

# Non-Parametric Spherical Jeans Mass Estimation with B-splines

Nabeel Rehemtulla,<sup>1</sup>\* Monica Valluri,<sup>1</sup>† and Eugene Vasiliev<sup>2,3</sup>

<sup>1</sup>Department of Astronomy, University of Michigan, 1085 S. University Ave., Ann Arbor, MI 48109, USA

<sup>2</sup>Institute of Astronomy, University of Cambridge, Madingley road, Cambridge, UK, CB3 0HA

<sup>3</sup>Lebedev Physical Institute, Leninsky prospekt 53, Moscow, Russia, 119991

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

Spherical Jeans modeling is widely used to estimate mass profiles of systems from star clusters to galactic stellar halos to clusters of galaxies. It derives the cumulative mass profile,  $M(< r)$ , from kinematics of tracers of the potential under the assumptions of spherical symmetry and dynamical equilibrium. We consider the application of Jeans modeling to mapping the dark matter distribution in the outer reaches of the Milky Way using field halo stars. We present a novel non-parametric routine for solving the spherical Jeans equation by fitting B-splines to the 3-dimensional velocity and density profiles of halo stars. While most implementations of Jeans modeling assume parametric forms for these profiles, B-splines provide non-parametric fitting curves with analytical derivatives. Despite Jeans modeling's prevalence, there is little work quantifying the biases introduced in the resulting mass profile when considering observational errors on full 6D phase space coordinates. We validate our routine on several progressively more complex and realistic mock datasets. Our routine recovers the mass profiles of equilibrium systems with flattened halos or a stellar disk and bulge excellently ( $\leq 4\%$  error at most radii). Tests with non-equilibrium, Milky Way-like galaxies from the Latte cosmological simulations perform quite well ( $\leq 15\%$  error out to 100 kpc). Lastly, when we impose observationally motivated selection functions and errors on the phase space coordinates of tracer particles drawn from cosmological simulations, our code still yields cumulative mass estimates within  $\sim 30\%$  error out to 100 kpc. Our results with observationally motivated mock datasets from cosmological simulations compared to the equilibrium models implies that spherical Jeans modeling is more sensitive to incomplete samples due to survey selection functions and deviations from dynamical equilibrium than deviations from sphericity.

**Key words:** – galaxies: halo – galaxies: kinematics and dynamics – galaxies: structure

## 1 INTRODUCTION

The European Space Agency's *Gaia* satellite (Perryman et al. 2001; Lindegren et al. 2016; Gaia Collaboration et al. 2016, 2018), launched in 2013, has released 3D positions and proper-motions for billions of Milky Way stars. In conjunction with line-of-sight velocities and chemical abundances for stars obtained with ground based spectroscopic surveys (e.g. RAVE, Steinmetz et al. 2006; LAMOST, Hu & Jiang 2005; APOGEE, Allende Prieto et al. 2008; GALAH, De Silva et al. 2015; Gaia-ESO, Gilmore et al. 2012, DESI Allende Prieto et al. 2020), various components of the Milky Way's stellar halo are being characterized: field halo stars, stars in individual tidal streams, globular clusters, and satellites. The dynamics of these halo objects are important probes of the Milky Way's dark matter distribution and, in principle, allow us to determine several things about the Milky Way's dark matter halo, most fundamentally its *mass*

*density profile*. The mass of the Milky Way is fundamental quantity of interest for comparisons with cosmological simulations, but it is surprisingly poorly constrained. The availability of distances, radial velocities and even proper motions for huge numbers of individual stars, globular clusters and satellite galaxies has led to numerous efforts to determine the Milky Way halo parameters. Despite the availability of increasingly high quality data, the measurements of some basic properties of the dark matter halo of the Milky Way have not converged. Bland-Hawthorn & Gerhard (2016) summarize in their Table 8 estimates for  $M_{200}$  from 1999–2014 that range from  $0.55 - 2.62 \times 10^{12} M_\odot$ . A more recent compilation by Wang et al. (2020) considering only results obtained with *Gaia* DR2 data finds a similar range of values. Other methods using tidal streams, the velocities of satellite galaxies and field halo stars give a similarly large range of halo masses. This range of a factor of two in mass could be due to differences in the spatial distributions tracers, systematic differences between methods, and differences in the treatment of errors and perturbations due to the LMC (which has only recently been accounted for in this context by, Deason et al. 2021).

