where the system is self-gravitating the potential, \( \Psi \), and the density profile, \( \rho \), are a potential-density pair. This is not necessary, however, and \( \rho \) may be independent from \( \Psi \). For our DFs \( \rho \) is an exponentially truncated power law density profile of the form

\[ \rho(r) \propto r^{-\alpha} e^{-r/r_c} . \]  (9)

where \( \alpha \) is the inner power law slope and \( r_c \) is the truncation radius. The overall normalization of the density profile is unimportant because it only determines positions and kinematics of samples in a relative sense. We set \( \alpha = 3.5 \) and \( r_c = 30 \) kpc. Our choice of parameters is inspired by Mackereth & Bovy (2020), who find \( \alpha = 3.49 \) and \( r_c = 25 \) kpc fit a large sample of halo stars spanning both [Fe/H] and \([\alpha/Fe] \). These fits stand in broad agreement with prior results, both in terms of the power law slope, and the transition radius at which the density profile steepens (Deason et al. 2011; Xue et al. 2015; Iorio et al. 2018).

The potential in the calculation of these stellar-halo DFs is spherical. Because our adopted \( \text{MWPotential2014} \) model is non-spherical, we construct a spherically symmetric analog of this potential by calculating its enclosed mass within spherical shells as a function of Galactocentric radius \( M(< r) \). The spherical version of the potential is then defined as \( \Phi_{\text{MW}} = -\frac{GM(< c)}{r} \). While possible, it is difficult to create anisotropic DFs for spherical systems, such as the stellar halo, with a given density and anisotropy embedded within axisymmetric, non-spherical potentials (Posti et al. 2015; Williams & Evans 2015). Our approach, while approximate, is simple to implement and gives us greater control over the kinematic and spatial properties of our halo DFs. The act of sphericalizing the potential should have a minimal impact on our results, which focuses on kinematic properties which are sensitive to the total enclosed mass of the potential rather than increased mass near the disc plane.

We sample from DFs of this type using `galpy`, where they are implemented as the class `galpy.df.constantbetadf`. More details on our implementation of these spherical DFs and how we specifically sample from them can be found in Appendix A.

3 DF SAMPLES

In this section, we present the samples used to explore the observable kinematics of Milky Way stellar populations, and describe how we calculate their kinematics. We first outline the overall properties of our samples and then describe how we generate the two specific sets of samples that we use throughout the rest of the paper. Finally, we present the kinematics of each of these samples.

Broadly speaking our samples consist of phase space positions drawn from a combination of the distribution functions outlined in § 2.2 and 2.3, representing the disc (approximated as a thin plus thick disc) and stellar halo. We do not consider the stellar bulge for two reasons. First, the bulge does not contribute stellar density near the solar neighbourhood in any meaningful capacity, and therefore bulge stars are reasonably easy to remove from real stellar samples by cutting stars in angular proximity to the Galactic centre. This means one can be confident that the bulge does not represent a substantial source of contamination when studying halo populations, in contrast to the stellar disc. Second, the complexity of bulge kinematics due, for example, to the presence of the rotating Galactic bar means that it is challenging to realistically model bulge populations using DFs.

3.1 The Solar Samples

The first set of samples, which we refer to as Solar are located at the Sun’s position: \((R, z) = (8.178 \text{ kpc}, 20.8 \text{ pc})\) (Gravity Collaboration et al. 2019; Bennett & Bovy 2019). This set is designed to be instructive, and so having all samples share the same position simplifies the kinematics. Yet, because spatial gradients in the DF occur over multiple-kpc scales, this sample is also somewhat realistic in that it represents stellar samples confined to the \(\approx 1 \text{ kpc} \) region around the
Sun, a distance over which the halo DF varies little. This includes existing samples such as the Gaia radial velocity sample (see section 3.3 in Gaia Collaboration et al. 2020, for the most recent data) or GALAH (Martell et al. 2017; Buder et al. 2020). The Solar set of samples consists of six independent samples of stars, which differ by the value of \( \beta \) used for the stellar halo DF.

Each of the Solar samples consists of \( 10^5 \) thin disc orbits and \( 1.2 \times 10^4 \) thick disc orbits, sampled according to the procedure outlined in § 2.2. These numbers are chosen such that the ratio of thick to thin disc stars is approximately 1:8.5, which reflects the ratio of thick to thin disc stars expected at the solar circle (e.g. see the potential of McMillan 2017). Each of the samples also contains \( 10^5 \) orbits sampled from our spherical halo DF with a different value of \( \beta \). We are not concerned with the relative numbers of halo and disc stars, because we will differentiate between the two populations when studying their properties. The first five samples each use a different value of \( \beta \) from among \([0, -0.5, 0.5, 0.7, 0.9]\). The sixth sample has a composite stellar halo, also \( 10^5 \) orbits total, consisting of a 50:50 mixture of orbits from DFs with \( \beta = 0.5 \) and 0.9 respectively. We refer to this final sample as our fiducial model.

The ratio of the component DFs in this model follows from the recent Milky Way stellar halo mass measurements of Mackereth & Bovy (2020), who find that the ratio of mass in low and high eccentricity stars is approximately 50:50. These low and high eccentricity populations are interpreted as a traditional metal poor stellar halo with \( \beta \approx 0.2 - 0.4 \), and the more metal rich Gaia-Enceladus remnant which has been shown to be characterized by large anisotropies of \( \beta \) close to 0.9 (Belokurov et al. 2018; Fattahi et al. 2019; Iorio & Belokurov 2021). This model is broadly consistent with other results that have suggested that these two halo components exist in roughly equal proportion near the solar neighbourhood (Belokurov et al. 2018; Lancaster et al. 2019; Iorio & Belokurov 2021). By using \( \beta = 0.5 \) to represent the more isotropic component of the halo we err on the side of slightly larger \( \beta \), but this allows us to be confident we do not underestimate the degree of contamination from this component when analyzing their kinematics. This is similar to the reasoning behind our choice of slightly hotter disc DF parameters discussed in § 2.2.

Throughout the paper we will refer to the two constituent components of our fiducial model as “low-\( \beta \)” and “high-\( \beta \)” for the \( \beta = 0.5 \) and \( \beta = 0.9 \) halos respectively. While our fiducial model is inspired by literature results which employ metallicity and eccentricity as discriminators for the halo populations, we emphasize that we define our halo components using \( \beta \), and therefore they may not be perfectly faithful representations of a halo defined in another manner.

### 3.2 The Survey Samples

The second set of samples has physical positions which are matched to the locations of stars in the APOGEE DR16 (Ahumada et al. 2020) statistical sample, rather than occupying a single location. The statistical sample is, briefly, the subset of APOGEE stars which were selected randomly based on the 2MASS photometric catalogue according to the procedures outlined in Zasowski et al. (2013, APOGEE-1) and Zasowski et al. (2017, APOGEE-2). We refer to these samples with positions matched to APOGEE data as the Survey samples. These samples more realistically represent modern and upcoming datasets where Gaia data is augmented by spectroscopy, which allows for the calculation of accurate spectro-photometric parallaxes (e.g., Leung & Bovy 2019b) and radial velocities for a small number of stars (compared to the broader Gaia dataset) out to large (\(~20\) kpc) distances. Examples of these surveys include the contemporary APOGEE Ahumada et al. (2020) and H3 (Conroy et al. 2019) surveys, as well as the future WEAVE (Dalton et al. 2016), 4MOST (de Jong et al. 2019), DESI (Levi et al. 2019) and SDSS-V (Kollmeier et al. 2017) projects.

We use this APOGEE sample to select an approximate population of halo, thin and thick disc stars, whose positions and distances are used to generate samples in our model. To split the APOGEE sample into halo and disc populations we apply the following selection procedure. We start with the on-sky locations, \([\text{Fe/H}]\) and \([\text{Mg/Fe}]\) abundances of stars in the APOGEE DR16 statistical sample. We supplement this data with accurate spectro-photometric distances derived using the Bayesian neural network framework astroNN (Leung & Bovy 2019a,b) (note these are based on Gaia DR2 (Gaia Collaboration et al. 2018)). Our goal is to use this information to obtain a rough set of positions corresponding to three populations we seek to study: the halo, thin disc, and thick disc.

We perform a number of cuts on the APOGEE data to ensure both quality and that the underlying stellar populations correctly reflect our model DFs. The cuts are as follows. The log \( g \) uncertainty is less than 0.1, which removes many dwarfs with poorly constrained properties. The relative distance uncertainty, \( \sigma_d / d \), is less than 0.2. We remove fields near the Galactic bulge, which we define as those with \(-20^\circ < \ell < 20^\circ\) and \(|b| < 20^\circ\). We also remove any fields which contain distinct groups of stars from a globular cluster. We first identify which APOGEE fields contain the on-sky location of a known globular cluster from the catalogue of Harris (1996, The December 2010 version). We then examine the distribution of \([\text{Fe/H}]\) abundances and radial velocities for the stars in the field, in particular those within two tidal radii of the globular cluster location. Fields where there is an obvious enhancement of stars with the appropriate \([\text{Fe/H}]\) and radial velocities near the on-sky location of the globular cluster are removed from consideration. Figure 1 shows the spatial distribution of the final selection of APOGEE stars. Our sample spans a wide range of the Milky Way, nearly 15 kpc in each direction from the location of the Sun.

We separate the APOGEE stars into three categories: halo, thin disc, and thick disc using their \([\text{Fe/H}]\) and \([\text{Mg/Fe}]\) abundances as well as their velocities. We stress that the division into a thin and thick disc is purely an approximation to the more complex structure of the disc, but one that is good enough for our main purpose of investigating the stellar halo. Figure 2 shows these abundances for our final APOGEE sample, as well as the boundaries we adopt to separate each stellar population. We choose these boundaries based on the number density of APOGEE stars in the abundance space, as well as their eccentricities (which color-codes the points in the top panel of Figure 2). The thin-thick disc boundary follows the upper edge of the low-[Mg/Fe] disc locus, while the disc-halo boundary follows the low-[Fe/H] edge of the disc locus and approximately matches the transition from low to high eccentricity. We also add to the halo sample any stars lying in the thin or thick disc chemical regions which have net velocities with respect to their local circular velocity frame which are greater than the magnitude of their local circular velocity: \( |\vec{v} - \vec{v}_{\text{circ}}(R)| > |\vec{v}_{\text{circ}}(R)| \). We find that this efficiently attributes many higher metallicity stars with genuine halo kinematics to the halo sample. We examine the energy and angular momentum plane, as well as the tangential and radial velocity plane for each of these selections to ensure that they reflect what we expect from the kinematics of the respective populations. Overall, however, the accuracy of this separation is not a major concern, because this only sets the positions based on which velocities will be sampled from the individual DFs.

The final separated samples consist of 124,879 thin disc stars,
3.3 Sample Kinematics

For both the Solar and Survey set of samples we calculate kinematic quantities as well as actions using galpy (Bovy 2015). Actions are calculated using the “Stäckel fudge” method of Binney (2012). In this approach our Milky Way potential is locally approximated for each individual sample orbit as a Stäckel potential using a focal length estimated with equation (9) of Sanders (2012). Eccentricities are similarly calculated using a variation of the “Stäckel fudge” method as described in Mackereth & Bovy (2018). These actions and eccentricities are calculated in the axisymmetric MWPotential2014 potential.

Potentials like that of the Milky Way are well-represented by Stäckel potentials (Dejonghe & de Zeeuw 1988), and the “Stäckel fudge” approximation gives more accurate action estimates for general orbits when compared with the spherical and adiabatic approximations, which are only good approximations when the potential is close to spherical, and the orbit is confined near the symmetry plane of an axisymmetric potential respectively. As shown by (Bovy 2015) the variation in Stäckel method action estimates throughout an orbit is less than about 2 per cent for a disc-like orbit with reasonable radial and vertical extent. Mackereth & Bovy (2018) find that the Stäckel approximation performs just as well as direct orbit integration for essentially all halo orbits and most disc-plane orbits which do not venture near the center of the galaxy (where resonances tend to complicate the orbit structure). We test the constancy of this approximation by calculating focal lengths and actions for a few randomly selected APOGEE halo stars at multiple points along their orbits. We find the typical median absolute deviation to be 7 kpc km s\(^{-1}\) for the radial actions and 3 kpc km s\(^{-1}\) for the vertical action \((L_z)\) is obviously invariant along the orbit). The typical median absolute deviations expressed relatively are 2 per cent for the radial actions and 1 per cent for the vertical action, which is similar to the findings of Bovy (2015).

