The Astrophysical Journal, 926:188 (15pp), 2022 February 20 https://doi.org/10.3847/1538-4357/ac4243

Substructure at High Speed. I. Inferring the Escape Velocity in the Presence of Kinematic Substructure

Lina Necib[1,2,3,4,5] and Tongyan Lin[6]
[1] Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA; lnecib@mit.edu
[2] Center for Cosmology, Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA
[3] Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA
[4] Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[5] The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[6] Department of Physics, University of California San Diego, La Jolla, CA 92093, USA
Received 2021 February 23; revised 2021 December 10; accepted 2021 December 11; published 2022 February 24

Abstract

The local escape velocity provides valuable inputs to the mass profile of the galaxy, and requires understanding the tail of the stellar speed distribution. Following Leonard & Tremaine, various works have since modeled the tail of the stellar speed distribution as $\propto(v_{\rm esc}-v)^k$, where $v_{\rm esc}$ is the escape velocity, and k is the slope of the distribution. In such studies, however, these two parameters were found to be largely degenerate and often a narrow prior is imposed on k in order to constrain $v_{\rm esc}$. Furthermore, the validity of the power-law form can breakdown in the presence of multiple kinematic substructures or other mis-modeled features in the data. In this paper, we introduce a strategy that for the first time takes into account the presence of kinematic substructure. We model the tail of the velocity distribution as a sum of multiple power laws as a way of introducing a more flexible fitting framework. Using mock data and data from FIRE simulations of Milky Way-like galaxies, we show the robustness of this method in the presence of kinematic structure that is similar to the recently discovered Gaia Sausage. In a companion paper, we present the new measurement of the escape velocity and subsequently the mass of the Milky Way using Gaia eDR3 data.

Unified Astronomy Thesaurus concepts: Milky Way dynamics (1051); Galaxy stellar halos (598)

1. Introduction

Evidence of the theory of hierarchical galaxy formation (White & Rees 1978) has been abundant in recent years. The Gaia mission (Lindegren et al. 2016; Gaia Collaboration et al. 2018, 2020) in particular has found evidence of multiple mergers in the Milky Way (see e.g., Helmi 2020 for a review). The Milky Way, and in particular the stellar halo, is a graveyard of disrupted substructures such as streams (Helmi et al. 1999; Belokurov et al. 2006, 2007; Naidu et al. 2020), clumps (Diemand et al. 2008; Myeong et al. 2018a), tidally disrupted dwarf galaxies (e.g., Belokurov et al. 2006; Zucker et al. 2006; Niederste-Ostholt et al. 2009; Kirby et al. 2013; Collins et al. 2017; Simon et al. 2017), and debris flow (Belokurov et al. 2018; Helmi et al. 2018; Deason et al. 2018; Lancaster et al. 2019; Myeong et al. 2018b; Necib et al. 2019a). The Gaia mission (Lindegren et al. 2016; Gaia Collaboration et al. 2018) has shed light on some of these substructures, and in particular led to the identification of a large debris flow called the Gaia Sausage[7] (Belokurov et al. 2018), or Gaia Enceladus (Helmi et al. 2018). Such a structure extends to $\sim$30 kpc, including stars on highly eccentric orbits. It is most likely the product of a merging satellite of a stellar mass $10^{8-9}M_{\odot}$ that was disrupted at about redshift $z\sim1$–3 (Deason et al. 2018; Lancaster et al. 2019; Myeong et al. 2018b).

In light of these findings, we must revisit the methods built for investigating properties of the Milky Way, and specifically in inferring the local escape velocity. Determining the escape velocity is important as it is used to (1) constrain the total mass of the Milky Way, (2) predict signals for dark matter (DM) direct detection, (3) and build the DM density profile of the Milky Way. For instance, measurements of the escape velocity and the circular velocity can be used to constrain the potential of the Milky Way, assuming some spatial distribution of the disk, bulge, and DM. This has been done extensively in the literature (e.g., Smith et al. 2007; Piffl et al. 2014; Williams et al. 2017; Monari et al. 2018; Deason et al. 2019) in order to obtain a measurement of the Milky Way mass. However, the presence of many velocity substructures can affect our measurement of the escape velocity and thus the Milky Way mass. In this work, we aim to build a robust strategy for determining the escape velocity accounting for such substructure. In particular, we build a pipeline that incorporates, for the first time, multiple components in modeling the tail of the stellar velocity distribution.

The majority of previous studies of the escape velocity are based on Leonard & Tremaine (1990), which model the tail of the stellar velocity distribution as

$$g(v|v_{\rm esc},k)\propto(v_{\rm esc}-v)^k\Theta(v_{\rm esc}-v)\qquad v>v_{\rm min},\qquad(1)$$

where v is the speed in Galactocentric coordinates and the two fitting parameters are the escape velocity $v_{\rm esc}$ and the slope of the distribution k. This model is applied to stars with speeds greater than the threshold $v_{\rm min}$, and it is assumed that the approximation holds for $v_{\rm min}$ well above the local rotation speed.

Following Leonard & Tremaine (1990), studies have inferred the local escape velocity using line-of-sight velocities with the Radial Velocity Experiment (Smith et al. 2007; Piffl et al. 2014)

[7] In the remainder of this paper, we will refer to this substructure as the Sausage.

and Sloan Digital Sky Survey (Williams et al. 2017), and then using 3D velocities from Gaia (Monari et al. 2018; Deason et al. 2019). In all of these studies, there is a large degeneracy between v_{esc} and k, as will be discussed further below. The degeneracy leads to rather large error bars on v_{esc} and subsequently large error bars on the estimated mass of the Milky Way.

In order to overcome these large error bars, many of these works argued for narrow priors on the slope k; the arguments for small values of k were violent relaxation or collisional relaxation, both leading to $k \leqslant 2$ (Leonard & Tremaine 1990). Meanwhile, Piffl et al. (2014) (and subsequently Monari et al. 2018) used cosmological simulations based on the Aquarius suite (Springel et al. 2008; Scannapieco et al. 2009) to argue for a prior $k \epsilon$ [2.3, 3.7], while Deason et al. (2019) used the Auriga simulation (Grand et al. 2017) to argue that for mergers resembling the Sausage, k should be small, and therefore $k \epsilon$ [1, 2.5]. More recently, Koppelman & Helmi (2021) used a much larger sample of stars with only proper motion measurements to reduce the degeneracy. However, using simulations, Koppelman & Helmi (2021) estimated a possible 10% bias due to the fact that the tail of the distribution is not necessarily populated all the way up to v_{esc}.

These studies illustrate some of the difficulties in using Equation (1) to model the tail. The degeneracy in the parameters v_{esc} and k is due to the fact that a higher v_{esc} can be partially compensated by a higher slope k in the shape of the distribution. Because there are very few stars near v_{esc}, a fit to Equation (1) could then easily lead to biased results if the model is not a good description of the data over the entire range of speeds. This could be the result of additional unbound stars, a mismodeled or unmodeled component, measurements with large errors that contaminate the data set, or selection effects in data sample. A second related issue is that there is no precise definition of the *tail* where the model is expected to be a good description. For example, Grand et al. (2019) studied numerical simulations and found that those distributions deviate from Equation (1) due to the presence of substructure, often leading to underestimates of the Milky Way mass.

In this paper, we argue for an approach that can more robustly determine the tail of the stellar speed distribution. This can take into account the presence of kinematic substructure, which we know to be present, as well as additional effects as described above. Given what is known about the Sausage, it is likely that a large fraction of stars in the tail of the distribution can be attributed to substructure; as argued by Deason et al. (2019), it will have a different slope k than the rest of the stellar halo. Including substructure is thus physically motivated. The tail of the distribution is then the sum of (at least) two distributions and might not be well described by a single power law for low v_{min}. Not including substructure could then bias v_{esc} measurements.

To address these points, we build a pipeline where we add additional components to the velocity distribution, also modeled as in Equation (1) but with a new slope k_S. While this is motivated by substructure, it can also be viewed as a more flexible model for the steeply falling speed distribution. We show how v_{esc} can be obtained more robustly by performing tests on the data as a function of v_{min} and the number of bound components. For instance, it is expected that a single component will be adequate for large enough v_{min}. Performing these tests can ensure that the fit for v_{esc} is not biased by structure in the speed distribution at lower speeds. In all of these tests, it is important that we keep the priors for all parameters as wide as possible so we are not artificially shaping the results.

In this work, we present the pipeline and perform analyses with mock data sets containing kinematic substructure. We also apply our pipeline to a Milky Way realization from the LATTE suite Wetzel et al. (2016) based on the FIRE code (Hopkins 2015; Hopkins et al. 2018). This paper is organized as follows. In Section 2, we first discuss in more depth the substructure motivation for including additional velocity components and illustrate the main points. The details of the pipeline are provided in Section 3. We test the pipeline on mock data sets containing substructure in Section 4 and on the Milky Way simulations in Section 5. We compare results when one or more bound components are used in the fit. In a companion paper (Necib & Lin 2022), we apply this method on Gaia eDR3 data for stars in the local neighborhood ([7,9] kpc in Galactocentric distance), to present the most robust estimate of the local escape velocity, from which we deduce the mass of the Milky Way.

2. Motivation from Substructure

In this section, we discuss two broad motivations for including additional velocity components in modeling the tail of the velocity distribution. First, kinematic substructure—the Sausage—is present in the Milky Way (Belokurov et al. 2018; Helmi et al. 2018) and can comprise a large fraction of stars (Necib et al. 2019a). However, based on empirical studies of the kinematic properties of the Sausage from Necib et al. (2019a), it is not obvious what the substructure slope k and fractional contribution to the tail of the velocity distribution should be. Rather than using simulations as a prior on the slope, we prefer to obtain independent information about the kinematic substructure from the data.

Second, not accounting for this substructure can lead to biases in Milky Way mass estimates (Grand et al. 2019). In particular, the choice of a low $v_{min} = 300$ km s^{-1} is common in the literature as it increases the statistics. For such low v_{min} compared to an expected escape velocity $v_{esc} \sim 500$ km s^{-1}, there may be contributions from multiple kinematic structures, including, for example, the Sausage. Not accounting for these structures can then pull the fit toward larger v_{esc}, depending on v_{min}. By performing a two-component fit over different v_{min}, we can demonstrate the robustness of the posterior distributions on v_{esc} and hence on Milky Way mass estimates.

2.1. The Presence of Substructure in the Tail

The Milky Way recently underwent a major merger, the Sausage (Belokurov et al. 2018; Helmi et al. 2018), which was discovered through its distinct chemical and phase space properties. It was shown that the merger contributes about ~60%–72% of the non-disk stars in the local neighborhood (Necib et al. 2019a, 2020), which means that it would be expected to strongly shape the tail of the stellar speed distribution.