\* E-mail: nabeelr@umich.edu

† E-mail: mvalluri@umich.edu

## 2 Rehemtulla et al.

The absence of a large population of halo tracers with precision 6-D phase space information led over the past two decades to the development of a number of sophisticated techniques that incorporate the observational uncertainties via forward modeling. Starting with assumed parametric forms for the gravitational potential and distribution functions of tracers (star or globular clusters), these techniques impose observational errors and selection functions on the models to derive, using Bayesian inference, the best fit estimates and their confidence intervals for a variety of potential parameters, including the mass of the halo. The most recent example of such distribution function modeling applied to Gaia EDR3 data (Deason et al. 2021) is based on a sample of 665 halo stars (98% with full 6D phase space data) in the distance range 50-100 kpc (including K giant stars, Blue Horizontal Branch (BHB) stars, RR Lyrae stars and Blue Stragglers) and attempts to correct for the dynamical effect of the LMC.

We are on the verge of an era where the numbers of stars with full 6D phase space data is going to see a dramatic increase from hundreds to hundreds of thousands or even millions. The Dark Energy Spectroscopic Instrument (DESI) (DESI Collaboration et al. 2016a,b; Allende Prieto et al. 2020) is a 5000 fiber spectrograph on the Mayall 4-meter telescope at Kitt Peak National Observatory, Arizona, and is expected to obtain spectra of 5-8 million Gaia stars (down to  $G = 19.5$ ) in the next five years. In addition, the 4MOST de Jong et al. (2012) and WEAVE Dalton et al. (2014) surveys are similar multi-fiber spectrographic surveys which will together obtain spectra for millions of stars. These spectroscopic surveys will deliver radial velocities and spectro-photometric distances to stars out to 100 kpc, enabling – for the first time – the assembly of samples of stars with 6D phase space coordinates that are *orders of magnitude* larger than the best samples currently available.

In this work we take a step back and implement a new version of one of the oldest and conceptually simplest dynamical modeling tools – the spherical form of the Jeans equations (Jeans 1915; Binney 1980; Binney & Tremaine 2008) – under the assumption that very soon we will have samples of  $10^4 - 10^5$  halo stars out to  $\sim 100$  kpc with full 6D phase space coordinates. Our goal in this paper is to assess the effects of various factors on the derived cumulative mass profiles obtained with the spherical Jeans equation: (a) the underlying assumptions of spherical symmetry, (b) dynamical equilibrium, (c) realistic observational errors on proper motions, radial velocities and distances to halo stars and (d) the effects of limited survey volumes inherent to all surveys.

Most previous implementations of the spherical Jeans equation, applied to the Milky Way halo either bin the data (Kafle et al. 2018) and/or use analytic functions (e.g. power laws) to describe the density and velocity profiles and velocity anisotropy of tracer stars (Gnedin et al. 2010). Binning the data necessitates the computation of numerical derivatives which can be noisy (Kafle et al. 2018). For this reason, the use of analytic functions has been favored because they are relatively easy to implement and, in some cases, align well with theoretical predictions. For example, Gnedin et al. (2010) assign a power law relation to the tracer density profile of BHB stars beyond 25 kpc, which is broadly consistent with predictions for the “broken power-law” density profile of the stellar halo from cosmological simulations (Bullock & Johnston 2005; Johnston et al. 2008; Cooper et al. 2011). Other studies use parametric fits to the velocity distribution and velocity anisotropy profile. These approaches have their limitations since parametric fitting curves may not adequately represent the true velocity or density profiles. Many studies construct parametric fits from binned data, which is sensitive to how bins are chosen. Since the velocity anisotropy profile, in particu-

lar, requires accurate determination of proper motions (which have only become available in large numbers in the past 3 years, thanks to Gaia), it has historically been common to assume either a constant value or a few functional forms (e.g. Battaglia et al. 2005), or to use functional forms motivated by cosmological simulations (Xue et al. 2008; Gnedin et al. 2010).