We consider six kinematic spaces which are commonly used throughout the literature for both structure identification and classification, particularly of halo populations. The first two use only velocity information, while the latter four rely on conserved quantities calculated using the underlying potential. For each space we provide a few recent example studies which have made use of it, but the list is by no means exhaustive. The first is the \(v_R - v_T\) plane (e.g. Belokurov et al. 2018; Fattahi et al. 2019; Lancaster et al. 2019; Mackereth et al. 2019a; Belokurov et al. 2020; Feuillet et al. 2020), which is useful since disc and halo populations occupy distinct loci defined by their mean \(v_T\), and disc populations have much lower tangential and radial velocity dispersions than halo populations. Additionally...
the degree of radial bias in a halo population is easy to gauge by the radial extent of its velocity ellipsoid. The second space is the Toomre diagram (e.g. Hawkins et al. 2015; Helm et al. 2018; Koppelman et al. 2019b; Feuillet et al. 2020; Cordoni et al. 2020) which plots the quadrature\(^{\frac{1}{2}}\text{\(v_R^2 + v_T^2\)}\) of the \(v_R\) and \(v_T\) velocities (also known as \(v_{\text{perp}}\)) as a function of \(v_T\). Here we use cylindrical velocities, yet the Toomre diagram is also often expressed using \(UWV\) velocities. In samples that span a large range in Galactic azimuth, the \(v_R\) and \(v_T\) velocities project onto \(U\) and \(V\) in a manner that renders the diagram less useful. The advantage of the Toomre diagram is that it factors in \(v_z\) in addition to \(v_R\), further isolating halo populations thanks to the relatively cold vertical velocity kinematics of the thin and thick disc populations.

The next three spaces all use angular momentum about the disc rotation axis, \(L_z\), as the abscissa. \(L_z\) is an effective coordinate since disc stars tend to have characteristically large, positive values. Additionally, because it is one of three actions, it is conserved under adiabatic conditions. The third space is the \(E - L_z\) plane (e.g. Helmi et al. 2018; Koppelman et al. 2019b; Naidu et al. 2020; Feuillet et al. 2020; Cordoni et al. 2020; Horta et al. 2021), which is perhaps the most widely used kinematic space for structure identification in the \(Gaia\) era. This is due to the relative ease with which these parameters are calculated, their approximate conservation through cosmic time, the fact that the disc occupies a compact locus at high \(L_z\) and moderate \(E\), and that the energy encodes both radius and velocity information. We plot all energies offset by the potential of \(MW\) to maximally separate halo stars on tangential versus radial orbits. The fourth space is the root of the radial action, \(\sqrt{R_J}\), as a function of \(L_z\) (e.g. Tril et al. 2019; Feuillet et al. 2020; Matsumo et al. 2021; Horta et al. 2021). By using the radial action this space can very efficiently separate halo stars on tangential versus radial orbits. This is because the action labels the entire orbit while velocities, for example, will change throughout an orbit. We use the square root of the radial action to compactify the vertical axis for efficient visualization. The fifth space is eccentricity (e) as a function of \(L_z\) (e.g. Cordoni et al. 2020). While this specific two-dimensional space is not commonly used, eccentricity as a single parameter is often used as a metric to separate GE from the remainder of the stellar halo since it reflects the overall radial extent of an orbit (e.g. Belokurov et al. 2018; Mackereth & Bovy 2020; Naidu et al. 2020).

The final space is the “action diamond” (e.g. Vasilev 2019; Myeong et al. 2019; Monty et al. 2020; Cordoni et al. 2020; Naidu et al. 2020) which in our implementation plots \(J = J_R\) as a function of \(L_z\), the azimuthal action. But both quantities are scaled by the total action defined as \(J_{\text{tot}} = |J_R| + |L_z|\), which causes the appearance of a characteristic diamond-shaped boundary (Note that if the normalizing total action is \(J_{\text{tot}} = \sqrt{J_R^2 + L_z^2}\) the boundary is circular, e.g. Naidu et al. 2020). One advantage of this space is that the quantities are scaled by their total action, which means that disc populations whose action budget is dominated by \(L_z\) tend to cluster tightly in the right corner of the bounding diamond. For the same reason halo sub-populations which arise from, for example a single merger event, are likely to cluster at a single location defined by the actions of the progenitor. The downside of this space is that since all quantities are normalized by the total action, populations which differ by an overall magnitude in their actions, such as eccentric disc and halo populations, may appear indistinct. To combat this effect it is often advisable to use this space only with a sample confined to a narrow range of radii or \(L_z\).

It is natural to consider whether higher dimensional spaces would be more appropriate for this type of study. While we have mentioned above many unique quantities that describe stellar orbits, they are not all completely independent. We believe that there are three broad classes of quantities here that we consider here. First are quantities that describe tangential motion, which includes \(v_T\) and \(L_z\) (whether it is normalized by \(J_{\text{tot}}\) or not. The second are quantities which describe the magnitude of radial motion, which include \(v_R\), \(v_{\text{perp}}\), energy, and \(J_R\). The third are scaled quantities which describe the radiality of an orbit, which include the normalized quantity \((L_z - J_R)/J_{\text{tot}}\) and eccentricity. Within each of these categories is a modest to large amount of degeneracy. For example a star with high radial velocity likely also has a large \(v_{\text{perp}}\), energy, and \(J_R\). On the other hand, a star with large radial velocity may not necessarily be on a particularly radial orbit if its tangential velocity is also large. This suggests that the largest dimensional space which would be useful in a coherent sense would be three dimensional. We will explore the possibility of using higher dimensional spaces to select high purity samples of stars from high-\(\beta\) populations in § 5.1. For now we will continue to work with two dimensional spaces for ease of visualization.

Figure 3 shows these six kinematics spaces for the \(Solar\) samples. The disc samples, both thin and thick together, are shown as the black contours, while the halo samples are shown as blue points. Each of the columns shows one of the six different halo models. Figure 4 is the same as Figure 3, but shows the kinematics for the \(Survey\) samples instead. The disc sample contours are constructed by binning the data with 40 bins across each dimension in each space. The contour levels are then placed at 1, 10, and 100 samples per bin.

4 EXPLORING THE KINEMATIC PROPERTIES OF OUR SAMPLES

In this section, we examine in detail each of the kinematic spaces we presented in the previous section. We first provide a qualitative overview of the kinematic spaces, especially with regards to how kinematics change with variations in the value of \(\beta\) of the stellar halo DF.

4.1 A Qualitative Assessment of the Kinematic Spaces

4.1.1 The Solar Samples

We first take a closer look at the \(Solar\) sample kinematics from Figure 3. Starting with the \(v_R - v_T\) velocities, we can immediately see a strong degree of evolution in the morphology of the halo samples as \(\beta\) changes. Specifically, as \(\beta\) increases the dispersion along the \(v_R\) axis increases and the dispersion along the \(v_T\) axis decreases, broadly in accordance with the definition of \(\beta\).

One subtle feature is that as \(\beta\) increases the density of samples along the \(v_T = 0\) line forms a broad plateau at \(v_R = 0\), rather than the distinct density peak that can be seen when \(\beta = 0\) for example. This behaviour is the result of a tendency towards bimodality in the \(v_R\) direction for high-\(\beta\) samples, since stars are on more radial orbits and therefore preferentially have non-zero radial velocities (with equal preference to positive and negative velocities). Examining the fiducial model highlights why it is important to keep this fact in mind. When low and high-\(\beta\) populations are shown together, it may appear to be a superposition of two centrally concentrated ellipsoids, yet in reality the lower-\(\beta\) population provides the peak at \((v_R, v_T) = (0, 0)\) while the radially biased population tends to more evenly populate the low and high radial velocity region. For much higher \(\beta\) or when the density profile allows for apocenters at much larger Galactocentric radii, one will also expect to observe a genuine bimodality in
Figure 3. Kinematics of the six Solar samples. Each of the columns shows one of the subsets of the Solar sample, which vary by the $\beta$ of the halo component (labelled). The blue points show the halo samples while the black contours show the disc samples, combining both thin and thick. Top row: Radial and tangential velocities; second from top: Toomre diagram; third from top: energy and $L_z$; third from bottom: $\sqrt{\mathcal{J}_R}$ and $L_z$; second from bottom: eccentricity and $L_z$; bottom row: action diamond. The units (unlabelled) are as follows: velocities are in km s$^{-1}$, actions and angular momenta are in kpc km s$^{-1}$, and energies are in km$^2$ s$^{-2}$. 

$\beta = -0.5$, $\beta = 0.5$, $\beta = 0.7$, $\beta = 0.9$, and fiducial.
the radial velocity direction, with a distinct lack of stars with zero velocity.

This is noteworthy when modelling the local stellar halo velocity field using, for example, a Gaussian mixture model. One would expect a two-component model to fit the radially anisotropic halo better than a single component. This assessment is supported by the results of Belokurov et al. (2018) and Fattahi et al. (2019), who find that using two Gaussian components for the local halo (i.e. one for the metal poor, more isotropic halo, and one for the radially anisotropic halo) produces best-fits which have systematically negative model residuals near $v_r = 0$ (note they use spherical polar coordinates), and positive model residuals towards larger radial velocities. This

Figure 4. Same as Figure 3 but for the six Survey samples.
indicates that the single component fit to the radially anisotropic population expects more stars near \( v_\rho = 0 \) and fewer at large radial velocities, in accordance with our expectations. Other studies which have used symmetric, bimodal Gaussians have found good fits to data (Lancaster et al. 2019; Necib et al. 2019).

The Toomre diagram shows evolution with increasing \( \beta \) in a manner similar to the \( r + v_T \) plane. As the radial anisotropy of the halo population increases the distribution elongates along the \( v_{\text{perp}} \) axis, and narrows along the \( v_T \) axis. The inclusion of \( v_z \) information in addition to \( v_R \) in the Toomre diagram does tend to separate the halo samples from the disc samples somewhat more than in the \( v_R - v_T \) plane. But the effect is not as pronounced as might be expected, likely because many disc samples have roughly comparable vertical and radial velocities, and so their extent along the \( v_{\text{perp}} \) axis also increases. In both of these first two kinematic spaces, the disc occupies similar loci at \( v_T \approx 200 \text{ km s}^{-1} \) with zero net radial or vertical velocity. It is important to note, however, that even in this population of disc samples located at the solar position there are a marginal yet noticeable number belonging to the thick disc with near-0 \( v_T \). In the Toomre diagram even though the halo samples are more distinct, the maximum perpendicular velocities of the hottest thick disc samples approaches 200 km s\(^{-1}\).

The \( E - L_z \) plane in the third row of Figure 3 exhibits the now famous central enhancement of samples with \( L_z \approx 0 \) for large \( \beta \). It is this feature which is most characteristic of the GE population as first shown by Helmi et al. (2018). One of the reasons this space has been so frequently in the Gaia era is that both \( E \) and \( L_z \) are approximately conserved integrals of motion, and therefore the appearance of stellar (sub)structures in this space often point to a common ancestry for the constituent stars. Another key feature of this space which we will explore in more depth later is the distinct parabolic shape of the overall distribution. This boundary is set by the fact that a given energy and radius the magnitude of the velocity is fixed. This results in a bounding value of \( L_z \) which occurs when the velocity is wholly oriented in the direction of \( v_T \). The equation of the bounding curve is: \( E = L_z^2/(2R^2) + \Phi(R) \). Because the samples in the Solar set are all at a fixed radius the boundary is obvious in Figure 3. There is also a clear gradient in energy, with the density of samples decreasing as the energy increases. This is intrinsic to the energy-dependence of spherical DFs (see chapter 4.3 in Binney & Tremaine 2008) and is therefore observed even at a fixed radius. As the radius changes, the energy of the apex of the bounding parabola also changes. This emphasizes that the relationship between radius and energy is complicated, a feature which we will discuss more later.

The radial action, \( J_R \), is perhaps the most natural variable for isolating stars from populations with high radial anisotropy, because it labels the entire underlying orbit rather than being sensitive to the phase of the orbit such as velocities are. We see in Figure 3 that there is an approximately inverse relationship between the magnitude of \( L_z \) and the radial action. This occurs because as angular momentum increases at a fixed position, less of the velocity budget is necessarily contained in radial motions. The relationship is not perfect, however, for two reasons. First, this kinematic space does not account for vertical motions, and second there is a range of energies in the Solar samples as we have just discussed. As \( \beta \) increases a distinct cusp forms around \( L_z = 0 \) spanning a range of \( \sqrt{\mathcal{J}_R} \). The fact that this cusp has a distinct minima (around \( \sqrt{\mathcal{J}_R} \approx 25 \text{ kpc}^{1/2} \text{ km}^{1/2} \text{ s}^{-1/2} \)) in Figure 3 reflects the fact that at the solar circle, the only orbits that exist with \( L_z = 0 \) must either have a certain degree of radial motion, or otherwise must be on a circular orbit exactly over the Galactic pole. The cusp spans a range of \( \sqrt{\mathcal{J}_R} \) for the same reason that the inverse relationship between \( \sqrt{\mathcal{J}_R} \) and \( L_z \) is imperfect at non-zero \( L_z \); there is a range of energies in the samples and vertical motions are not accounted for.