To illustrate this, in Figure 1 we plot the speed distributions of the stellar halo and the Sausage (Necib et al. 2019a), normalized by their relative fractions, where we assume that the Sausage comprises 70% of the total distribution.[8] These are the best-fit distributions that have been built by modeling Galactocentric velocities and metallicity measurements from a cross match of

[8] https://linoush.github.io/DM_Velocity_Distribution/

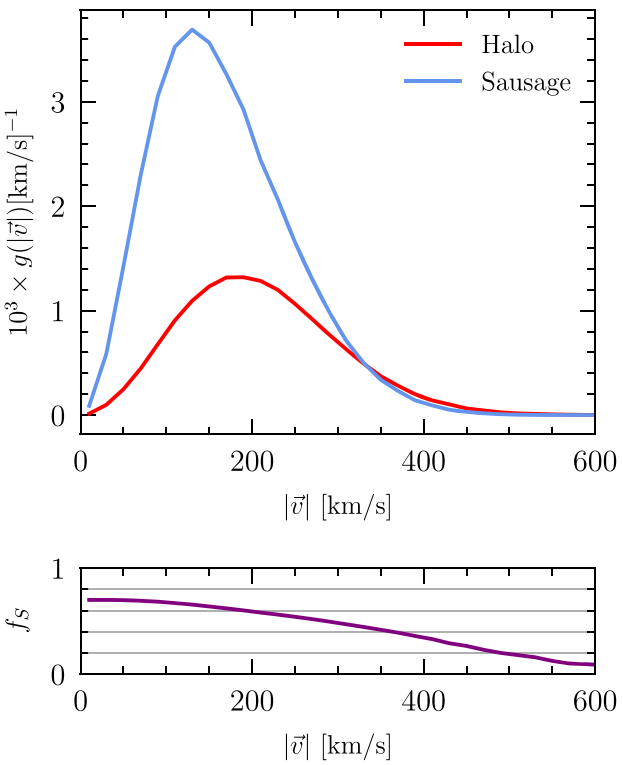

Figure 1. Speed distributions of the halo and Sausage components from Necib et al. (2019a), assuming a total Sausage fraction of 70% of the sum of the distributions. (Top) Full speed distributions. (Bottom) Fractional contribution of the Sausage distribution for all speeds above $|\vec{v}|$ (see Equation (2)).

Gaia DR2 and the Sloan Digital Sky Survey (Ahn et al. 2012) using a Gaussian mixture model with halo, Sausage, and disk components. While the halo and disk were modeled as three-dimensional Gaussian distributions in spherical Galactocentric coordinates, the Sausage was modeled as a sum of two Gaussians with opposite-sign means and equal dispersions in v_r, and single Gaussians in v_θ and v_ϕ.

In Figure 1, we plot the full speed distributions of the halo and Sausage components. In the bottom panel, we plot the integrated ratio of the Sausage distribution to that of the sum of the halo and Sausage distributions:

$$f_S(|\vec{v}|) = \frac{\int_{|\vec{v}|}^{\infty} \boldsymbol{g_S}(v')dv'}{\int_{|\vec{v}|}^{\infty} \boldsymbol{g_S}(v')dv' + \int_{|\vec{v}|}^{\infty} \boldsymbol{g_H}(v')dv'}, \quad (2)$$

where $\boldsymbol{g_S}$ and $\boldsymbol{g_H}$ are the speed distributions of the Sausage and the halo, respectively. The Sausage component peaks at lower velocities than the halo component, and is a smaller fraction of the high-speed stars. This is suggestive of a two-component component where the slope of the Sausage is larger[9] than that of of the halo ($k_S > k$), corresponding to a sharper drop at higher velocities. Of course, the distributions in Figure 1 are the result of a fit to a Gaussian mixture model; they therefore are not tuned for accuracy of the tails, nor do they truncate at $v_{\rm esc}$. Nevertheless, based on this, one can see that for the typical $v_{\rm min} = 300\ \mathrm{km\ s^{-1}}$ used in analyses, both components will contribute an $O(1)$ fraction of the stars.

[9] Using some intuition, $k \to 0$ is a Heaviside function that is truncated at $v = v_{\rm esc}$, while $k \to \infty$ is a sharply falling function.

Because the contribution of the Sausage for a given $v_{\rm min}$ is a priori unknown but likely sizable, single-component fits to the tail of the velocity distribution might fail to describe the data accurately and bias results. Based on studies of the Sausage so far, there is not enough information to set a strong prior on the slope or fractional contribution of the Sausage. Deason et al. (2019) adopts a prior on $k \in [1, 2.5]$ based on simulations where the tail of the velocity distribution is dominated by a substructure like the Sausage, which might not be true in the case of the Milky Way. Instead, we find the opposite behavior in Figure 1. Meanwhile, other works have argued for and used different priors based on simulations, with Smith et al. (2007) using $k \in [2.7, 4.7]$ and Piffl et al. (2014) arguing that one expects $k \in [2.3, 3.7]$. The choice of prior thus largely depends on the merger history of the simulations considered. For a fit with degenerate parameters, the results are then strongly molded by the priors, and could lead to incorrect results, as we discuss in Section 5. In this work, we prefer to remain agnostic as to the interpretation of each velocity component. Instead, we will show how including additional components in the fit allows an independent way to measure properties of the substructure.

2.2. *Influence of Substructure on Single-component Fits to the Tail*

Having argued that there could be a large substructure component, we now show how results could be affected if this substructure is not included in the model. While we defer the detailed discussion of our pipeline to Section 3, here we show the results of some analyses on mock data to illustrate the main ideas.

We generate two sets of mock data, one with a single bound distribution drawn from Equation (1) and another containing two velocity distributions with different slopes k but a common $v_{\rm esc}$. For the single bound distribution, we assume $k = 2.5$ and $v_{\rm esc} = 500\ \mathrm{km\ s^{-1}}$. For the mock data set with substructure, we assume a substructure fraction of $f_S = 0.6$ and substructure slope $k_S = 1$, while the other component has $k = 3.5$. In both cases, we also include an unbound outlier population, which is a fraction $f = 0.01$ of the total stars and described by a Gaussian with dispersion $\sigma_{\rm out} = 1000\ \mathrm{km\ s^{-1}}$. To simulate a realistic data set, we take $v_{\rm min} = 300\ \mathrm{km\ s^{-1}}$ and smear the true speeds of the stars with a random Gaussian error of 5%, which will be further discussed later.

In Figure 2, we show the result of fitting both these data sets to a model with a single bound component. The model also includes the outlier component and accounts for the errors on the velocity measurements, as will be described in the next section. When the mock data contains a single component (left), we find that the fit actually underpredicts $v_{\rm esc}$ somewhat, although this is only one mock data realization.

However, when the mock data contains two components (right panel of Figure 2), we find two important impacts on the fit for $v_{\rm esc}$: the best-fit value of $v_{\rm esc}$ is biased higher, and the degeneracy between k and $v_{\rm esc}$ is larger, with larger error bars on $v_{\rm esc}$ and k. To understand this behavior, we first observe that the best-fit slope k is in between the values of the individual slopes of the two components; the best-fit value is $k = 2.32^{+0.29}_{-0.24}$, while the true values are $k_S = 1$ and $k = 3.5$. This is because the tail of the distribution will be described by the component with lower slope (here $k = 1$), while the stars at lower speeds will drive the preferred slope to higher values. Because of the correlation between the effects of increasing k

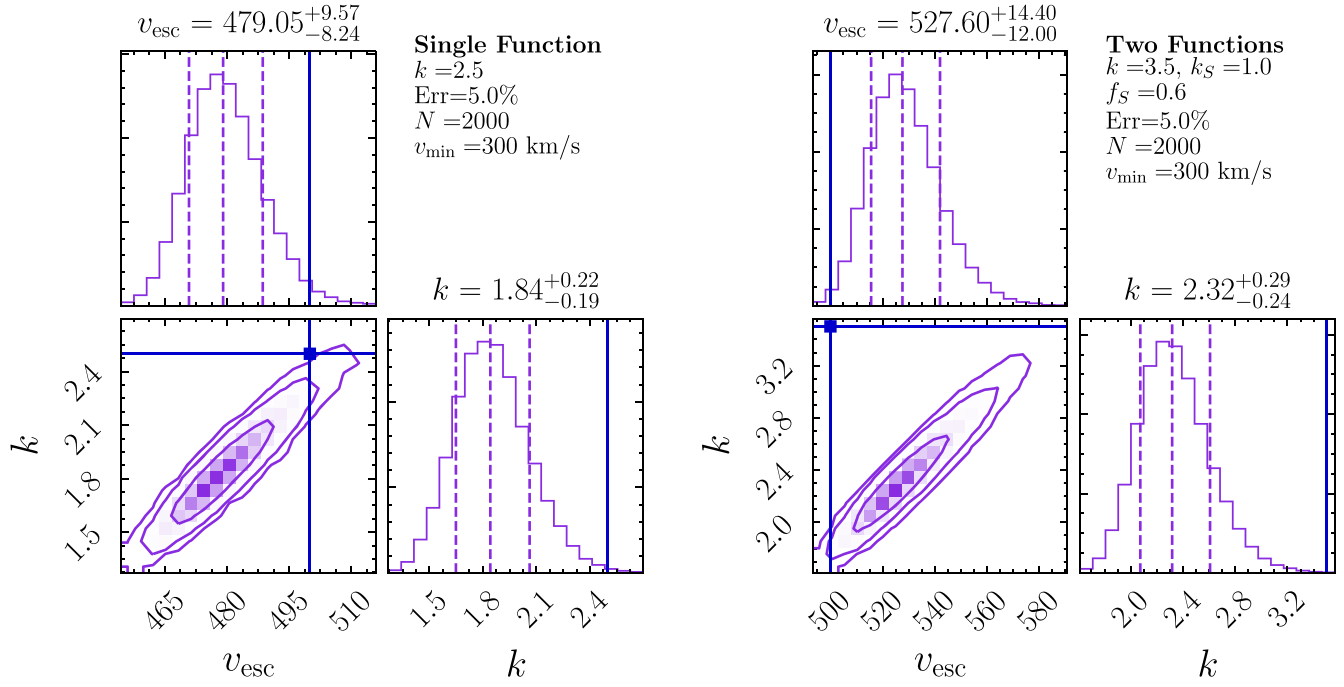

Figure 2. Corner plots of the fit to a single bound component. The 2D contours are the 68%, 95%, and 99% containment regions. The dashed lines in the 1D posteriors are the median and 1σ containment regions. (Left) We use mock data that includes a single component with true values $v_{esc} = 500$ km s^{-1} and $k = 2.5$, which are indicated by the blue lines. The stars were smeared with a random Gaussian error of 5%. (Right) We use mock data that contains two bound components, as described in more detail in the text. In this case, there is a larger degeneracy between k and v_{esc} compared to the single-component data and v_{esc} is biased toward larger values.

and v_{esc}, the escape velocity will also be driven to larger values to better fit the tail.

As discussed in Section 1, other works have set strong priors on the slope k in order to reduce the degeneracy seen between k and v_{esc}. In general, these priors strongly shaped the posterior distribution, with the slope parameter k piling up at the edge of the priors. This can potentially lead to incorrect results, especially if the underlying distribution is not well described by a single power law. Indeed, Deason et al. (2019) found that the choice of the priors affects the end result of the Milky Way mass. The example in Figure 2 shows that a degeneracy in k and v_{esc} could be partially due to the presence of multiple stellar components. As we will show in detail throughout this work, including multiple components properly leads again to a robust fit. The goal of this paper is thus to build a robust method that can account for kinematic substructure and other mismodeled features.