In the context of external dwarf spheroidal galaxies, sophisticated Bayesian methods have been developed to derive parametric fits to the tracer density, line-of-sight velocity distributions and velocity anisotropy profiles (e.g. GravSphere, by Read & Steger 2017). Diakogiannis et al. (2017) have developed a hybrid approach and use B-splines to construct the radial velocity dispersion profile  $\sigma_r^2$ , while using a parametric mass model. They apply their method to line-of-sight velocities in the Fornax dwarf spheroidal galaxy. There have been several recent studies to assess how well the spherical Jeans equation performs on dwarf spheroidal galaxies when the various underlying assumptions are broken (e.g. spherical symmetry, dynamical equilibrium). For example, Evslin & Del Popolo (2017) use tracer particles drawn from N-body simulations to evaluate the fourth-order spherical Jeans equation when applied to dwarf spheroidal galaxies. Genina et al. (2020) assess how GravSphere fares on mock dwarf spheroidal galaxies drawn from cosmological  $\Lambda$ CDM and SIDM simulations. El-Badry et al. (2017) show that, using mock data sets from cosmological hydrodynamical simulations, dwarf galaxies with recent episodes of star formation and feedback are not in dynamical equilibrium and Jeans modeling results in an overestimate of the mass of the dwarf galaxy. We do not discuss these methods for modeling external dwarf galaxies any further since the nature of the data, both in sample size and phase space dimensionality (typically only position on the sky and line-of-sight velocities are available for  $< 5000$  stars) are completely different from the situation considered in this work.

There have been several recent works done to understand the errors introduced when breaking various assumptions of the spherical Jeans equation (Kafle et al. 2018; Wang et al. 2018). This is pertinent because real potentials of Milky Way-like disk galaxies are not spherically symmetric nor are they in dynamical equilibrium. Prior to applying it to real data, it is important to assess the modeling biases introduced by any spherical Jeans modeling code. In addition, it is helpful to have a quantitative understanding the effect that breaking these assumptions has on the resulting mass profile.

We present a novel B-spline based routine for performing spherical Jeans modeling. We construct a diverse suite of mock datasets to validate our routine and quantify the errors in the derived cumulative mass estimate when the standard assumptions are broken. The mock datasets in our suite include smooth, self-consistent equilibrium distribution functions generated using Agama (Vasiliev 2019) – an efficient all-purpose galactic modeling package – and mock stellar halos generated from three Latte (FIRE-2) cosmological hydrodynamic simulations (Wetzel et al. 2016). The results from each mock dataset provide insights into how the quality of the Jeans mass estimate changes with the geometry of the system and with its dynamical state.

For the Latte simulations, we also test different methods for identifying halo stars including by imposing different metallicity cuts to select metal poor star particles. We also consider the effect of imposing survey selection functions. Specifically, we impose a Gaia-like magnitude limit of  $G = 20.7$  and exclude star particles outside the fiducial footprint of the Dark Energy Spectroscopic Instrument (DESI) (DESI Collaboration et al. 2016a,b; Allende Prieto et al. 2020). Finally, we consider the effects of imposing

observational errors on proper-motions, line of sight velocities and distances to tracer stars. Once DESI data are available, we will apply our Jeans routine to Gaia proper motions and DESI radial velocities to construct a mass profile of the Milky Way out to  $\sim 80 - 100$  kpc.

## 2 ANALYSIS METHODS

The Jeans equations are obtained from the collisionless Boltzmann equation by computing second order moments of the distribution function (Binney & Tremaine 2008). Under the assumption of spherical symmetry, they reduce to a single equation which can be used to derive the enclosed mass at any radius  $M(< r)$ , from which the full gravitational potential  $\Phi$  can be obtained using  $\Phi(r) = \frac{GM(<r)}{r}$ . The form of the spherical Jeans equation we use here is given by,

$$M(< r) = -\frac{rv_r^2}{G} \left( \frac{d \ln \rho}{d \ln r} + \frac{d \ln v_r^2}{d \ln r} + 2\beta \right). \quad (1)$$