The relationship between eccentricity and \( L_z \) is similar in spirit to the relationship between \( \sqrt{\mathcal{J}_R} \) and \( L_z \). As \( L_z \) increases at a fixed location, eccentricity must necessarily decrease because orbits become less radial. This lends a characteristic v-shape to the samples in this space. Unlike \( \sqrt{\mathcal{J}_R} \), eccentricity is a “scaled” parameter, it depends only on the relative values of the orbit’s pericenter and apocenter. Therefore, we see that disc samples appear less distinct from halo samples than in other kinematic spaces. From an intuitive perspective, a halo star with pericenter at the solar circle could have the same eccentricity as a disc star with apocenter at the solar circle, yet these two stars would have very distinct radial actions. A second curiosity of eccentricity is the fact that as orbits become more radial, their eccentricity approaches 1 asymptotically slowly. This means that as \( \beta \) increases samples bunch up close to \( L_z = 0 \) and \( e = 1 \). This is markedly different behaviour than most other kinematic spaces we have presented so far, where the kinematic properties of radially-biased samples have a distinct magnitude. Here, the only distinct property of the radially biased population is its sheer density at high eccentricity. We will discuss how this impacts the use of eccentricity as a parameter to select high-\( \beta \) populations later.

The action diamond is perhaps the least used among the kinematic spaces shown here, with a similar implementation first being used by Vasiliev (2019) (to our knowledge, although they recognize that the space is analogous to visualisations of three-dimensional action spaces presented by Binney & Tremaine 2008). Its most common use thus far has been to study globular cluster kinematics and their potential association with larger halo structures (Vasiliev 2019; Myeong et al. 2019). But it still offers useful information for the types of aggregate stellar populations we study here. Understanding the action diamond is intuitive. The right and left points of the diamond (\( |L_z| = J_{\text{tot}} \)) are in-plane prograde and in-plane retrograde orbits respectively. The top point (\( J_z = J_{\text{tot}} \)) is a polar orbit, and the bottom point (\( J_R = J_{\text{tot}} \)) is a radial orbit. The bottom-right and bottom-left edges (\( J_{\text{tot}} \) wholly contained in \( L_z \) and \( J_R \)) are prograde and retrograde in-plane orbits. The top-right and top-left edges (\( J_{\text{tot}} \) wholly contained in \( L_z \) and \( J_z \)) are prograde and retrograde out-of-plane orbits.

One of the most noticeable features of the action diamond is that both disc and halo populations tend to be biased towards the bottom half of the diamond, which reflects the fact that both populations tend to have modestly hotter radial versus vertical kinematics. Indeed, we see what is perhaps the most subtle transformation in morphology from the low to high-\( \beta \) halo population, where the samples appear to concentrate more towards the bottom point of the diamond rather than the left and right points. This results from the dominance of the radial action over the vertical action in the high-\( \beta \) sample. The disc also hugs the bottom-right edge of the diamond, but the action budget is dominated by \( L_z \), meaning it is mainly confined to the right point of the diamond. Similarly to eccentricity this is also a “scaled” space, so two orbits with different absolute kinematic properties may occupy the same point. So, while the strength of the action diamond lies in the fact that distinct types of orbits occupy different loci in the space, the weakness is that confusion between two populations that may have different absolute kinematics, but similar types of orbits is likely. This will become more evident when examining the Survey sample set.
4.1.2 The Survey Samples

The Survey samples differ from the Solar samples in three key respects with regards to the six kinematic spaces that we consider. First, the sample positions span a range of Galactocentric radii, and so in some sense it is useful to consider the kinematic profiles as a superposition of many of the individual samples seen in Figure 3, but determined at different radii. Second, the samples are perturbed by Gaia and APOGEE errors. Third, and perhaps most important, is that the positions underlying the samples are subject to the complicated APOGEE DR16 selection function. It is through these three lenses that we will assess the kinematics of the Survey sample set in Figure 4.

First, one obvious change to all of the kinematic spaces is the increased range over which the disc populations are observed, while in the Solar samples the disc populations tended to be confined to a narrow region. This is due mostly to the first two factors mentioned above: the increased radial range over which the DFs are sampled and the addition of realistic uncertainties. Furthermore the discrepancy in the number of disc samples versus halo samples (roughly 66:1) exacerbates the degree to which the disc populations extend into the stellar halo kinematic space with meaningful density.

In the $v_R - v_T$ plane, the overall morphology of the halo samples stays the same, with bimodality along the $v_R$ axis increasing as $\beta$ increases. The impact of the realism steps on the data is clear, particularly in how the Survey samples do not cluster so tightly on the $v_T = 0$ line, especially in the samples with higher $\beta$, as the Solar samples do. Overall our fiducial model looks realistic in its representation of the nearby Milky Way $v_R - v_T$ plane (c.f. figures 2 and 1 in Belokurov et al. 2018; Fattahi et al. 2019, respectively, although note they use spherical coordinates), a testament both to the validity of our models and the steps we take to add realism to our DF samples. The Toomre diagram changes in much the same way as the $v_R - v_T$ plane: the overall morphology is the same, but the addition of uncertainties has softened the cusp in the halo samples along the $v_T = 0$ line. In both these planes, the disc population now extends down to near $v_T = 0$, which reflects the addition of samples from the inner galaxy where the angular momentum is lower and the radial velocity dispersion is higher. The extent of the disc in the $v_R$ and $v_{\perp}$ dimensions does not increase as drastically, because only the increased radial and vertical velocity dispersions in the inner Galaxy serve to modify its extent.

From a qualitative standpoint, the $E - L_z$ plane changes in much the same way as the velocity planes. The cuspy overdensity of the high-$\beta$ populations is softened, and the hottest thick disc stars begin to impinge on this characteristic region, yet leave a discernable overdensity of halo samples visible just beyond their boundary. The $E - L_z$ plane also shows a distinct change in morphology of the bounding region of the kinematic space, which we mentioned in the previous section is due to there being a fixed velocity budget at a given energy and radius. For the Survey sample, where stars at many radii are included, the bounding region is a superposition of many of the aforementioned parabolic curves. These bounding regions tend to get wider as radius increases since the potential scales inversely with distance from the Galactic center. For inner Galaxy samples the bounding parabola is narrow with minima at very negative energies. Conversely, for outer Galaxy stars, the bounding parabola is wide with a minima at higher energy. The superposition of these bounding curves results therefore in a funnel shape which narrows as the energy is decreased. An example of this morphology is shown in Figure 5 where the in-plane bounding parabolas for MWPotential2014 corresponding to radii 2-15 kpc are shown.

A particularly interesting feature of the Survey sample $E - L_z$ plane is that there appears to be a region of maximal density which begins at $E \sim -1.3 \times 10^5$ km$^2$ s$^{-2}$ and follows an approximate parabolic shape towards larger energies as $|L_z|$ increases. These are most obvious in the $\beta = 0$ and 0.5 models, but are also noticeable at higher $\beta$. Below this energy there are halo samples, yet they are observed in much lower numbers than at this overdensity. The interpretation of this overdensity is that the majority of the samples in the Survey set lie near the Sun, and therefore they tend to be roughly confined by the bounding parabola defined at that location. The occurrence of overdensities in $E - L_z$ space depending on the underlying radial distribution of samples is something we will explore in the next subsection.

The $\sqrt{J_R} - L_z$ plane no longer has any indications of the inverse relationship between $J_R$ and $L_z$ observed in the Solar samples and discussed above. Likely the combination of the radial range as well as realistic uncertainties causes all traces of the relationship to disappear. Instead now the halo samples manifest as a broad cloud that is slightly biased either towards $L_z$ or $\sqrt{J_R}$ depending on the value of $\beta$. The halo samples of the most radially biased models still do exhibit a diffuse cusp along the $L_z = 0$ line. The differences in the $e - L_z$ plane tell a similar story. There is no longer as sharp an inverse relationship between eccentricity and $L_z$ as was observed in the Solar samples, yet the relationship is at least somewhat apparent in contrast with the $\sqrt{J_R} - L_z$ plane. The higher $\beta$ models still tend to simply cluster more tightly around $L_z = 0$ and $e = 1$, similar to the Solar samples. Notably, the disc samples extend all the way to the high-$\beta$ overdensity, and in general completely dominate the prograde half of the plane.

In the action diamond, the disc populations transition from occupying a narrow strip near the prograde corner to dominating the lower right quadrant of the diamond. This large change in area of the disc populations demonstrates a weakness in the action diamond (indeed one shared by eccentricity), which is that “scaled” kinematic quantities often confuse orbits that only broadly share certain properties. In this case, when the vertical and radial actions of the hottest thick disc stars become of order comparable with their $L_z$, they migrate towards the center and bottom corner of the diamond overlapping with the halo samples, even though the magnitude of their actions can be quite different from those of stars in the halo. The halo samples have the same qualitative trends as are seen in the Solar samples. As
As β increases, the samples concentrate near the bottom of the figure. Additionally of note is that very few ergodic samples occupy this region of the diagram, they are more uniformly spread throughout the space and even biased towards the high-$L_z$ corners. This means the bottom corner is likely a good region to isolate radially biased halo populations from ergodic populations.

4.2 The Impact of the Underlying Radial Distribution of Samples on the $E - L_z$ plane

In the previous section, we noted that the Survey samples exhibit an interesting change in morphology in the $E - L_z$ plane when compared with the Solar samples. In the $E - L_z$ plane of the latter (Figure 3) the samples were confined by a bounding parabola defined by the escape velocity as a function of changing energy at a fixed Galactocentric radius (this being the location of the Sun in the Solar sample). Sample density decreases as energy increases moving away from the bounding parabola. The radially biased samples have a higher energy cusp around $L_z = 0$, leading to different overall morphology but the same trend still holds. This phenomenon lead us to interpret the apparent overdensity in the $E - L_z$ plane of the low-β Survey samples around $E \sim -1.3 \times 10^5$ km$^2$ s$^{-2}$ as being caused by the concentration of the samples near the Sun.

While this observation is not particularly unexpected, it suggests a question: what would happen if the intrinsic radial distribution of samples (i.e. the APOGEE DR16 data which sets the approximate positions of our samples) took a more complicated form? We can investigate this using APOGEE DR16 by augmenting the Survey sample with additional samples directed towards the Galactic bulge. Recall that we do not consider any stars for inclusion in the Survey sample if they have $-20^\circ < \ell < 20^\circ$ and $|b| < 20^\circ$. Here, just to investigate the radial distribution of samples, we soften that requirement and allow stars lying in this angular range with Galactocentric radius greater than 3 kpc to be included into the Survey sample set. Note that this distance from the Galactic center is approximately equal to the distance from the Galactic center enforced by the $20^\circ$ angular cut we use to define the Survey sample - i.e. $\tan(20^\circ) \times 8.178$ kpc $\approx 3$ kpc. This radius is also the radius at which the density of the bulge component of MWPotential2014 drops below the density of our stellar halo model when it is normalized to have the mass found by Mackereth & Bovy (2020). This is all to say that we are confident we do not introduce many bulge stars into the sample by including these new data. Separating halo stars from disc stars, sampling kinematics from the DF, and the perturbation by uncertainties are all done in the same manner as described in § 3. This results in 497 new halo samples.

The top panel of Figure 6 shows the radial distribution of the Survey samples, the additional samples in the direction of the Galactic center, and their combination. Note that a number of the additional Galactic center samples appear at radii less than 3 kpc. This can be attributed to the perturbation by the APOGEE and Gaia uncertainties. The overall distribution of radii is clearly bimodal, with peaks around 3.5 and 8 kpc, and a decrease at 6 kpc. The inner peak is driven mostly by the additional samples from the inner Galaxy pointings, but hints of the overdensity are clearly present in the Survey sample. The occurrence of this bimodality is undoubtedly caused by the APOGEE selection function, and specifically the location of the survey pointings.

The bottom-left panel of Figure 6 shows the $E - L_z$ plane for the fiducial model of the Survey sample plus the additional samples towards the Galactic center, 2,695 samples in total (halo samples only, no disc). It is immediately clear that there is an underdensity between $-1.25$ and $-1.5 \times 10^5$ km$^2$ s$^{-2}$. This is confirmed by examining the bottom right panel, a histogram of energies, which highlights the overdensity at the aforementioned location. It does not perfectly highlight the overdensity because, as we have established, the density fluctuations occur on roughly parabolic contours, but nonetheless its signature is there. Overlaid on the figure are three black lines that are sections of the bounding parabolas corresponding to the three radii marked with grey dashed lines in the top panel: 3.5, 6, and 8 kpc. Three black lines in the bottom panel show sections of the bounding parabolas for those corresponding radii, roughly marking the boundaries between over and underdensities in the $E - L_z$ plane.