3. Analysis

3.1. Multicomponent Pipeline

We now present the likelihood function for an unbinned analysis on a sample of stars with minimum observed velocity v_{min}. This analysis will involve either a single bound component, or two bound components, where by *bound* component we mean a stellar velocity distribution following Equation (1).

The true velocity distribution is smeared out by the measurement error, modeled by a one-dimensional Gaussian. We define for each star α the probability to observe it at velocity v_{obs}

$$p_\alpha(v_{obs}|v) = \frac{1}{\sqrt{2\pi\sigma_{v,\alpha}^2}} \exp\left(-\frac{(v - v_{obs})^2}{2\sigma_{v,\alpha}^2}\right), \quad (3)$$

where v is the true velocity of the star, and $\sigma_{v,\alpha}$ is the observed measurement error. Then the likelihood for star α to be drawn from the distribution defined by Equation (1) is given by

$$\tilde{p}_\alpha(v_{obs}|v_{esc}, k) = C_\alpha(k, v_{esc}) \times \int_0^\infty dv\ (v_{esc} - v)^k\, p_\alpha(v_{obs}|v)\Theta(v_{esc} - v), \quad (4)$$

where the lower limit of the integration region is 0 to account for stars with *true* speed below v_{min}. The factor $C_\alpha(k, v_{esc})$ leads to a normalized PDF in the *data* region $[v_{min}, \infty]$, meaning

$$\int_{v_{min}}^\infty dv_{obs}\, \tilde{p}_\alpha(v_{obs}|v_{esc}, k) = 1. \quad (5)$$

Note that the power law in Equation (1) should approximately describe stars even with velocities below v_{min}, since some of these stars may be observed above v_{min} due to measurement error. The normalization factor is

$$C_\alpha(k, v_{esc}) = 2\left[\int_0^{v_{esc}} dv (v_{esc} - v)^k \times \left(1 + \mathrm{erf}\left(\frac{v - v_{min}}{\sqrt{2}\,\sigma_{v,\alpha}}\right)\right)\right]^{-1}. \quad (6)$$

Studies such as those of Piffl et al. (2014) and Deason et al. (2019) used bootstrapping methods to take into account the error distributions. They resampled the stars within their error bars to quantify the errors on the final values of the escape velocity. In this paper, we account for individual errors on all stars in the likelihood function. Although our method is more computationally intensive, we forward model all errors to obtain posterior distributions. A similar treatment of the errors was used by Koppelman & Helmi (2021) (with the difference that ours includes an outlier model and a second component).

Beyond the bound component, we also expect a small fraction of the stars to be either ejected or on unbounded orbits (e.g., Hattori et al. 2018). To account for such stars, we use an outlier model similar to that of Williams et al. (2017), where

$$p_{\alpha}^{\rm out}(v_{\rm obs}) = \mathcal{A}\exp\left(-\frac{v_{\rm obs}^2}{2[\sigma_{\rm out}^2+\sigma_{v,\alpha}^2]}\right). \quad (7)$$

Unlike Williams et al. (2017), which fixed the value of $\sigma_{\rm out} = 1000\,{\rm km\ s^{-1}}$, we marginalize over the dispersion $\sigma_{\rm out}$ of the outlier model as well as its fraction f. We also add in quadrature the measurement error of a particular star, although we expect it to be subdominant to $\sigma_{\rm out}$. We then normalize Equation (7) over the data region $[v_{\rm min}, \infty]$, and obtain

$$\mathcal{A}^{-1} = \sqrt{\frac{\pi}{2}}\sqrt{\sigma_{\rm out}^2+\sigma_{v,\alpha}^2}\ {\rm erfc}\left(\frac{v_{\rm min}\sqrt{\sigma_{\rm out}^2+\sigma_{v,\alpha}^2}}{\sqrt{2}}\right). \quad (8)$$

The likelihood per star α for a single bound component is therefore

$$\mathcal{L}_{\alpha}^{1} = (1-f)\tilde{p}_{\alpha}(v_{\rm obs}^{\alpha}|v_{\rm esc}, k) + f p_{\rm out}(v_{\rm obs}^{\alpha}|\sigma_{\rm out}), \quad (9)$$

while for two bound components it is

$$\mathcal{L}_{\alpha}^{2} = (1-f)[f_S\tilde{p}_{\alpha}(v_{\rm obs}^{\alpha}|v_{\rm esc}, k_S) + (1-f_S)\tilde{p}_{\alpha}(v_{\rm obs}^{\alpha}|v_{\rm esc}, k)\,] + f p_{\alpha}^{\rm out}(v_{\rm obs}^{\alpha}|\sigma_{\rm out}), \quad (10)$$

where the slopes of the components are k and k_S, and the fraction of the second component is labeled as f_S. This can be generalized to n components. The total log likelihood is given by

$$\log\mathcal{L}^{i} = \sum_{\alpha}\log\mathcal{L}_{\alpha}^{i}, \quad (11)$$

with $i = \{1, 2\}$ the number of bound components assumed in the analysis. In what follows, we will refer to the two analyses as the *single-component* and *two-component* fits, by which we are discussing the bound components. In this work, we adopt the same terminology as in Necib et al. (2019a), where we call the relaxed component the halo, and the kinematic substructure the Sausage. We emphasize that we do not know a priori which component corresponds to which value of k.

We use the Markov Chain Monte Carlo *emcee* (Foreman-Mackey et al. 2013) to find the best-fit parameters, using 200 walkers, 500 steps for the burn-in stage, and 2000 steps for each run. The parameters are initialized to random small deviations from the following values: the escape velocity $v_{\rm esc} = 550\,{\rm km\ s^{-1}}$, the slope $k = 2.2$, the fraction $f = \log(10^{-5})$, the outlier dispersion $\sigma_{\rm out} = \log(800\,{\rm kms^{-1}})$, the Sausage's slope $k_S = 0.7$, and the Sausage's fraction $f_S = 0.8$. We next describe the parameters and priors.

3.2. Priors

The parameters of the single-component fit are the escape velocity $v_{\rm esc}$, the slope k, the fraction of the outlier distribution f, and the dispersion of the outliers $\sigma_{\rm out}$. For the two-component likelihood function, Equation (10), we add the slope of the second component, k_S, and its relative fraction with respect to the bound components, f_S. Without loss of generality, we assume $k_S < k$, but we remain agnostic as to the physical interpretation of each component.

We list these parameters in Table 1 along with their priors. The prior on the escape velocity is taken to be $1/v_{\rm esc}^2$, corresponding to a uniform prior in $1/v_{\rm esc}$. Note that our choice is slightly unconventional, and most other authors such as Deason et al. (2019) have taken a prior that is uniform in $\log v_{\rm esc}$. However, given the narrow posteriors we will obtain, this choice will not impact results significantly. The fraction and dispersion of the outlier distribution are taken to have uniform prior in $\log f$ and $\log\sigma_{\rm out}$, respectively, given the large ranges that they might span. The fraction of the second component f_S and slopes k, k_S are all taken to have uniform prior in their respective ranges. The default prior range on the slopes is taken to be very wide,[10] where $k, k_S \in [0.1, 15]$, as our goal to avoid shaping the distribution with restrictive priors.

Table 1
List of the Priors Used in the Analysis

Parameter	Prior Range	Prior
$v_{\rm esc}$	$[v_{\rm min}, 1000]\ {\rm km\ s^{-1}}$	$1/v_{\rm esc}^2$
k	[0.1, 15]	1
f	$[10^{-6}, 1]$	$1/f$
$\sigma_{\rm out}$	$[3, 3000]\ {\rm km\ s^{-1}}$	$1/\sigma_{\rm out}$
k_S	$[0.1, k]$	1
f_S	[0, 1]	1

3.3. Akaike Information Criterion

We will run the pipeline of Section 3.1 with a single-component model, as is standard in the literature, as well as the two-component model. In order to compare fits with the two models, we compute the Akaike information criterion (AIC) of each distribution, where the AIC is defined as (Akaike 1974)

$${\rm AIC} = 2s - 2\log(\hat{\mathcal{L}}), \quad (12)$$

where s is the number of parameters of the fit, and $\log(\hat{\mathcal{L}})$ is the maximum log likelihood of the fit. We compare the AIC of the single- and two-component fits to the data, where the one with the lower value of AIC is preferred. Alternative functions, for example, the Bayes information criterion (BIC) can also be used (see, e.g., Wit et al. 2012), but the AIC provided the most robust results when applied to simulations. The difference is in the way that it penalizes the added number of parameters, where AIC penalizes the models as $2s$, while BIC penalizes them as $s\log(n)$, with n the total number of data points in the set (Vrieze 2012).

In this analysis, we will be evaluating

$$\Delta{\rm AIC} = {\rm AIC}_2 - {\rm AIC}_1, \quad (13)$$

where ${\rm AIC}_i$ is the AIC of the single ($i = 1$) or double ($i = 2$) component fit.

4. Results with Mock Data

In this section, we present fit results analyzing mock data that includes a halo component, a component due to a Sausage-like merger, and an outlier distribution. The bound components are drawn from true distributions following Equation (1). We use this analysis to validate our pipeline and explore how well the true $v_{\rm esc}$ can be recovered in a fit, depending on different choices for $v_{\rm min}$ and on the number of bound components in the fit.