The velocity anisotropy parameter  $\beta$  is defined as

$$\beta = 1 - \frac{\overline{v_\theta^2} + \overline{v_\phi^2}}{2\overline{v_r^2}}. \quad (2)$$

The terms  $\overline{v_r^2}$ ,  $\overline{v_\phi^2}$  and  $\overline{v_\theta^2}$  are the means of the squares of the radial, azimuthal, and polar velocities respectively;  $\rho$  is the number density of tracer (stars, globular clusters, etc.);  $r$  is the spherical galactocentric radius;  $G$  is the gravitational constant.  $\beta < 0$  indicates a tangentially biased velocity distribution,  $\beta = 0$  indicates velocity isotropy, and  $\beta > 0$  indicates a radially biased velocity distribution. Note that in the above we define  $\beta$  in terms of  $\overline{v_r^2}$ ,  $\overline{v_\phi^2}$  and  $\overline{v_\theta^2}$  rather than in terms of the velocity dispersion components,  $\sigma_r^2$ ,  $\sigma_\phi^2$ , and  $\sigma_\theta^2$ , since the former are more suitable for B-spline evaluation and these quantities are more relevant from the dynamical standpoint. Note that we do not assume zero mean tangential streaming velocity ( $\overline{v_\phi} = \overline{v_\theta} = 0$ ) as is often done.

The spherical Jeans equation can be solved in two complementary regimes: (1) solving a “dynamical inverse problem” where one measures the dispersion of the line-of-sight velocities ( $v_{\text{los}}$ ) directly from observations (with errors) and then “inverts” it to derive the radial velocity dispersion and anisotropy profiles (or assume  $\beta$ ), and then infer the potential or cumulative mass distribution (Eq. 1); (2) “forward modeling” on the other hand assumes some parametric form for the gravitational potential and tracer distribution function and then computes the expected velocity dispersion and anisotropy profiles and convolves them with expected observational uncertainties and selection functions and compares with observations.

Although dynamical inverse modeling was the original formulation of the method, inverse problems in general have difficulty dealing properly with error and noise from small samples. Furthermore, it has been shown that in the context of the Milky Way halo, inferring the anisotropy profile from the line-of-sight velocity distribution results in a biased profile at galactocentric distances beyond  $\sim 20$  kpc (Hattori et al. 2017). Most current applications therefore follow the second route because it is easier to account for observational errors. In this work, we assess how well the “inverse modeling” route functions when full 6D phase space coordinate information is available (i.e.  $\beta$  does not need to be inferred from  $v_{\text{los}}$  alone), for samples that are orders of magnitude ( $>10^4$ ) larger than have previously been available. This method gives the potential (more precisely the cumulative mass) more directly, but also suffers from errors more directly. In future we aim to modify the

B-spline fitting of the  $\overline{v^2}$  profiles while accounting for observational errors as has been done for the measurement of Gaia proper-motion dispersions in globular clusters (Vasiliev & Baumgardt 2021).

Our goal in this work is to assess if the enormous increase in sample size mitigates the effects of only partially accounting for the observational errors.

Each component of the spherical Jeans equation (Eq. 1) is computed using B-splines, a category of piecewise polynomial function defined by an array of grid points called knots. We choose our polynomials to be cubic functions to fit a smooth curve between successive pairs of knots. We specifically perform penalized spline regression to fit B-splines to  $v_r^2$ ,  $v_\phi^2$ ,  $v_\theta^2$  and perform a penalized spline density estimate to calculate  $\ln(\rho)$ . This computation is done with the routines implemented in Agama (Vasiliev 2019), and more information on penalized spline regression and penalized spline density estimation can be found in appendix A.

There is no universally optimal procedure for the determination of the knots’ parameters – their count, spacing, minimum bound, and maximum bound. Diakogiannis et al. (2017) use a sophisticated but more computationally intensive evolutionary modeling algorithm to optimize and adapt their choice of knots when applying the JEAnS code to dwarf spheroidal galaxies. In contrast, we empirically determine our knot configurations, placing significant emphasis on keeping their parameters as consistent as possible to mitigate overfitting.