5 THE PURITY OF POPULATIONS OF INTEREST

One of the most useful questions we might ask is whether we can use our samples to identify which kinematic spaces are best for identifying certain populations that we are interested in separating from a background of contaminants. For the purposes of this work we focus on separating the high-β populations from low-β and disc populations.
5.1 Separating High-$\beta$ Populations from Low-$\beta$ and Disc Populations

The separation of high-$\beta$ populations from low-$\beta$ populations is currently of special importance, because the two dominant inner-Milky Way halo populations are the in-situ low-$\beta$ component and the high-$\beta$ GE component (Belokurov et al. 2018; Helmi et al. 2018; Iorio & Belokurov 2021). While these two populations do have somewhat unique chemistry (Haywood et al. 2018), it is not so distinct that abundances are an efficient means of separating them, such as is the case for separating the disc populations both from one another and from the halo. Building on this sentiment, we do not even know a priori the detailed chemical profile of the in-situ and GE components so that we might try and separate them using their abundances. Their individual chemistries can only be effectively learned if the populations can be separated by some other means. Additionally, detailed mass modelling in the halo requires some means of tagging different populations of interest. For example Mackereth & Bovy (2020) divide the halo into mono-abundance populations. They use a simple eccentricity cut to separate ergodic and high-$\beta$ populations. We here will delve deeper into the methods for kinematically separating these two major halo populations.

While in the previous section, our focus was on faithfully reproducing the types of samples observed in modern datasets, here we must focus on numerical accuracy. We therefore increase the number of samples in the Survey sample set so that the small number of halo stars used to generate the model (2,198) does not impact the fidelity of our results. We do this by generating one hundred different velocity samples at each APOGEE halo star location, rather than just one, for each halo DF with a different $\beta$. This “augmented” Survey sample set is otherwise processed exactly the same as described in § 3.2 and 3.3. For the disc we increase the number of samples by a factor of ten in the same manner. We use these augmented Survey samples for the remainder of the paper, and the Solar samples remain unchanged.

We use our fiducial halo model to investigate the purity of ergodic and high-$\beta$ populations; recall that the fiducial halo model consists of a 50:50 mixture of $\beta = 0.5$ and 0.9 populations. Figures 7 and 8 show the purity of the high-$\beta$ component of the fiducial model with respect to the low-$\beta$ component for the Solar and Survey sample sets respectively. Purity is calculated as $N_{\beta=0.9}/(N_{\beta=0.9} + N_{\beta=0.5})$, and the data is binned with 40 bins across each dimension of each kinematic space. Overlay on each kinematic space is the lowest contour of the disc samples from Figures 3 and 4 to indicate where the disc may contribute contaminants to the purity measurement.

Qualitatively, the purity of high-$\beta$ samples follows intuitively from Figures 3 and 4. In the $v_R - v_T$ and Toomre diagrams the purity is highest along the $v_T = 0$ axis and at large values of $v_R$ and the $v_{\text{perp}}$ respectively. In the $E - L_z$ plane, we see that the purity is highest when $L_z$ is zero and at modest to high energies. The $\sqrt{v_R^2 - L_z^2}$ plane behaves similarly, with a cloud of high-purity at large values of the radial action near $L_z = 0$. Eccentricity is peculiar, as we have already noted in that being a “scaled” parameter the radially biased populations do not occupy a distinct region of the space. Rather they tend to cluster with greater density near $L_z = 0$ and $e = 1$. This behaviour is necessary to keep in mind when evaluating this space on the merits of high-$\beta$, because one might assume that the best place to isolate radially biased populations is near $e = 1$ but at higher $L_z$. In reality, the bulk of the sample lies at $L_z = 0$ but simply exists at lower purity because the ergodic sample also populates that region.

The action diamond exhibits the highest purity at the bottom corner where the radial action dominates the action budget. In particular in the Solar sample, the region of high-purity is only seen in the few pixels around the apex of the diamond.

When evaluating how each kinematic space performs in separating populations, it is important to keep in mind the overall density of halo stars. Many regions of each space exhibit spuriously high purity where there are only a few halo samples. We therefore consider only bins with greater than 5 samples in them, and render those bins with fewer than 5 samples grey in Figures 7 and 8. With this in mind, we have selected elliptical regions in each space where both the purity is large and the density of high-$\beta$ halo samples is reasonable. These are shown as the dashed black lines in Figures 7 and 8. The parameters of the selection ellipses are given in Table 2. We choose these simple shapes to represent the high-$\beta$ regions because one goal in this analysis is to emphasize applicability to real data. Shapes which are too specific to the populations here may provide errant results when applied to other data where the kinematic properties may be subtly different (e.g., if either the spatial selection or the potential used to derive kinematics changes). Instead, we pick regions which would be mostly insulated from the peculiarities of a specific data set. In each of the six kinematic spaces at least one of the dimensions probes the circularity of the orbit in the disc plane. Following this, each of the selection regions primarily picks out samples with $L_z$ or $v_T$ approximately equal to zero (the bottom corner of the action diamond necessarily has $L_z = 0$). The selections then pick out extreme values of the other quantity, which generally trends with the degree of radial bias in a population (specifically $v_R$, $v_{\text{perp}}$, $E$, and $\sqrt{v_R^2 - L_z^2}$). The selection regions of the “scaled” spaces tend to be concentrated at the extremas of the scaled quantity (eccentricity and $(J_2 - J_R)/J_R$).

| Space | Center [x,y] | SMA [x,y] | C | $P_{\text{halo}}$ | $P_{\text{disc}}$ |
|-------|--------------|-----------|---|----------------|-----------------|
| Solar sample | $v_R - v_T$ | [±260, 0] | [140,20] | 0.29 | 0.79 |
| | Toomre | [0, 200] | [20,200] | 0.72 | 0.74 |
| | $E - L_z$ | [0, -0.9] | [200,0.45] | 0.72 | 0.74 |
| | $\sqrt{v_R^2 - L_z^2}$ | [0, 45] | [250,20] | 0.69 | 0.78 |
| | $e - L_z$ | [0, 1] | [500,0.025] | 0.62 | 0.83 |
| Action Diamond | [0, -1] | [0.1,0.1] | 0.62 | 0.85 |
| Survey sample | $v_R - v_T$ | [±290, 0] | [110,35] | 0.12 | 0.76 |
| | Toomre | [0, 300] | [35,120] | 0.15 | 0.71 |
| | $E - L_z$ | [0, -1] | [300,0.5] | 0.74 | 0.66 |
| | $\sqrt{v_R^2 - L_z^2}$ | [0, 45] | [300,20] | 0.52 | 0.76 |
| | $e - L_z$ | [0, 1] | [500,0.025] | 0.61 | 0.82 |
| Action Diamond | [0, -1] | [0.08,0.3] | 0.39 | 0.86 | 0.85 |

It still interesting to ask, however, what is the tradeoff between...
Figure 7. Purity of high-β halo samples with respect to low-β samples for the Solar sample set. The kinematic spaces and units are as described in Figure 3. The grey bins are those with fewer than 5 samples in them. In each of the panels the dashed black ellipse(s) show the selection ellipses for the high-β sample. The solid black contour is the lowest contour of disc stars from Figure 3.

Figure 8. Same as Figure 7 but for the augmented Survey sample set. Here, the solid black contour shows ten samples per bin calculated in the same manner as Figure 4.
increasing or decreasing the size of the selection ellipses? Intuitively, a larger selection boundary would include more high-β samples, but would also include more low-β samples which would decrease purity. To explore this, we make purity-completeness relations for each of the kinematic spaces for both the Solar and Survey sample sets. To do this we consider two sets of scaling factors: \((x_1, x_2)\) where each factor takes the form \(x_1 = 2^y\). For each of the two scaling factors we consider 25 values of \(y\) uniformly spaced on the interval \([-2, 2]\), giving a grid of 625 total combinations. For each of these 625 unique pairs of factors we multiply them to the semi-major axes in Table 2 to generate a new selection ellipse. We then calculate the high-β purity and completeness according to this new selection ellipse.

To generate an average purity-completeness relationship we bin the 625 purity measurements in completeness and calculate the mean purity in each bin. The results are shown for the Solar and Survey sample sets in Figures 9 and 10 respectively. Each of the panels highlights one kinematic space (labelled), the curves for the other kinematic spaces are shown faintly behind to emphasize differences. In the panels of Figure 10 there is a solid and dashed line that show the purity calculated only considering halo samples (solid) and the purity calculated including both halo and disc samples (dashed). Additionally, in each panel there is a colored circle which shows the purity and completeness of the default boundaries shown in Figures 7 and 8.

For the Solar samples, the purity-completeness relations are similar for quantities which are similar to one another. \(v_R - v_T\) and the Toomre diagram are similar, \(E \perp L_z\) and \(\sqrt{J_R} - L_z\) are similar, and \(e \perp L_z\) and the action diamond are similar. The action diamond appears to be the superior space, achieving a higher purity at a given completeness than any other kinematic space. At large completeness the other kinematic spaces are more-or-less tied with one another. The apparent bimodalities (i.e. the peaks near completeness of 0.6) in the purity-completeness relations for the absolute kinematic spaces is just an artifact of how we have visualized the results of this analysis.

In general, the linear trend at higher completeness in these spaces captures the variation in the rotation-sensitive quantity \((v_T\) and \(L_z)\) when the anisotropy-sensitive quantity \((v_R, v_{\perp}, E, \sqrt{J_R})\) is scaled to a large range. At lower completeness the same trend holds when the radial quantity is scaled to small values. These two trends emerge because of the logarithmic manner in which we generate the scaling factors. A general trend for these four spaces is as follows: at a given ellipse size for the anisotropy-sensitive quantity, a narrower selection for the rotation-sensitive quantity produces higher purity and lower completeness. For the action diamond and \(e \perp L_z\) the relations are simple because the high-β samples occupy a single locus in these spaces. In general the only takeaway from this analysis applied to the Solar samples is that the action diamond appears to outperform the other spaces. We are wary of making any other statements because the simplicity of our Solar samples clearly incites slightly pathological results in this specific analysis.

For the Survey samples the differences between the kinematic spaces are more interesting. The action diamond and \(e \perp L_z\) plane have very similar purity-completeness relations, with the action diamond having lower completeness and higher purity, and the \(e \perp L_z\) plane having higher completeness and lower purity. Both of these spaces have little contamination from disc samples (solid and dashed lines are similar). These two scaled spaces appear superior to the other four spaces under consideration. A close third though appears to be the \(\sqrt{J_R} - L_z\) plane, which has slightly lower purity, with the other three spaces at the lowest purities. The shapes of the purity-completeness relations for the velocity planes appears to be slightly peaked at modest completeness. These modest peaks represent selections that span the whole range of \(v_R\) and \(v_{\perp}\) respectively. With these selections you would greatly increase the completeness while keeping the purity relatively constant, however the purity including disc stars drops drastically. Overall the \(v_R - v_T\) plane achieves slightly higher purity than the Toomre diagram, likely confirming earlier suspicions that the inclusion of \(v_z\) in the Toomre diagram’s \(v_{\perp}\) actually serves to confuse between low- and high-β samples. Finally, the \(E \perp L_z\) plane achieves the lowest purities, slightly below the velocity planes, but it does so with quite high completeness. The best kinematic space appears to be the action diamond if purity is of interest. If completeness is of interest, the \(e \perp L_z\) plane achieves high completeness while also keeping the purity relatively high.

For both the Solar and Survey samples we also consider composite selections combining the selection ellipses of multiple kinematic spaces. Intuitively, one might expect that the best composite selection would be to combine a scaled space, such as the action diamond or \(e \perp L_z\) with \(\sqrt{J_R} - L_z\) or \(v_R - v_T\). The scaled space selects for stars on very radial orbits, while the other space selects for stars with radial motions of a large magnitude. We find only very minor improvements by combining the action diamond or \(e \perp L_z\) selections with the \(\sqrt{J_R} - L_z\) or \(v_R - v_T\) selections for the Survey samples.

![Figure 9](image_url)
The improvement in purity is at most 0.01, while it tends to reduce completeness by 0.1 to 0.2. For example, when combining the action diamond selection with the $\sqrt{J_R - L_z}$ selection purity remains the same (compared to just the action diamond), and completeness falls from 0.39 to 0.33. One positive aspect of combining selections is it does tend to erase any disparity between the purity with and without disc samples. The story is the same for the Solar samples. With all of this information in mind, the best composite selection (for the Survey samples) appears to be the union of the action diamond and $e - L_z$ selections. When compared to just the action diamond selection, the purity without (with) disc contaminants is 0.86 (0.86) and the completeness remains constant at 0.39. However this selection is essentially the same as using the action diamond alone.