Throughout this section, we work with fiducial samples of 2000 mock stars. The number of stars was chosen to be

[10] We have also verified that our results are unchanged with the prior range $k, k_S \in [0.1, 20]$

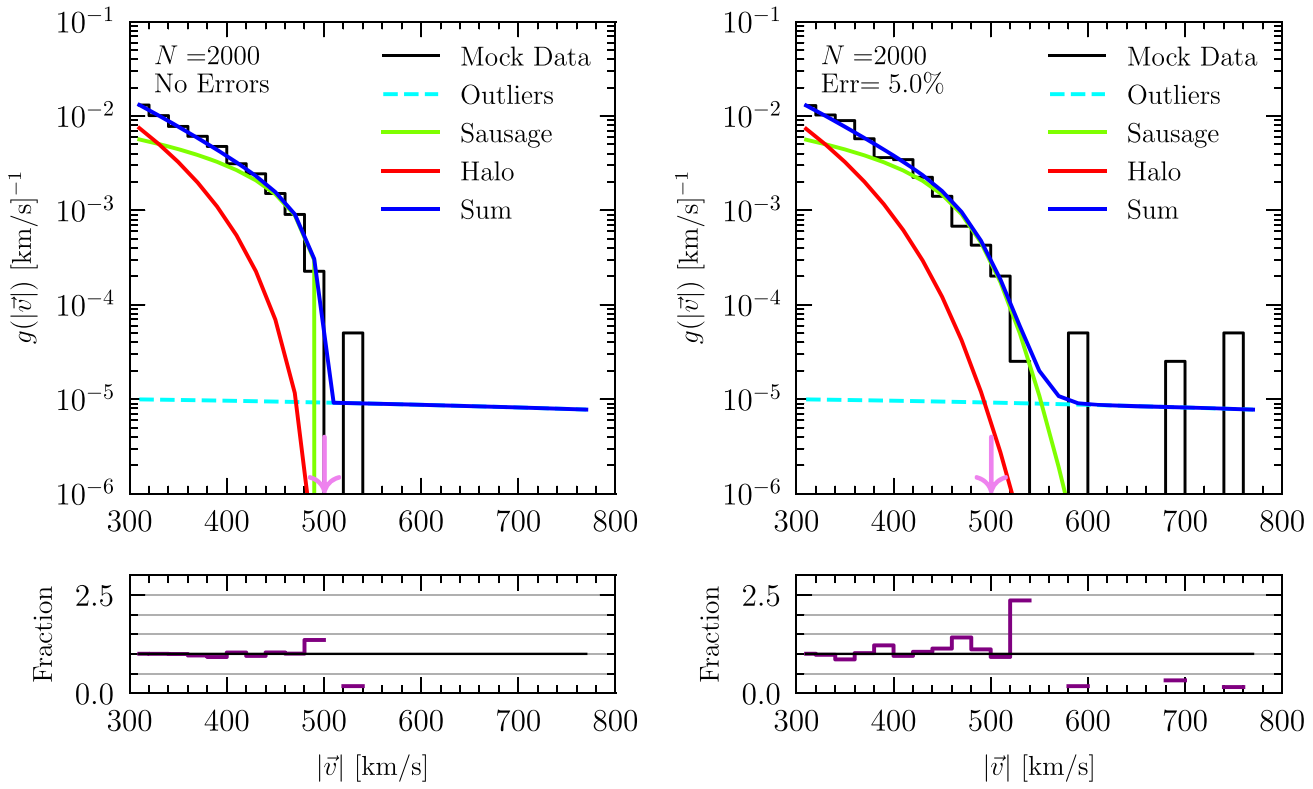

Figure 3. (Top) Mock data drawn from a mixture of two bound components with $v_{\rm esc} = 500$ km s^{-1}, and an outlier component with a fraction of 0.01 and a dispersion $\sigma_{\rm out} = 1000$ km s^{-1}. The halo component has $k = 3.5$ and the substructure fraction is $f_S = 0.6$, with $k_S = 1.0$. The colored curves are the true distributions. (Left) We assume no measurement errors. (Right) We assume a measurement error of 5% on the speeds. The true escape velocity is shown as a pink arrow. (Bottom) Ratio of the generated distribution to the true distribution. The generated stars follow the true distributions (accounting for errors), with fluctuations at high speeds due to the small outlier fraction. The missing bins are due to the lack of data in these bins.

comparable to that found in the Gaia data sample, and we assume $v_{\rm esc} = 500$ km s^{-1}, with the slopes $k = 3.5$ and $k_S = 1.0$. The fraction attributed to the Sausage is $f_S = 0.6$ for $v_{\rm min} = 300$ km s^{-1}. The outliers are sampled from a Gaussian distribution with zero mean, a dispersion $\sigma_{\rm out} = 1000$ km s^{-1}, and an associated fraction $f = 0.01$. From each fiducial sample, we generate a data set accounting for possible measurement errors, where we assume the measurement errors on stars are a percentage of their true velocity. More explicitly, for each star with a true speed $|\vec{v}|$, we sample each star's observed speed from a Gaussian distribution with a mean given by the true $|\vec{v}|$ and a dispersion $x \times |\vec{v}|$. Here, x is the percentage error, and we will consider as a representative value of $x = 5\%$. This case is closest to the Gaia data, discussed in more detail in Necib & Lin (2022).

We show a fiducial sample of stars in Figure 3, where in the left panel there are no measurement errors and in the right panel we include percentage errors of 5%. The pink arrow shows the true escape velocity at 500 km s^{-1}, but in the right panel, the tail of the distribution extends out beyond 500 km s^{-1} due to the errors. It is thus imperative that the likelihood takes into account the presence of such errors. The true distributions (solid curves) in the right panel are different from those in the left panel because we plot Equation (4) instead of Equation (1) in order to account for the presence of errors.

From Figure 3, we can immediately see that if the minimum velocity of the data sample $v_{\rm min}$ is too low, we might see more than a single distribution in the fit. The total distribution (blue) is dominated by the distribution with the lower k for high enough speeds ($|\vec{v}| \gtrsim 400$ km s^{-1}). Below these values, the presence of the second distribution starts to dominate and will affect the fit. Using a single distribution would not produce the correct fit and slope, as we will explicitly show. Nevertheless, in order to have sufficient statistics and to obtain a reasonable fit to the distribution, previous works[11] have used rather low values of $v_{\rm min} = 250$, 300 km s^{-1} and obtained a local escape velocity $v_{\rm esc} \sim 520$–580 km s^{-1} (Monari et al. 2018; Deason et al. 2019). It is not known a priori where the cut should be, such that a single power-law distribution is valid. Therefore, in this paper, we will use different values of $v_{\rm min}$ on mock data to show how this can provide an additional handle on the robustness of the result.

We now proceed by implementing the analysis outlined in Section 3 for different sets of minimum velocities, with $v_{\rm min} \in [300, 325, 350, 375, 400]$ km s^{-1}. When increasing $v_{\rm min}$, we hold the data set fixed, which leads to the number of stars decreasing with $v_{\rm min}$. For example, for the sample with no errors, the number of stars in the sample is 2000 stars for $v_{\rm min} = 300$ km s^{-1}, while it drops to 408 stars for $v_{\rm min} = 400$ km s^{-1}. The Gaia data behaves similarly; thus, we can account for the effect of decreasing statistics with increased cutoff velocity in this manner.

We show the results of the fit as a function of $v_{\rm min}$ in Figure 4 for 10 different mock data sets. In the top panel, the red and blue

[11] Earlier studies performed this analysis with just the line-of-sight velocity measurements, so in order to have a fair comparison, we only compare to Monari et al. (2018) and Deason et al. (2019) as they have used three-dimensional velocity measurements.

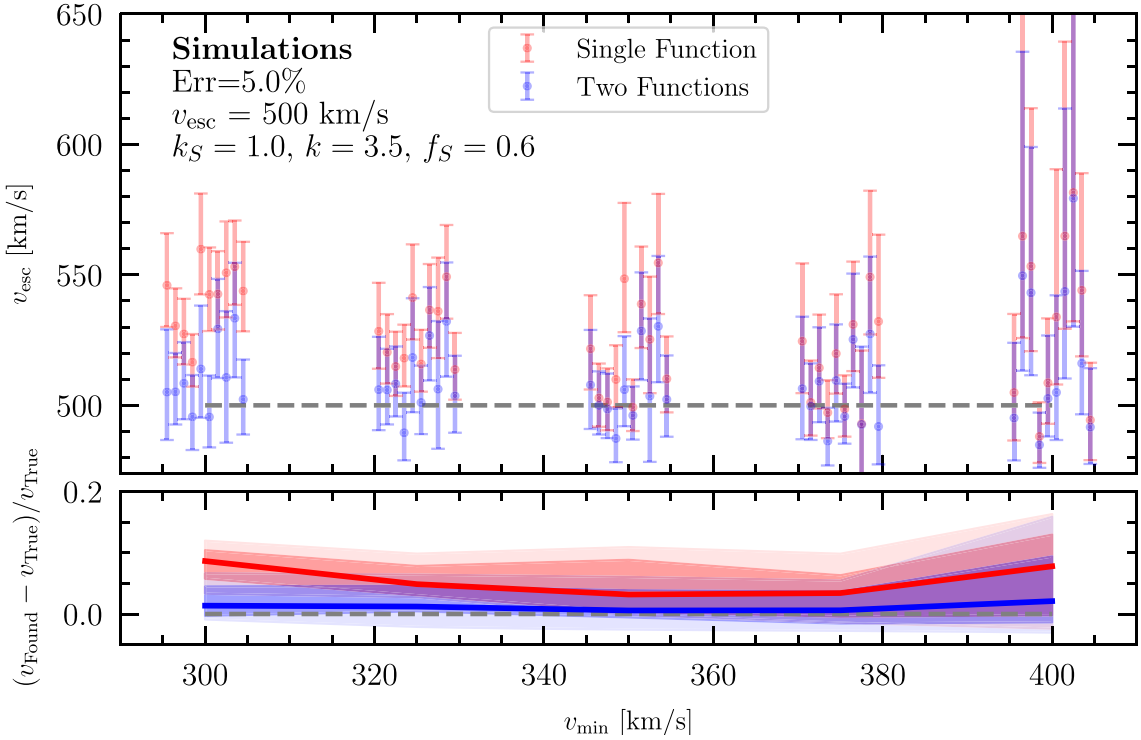

Figure 4. Analyses of generated data with 5% errors, assuming the fiducial two-component data set used throughout this section. (Top) For each $v_{min} \in [300, 325, 350, 375, 400]$ km s^{-1}, we show on the vertical axis the obtained v_{esc} and one standard deviation error bars for the single-component fit (red) and the two-component fit (blue). The true value is indicated by the dashed gray line. (Bottom) The solid lines show the average bias in the single-component and two-component fits obtained by comparing the best-fit v_{esc} with the true value. The darker bands are $\sim 1\sigma$ regions for the mock analyses, and the lighter shaded region shows the maximum and minimum values in the set.

data points are from fits to one or two bound components, respectively, while the dashed gray line is the true v_{esc} value. We begin by considering the sample with $v_{min} > 300$ km s^{-1}. In the single-component fit, v_{esc} is biased toward larger values and the bottom panel shows that the average best-fit v_{esc} is high by about 10% relative to the true value. The v_{esc} results for the single-component fit are consistent with the behavior discussed in Figure 2, where the single-function fit tends to find an average of slopes for both components and higher v_{esc}. Meanwhile, for the two-component fit we find the best fit is consistent with the correct value of $v_{esc} = 500$ km s^{-1}. For larger v_{min}, as shown in Figure 4, the average bias in the single-component fit becomes smaller and the two analyses converge on the correct value. Even though the bias in the single component persists at a level of 5%–10% for $v_{min} = 400$ km s^{-1}, the spread becomes much larger and the range of best-fit values for single- and two-component fits is similar. This is the expected behavior since the speed distribution for larger $|\vec{v}|$ is then dominated by a single component.

In Figure 5, we show an example of one of the data sets with $v_{min} = 300$ km s^{-1}, along with the best-fit distribution from both the one-component and two-component fits. Note that in plotting the best fit (solid lines) and true distributions (dashed lines), we have taken into account the presence of errors. To do so, we average over the errors of the stars in the sample, and use Equation (4) with these values. While there appear to be good fits in both panels, with only a single component (right panel) the fit has to account for both the lower speed stars near $|\vec{v}| \sim 300$ km s^{-1} with a larger k, as well as the slope and cutoff near v_{esc}. Doing so leads to overestimating the slope (since it is averaging the slopes of the two distributions) and the escape velocity, with best-fit $v_{esc} = 546^{+20}_{-16}$ km s^{-1}. Because of the limited statistics of stars near v_{esc}, the fit is largely influenced by the distribution at lower speeds. While the single-component fit appears to perform well, the preference for higher v_{esc} and k is revealed in the corner plot in Figure A1 as well as in Figure 4.