In these fits, we use logarithmically spaced knots in radius (equally spaced in  $\ln r$ ) because we find that this best captures the radial distribution of tracer particles. The number of knots varies from  $N_{\text{knots}} = 6$  for datasets described in Section 3.1 to  $N_{\text{knots}} = 4$  for the datasets described in Sections 3.2 and 3.3. The minimum and maximum radial bound of the knots is similarly dataset dependent, but we have arranged them such that the knots are identical across all datasets presented in a given figure. When constructing our B-spline fits, we only use tracers whose radii lie within the extent of the knots. This is not required when using B-splines, but we find that this produces the best results, particularly at large radii.

B-splines are an attractive choice for multiple reasons: namely, they provide analytical derivatives and non-parametric fits. The radial derivative terms in Eq. 1 on  $\ln(\rho)$  and  $\ln(v_r^2)$  are computed trivially by utilizing the B-splines’ analytical derivatives. Further, the B-spline’s non-parametric nature means that they do not restrict the velocity or density profiles to a specific shape as power-laws would.

Each of the mock datasets described in Section 3 are provided as input to our routine and the resulting  $M(< r)$  is computed using Eq. 1. This estimated mass profile is compared with the system’s true mass profile over the relevant radial range. The true mass profile is computed for the entire radial extent of the system by cumulatively summing the masses of all particles over their Galactocentric radius.

## 3 MOCK DATA

We now describe the various mock datasets used to validate our B-spline-based routine.

In Section 3.1 we describe the equilibrium mock datasets that were generated using Agama. These come in two categories: “Halo-alone” models (hereafter HA-models) and “Halo-Disk-Bulge” models (hereafter HDB-models). The purpose of these tests is verify the accuracy of the routine and assess the effects of realistic deviations from sphericity and, in the case of the HDB-models, the effect of contamination of halo tracers by disk particles.

In Section 3.2 we describe more realistic mock datasets generated from star particles drawn from three galaxies ( $m12f$ ,  $m12i$ , and  $m12m$ ) from the Latte cosmological hydrodynamic zoom-in simulations (Wetzel et al. 2016). These models introduce dramatically more complexity and realism, including the challenge of accurately separating disk and halo stars, and the effects of halo substructure (streams and clustered halo stars) and disequilibrium arising from tidal interactions with nearby satellites.

In Section 3.3 we describe how we impose observational selection functions and errors onto the mocks described in 3.2. The selection functions are the fiducial DESI footprint and a Gaia-like magnitude limit, and errors are present on proper motions, line-of-sight velocities, and heliocentric distances. These mocks also introduce a dependence on the choice of solar position, which we vary between three locations in the disk.

### 3.1 Self-consistent equilibrium mock datasets

The *Agama* dynamics package is used to generate self-consistent equilibrium distribution functions with all particles having the same mass. In practice, the halo tracers provided to the Jeans routine are dark matter particles.

The HA-models contain only a spheroidal halo with axis ratio  $q$  varying from  $q = 1.0$  (spherical) to  $q = 0.6$ . *Agama* uses the double-power-law distribution function model from Posti et al. (2015) in which we tailor the mixing coefficients to produce the desired oblate axisymmetric shape. Changing the mixing coefficients also has the minor side-effect of unpredictably changing the velocity anisotropy. We verify the system has the desired axis ratio  $q$  by measuring it with E1 method described in Zemp et al. (2011). These halos are initialized with *Agama*'s Spheroid mass density function, with parameters  $(\alpha, \beta, \gamma) = (1, 4, 1)$ . With this specific choice of parameters, the Spheroid density function is exactly equivalent to the Hernquist (1990) density function. Once generated, the halo mass and scale radius are scaled to  $M_{\text{halo}} = 1.3 \times 10^{12} M_{\odot}$  and  $R_{\text{halo}} = 25$  kpc, roughly matching the Galactic virial mass and halo density break radius quoted in Bland-Hawthorn & Gerhard (2016).