This analysis does have a few restrictions. First, we impose the use of ellipses because, as we argued above, they are simple selection shapes that can be confidently applied to real data. Second, we fix the location of the ellipses in each kinematic space. We also explored an alternative method for gauging purity-completeness relations for the Solar and Survey sample sets which does not assume a shape for the selection region, and instead works on a pixel-by-pixel basis. The advantage of this approach is that it examines purity and completeness in an unbiased manner. The weakness is that any selection boundaries based on this individual pixel approach are highly sensitive to the exact kinematic configuration of these data. With this in mind we would not be confident using this method to define selection boundaries for extracting high-$\beta$ stars from real data. The ellipse based method, while imperfect, is sufficiently general that it should work well on real data.

5.2 Eccentricity as a Single Metric for Separating the High-$\beta$ and Low-$\beta$ Halo

In the analysis presented in the previous section, we found that good purity of high-$\beta$ halo samples with respect to low-$\beta$ halo samples and disc samples can be achieved using eccentricity only if a very narrow region around $e = 1$ and $L_z = 0$ is selected. Outside of this region confusion between high-$\beta$ samples and these contaminants is high. The reasons for needing this very specific selection criterion is as follows: first, that eccentricity is a “scaled” quantity, meaning two physically distinct orbits that share a common ratio of apocenter to pericenter have the same eccentricity; second, that as orbits become increasingly radial, eccentricity approaches unity asymptotically. Indeed, examining the $e - L_z$ planes in Figures 3 and 4 shows that a reasonable fraction of the sample in the ergodic halo models lies at eccentricity near 1.

In the literature, eccentricity is commonly used as a metric to both describe (Myeong et al. 2019; Mackereth et al. 2019a), and separate (Mackereth & Bovy 2020; Naidu et al. 2020) low- and high-$\beta$ halo populations. Oftentimes eccentricity is used by itself, as a single parameter without the help of $L_z$ or another parameter that selects for tangential orbits. In order to investigate how well these sorts of selection criteria work, we plot the eccentricity distribution function for the halo samples of each of the six models in the Survey sample set in the top panel of Figure 11. We see, as expected, that the peak of the distribution moves towards larger eccentricities as $\beta$ increases from $-0.5$ to 0.9. Interestingly, we see that the ergodic and $\beta = -0.5$ models have a non-negligible number of samples with very high eccentricity. The ergodic model in particular has a nearly uniform distribution of samples between eccentricities 0.4 and 1, with a slight peak approaching eccentricity of 1. This suggests that while high eccentricity cuts do a good job of preferentially picking out high-$\beta$ stars, there is a significant amount of contamination by the well-populated high-eccentricity tail of the low-$\beta$ populations.

To quantify this confusion, we plot the completeness and purity of the high-$\beta$ subsets with respect to the low-$\beta$ subset of the fiducial Survey sample in the bottom panel of Figure 11. These quantities are expressed as a function of a cutoff eccentricity, below which we do not consider any samples. The curves are constructed in this way: For each cutoff eccentricity we take the samples of the fiducial model which have greater eccentricity and calculate purity as $N_{\beta=0.9}/(N_{\beta=0.9} + N_{\beta=0.5})$ and completeness as $N_{\beta=0.9}/N_{\beta=0.9, tot}$ where $N_{\beta=0.9, tot} = 50$ per cent of the total number of halo samples in the fiducial model, as per its definition. For smaller eccentricity cutoffs $\leq 0.8$ the purity and completeness vary in an inverse, yet linear fashion. For higher values of the cutoff eccentricity the purity rises sharply while the completeness drops sharply.

Notably, when the eccentricity cutoff is 0.7 the purity of the high-$\beta$ sample is only 0.59, with the completeness being 0.91. This is a remarkable amount of confusion between these two samples. Recall that the high-$\beta$ purity of the entire fiducial sample is 0.5 by definition, and an eccentricity threshold of 0.7 does little to improve that. When the eccentricity cutoff is increased to 0.9, the purity rises to 0.7 and the completeness is reduced to 0.79. Even at the largest value we consider for the eccentricity cutoff, 0.99, the purity only reaches 0.88 (roughly in accordance with our variation of ellipses analysis in the
previous section), and the completeness is reduced to 0.49. Note that while this purity and completeness appear superior to those of the action diamond we are wary of advocating for a selection boundary at extreme eccentricity lest it be too specific to our models, but it may be indicative of a good option for certain analyses. Nonetheless these results clearly indicate that care need be taken when separating high-$\beta$ and ergodic populations on the basis of eccentricity alone.

6 DISCUSSION

The data and results we present in this paper are clearly useful to discuss in the context of a large fraction of the literature on Milky Way stellar dynamics done in the Gaia era, specifically those which focus on the stellar halo. Here we focus on discussing the takeaways we deem are most salient, and how they relate to the wide body of literature.

6.1 The Validity and Applicability of our Models

First, we must assess how good the models we have presented are at representing the Milky Way halo as we understand it to be today. From a theoretical perspective, our models are clearly a simplification. For example, it is now well-known that the Milky Way stellar halo is effectively described by a metallicity-dependent triaxial power law density profile (e.g., Mackereth & Bovy 2020). Our models at least have a power law index and cutoff radius which match observations. But more complicated models are limited by the availability of usable DFs. Additionally, the Milky Way stellar halo is unlikely to currently be in a completely relaxed state, with perturbations from nearby massive dwarf galaxies in various stages of merging (Garrow et al. 2020), and specifically the Large Magellanic Cloud (Erkal et al. 2019). Fortunately for us these effects are greatest in the outer halo, with dynamical times in the inner halo where our sample is concentrated being short enough such that these perturbers do not meaningfully impact the equilibria.

A better method of addressing the validity of our models is to take an empirical approach and compare the kinematics of our samples with real observed stellar kinematics. For example, when considering the fiducial Survey model, the $v_R - v_T$ planes of Belokurov et al. (2018); Fattahi et al. (2019); Naidu et al. (2020), the Toomre diagram of Helmi et al. (2018); Koppelman et al. (2019b), the $E - L_z$ plane of Helmi et al. (2018); Koppelman et al. (2019b); Naidu et al. (2020), and the action diamond of (normalized to be a circle, Naidu et al. 2020) all show similarities with the samples produced by our models. Note that there can and should be disimilarities between these kinematics, which arises from the use of different gravitational potentials to calculate the kinematic properties, and varying underlying positional distributions. But overall, on the basis of kinematic properties, our samples are largely consistent with what is seen in real observational data.

Another detail we must address is how well these types of DFs represent the high-$\beta$ stellar halo associated with the debris of GE. This is important given that we spent a great deal of time in § 5 presenting regions of phase space that uniquely identified high-$\beta$ samples with high purity, with the implication that if we applied these to real data they would serve to identify GE stars. Generalized DFs corresponding to merger debris located in a stellar halo have not been formulated. This is mostly because the orbit and mass ratio of the merging galaxy would play a major role in determining the resulting DF. It is reasonable to assume, however, that the corresponding DF should be somewhere between that of a thin tidal stream occupying a distinct locus in action space (e.g. Bovy 2014), and the spherical DFs we have used here. It should also be the case that longer timescales and larger mass ratio mergers tend to make the DF of the debris resemble the latter, rather than the former. With this in mind we believe it reasonable to assume that the GE merger debris should resemble a roughly spherical, halo-like DF. This assertion is supported by a wide range of evidence, including that: the GE debris is distributed over the whole sky (see Figure 12 in Helmi 2020), the GE debris can be described by a triaxial power law density profile (Mackereth & Bovy 2020), and that measured stellar ages indicate GE was accreted roughly 10 Gyr ago (Montalbán et al. 2020), giving the debris ample time to relax.

Additionally, we can look to simulations to further validate our models, and specifically to support our use of high-$\beta$ halo DFs to model the GE remnant. Amorisco (2017) presents the kinematic properties of debris from a number of simulated mergers parameterized by the mass ratio and orbit circularity of the infalling satellite. They show that the deposited stellar debris has a power law density profile, with some appearance of a cutoff radius around a few halo scale radii, and the degree of radial bias is high. These facts together support the use of a spherical, radially biased halo DF to model the remnants. Jean-Baptiste et al. (2017) present a series of 1:10 merger simulations which show that the $E - L_z$ distribution of the deposited stars is broadly consistent with our high-$\beta$ halo models. They show cusps around $L_z = 0$ and are roughly bound by extended funnel shaped boundaries. Most of these mergers have a significant degree
6.2 Selection Criteria for Gaia-Enceladus

In § 5, we explored the feasibility of using kinematics alone to separate high-β populations from low-β and disc contaminants within the framework of our fiducial model. We found that most kinematic spaces had regions of high purity, and using well-informed selections, we were able to achieve simultaneously high purity (≥ 0.7) and relatively high completeness (≥ 0.2). Our main goal in providing these assessments of purity and completeness is to inform and enable identification of GE stars on the basis of kinematics alone. Chemical and stellar properties of GE constituent stars are of significant interest currently, as they tell us a great deal about the progenitor dwarf galaxy. Since it is not obvious a priori what these properties are in the context of the broader stellar halo, it is necessary to develop accurate methods to extract GE stars from larger samples of halo and disc stars. It is for this reason that throughout this paper, we have placed more emphasis on achieving high purity when it comes to separating low- and high-β populations, than on completeness. In general we recommend using the action diamond, which achieves a high purity. If completeness is of interest we recommend the e − 𝐿_2 plane, which achieves high completeness while keeping purity high.

The selection criteria we present are based upon a number of key assumptions. First, we assume that the ratio of high-β to low-β stars in the halo is 1:1. While it is clear that these components exist in roughly equal proportion near the solar neighbourhood (Belokurov et al. 2018; Lancaster et al. 2019; Iorio & Belokurov 2021), the exact ratio is uncertain, and is subject to variation with radius (Iorio & Belokurov 2021). We explore possible variations in this ratio by repeating our analysis with high- to low-β ratios of 3:7 and 7:3. We find that this variation does not significantly change where we would place the selection ellipses for any kinematic space. As one might expect, the purity of the high-β selection varies with the fraction of high-β stars. For example, the purity without disc contamination of our action diamond selection applied to the Survey sample drops to 0.70 when the high- to low-β ratio is 3:7, and increases to 0.93 when it is 7:3.

Another assumption built into our model is the use of the MWPotential2014 potential of Bowy (2014). Notably, this potential has light dark matter halo (virial mass of 8 × 10¹¹ M☉) when compared to other Milky Way potentials used in the literature. We repeat our analysis using the potential of McMillan (2017), which has a heavier dark matter halo (1.2 × 10¹² M☉, virial mass) and a slightly larger circular velocity at the location of the sun (~ 233 km s⁻¹). We find next to no differences in our findings using this potential, save for an overall shift in the magnitude of the energies (reflecting the heavier potential) and a barely perceptible increase in the extent of the distribution of the radially sensitive parameters ( v_资源共享, v_资源共享, energy, and √(R)). This suggests that the selection of high-β halo stars is agnostic to any choice of reasonable Milky Way potential.

We also emphasize that these selection are derived using our Survey sample set, which is based on the positions of APOGEE DR16 data. While we would not expect the distribution of kinematic properties to change drastically were it based on another survey, there would undoubtedly be noticeable changes. We attempt to gauge how resilient our choice of selections are to a change in the radial distribution of the data by examining the kinematics of samples with R < 7 kpc, 7 > R > 9 kpc and R > 9 kpc. In terms of the halo samples, we see clear trends regarding the location of the high-purity regions in each kinematic space. The regions tend to have a wider extent in L_2 with increasing radius, while the extent of radially sensitive quantities (specifically E, e, √(R), and (L_2 − J_R)/(J_资源共享)) increases with decreasing radius. The velocity spaces appear largely unchanged. While it does not appear that the ideal center of the selection ellipses should change with varying radius, their ideal extents should change in a manner reflecting these observed trends. With regards to the disc sources, we observe a slightly different trend. Samples at smaller radii have L_2 distributions which lie closer to the L_2 = 0 / v_T = 0 line where the high-purity halo sample region lies. This would allow the high-β selection ellipses to stretch towards lower values of the radially sensitive quantity in each kinematic space. The extent of the radially sensitive quantities for the disc samples also changes, but not in a manner that impacts the regions of high halo purity. We do note that the purity and completeness of many of the selections we have discussed in this work does appear to change by up to ~ 20 per cent, but we consider it outside the scope of this work to delve into the radial trends of the completeness and purity for each kinematic space and simply leave this as a cautionary note.

In § 3.3 we asked whether higher dimensional spaces would be useful for studying the types of kinematic populations we consider in this work. There, we argued that there are really three broad classes of kinematic quantities that we consider here: those sensitive to tangential motions (v_T and L_2), those sensitive to radial motions (v_R, v_资源共享, energy, and √(R)), and those sensitive to the shape of an orbit (eccentricity and (J_z − J_R)/J_资源共享). Within each of these categories the parameters tend to communicate similar sorts of information, meaning their combination would not be useful and only serve to complicate by increasing dimensionality. For the most part this assertion has been borne out by our results. When we tested combining selections for spaces from within these same categories we find that, in general, purity will remain unchanged, while completeness will decrease (likely owing to the increased dimensionality). For example, when combining the selections for the v_R − v_T and Toomre spaces the completeness and purity without (with) disc contamination are 0.12 and 0.76 (0.76) respectively. This is essentially identical to the selection which uses just v_T − v_T, suggesting that every point within the smaller (less complete) v_R − v_T selection also lies within the larger (more complete) Toomre selection; their combination adds no new information. As another example, when √(R) − L_2 selection is combined with the Toomre selection the purity is the same as if one just used the √(R) − L_2 selection yet the completeness falls to 0.13 from 0.52.