In the right panel of Figure 5, we can see that with the two-component fit, one of the components may be poorly constrained, in this case the halo component. This is because the halo component dominates for a relatively small range ($|\vec{v}| \in [300, 350]$ km s^{-1}, as can be seen in the true distributions). The corner plot in Figure A2 shows that there is a degeneracy between the substructure fraction and the halo slope, where a larger halo slope can be compensated by increasing f_S. The fit, however, is finding the correct model of the tail of the speed distribution, and subsequently the correct escape velocity. The outlier distribution is also correctly recovered.

The example in this section illustrates that if there are kinematic substructures that not captured by the model, this can lead to biased results in v_{esc} that are sensitive to the arbitrary choice of v_{min}. In order to evaluate this possibility in real data, we next turn to a comparison of models with one or two bound components using the ΔAIC, introduced in Section 3.3. We compare the two-function fit to the single-function fit using Equation (13) for the different values of v_{min}. Negative values of the ΔAIC correspond to the two-component model being favored over the single-component model. Figure 6 shows the average ΔAIC (solid line) and range of values for the analyses on the 10 mock data sets. For $v_{min} < 350$ km s^{-1}, the two-component fit is on average favored, although there is a large range and for some realizations a single-component fit might perform just as well. For higher $v_{min} \geqslant 350$ km s^{-1}, the one-component fit can capture the tail

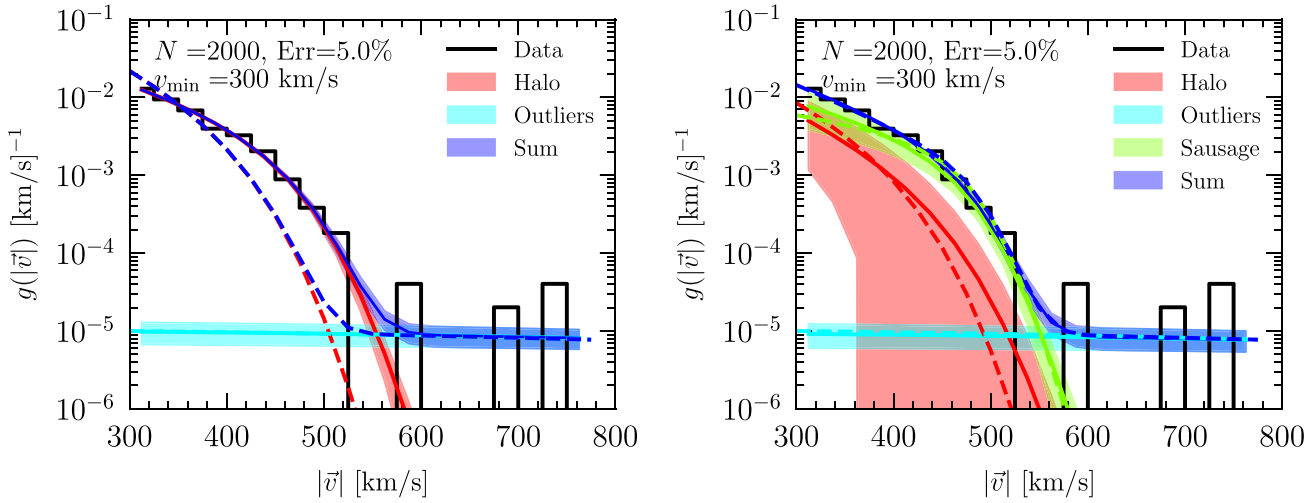

Figure 5. Best-fit speed distributions overlaid on a histogram of mock data, drawn from a two-component model and assuming 5% errors. The true distributions are shown in dashed lines, and the best fit in solid lines. The shaded regions are 68% containment regions obtained from the posteriors for each component. (Left) Best-fit distributions for a single-function fit. The functions shown are the halo (red) and outliers (yellow). The summed distribution is shown in blue. The true distribution shown in this case is that of the halo alone, normalized to the full distribution. (Right) Best-fit distributions assuming a two-function fit. The distributions shown are the halo (red), the Sausage (green), the outliers (yellow), and the summed distribution (blue).

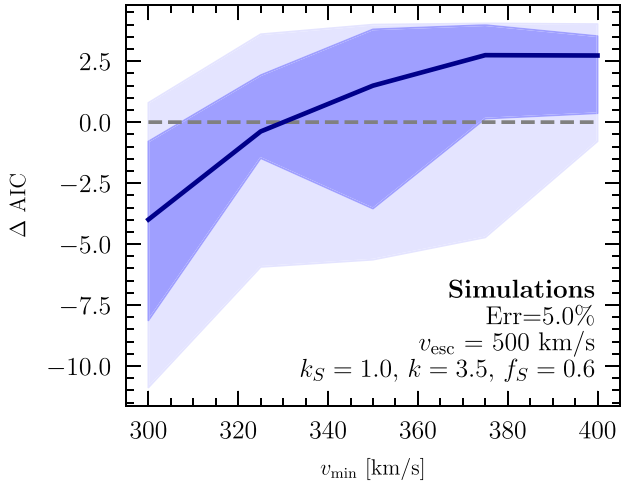

Figure 6. ΔAIC for the two-component fit compared with the one-component fit. Negative values indicate the two-component fit is favored. The solid line shows the average over the 10 mock analyses, the darker shaded region indicate the 16th and 84th percentiles, and the lighter shaded region shows the maximum and minimum values obtained in this set. For larger $v_{\rm min}$, single- and two-component fits give similar results and the ΔAIC reflects the penalty for introducing additional parameters.

well, and we see the average ΔAIC tends toward positive values since there is a penalty for extra model parameters. From Equation (13), we would expect $\Delta{\rm AIC} = 4$ if the maximum likelihoods were exactly equal. Depending on the data set, the need for a multicomponent model could be revealed by performing the fit as a function of $v_{\rm min}$ and evaluating the ΔAIC. This is the strategy we adopt when studying the Gaia eDR3 results in Necib & Lin (2022).

4.1. Estimating the Substructure Fraction

It is interesting to see how well we can reconstruct properties of the substructure component, as it could offer independent information on the Milky Way merger history from other studies. Here, we consider whether it is possible to extract the fraction of the non-outlier distribution associated with the Sausage, f_S. Note that the parameter f_S in a given analysis is not exactly the same as the value of $f_S = 0.6$ used in generating the mock data set. This is because the fraction changes as a function of $v_{\rm min}$, which can be seen for example in Figure 3. We thus compute the true values of $f_S\,(v_{\rm min})$ by integrating the true distributions in the interval $[v_{\rm min}, \infty]$ km s^{-1} as shown in Equation (2). The distributions $\boldsymbol{g_S}$ and $\boldsymbol{g_H}$ are those with the average errors convolved in them, i.e., the distributions in the right panel of Figure 3.

In Figure 7, we show the recovered posteriors of the fraction f_S along with the true values for $f_S\,(v_{\rm min})$, indicated by dashed lines. We find that the posteriors on f_S are quite wide and it is difficult to extract a precise value, though the results are consistent with the true values. The difficulty in obtaining f_S for these mock data sets is because the halo component is difficult to constrain when it dominates for only a narrow range of speeds. Of course, this statement depends on the differences between the slopes, but we find similar errors on f_S in the Gaia eDR3 analysis in Necib & Lin (2022).

5. Results with FIRE Simulations

In this section, we implement the previous analysis on realizations of the Milky Way. We use the hydrodynamic simulations[12] LATTE (Wetzel et al. 2016) as part of the FIRE simulations[13] (Hopkins 2015; Hopkins et al. 2018). We apply the analysis to the realization m12f, a Milky Way analog with stellar properties described in Sanderson et al. (2018). This galaxy was chosen as it underwent multiple mergers past redshift 1, the most dominant of which is a merger of peak mass $M_* = 1.5 \times 10^{11} M_\odot$ between redshifts 0.12 and 0.39 (Necib et al. 2019b). The stars of this merger have a distinct velocity distribution that peaks at high speeds at redshift zero at the location of the Sun (Necib et al. 2019b), which will affect

[12] https://ananke.hub.yt/
[13] https://fire.northwestern.edu

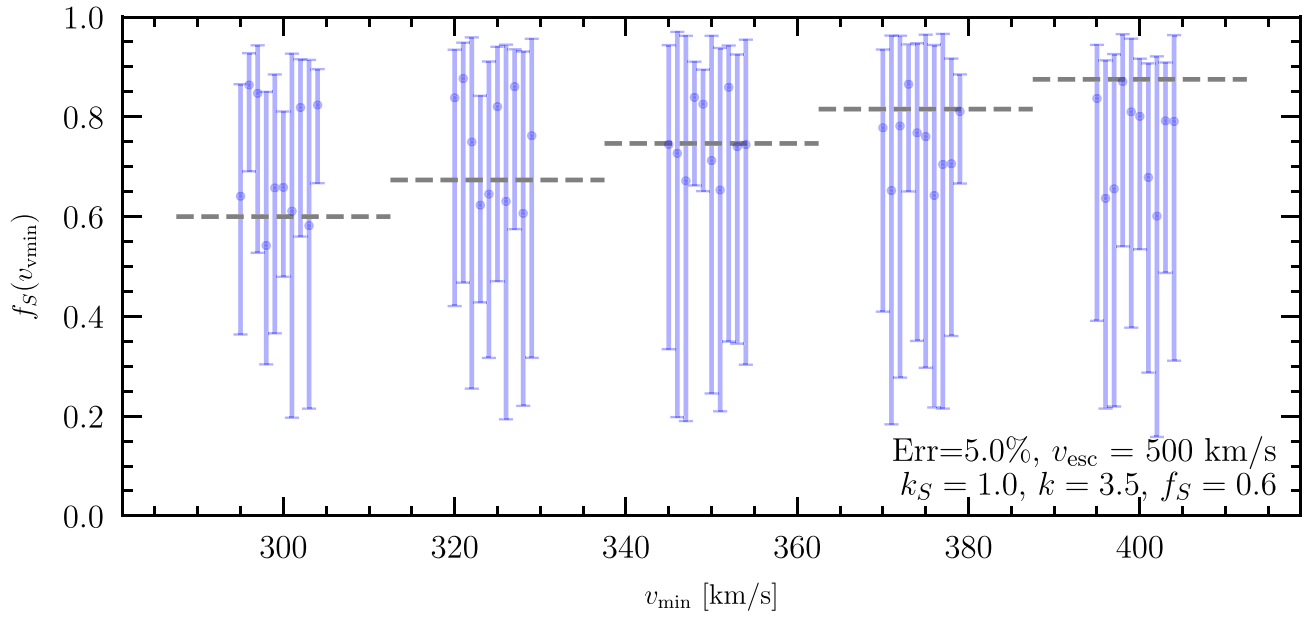

Figure 7. True vs. recovered substructure fraction posteriors for the 10 mock data sets considered throughout Section 4. The true fractions are shown as dashed lines, and are ordered from lightest to darkest corresponding to the values of $v_{\rm min} = [300, 325, 350, 375, 400]$ km s^{-1}.

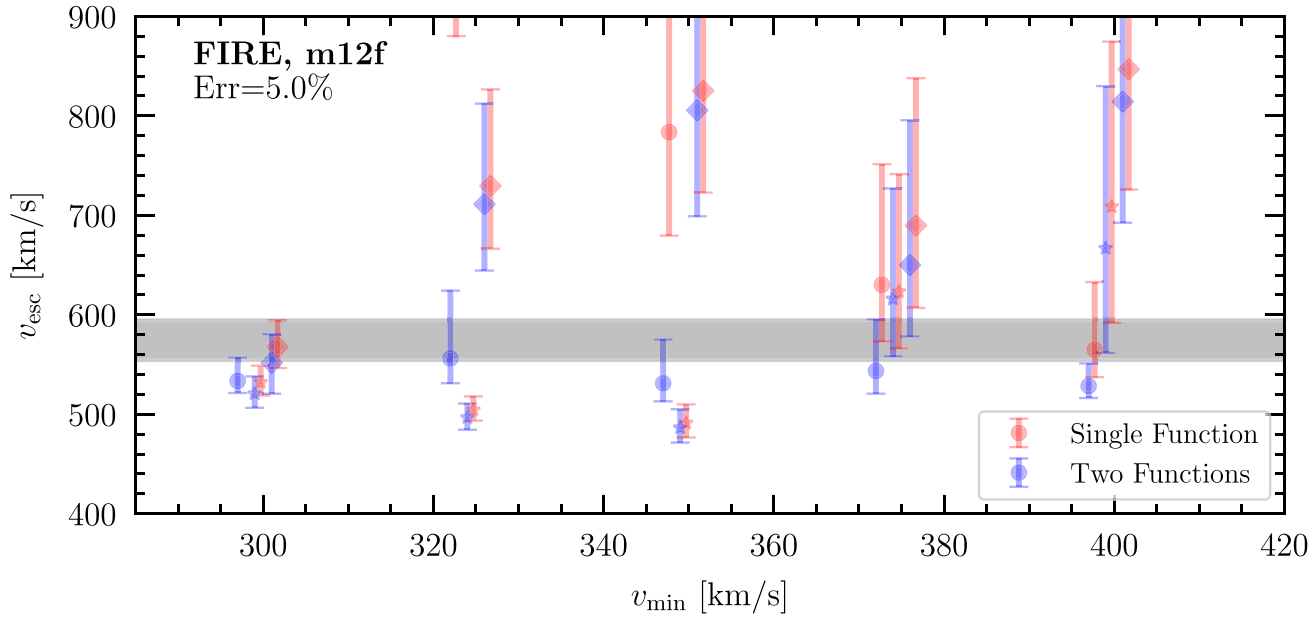

Figure 8. Similar to Figure 4, but for the FIRE simulation m12f. Best-fit values and one standard deviation for the results of the fits of the escape velocity $v_{\rm esc}$, for different values of $v_{\rm min}$. The circle, stars, and diamond makers correspond to the three positions of the Sun in the simulation (LSR0, LSR1, and LSR2). The gray band is the true escape velocity, found through computing the minimum and maximum escape velocity from the potential of the stars in the selected [7, 9] kpc region. The number of stars has been capped at 2000 so it is comparable with the data selection, and a percentage error of 5% has been added to the velocities of the stars, similar to the treatment in Section 4.

the tail of the velocity distribution that we fit in this paper. We include a 5% measurement error, similar to the discussion in Section 4.

Figure 8 shows the results of single- and two-function fits to the simulated data, for three different solar positions (LSR0, LSR1, and LSR2). The true escape velocity (gray band) is obtained by computing the potential over the particles in the region of interest ($r \epsilon [7.9]$ kpc) and computing the corresponding escape velocity using $\sqrt{2|\Phi(r) - \Phi(2r_{\rm vir})|}$, with the $r_{\rm vir} = 293.2$ kpc, as stated in Sanderson et al. (2018). The band corresponds to the minimum and maximum inferred escape velocity in this region. We find that the two-function fit is stable and consistent with the true $v_{\rm esc}$ for all $v_{\rm min}$. For $v_{\rm min} = 300$ km s^{-1}, the result for $v_{\rm esc}$ is actually lower than the true value by roughly 10%, similar to what was found in Grand et al. (2019) and Koppelman & Helmi (2021).

In contrast, the results for the single-function fit are more variable across the different solar positions. In particular, for LSR0 and $v_{\rm min} = 300$ km s^{-1} the fit gives $v_{\rm esc} = 970^{+22}_{-43}$ km s^{-1}, beyond the limits of the plot. The posterior is in fact peaked at the upper range on the prior, $v_{\rm esc} = 1000$ km s^{-1}, as

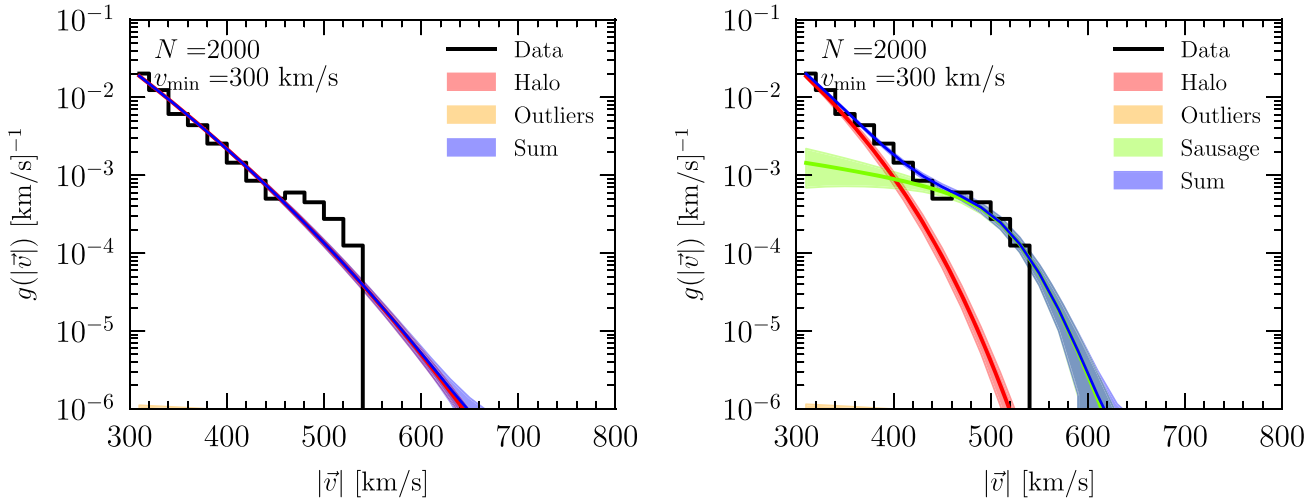

Figure 9. Similar to Figure 5, but for the FIRE simulation m12f, and the solar position LSR0. (Left) is the single-function fit, while (right) is the two-function fit.

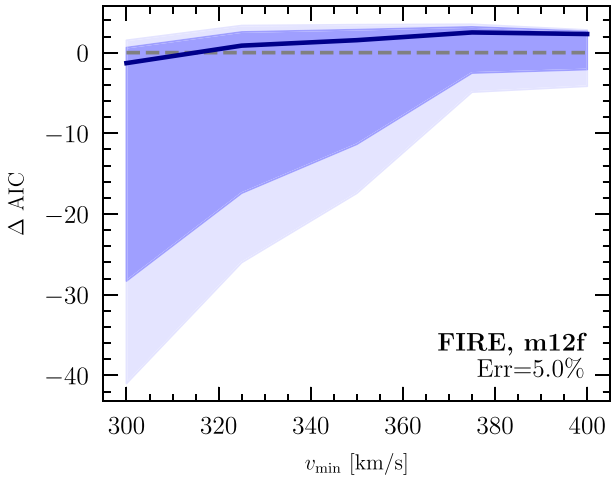

Figure 10. Similar to Figure 6, but for the FIRE simulation m12f.

can be seen in the corner plot of Figure A3 (by comparison, Figure A4 shows the corner plot for the two-function fit for the same underlying distribution). The stellar velocity distribution and best-fit distributions for this solar position are shown in Figure 9. We see for this case that a single power law is not able to fit the data well, with an excess of high-speed stars near the tail in the left panel. These high-speed stars can be fit with a substructure component with low slope $k_S = 0.70^{+0.68}_{-0.38}$, consistent with the expectation for the Sausage.

Figure 10 shows the model comparison between single- and two-component fits, obtained from evaluating the ΔAIC as a function of $v_{\rm min}$. There is a large spread in results, with LSR0 showing a strong preference for a two-component fit for $v_{\rm min} \leqslant 350$ km s^{-1}. For the other LSRs, there is little or no preference for the two-component fit. For these, the inferred $v_{\rm esc}$ is similar whether one or two components is included, although the best-fit $v_{\rm esc}$ may be higher or lower than the true value by up to 2σ–3σ, depending on $v_{\rm min}$. Thus, we see that even within one simulated galaxy, results can vary significantly by location, which is expected since the Sausage is not a fully mixed structure.

6. Conclusions

In order to extract sensible results for the local stellar escape velocity, recent studies have taken low values for $v_{\rm min}$, where it is not clear if the power-law distribution of Leonard & Tremaine (1990) holds, and imposed artificial priors on the slope of the speed distribution k. These choices can shape the measurement of the escape velocity and the mass of the Milky Way. In this paper, we developed a more general framework to fit the escape velocity by accounting for deviations from the single power-law distribution. In particular, we account for the presence of multiple kinematic substructure components in the speed distribution for the first time. Our pipeline also accounts for individual errors on stellar speeds in a forward model, as well as the outlier distribution.

A second kinematic component in the speed distribution is motivated by the presence in our local neighborhood of (at least) a second kinematic structure besides the stellar halo, called the Gaia Sausage. To account for this, studies of the tail of the speed distribution either need to increase the minimum velocity $v_{\rm min}$ above which we define a tail, or make sure that we have the correct number of components in the model. Thus, including a second kinematic component is physically motivated.

To model the presence of this substructure, we introduce a second bound component following Equation (1), with the same escape velocity but a new slope for the tail k. The approach can also be generalized to include more components. We then fit for the escape velocity, the slopes of the structures, and their fractional contributions. The fit is repeated with different numbers of components, and different definitions of the tail of the distribution (i.e., different values of $v_{\rm min}$).

Using mock data with multiple components, we found that our pipeline can reconstruct $v_{\rm esc}$ in the presence of substructure, and is robust to the presence of observational errors. We have shown how a single-component fit could be biased for low $v_{\rm min}$, while at high $v_{\rm min}$ the result of the single and two-component fits should converge. Using the ΔAIC to compare models, we show that multiple components are preferred at low $v_{\rm min}$, while a single function is sufficient as $v_{\rm min}$ increases.