When combining the best quantities from each of the three categories: the action diamond and the √(R) − L_2 plane, the purity is the same as the combination of the action diamond with the e − L_2 plane (0.86), but the completeness falls to 0.33 (again, likely due to increased dimensionality). As mentioned above, we do find that these composite selections tend to erase any disparity between the purity with and without disc contamination. This improvement is marginal however, the purity including disc contamination of the action diamond, for example, only increases from 0.85 to 0.86 when combining it with √(R) − L_2. We do note that combining the action diamond and e − L_2 selections also increases the purity including disc contamination to 0.86, but keeps the completeness at 0.39. These results indicate that higher dimensional spaces do not appear to do a better job of separating high- and low-β halo populations. They do appear to help in separating halo and disc populations, however, yet the selections we create in the scaled kinematic spaces (action diamond and e − L_2) are relatively immune from disc contamination to begin with. This result makes sense when considering that disc and halo stars do genuinely have a large difference in the magnitude of their radial motions, whereas it is characteristically the shape of...
the orbit which distinguishes a star from a low- versus high-β population. In other words, there are very few stars in halo populations which have radial motions of a large magnitude, yet their orbit is not radial in shape, and so two dimensional spaces suffice to separate the two populations.

A final comment, we recognize that our choice of selection regions in each kinematic space is subjective, even though we have endeavoured to provide what we believe are the best choices. We are explicit though in that we favour selections which give a higher purity at a modest expense in completeness, and we also prefer selections which have minimal contamination from disc populations. Nonetheless, other selections which prioritize, for example, completeness over purity may be equally as valid or better for other use cases. In an effort to communicate this tradeoff we plot the purity and completeness as a function of ellipse scale parameters. Figure 12 shows the performance of the Survey sample action diamond selection as the selection ellipse semi-major axes are modified. Purity, purity including disc samples, and completeness are all shown. As the selection ellipse becomes larger the purity decreases while the completeness increases, as expected.

### 6.3 Comparison with Literature Selection Criteria

There already exist selection criteria used in the literature to define the boundaries of GE for chemical and stellar property assays. Here we compare our results with some of the more common selections. For reference, our benchmark selection is the union of our action diamond selection ellipse and \( e - L_z \) selection ellipse (although as we have mentioned this is very similar to the action diamond selection alone). For the Survey sample (the most apt sample set for comparison), the purity of this selection without (with) disc contamination is 0.86 (0.86) and the completeness is 0.39. In order to compare other selections in the fairest manner we evaluate purity and completeness for them using our fiducial Survey model. We will quote purity both with and without disc contaminants, but it is important to keep in mind that disc stars may also be efficiently removed from a sample using abundances.

First, we consider the selection criterion of Feuillet et al. (2020), which is also used in Matsuno et al. (2021), and is defined as \( \sqrt{R} > 30 \text{ kpc} \times \text{s}^{-1/2} \times \text{km}^{-1/2} \) and \( |L_z| < 500 \text{ kpc km s}^{-1} \). The completeness and purity without (with) disc contaminations of this selection are 0.31 and 0.72 (0.70). Second is the selection of Myeong et al. (2019), also used by Monty et al. (2020) and Cordoni et al. (2020), which is defined as \( (J_\phi/J_{\text{tot}}) < 0.07 \) and \( (J_z - J_R)/J_{\text{tot}} < -0.3 \). The completeness and purity without (with) disc contaminations of this selection are 0.58 and 0.82 (0.82). A final selection criterion to consider is that of Naidu et al. (2020), based off of Belokurov et al. (2018) and similar in spirit to Mackereth & Bovy (2020), which is defined as \( e > 0.7 \). We note specifically that Naidu et al. (2020) apply this criterion to their data only after removing a subset of the halo corresponding to the Sagittarius dwarf and the hot thick disc. The completeness and purity without (with) disc contaminations of this selection are 0.91 and 0.59 (0.37), as per § 5.2.

Our intent here is not to critique the selection criteria of these studies, but instead to provide an unbiased characterization of their performance in the context of our models. What we see from this comparison, and indeed our study in general, is that when one is kinematically separating high-β populations from ergodic populations there is a tradeoff between purity and completeness. If one’s interest is in maximal purity of a high-β population, then the selection criteria we have provided are best, specifically the union of our action diamond and \( e - L_z \) selections.

![Figure 12. Selection performance as a function of ellipse scale parameters for the Survey sample and action diamond space. The three panels show purity (top), purity including disc samples (middle), and completeness (bottom). The scaling factors modify the default action diamond selection ellipse for the Survey sample from Table 2.](image-url)

### 6.4 Structure Identification

One of the most remarkable emergent features of our Survey models was the appearance of artificial substructure in the distribution of halo samples in the \( E - L_z \) plane. The substructure is symmetric about the \( L_z = 0 \), and appears as a small overdensity of stars at low energies. The fact that this feature appears in our simple halo models defined by a single truncated power law density profile embedded in a spheroidalized Milky Way emphasizes that it is entirely spurious. We thoroughly explored this phenomenon in § 4.2, and found it to
be caused by a bimodality in the underlying radial distribution of APOGEE DR16 likely halo stars. Many stars are concentrated near the location of the Sun, about 8 kpc from the center, and many are concentrated also at about 3.5 kpc in radius. The cause of this bimodality is twofold. First, many pointings glance the Galactic bulge (see Figure 1), which means there is a concentration of halo stars just on its outskirts which share a galactocentric radius, but have differing galactocentric azimuth. Second, pointings towards the Galactic bulge contribute the most halo stars right on the outskirts of the bulge, as opposed to near the Sun, since observed volume scales with the square of the distance from us, and the density of halo stars should be centrally concentrated.

This is a particularly pertinent result given that the $E - L_z$ plane is often used as a means of structure identification (e.g. Helmi et al. 2018; Koppelman et al. 2019b; Horta et al. 2021), and therefore we strongly recommend that authors consider the selection biases present in their samples when attempting to verify the nature of newly discovered substructures. This is especially important for any survey with a non-trivial spatial selection function, which includes essentially all modern spectroscopic surveys. Furthermore, while the density fluctuations we observe are symmetric about the $L_z = 0$ axis, we have the clarity of being able to separate halo stars from disc stars. This is not a luxury afforded to most samples, which will also contain contamination from disc stars. These contaminants would potentially hide the symmetric nature of structure that arises thanks to the types of radial selection biases we explore here. This could result in, for example, the density fluctuations being observed in the lower density retrograde half of the plane but not the prograde half where they could be masked by disc contaminants.

We specifically highlight the case of the newly discovered Heracles structure in the inner Galactic halo by Horta et al. (2021) for two reasons. First, they also use the APOGEE DR16 data that underpins our models. Second, the structure they identify shares many features in common with the artifacts we find. This is to say that Heracles is roughly symmetric about the $L_z = 0$ axis, and it appears as a separate overdensity of stars at lower energies than the primary group of halo stars. We suggest that the structure be further investigated to confirm that it does not appear as a result of the APOGEE selection function.

In a similar vein, a potentially useful application of the sorts of models we have presented here, is to use them for the differential identification of substructures in both the disc and stellar halo. Because our models use simple, smooth underlying potentials, they can be used to generate phase space realizations of those smooth potentials which can be compared with observations. The advantage of this approach is that it would be relatively simple to account for the types of selection effects we just highlighted, because phase space need only be sampled at the locations where the real data are found. A major challenge of this approach is determining the impact of assuming the nature of the underlying potential, but this could be circumvented to a degree by comparing many different potentials.

7 CONCLUSION

In this work, we have demonstrated a framework for assessing the kinematic properties of the types of composite stellar populations which are observed by modern spectroscopic surveys complemented by Gaia kinematics. Our methodology relies on a simple, DF-based approach to generate mock kinematics for both solar-vicinity and APOGEE DR16 data using galpy, and would be simple to replicate for other surveys. We therefore suggest that these sorts of models be created for many different surveys in order that their unique kinematic distributions and underlying selection effects can be properly gauged.

We provide a general overview of many kinematic spaces commonly used in the study of the Milky Way, and specifically the stellar halo. In each kinematic space, we show two sets of models: the Solar models which are simple and mimic stellar samples concentrated near the Sun, and the Survey models which are more complicated and designed to mimic samples observed with spectroscopic surveys. Each model uses a stellar halo parameterized by the anisotropy, $β$, which is a useful quantity for describing the two dominant stellar halo populations in the Milky Way: the low-$β$ metal poor halo and the high-$β$ GE remnant.

In each kinematic space we use our models to show where disc populations, low- and high-$β$ halo populations all lie. We use purity as a metric to specifically assess each space on how well it separates low-$β$ halo populations from high-$β$ populations, finding that the scaled action diamond, the $\sqrt{L_R} - L_z$ plane, and $e - L_z$ plane achieve highest performance, with a focus on purity. These results assume a 1:1 ratio of low- and high-$β$ populations in the stellar halo, and also use APOGEE DR16 data as a set of positions for the phase space samples. Despite these constraints we believe our choice of selection regions in each kinematic space is resilient to changes in these and other underlying assumptions. While the location of the best selection regions is resilient to changes in assumptions, the specific purity and completeness values are liable to change. We therefore recommend that other authors generate their own completeness and purity corrections for their data if the specific values are needed. We briefly explored the effects of survey selection functions on kinematic properties of observed samples, finding that they can result in the appearance of specious structures.

We summarize this work with a number of recommendations for authors to consider as they explore current and future datasets:

- The scaled action diamond, $e - L_z$ plane, and $\sqrt{L_R} - L_z$ plane are best for separating high- and low-$β$ populations. We provide a number of simple-to-implement selection criteria (see Table 2) for separating high- and low-$β$ populations in each kinematic space we study.
  - If high purity is most important when separating low- and high-$β$ populations, then we recommend using our action diamond selection. It achieves a purity without (with) disc contamination of 0.86 (0.85) and completeness of 0.39. Combining the action diamond selection with the $e - L_z$ selection increases the purity with disc contamination to 0.86, while keeping the completeness constant at 0.39.
  - If high completeness is most important when separating low- and high-$β$ populations, then we recommend using our $e - L_z$ selection. It achieves a purity without (with) disc contamination of 0.82 (0.82) and completeness of 0.61. We also show the purity and completeness of a wide range of selection ellipse sizes for the action diamond, to illustrate how completeness may be increased for that selection given a tradeoff in purity.
  - If invariance to choice of Milky Way potential is most important when separating low- and high-$β$ populations, then we recommend using the $v_R - v_T$ plane, which outperforms the Toomre diagram.
  - In general, the $E - L_z$ plane and Toomre diagram do not do as good a job of separating low- and high-$β$ populations. Additionally, we recommend that caution be exercised using eccentricity by itself for this purpose. Commonly used eccentricity cuts for the radially anisotropic halo, such as $e > 0.7$ only result in a sample with purity of 0.59. Only extreme eccentricity cuts ($e > 0.95$) achieve good purity.
  - Be aware of all selection biases in the underlying sample, which
can cause spurious structures to appear in kinematic spaces with explicit dependencies on radius, specifically in the $E - L_z$ plane which could be easily mistaken for real substructure.

ACKNOWLEDGEMENTS

JMM and JB acknowledge financial support from NSERC (funding reference number RGPIN-2020-04712) and an Ontario Early Researcher Award (ER16-12-061). This work has made use of data from the European Space Agency (ESA) mission *Gaia* ([https://www.cosmos.esa.int/gaia](https://www.cosmos.esa.int/gaia)), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, [https://www.cosmos.esa.int/web/gaia/dpac/consortium](https://www.cosmos.esa.int/web/gaia/dpac/consortium)). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is [www.sdss.org](http://www.sdss.org).

DATA AVAILABILITY

The APOGEE DR16 data used in this article are available at: [https://www.sdss.org/dr16](https://www.sdss.org/dr16). The *Gaia* data used in this article are available at: [https://gea.esac.esa.int/archive/](https://gea.esac.esa.int/archive/).