We also applied our pipeline to mock data obtained from hydrodynamic simulations, focusing on a particular realization of a galaxy with similar features and merger history as the Milky Way. We find that a two-component fit performs well and is

consistent with the true v_{esc}. A single-component fit gives more variable results: at some solar positions, it gives results consistent with the true values, but we found that it can also sometimes give highly biased and degenerate fit results. This further motivates goodness-of-fit comparisons on real data with multiple substructure components.

In a companion paper, Necib & Lin (2022), we use the method outlined here to measure the escape velocity of the Milky Way. We apply the single- and two-function fits over the five values of v_{min}, as was done on the mock data sets. There we show that a two-component fit does provide a somewhat better fit to Gaia data, allowing us to extract a robust escape velocity of $v_{esc} = 445^{+25}_{-8}$ km s^{-1}. The corresponding Milky Way mass inferred from these results is $M_{200} = 4.6^{+1.5}_{-0.8} \times 10^{11} M_{\odot}$ for the two-component fit.

We are grateful to I. Moult for early discussions and collaboration on the project, and to M. Lisanti for helpful feedback. We would also like to thank L. Anderson, A. Bonaca, G. Collin, A. Deason, P. Hopkins, A. Ji, and J. Johnson for helpful conversations.

This work was performed in part at Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. L.N. is supported by the DOE under Award No. DESC0011632, the Sherman Fairchild fellowship, the University of California Presidential fellowship, and the fellowship of theoretical astrophysics at Carnegie Observatories. T.L. is supported by an Alfred P. Sloan Research Fellowship and Department of Energy (DOE) grant DE-SC0019195.

Software: Astropy (Astropy Collaboration et al. 2013, 2018), corner.py (Foreman-Mackey 2016), emcee (Foreman-Mackey et al. 2013), IPython (Pérez & Granger 2007), Galpy (Bovy 2015).

Appendix A
Corner Plots

In this appendix, we show example corner plots of the fits of the main text.

A.1. Mock Data

In Figures A1 and A2, we show the corner plots for mock data generated with 5% observational errors as discussed in Section 4. The mock data set is the same as that shown in Figure 5.

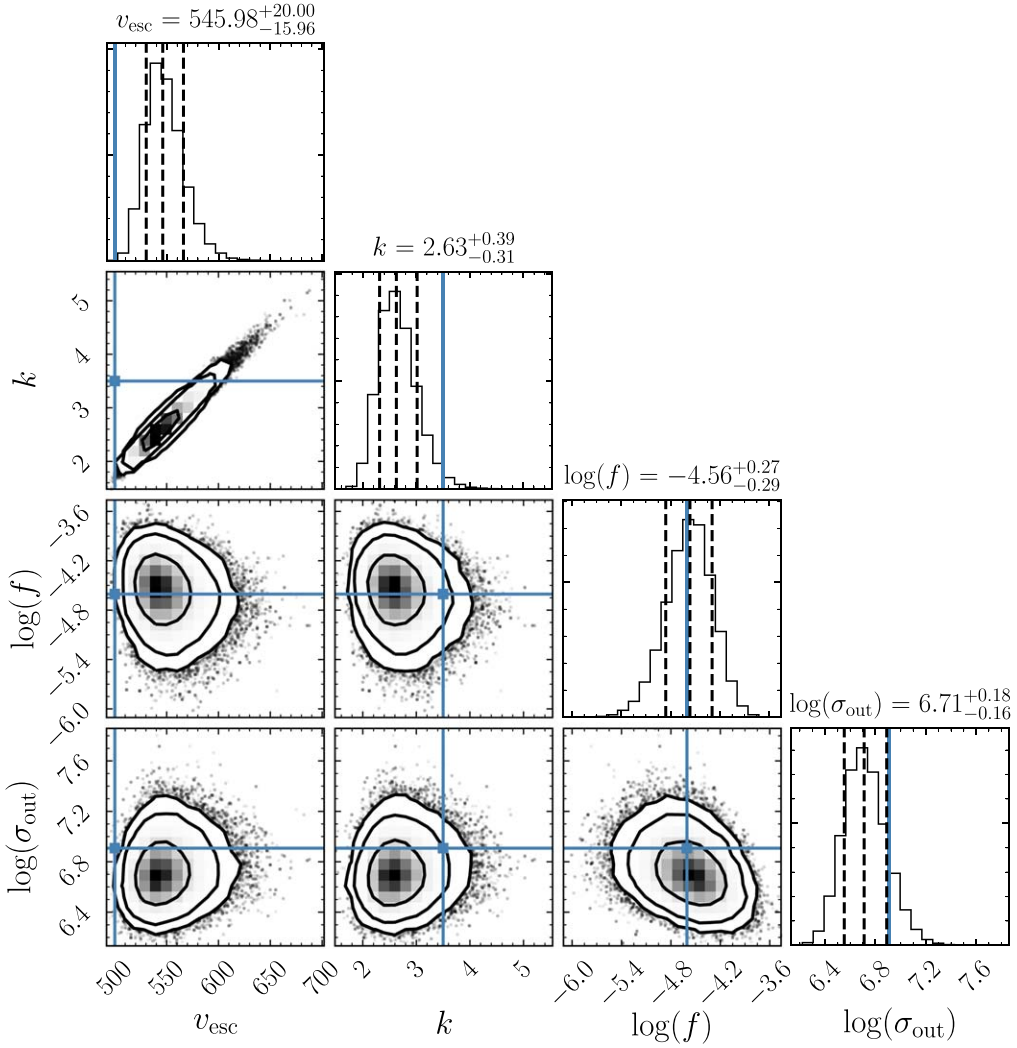

Figure A1. Corner plot of the run with mock data assuming one component and $v_{min} = 300$ km s^{-1}. The escape velocity is biased toward larger values compared to the true $v_{esc} = 500$ km s^{-1}. The true k shown here is that of the component with $k = 3.5$. See discussion in Section 4.

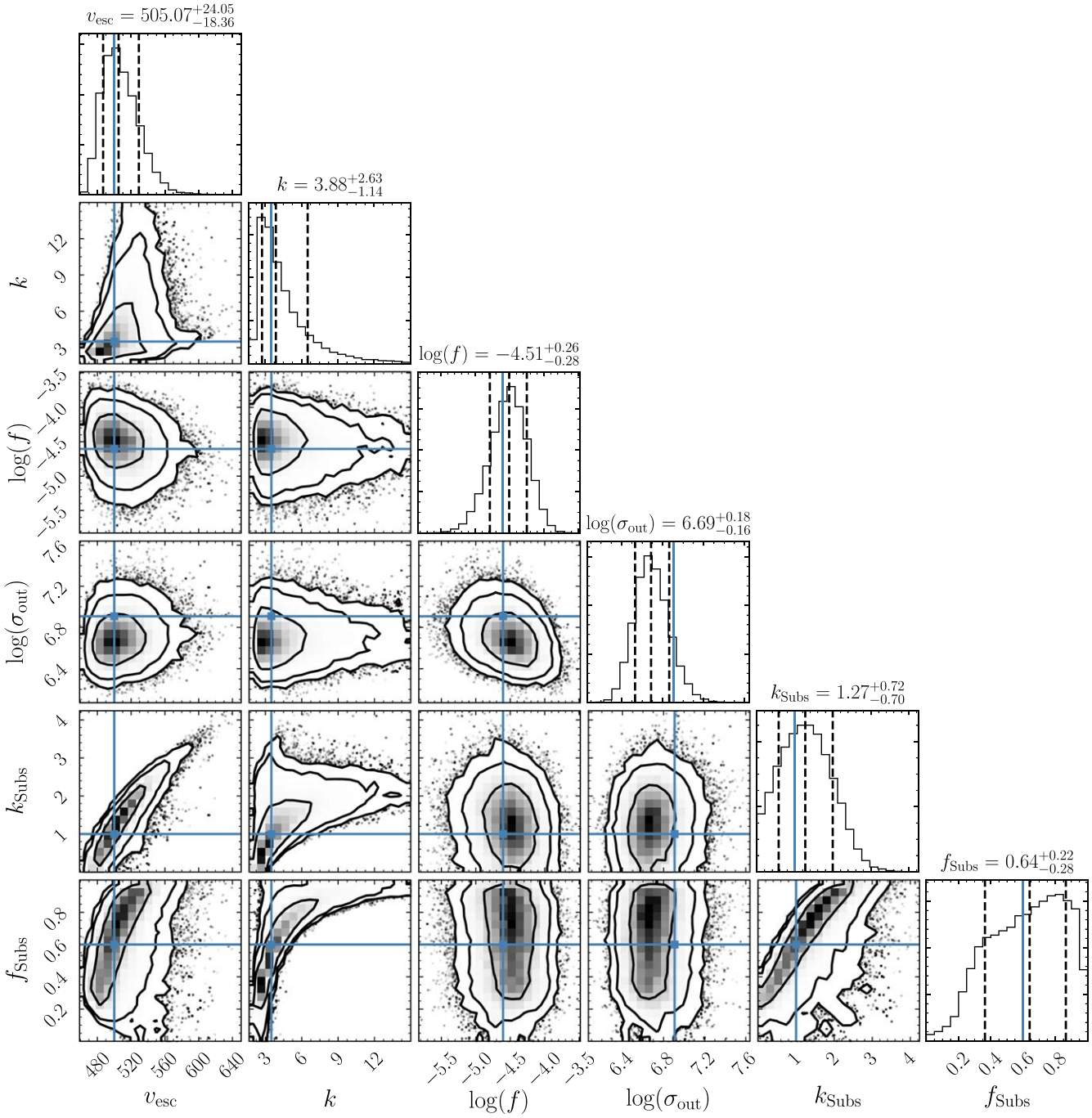

Figure A2. Corner plot of the run with mock data assuming two bound components, $v_{\rm min} = 300$ km s^{-1}, and percentage errors of 5%. See discussion in Section 4.

A.2. FIRE Data

In Figures A3 and A4, we show the corner plots for simulated FIRE data at the solar position associated with LSR0. The data is generated with 5% observational errors as discussed in Section 5. The data set is the same as that shown in Figure 9.