REFERENCES

Ahumada R., et al., 2020, ApJS, 249, 3
Amorisco N. C., 2017, MNRAS, 464, 2882
An J. H., Evans N. W., 2006, AJ, 131, 782
Bae M., Dejonghe H., 2002, A&A, 393, 485
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Belokurov V., Sanders J. L., Fattahi A., Smith M. C., Deason A. J., Evans N. W., Grand R. J. J., 2020, MNRAS, 494, 3880
Bennett M., Bovy J., 2019, MNRAS, 482, 1417
Binney J., 2010, MNRAS, 401, 2318
Binney J., 2012, MNRAS, 426, 1324
Binney J., McMillan P., 2011, MNRAS, 413, 1889
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press, Princeton, NJ
Bovy J., 2014, ApJ, 795, 95
Bovy J., 2015, ApJS, 216, 29
Bovy J., Rix H.-W., Liu C., Hogg D. W., Beers T. C., Lee Y. S., 2012, ApJ, 753, 148
Bradbury J., et al., 2018, JAX: composable transformations of Python+NumPy programs, [http://github.com/google/jax](http://github.com/google/jax)
Buder S., et al., 2020, arXiv e-prints, p. arXiv:2011.02505
Conroy C., et al., 2019, ApJ, 883, 107
Cordoni G., et al., 2020, MNRAS,
Cuddeford P., 1991a, MNRAS, 253, 414
Cuddeford P., 1991b, MNRAS, 253, 414
Cui X.-Q., et al., 2012, Research in Astronomy and Astrophysics, 12, 1197
Dalton G., et al., 2016, in Evans C. J., Simard L., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI. p. 99081G, doi:10.1117/12.2231078
Deason A. J., Belokurov V., Evans N. W., 2011, MNRAS, 416, 2903
Deason A. J., Belokurov V., Sanders J. L., 2019, MNRAS, 490, 3426
Dejonghe H., de Zeeuw T., 1988, ApJ, 329, 720
Di Matteo P., Haywood M., Lehnert M. D., Katz D., Khoperskov S., Snaith O. N., Gómez A., Robichon N., 2019, A&A, 632, A4
Eddington A. S., 1916, MNRAS, 76, 572
Erkal D., et al., 2019, MNRAS, 487, 2685
Evans N. W., An J. H., 2006, Phys. Rev. D, 73, 023524
Fattahi A., et al., 2019, MNRAS, 484, 4471
Feuillet D. K., Feltzing S., Sahhholdt C. L., Casagrande L., 2020, MNRAS, 497, 109
Forbes D. A., 2020, MNRAS, 493, 847
Gaia Collaboration et al., 2016, A&A, 595, A1
Gaia Collaboration et al., 2018, A&A, 616, A1
Gaia Collaboration Brown A. G. A., Vallenari A., Prusti T., de Bruijne J. H. J., Babusiaux C., Biemann M., 2020, arXiv e-prints, p. arXiv:2012.01533
Garro W., Webb J. J., Bovy J., 2020, MNRAS, 499, 804
Gilmore G., et al., 2012, The Messenger, 147, 25
Gravity Collaboration et al., 2019, A&A, 625, L10
Harris W. E., 1996, AJ, 112, 1487
Hawkins K., Jofré P., Masseron T., Gilmore G., 2015, MNRAS, 453, 758
Haywood M., Di Matteo P., Lehnert M. D., Snaith O., Khoperskov S., Gómez A., 2018, ApJ, 863, 113
Helmi A., 2020, ARA&A, 58, 205
Helmi A., White S. D. M., de Zeeuw P. T., Zhao H., 1999, Nature, 402, 53
Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
Hernquist L., 1990, ApJ, 356, 359
Horta D., et al., 2021, MNRAS, 500, 1385
Iorio G., Belokurov V., 2021, MNRAS, 502, 5686
Iorio G., Belokurov V., Erkal D., Koposov S. E., Nipoti C., Fraternali F., 2018, MNRAS, 474, 2142
Jean-Baptiste T., Di Matteo P., Haywood M., Gómez A., Montuori M., Combes F., Semelin B., 2017, A&A, 604, A106
King I. R., 1966, AJ, 71, 64
Kollmeier J. A., et al., 2017, arXiv e-prints, p. arXiv:1711.03234
Koppelman H., Helmi A., Veljanoski J., 2018, ApJ, 860, L11
Koppelman H. H., Helmi A., Massari D., Roelenga S., Bastian U., 2019a, A&A, 625, A5
Koppelman H. H., Helmi A., Massari D., Price-Whelan A. M., Starkenburg T. K., 2019b, A&A, 631, L9
Kruijssen J. M. D., et al., 2020, MNRAS, 498, 2472
Lancaster L., Koposov S. E., Belokurov V., Evans N. W., Deason A. J., 2019, MNRAS, 486, 378
Leung H. W., Bovy J., 2019a, MNRAS, 483, 3255
Leung H. W., Bovy J., 2019b, MNRAS, 489, 2079
Levi M., et al., 2019, in Bulletin of the American Astronomical Society, p. 57 (arXiv:1907.16868)
Lian J., et al., 2020, MNRAS, 494, 2561
Mackert H. T., et al., 2019a, MNRAS, 482, 3426
Mackert H. T., et al., 2019b, MNRAS, 489, 176
Majewski S. R., et al., 2019, in American Astronomical Society Meeting, p. 286 (arXiv:1907.10688)
Martell S., Venn K. A., Lane J. M. M., Lokhorst D., Yong D., 2020, MNRAS, 497, 1236
Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235
Naidu R. P., Conroy C., Bonaca A., Johnson B. D., Ting Y.-S., Caldwell N., Zaritsky D., Cargile P. A., 2020, ApJ, 901, 48
Naidu R. P., et al., 2021, arXiv e-prints, p. arXiv:2103.03251
APPENDIX A: ERGODIC AND ANISOTROPIC DISTRIBUTION FUNCTIONS FOR DARK-MATTER HALOS AND THEIR STELLAR HALOS IN galpy

To model the kinematics of the stellar halo and spherical galactic components more generally, we have implemented various spherical distribution functions in the galpy code. In this Appendix, we provide a brief overview of the implemented functionality and its mathematical basis that serves as a reference for the code. We consider distribution functions \( f(E, L) \) embedded in a spherical potential \( \Phi(r) \) that are functions of the specific energy \( E \) and the total specific angular momentum \( L \) and that have either constant orbital anisotropy \( \beta \) or anisotropy of the Osipkov-Merritt type. Such models are standard textbook material (e.g., Binney & Tremaine 2008), but for completeness we review the main theory behind them here.

A1 Spherical coordinates

We work in spherical coordinates \((r, \phi, \theta)\), with \(\theta\) the polar angle measured from the pole and \(\phi\) the azimuthal angle. In terms of the usual cartesian coordinates \((x, y, z)\), the spherical coordinates are given by

\[
\begin{align*}
    r &= \sqrt{x^2 + y^2 + z^2}, \\
    \phi &= \arctan(y/x), \\
    \theta &= \pi/2 - \arctan(z/\sqrt{x^2 + y^2}),
\end{align*}
\]

and the inverse transformation is

\[
\begin{align*}
    x &= r \sin \theta \cos \phi, \\
    y &= r \sin \theta \sin \phi, \\
    z &= r \cos \theta.
\end{align*}
\]

In the spherical coordinate frame, the velocities \( (v_r, v_\phi, v_\theta) \) are given in terms of the cartesian velocities \( (v_x, v_y, v_z) \) as

\[
\begin{align*}
    v_r &= v_x \sin \theta \cos \phi + v_y \sin \theta \sin \phi + v_z \cos \theta, \\
    v_\phi &= -v_x \sin \phi + v_y \cos \phi, \\
    v_\theta &= v_x \sin \theta \cos \phi + v_y \sin \theta \sin \phi - v_z \cos \theta, \\
\end{align*}
\]

(A3)

with inverse transformation

\[
\begin{align*}
    v_x &= v_r \sin \theta \cos \phi - v_\phi \sin \phi - v_\theta \cos \theta \cos \phi, \\
    v_y &= v_r \sin \theta \sin \phi + v_\phi \cos \phi + v_\theta \cos \theta \sin \phi, \\
    v_z &= v_r \cos \theta + v_\theta \sin \theta. \\
\end{align*}
\]

(A4)

When working with the velocities, we will also express them using a spherical coordinate system \((r, \psi, \eta)\) of their own, writing the velocities as

\[
\begin{align*}
    v_r &= r \cos \eta, \\
    v_\psi &= r \sin \eta \cos \psi, \\
    v_\eta &= r \sin \eta \sin \psi.
\end{align*}
\]

(A5)

In these coordinates, the specific energy is

\[
E = \Phi(r) + \frac{v^2}{2},
\]

(A6)

while the total specific angular momentum is given by

\[
L = r v \sin \eta.
\]

(A7)

A distribution function \( f(E, L) \) is therefore only a function of \((r, v, \eta)\) and does not depend on the two spatial angles \((\phi, \theta)\) or the velocity angle \(\psi\).

The orbital anisotropy \(\beta\) for our DFs is defined as

\[
\beta = 1 - \frac{\sigma_{r}^2 + \sigma_{\phi}^2}{2\sigma_{\psi}^2} = 1 - \frac{\sigma_{r}^2}{2\sigma_{\psi}^2},
\]

(A8)

where \(\sigma_{r,\phi,\psi}\) are the velocity dispersions in the radial, azimuthal, and polar directions and all quantities are measured at a given radius \(r\); the second equality uses the velocity dispersion in the tangential velocity \(v_\psi = r \sin \eta\).

A2 Distribution functions

As is usual, we work in terms of the relative potential \(\Psi = -\Phi + \Phi(\infty)\) and the specific binding energy \(E = -\Phi + \Phi(\infty) = \Psi - \frac{1}{2} v^2\). Bound orbits then have \(0 < E < E_{\text{max}}\), where \(E_{\text{max}}\) is the maximum binding energy given by \(E_{\text{max}} = \Psi(0) - \Phi(0) + \Phi(\infty)\).

The first type of spherical DF that we consider are the ergodic DFs

\[
f(E, L) \equiv f(E) .
\]

(A9)

Because \(E\) does not depend on \(\sin \eta\) (Equation A6), such DFs have \(\beta = 0\) and are therefore isotropic. For a given density \(\rho(r)\) embedded in a spherical potential \(\Psi\), the DF is given by the Eddington inversion (Eddington 1916)

\[
f(E) = \frac{1}{\sqrt{8\pi}^2} \frac{d}{dE} \int_0^E d\Psi \, \frac{1}{\sqrt{\Psi - E}} \frac{d\rho}{d\Psi},
\]

(A10)

or, equivalently,

\[
f(E) = \frac{1}{\sqrt{8\pi}^2} \left[ \int_0^E d\Psi \, \frac{1}{\sqrt{\Psi - E}} \frac{d^2\rho}{d\Psi^2} + \frac{1}{\sqrt{\Psi}} \frac{d\rho}{d\Psi} \right]_{\Psi=0},
\]

(A11)

Note that the density \(\rho(r)\) does not have to be related to the potential

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Necib L., Lisanti M., Belokurov V., 2019, ApJ, 874, 3
Niederste-Ostholt M., Belokurov V., Evans N. W., Peñarrubia J., 2010, ApJ, 712, 516
Nissen P. E., Schuster W. J., 2010, A&A, 511, L10
Osipkov M. P., 1979, Soviet Astronomy Letters, 5, 42
Posti L., Binney J., Nipoti C., Ciotti L., 2015, MNRAS, 447, 3060
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2007, Numerical Recipes 3rd Edition: The Art of Scientific Computing, 3 edn. Cambridge University Press, USA
Sanders J., 2012, MNRAS, 426, 128
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Steinmetz M., et al., 2020, AJ, 160, 82
Trick W. H., Coronado J., Rix H.-W., 2019, MNRAS, 484, 3291
Vasiliev E., 2019, MNRAS, 484, 2832
Widrow L. M., 2000, ApJS, 131, 19
Williams A. A., Evans N. W., 2015, MNRAS, 448, 1360
Xue X.-X., Rix H.-W., Ma Z., Morrison H., Bovy J., Sesar B., Janesh W., 2015, ApJ, 809, 144
Yanny B., et al., 2009, AJ, 137, 4377
Yuan Z., Chang J., Beers T. C., Huang Y., 2020, ApJ, 898, L37
Zasowski G., et al., 2017, AJ, 154, 198
de Jong R. S., et al., 2019, The Messenger, 175, 3
van der Marel R. P., 2006, in Livio M., Brown T. M., eds, Vol. 17, The Local Group as an Astrophysical Laboratory. pp 47–71 (arXiv:astro-ph/0404192)
The energy density can be obtained from DFs of the form 

$$f(E, L) \equiv L^{-2\beta} f_1(E).$$  

(A12)

As shown by Cuddeford (1991b) (see also An & Evans 2006 for a clearer statement of this result), for a given density $\rho(r)$ embedded in a gravitational potential $\Psi$, the DF is given by the inversion

$$f(E, L) = \frac{2^\beta L^{-2\beta}}{(2\pi)^{3/2} \Gamma(1-\alpha) \Gamma(1-\beta)} \times \left[ \int_0^\infty \frac{d\Psi}{(E - \Psi)^{\alpha}} \frac{d^{m+1}}{d\Psi^{m+1}} \left( \rho(\Psi) [r(\Psi)]^{2\beta} \right) \right] \left( \frac{1}{E^\alpha} \right),$$  

(A13)

where $m = [3/2 - \beta]$ and $\alpha = 3/2 - \beta - m$. For half-integer values of $\beta$, the DF is simply given by

$$f(E, L) = \left( \frac{-2\beta}{2\pi^2} \right) \frac{d^{3/2-\beta}}{d\Psi^{3/2-\beta}} \left( \rho(\Psi) [r(\Psi)]^{2\beta} \right), \quad 1/2 - \beta \in \mathbb{N}. $$  

(A14)

where $(-2\beta)! = 2^{1/2-\beta} \Gamma(1-\beta)/\sqrt{\pi}$.

As first proposed by Osipkov (1979) and Merritt (1985), DFs for a given density $\rho$ embedded in a gravitational potential $\Psi$ with radially-varying anisotropy of the form

$$\beta(r) = \frac{1}{1 + r^2_\alpha/r^2},$$  

(A15)

can be obtained from DFs of the form

$$f(E, L) = f(Q),$$  

(A16)

where

$$Q = E - \frac{L^2}{r_\alpha^2},$$  

(A17)

with $r_\alpha$ the anisotropy radius (the radius where $\beta = 1/2$). For a given density $\rho$ embedded in a potential $\Psi$, this DF is given by an Eddington-like inversion

$$f(Q) = \frac{1}{\sqrt{8} \pi^2} \int_0^Q d\Psi \frac{1}{\sqrt{Q - \Psi}} \frac{d^2}{d\Psi^2} \left[ \rho(\Psi) \left( 1 + r^2_\alpha/[r^2] \right) \right] + \frac{1}{\sqrt{Q}} \frac{d}{d\Psi} \left[ \rho(\Psi) \left( 1 + r^2_\alpha/[r^2] \right) \right] \bigg|_{\Psi = 0}. $$  

(A18)

This can also be written as

$$f(Q) = f^{\beta=0}(Q) + \frac{1}{\sqrt{8} \pi^2 r_\alpha^2} \left( \int_0^Q d\Psi \frac{1}{\sqrt{Q - \Psi}} \frac{d^2}{d\Psi^2} \left[ \rho(\Psi) r^2_\alpha/(\Psi) \right] \right) + \frac{1}{\sqrt{Q}} \frac{d}{d\Psi} \left[ \rho(\Psi) r^2_\alpha/(\Psi) \right] \bigg|_{\Psi = 0}. $$  

(A19)

where $f^{\beta=0}(Q)$ is the ergodic DF from Equation (A11) evaluated at $E = Q$. Because the square brackets in the second term do not depend on $r_\alpha$, Osipkov-Merritt DFs for a given $(\rho, \Psi)$ pair for all values of $r_\alpha$ can be efficiently computed from two one-dimensional quadratures that do not depend on $r_\alpha$.

### A3 Numerical implementation

All of the DFs given in the previous subsection can be computed through one-dimensional quadratures over the relative potential $\Psi$. The only exception to this are the half-integer-$\beta$ constant-anisotropy DFs from Equation (A14), which do not require a numerical quadrature and are therefore straightforward to compute. The numerical integrals are difficult to compute for two reasons: (i) the integral is over $\Psi$, but the density is expressed as a function of $r$ and the integrand is therefore most easily evaluated as a function of $r$, and (ii) the integrand diverges integrably at the upper limit. To deal with difficulty (i), we rewrite all integrals to be over $r$ rather than $\Psi$; this requires the Jacobian $d\Psi/dr$, which is simply equal to the gravitational field. We also need to transform the limits of the integration: the lower limit of $\Psi = 0$ is the maximum radius $r_{\max}$ of the system and is typically infinity, the upper limit $\Psi = \bar{E}$ is transformed to $r$ by

**Figure A1.** Ergodic distribution function $f(E)$ for a power-law density profile $\rho(r) = \rho_{0,0}(r/a)^{-\alpha}$ embedded in an NFW potential $\Phi(r) = -4\pi G \rho_0 a^2 \ln(1 + r/a)/(r/a)$ in terms of the dimensionless binding energy $E/(4\pi G \rho_0 a^2)$.

**Figure A2.** Energy part $f_1(E)$ of the self-consistent distribution function for an NFW profile with constant anisotropy $\beta$ for a dense grid in $\beta$. At $\beta > 0.5$, the distribution function becomes increasingly unphysical.
interpolating the function \( r(\Psi) \). Because the integral after this transformation is well-behaved, the numerical error is easily dominated by how well this interpolation works and we therefore use a fine, logarithmically-spaced grid of \( r \) values to determine \( r(\Psi) \). After the transformation, the integral extends from \( r(E) \) to \( r = r_{\text{max}} \) and when \( r_{\text{max}} = \infty \), this means that special handling of the integral is required at both limits: dealing with the integrable divergence at one end and an infinite limit at the other end. To deal with this, we split the integral at twice the lower limit and then use standard transformation methods to deal with the divergence and the infinite limit (e.g., Press et al. 2007).

The implementation of the Eddington inversion in Equation (A10) requires the second derivative of the density with respect to \( \Psi \), while the constant-anisotropy DFs require derivatives of \( \rho r^2 \beta \) and the Osipkov-Merritt DFs need derivatives of \( \rho r^\beta \). To compute these, we first of all rewrite them all as derivatives with respect to \( r \) using the chain rule; this introduces derivatives of the relative potential that can be expressed in terms of the gravitational field and higher order derivatives of \( \Psi \). For ergodic and Osipkov-Merritt DFs we only need at most two derivatives of either \( \rho \) or \( \Psi \) and to compute these, we explicitly add the first and second derivative of the density to all \textsc{galpy} spherical potentials; second derivatives of the potential already existed. Depending on how low \( \beta \) goes, we need arbitrary derivatives of \( \rho \) and \( \Psi \) with respect to \( r \). Rather than computing these by hand, we use \textsc{jax} (Bradbury et al. 2018), a Python library that implements automatic differentiation of arbitrary functions built using \textsc{numpy} functions and more and that as such can compute arbitrary derivatives of functions to machine precision. To use \textsc{jax} in our case, we implement \( d(\rho r^{2\beta})/dr \) as well as the radial field \( d\Psi/dr \) using \textsc{jax}'s version of \textsc{numpy}. Note that for many potentials, the density derivative or the radial field does not involve special functions and it is then not actually necessary to load \textsc{jax}'s \textsc{numpy} implementation. The \textsc{jax} derivatives are also used to compute the necessary derivatives in the half-integer case of Equation (A14).

As an example of our numerical implementation, we display the ergodic DF \( f(E) \) calculated using Equation (A10) for a power-law density profile \( \rho(r) = \rho_{\text{PL},0}(r/a)^{-\alpha} \) that is embedded in an NFW potential with scale parameter \( a \)

\[
\Phi(r) = -\frac{4\pi G \rho_0 a^2}{\ln(1 + r/a)} \ln(1 + r/a)/r/a \tag{A20}
\]

in Figure A1 for different values of the power-law exponent \( \alpha \). As expected, more centrally-concentrated profiles with higher values of \( \alpha \) have more stars on tightly-bound orbits. All of these DFs are physical, that is, do not contain negative regions.

A further example of our numerical implementation is given in Figure A2, which shows the \( f_1(E) \) part of the self-consistent DF for an NFW profile with constant anisotropy \( \beta \) for a dense grid in \( \beta \). The smoothness of the variation of the DF’s shape with \( \beta \) provides a good indication of the high numerical fidelity of our implementation. At \( \beta > 0.5 \), an increasingly large region at high binding energy is negative, indicating that there is no physical DF with such high constant anisotropy of the form of Equation (A12).

### A4 Sampling the distribution functions

To be able to sample six-dimensional phase-space coordinates for a wide range of spherical DFs, we have implemented exact random sampling for all of the DFs described in this Appendix. For all DFs, we sample positions \((r, \phi, \theta)\) by sampling \( r \) from the cumulative mass profile, using inverse transform sampling, and sampling random angles on the sphere \((\theta, \phi)\) keeping in mind that while the distribution of \( \phi \) is uniform on \( \phi \in [0, 2\pi] \), the distribution of \( \theta \) is \( p(\theta) \propto \sin \theta \) on \( \theta \in [0, \pi] \). Here and below, when the inverse cumulative distribution used in inverse transform sampling cannot be computed analytically, we calculate it using interpolating the inverse transform computed on a grid.

Sampling velocities necessarily depends on the anisotropy profile of the DF. For all DFs, we sample velocities in the spherical velocity coordinate system \((v, \psi, \eta)\) from Equation (A5). For DFs that only depend on \( E \) and \( L \), the distribution of the angle \( \phi \) is uniform on \( \phi \in [0, 2\pi] \). For ergodic DFs, the distribution of \( \eta \) is \( \rho(\eta) = \sin \eta \), which can be sampled exactly using inverse-transform sampling using the fact that the cumulative DF can be inverted analytically. For DFs with constant \( \beta \), we have that \( \rho(\eta) \propto \sin^{1-2\beta} \eta \), which can also be sampled using inverse transform sampling (although while the cumulative DF can be expressed using special functions, it cannot be inverted analytically). We then sample the magnitude of the velocity from \( p(v | r) \propto v^2 f(E) \) or \( p(v | r) \propto v^2 f_1(E) \) for non-zero, constant \( \beta \) using inverse transform sampling between zero and \( v = v_{\text{esc}}(r) \).

Exactly sampling Osipkov-Merritt-type DFs is slightly more difficult and has to our knowledge not been discussed before. We first sample the velocity angle \( \eta \) from \( p(\eta | r) \propto \sin \eta \left[ 1 + (r/a)^2 \sin^2 \eta \right]^{-3/2} \) using inverse transform sampling. We then proceed to sample \( v' = v \sqrt{1 + (r/a)^2 \sin^2 \eta} \) from \( p(v' | r) \propto v^2 f(Q) \). Remarkably, the cumulative distribution of \( \cos \eta \) can be analytically inverted, so sampling it using inverse transform sampling is straightforward.

In practice, we interpolate the inverse cumulative distributions for \( v \) or \( v' \) in both the quantile and \( r \) using two-dimensional interpolation. After this and other interpolations are set up, sampling the DFs is fast, achieving rates of \( \approx 10^8 \) points per second.

### A5 Special cases and new ergodic and Osipkov-Merritt NFW distribution functions

Many analytic solutions for ergodic and anisotropic spherical DFs exist. Of these we have currently implemented:

- The ergodic DF of a Hernquist profile (Hernquist 1990);
- The Osipkov-Merritt-type DF of a Hernquist profile (Hernquist 1990);
- The constant \( \beta \) DF for a Hernquist profile for arbitrary \( \beta \) (Baes & Dejonghe 2002);
- The ergodic DF of a Plummer profile \( f(E) \propto E^{-7/2} \) (e.g., Binney & Tremaine 2008);
- The King (1966) DF (see also Michie 1963).

Because NFW profiles describe the spatial distribution of simulated dark-matter halos well, self-consistent NFW DFs are of special interest. However, there is no known analytic form for the ergodic or general anisotropic DFs (however, see Evans & An 2006 for a notable exception, the DF with constant \( \beta = 1/2 \)). Approximate ergodic and Osipkov-Merritt-type DFs for the NFW profile were presented by Widrow (2000). Using the high-fidelity numerical calculation obtained using our code of the ergodic and Osipkov-Merritt-type NFW DFs, we have found more complex, but more accurate fitting func-


\[ f(Q) = f^{B=0}(Q) \times (4\pi G)^{3/2} a^3 \rho_0^{1/2} \]

\[
= \left( \frac{a}{r_a} \right)^2 \frac{1}{Q^{2/3}} \left( \frac{1 - Q}{\ln Q} \right)^2 \times \left\{ -0.995895790138335 Q^8 + 4.29052661245253 Q^7 \\
- 7.606946709185 Q^6 + 7.03132348658788 Q^5 \\
- 3.6920719890718 Q^4 + 0.831302363461598 Q^3 \\
- 0.217968733177408 Q^2 - 0.0408426627412238 Q \\
+ 0.0802975743915827 \right\},
\]

(A22)

where \( Q \) is evaluated in its dimensionless form \( Q/(4\pi G \rho_0 a^2) \) and \( f^{B=0}(Q) \) is the ergodic DF from Equation (A21) evaluated at \( E = Q \). Unlike the forms given by Widrow (2000), our new form allows for an arbitrary value of \( a/r_a \). How well the approximate DF recovers the density and radial velocity dispersion profile depends on \( a/r_a \), but the density is typically recovered to within a few percent out to 100\( r_a \) and the velocity dispersion to better than one percent.

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.