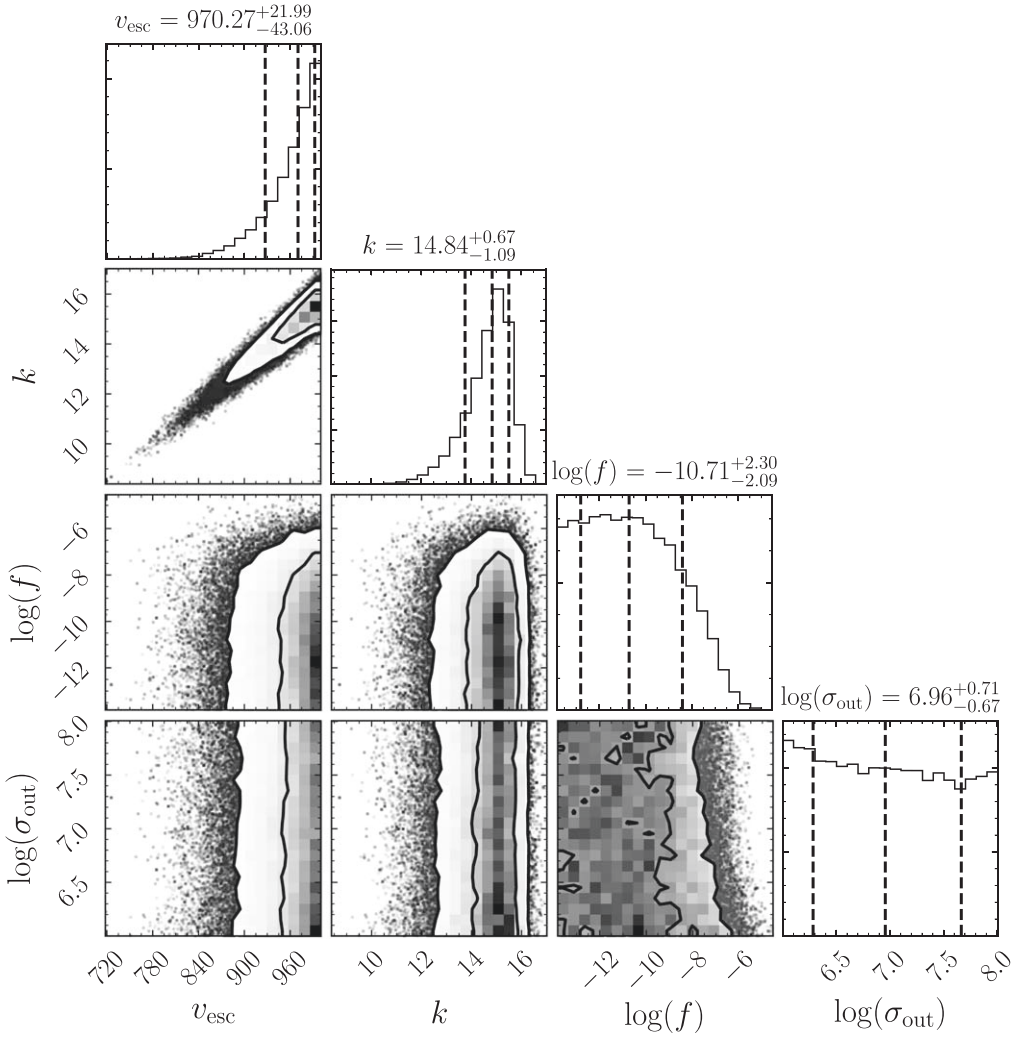

Figure A3. Corner plot of the run with FIRE data assuming one component and $v_{\rm min} = 300$ km s^{-1}. The escape velocity is biased toward much larger values compared to the true $v_{\rm esc}$ km s^{-1} and the posterior in fact is piled up at the upper bound of the prior range.

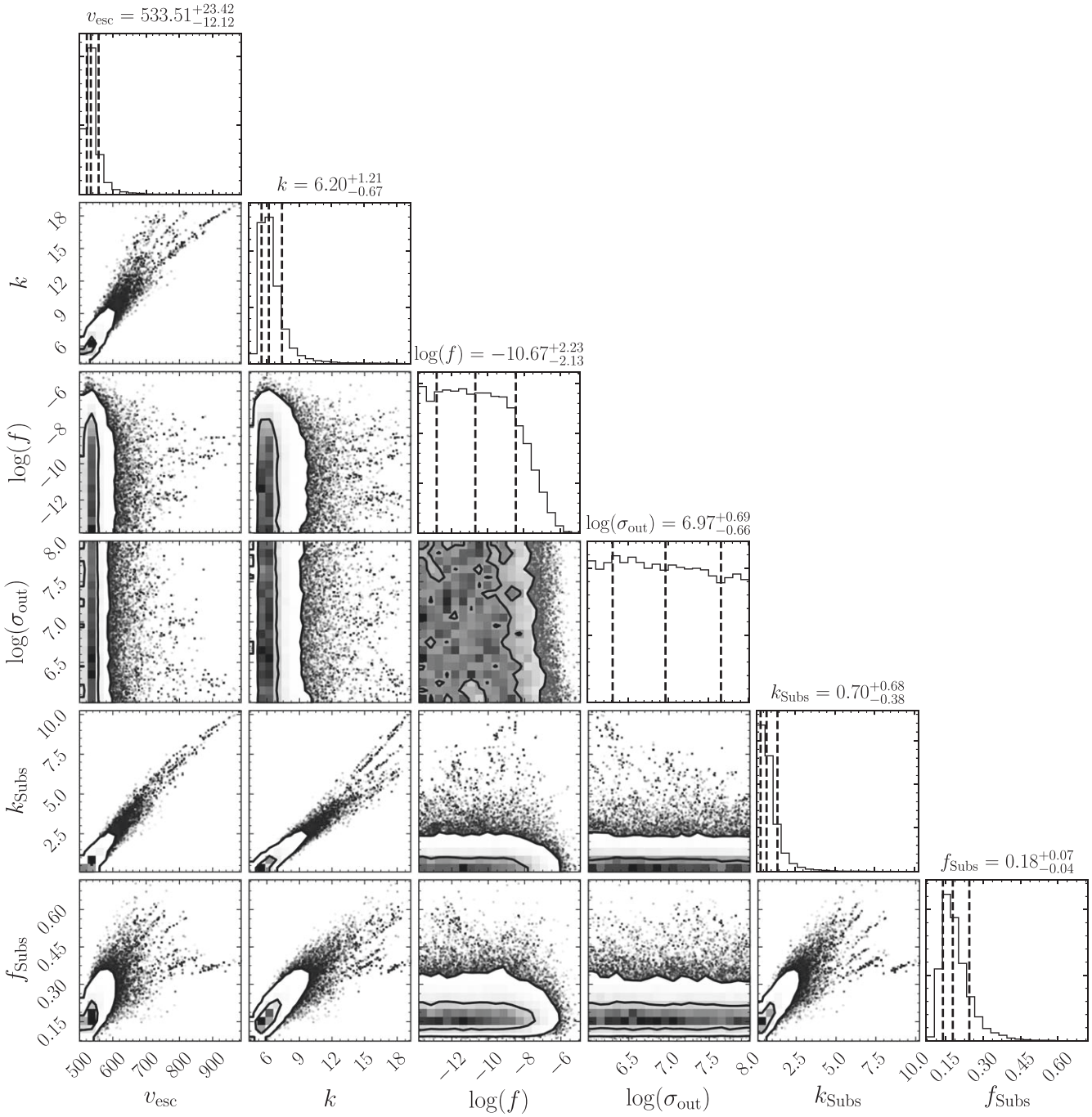

Figure A4. Corner plot of the run with FIRE data assuming two bound components, $v_{\rm min} = 300$ km s^{-1}, and percentage errors of 5%.

ORCID iDs

Lina Necib https://orcid.org/0000-0003-2806-1414
Tongyan Lin https://orcid.org/0000-0003-4969-3285

References

Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, ApJS, 203, 21
Akaike, H. 1974, ITAC, 19, 716
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611
Belokurov, V., Evans, N. W., Irwin, M. J., et al. 2007, ApJ, 658, 337
Belokurov, V., Zucker, D., Evans, N., et al. 2006, ApJ, 642, L137
Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJL, 647, L111
Bovy, J. 2015, ApJS, 216, 29
Collins, M. L. M., Tollerud, E. J., Sand, D. J., et al. 2017, MNRAS, 467, 573
Deason, A. J., Belokurov, V., Koposov, S. E., & Lancaster, L. 2018, ApJ, 862, L1
Deason, A. J., Fattahi, A., Belokurov, V., et al. 2019, MNRAS, 485, 3514
Diemand, J., Kuhlen, M., Madau, P., et al. 2008, Natur, 454, 735
Foreman-Mackey, D. 2016, JOSS, 1, 24
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2020, A&A, 649, A1
Grand, R. J. J., Deason, A. J., White, S. D. M., et al. 2019, MNRAS, 487, L72
Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179
Hattori, K., Valluri, M., Bell, E. F., & Roederer, I. U. 2018, ApJ, 866, 121
Helmi, A. 2020, ARA&A, 58, 205
Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, Natur, 563, 85
Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, Natur, 402, 53
Hopkins, P. F. 2015, MNRAS, 450, 53
Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800
Kirby, E. N., Boylan-Kolchin, M., Cohen, J. G., et al. 2013, ApJ, 770, 16
Koppelman, H. H., & Helmi, A. 2021, A&A, 649, A136
Lancaster, L., Koposov, S. E., Belokurov, V., Evans, N. W., & Deason, A. J. 2019, MNRAS, 486, 378
Leonard, P. J. T., & Tremaine, S. 1990, ApJ, 353, 486
Lindegren, L., Lammers, U., Bastian, U., et al. 2016, A&A, 595, A4
Monari, G., Famaey, B., Carrillo, I., et al. 2018, A&A, 616, L9
Myeong, G. C., Evans, N. W., Belokurov, V., Amorisco, N. C., & Koposov, S. E. 2018a, MNRAS, 475, 1537
Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018b, ApJ, 863, L28
Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, ApJ, 901, 48
Necib, L., & Lin, T. 2022, ApJ, 926, 189
Necib, L., Lisanti, M., & Belokurov, V. 2019a, ApJ, 874, 3
Necib, L., Lisanti, M., Garrison-Kimmel, S., et al. 2019b, ApJ, 883, 27
Necib, L., Ostdiek, B., Lisanti, M., et al. 2020, ApJ, 903, 25
Niederste-Ostholt, M., Belokurov, V., Evans, N. W., et al. 2009, MNRAS, 398, 1771
Pérez, F., & Granger, B. E. 2007, CSE, 9, 21
Piffl, T., Scannapieco, C., Binney, J., et al. 2014, A&A, 562, A91
Sanderson, R. E., Garrison-Kimmel, S., Wetzel, A., et al. 2018, ApJ, 869, 12
Scannapieco, C., White, S. D. M., Springel, V., & Tissera, P. B. 2009, MNRAS, 396, 696
Simon, J. D., Li, T. S., Drlica-Wagner, A., et al. 2017, ApJ, 838, 11
Smith, M. C., Ruchti, G. R., Helmi, A., et al. 2007, MNRAS, 379, 755
Springel, V., Wang, J., Vogelsberger, M., et al. 2008, MNRAS, 391, 1685
Vrieze, S. I. 2012, Psychol. Methods, 17, 228
Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, ApJL, 827, L23
White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341
Williams, A. A., Belokurov, V., Casey, A. R., & Evans, N. W. 2017, MNRAS, 468, 2359
Wit, E., Heuvel, E. v. d., & Romeijn, J.-W. 2012, Statistica Neerlandica, 66, 217
Zucker, D. B., Belokurov, V., Evans, N. W., et al. 2006, ApJL, 650, L41