The HDB-models have a spherical halo, a disk, and a spherical bulge. The HDB's initial halo density function adds an exponential cutoff to the HA density function but scales to the same mass  $M_{\text{halo}} = 1.3 \times 10^{12} M_{\odot}$  with the same scale radius  $R_{\text{halo}} = 25$  kpc. The disk is a combination of the thin and thick disk parameters quoted in Bland-Hawthorn & Gerhard (2016). Summing the thin and thick disk's stellar mass yields  $M_{\text{disk}} = 4.1 \times 10^{10} M_{\odot}$  and a mass weighted average of the radial scalelengths and scaleheights for the thin and thick disks yields  $R_{\text{disk}} = 2.5$  kpc and  $h_{\text{disk}} \approx 375$  pc<sup>1</sup>. The HDB's disk is initialized with *Agama*'s Disk density function (a double-exponential) with  $R_{\text{disk}}$  and  $h_{\text{disk}}$ . The spherical bulge mimics the axisymmetric bulge in McMillan 2017 by scaling with total mass  $M_{\text{bulge}} = 8.9 \times 10^9 M_{\odot}$ , scale radius  $R_{\text{bulge}} = 0.075$  kpc, and *Agama*'s Spheroid density function with  $(\alpha, \beta, \gamma) = (1, 1.8, 0)$ .

While we create variations of the HA-models with different axis ratios, the HDB-model halo remains spherical. However, we make additional HDB-models with other alterations. In one variant, we inject a random sample containing 1/4 of the system's disk particles into the halo tracer population to test the routine's sensitivity to contamination of the halo tracer population with disk particles.

<sup>1</sup> The exact mass-weighted value of the thin and thick disk scaleheights is 387 pc but we use 375 pc to generate our mocks. The difference is within the uncertainties of the current scale height measurements.

In another mock, we assess the effect of a radially varying velocity distribution for halo particles by creating a halo particle distribution with an anisotropic Osipkov-Merritt velocity dispersion profile (Osipkov 1979, Merritt 1985a, Merritt 1985b), using  $\beta(r = 0) = 0.2$  and anisotropy radius  $r_a = 100$  kpc.

The HA-models and HDB-models are generated with 300,000 particles in their halos, all of which are used in the Jeans estimation. The HDB-model with disk contamination has an additional 40,000 particles used as input, which come from the 160,000 particles comprising its disk.

### 3.2 Mock datasets from cosmological hydrodynamic simulations

The Latte cosmological simulations are performed using mesh-free hydrodynamics from the GIZMO code (Hopkins 2015) and the Feedback in Realistic Environments model 2 (FIRE-2) (Hopkins et al. 2014; Wetzel et al. 2016). The Latte simulations are a collection of simulated Milky Way-mass galaxies at uniquely high resolution that are well suited for generating mock Milky Way-like datasets. We use the  $z = 0$  public snapshots of the  $m12f$ ,  $m12i$ , and  $m12m$  galaxies released with the *Ananke* synthetic Gaia catalog (Sanderson et al. 2020). We create datasets for each galaxy by selecting only the star particles with metallicities  $[\text{M}/\text{H}] \leq -1.5$ . This is a proxy for selecting metal-poor accreted halo stars since the public FIRE snapshots do not contain information about whether stars were accreted from satellites or formed *in situ* in the host galaxy. Since information about whether a star was accreted is also not available *a priori* in real data this provides an additional level of realism to our mock datasets. We calculate the  $[\text{M}/\text{H}]$  ratio and the Hydrogen mass fraction ( $X$ ) from the mass fractions of He ( $Y$ ) and metals ( $Z$ ) provided in the public Latte data as follows:

$$X = 1 - (Y + Z) \quad (3)$$

$$[\text{M}/\text{H}] = \log(Z/X) - \log(Z_{\odot}/X_{\odot}) \quad (4)$$

These mocks and the mocks in Section 3.3 operate under the simplification considering each Latte star particle to be a single tracer star. We consider these stars to be RR-Lyrae stars (RRLs) because they serve as standard candles. Each metal poor halo star particle ( $\sim 7070 M_{\odot}$ ) can be considered to represent the entire stellar population of a single star cluster. Further, it is generally considered reasonable to assume each star cluster of this mass would result in at most a single RRL. After selecting tracers by metallicity, there are approximately 200,000 remaining particles in each dataset ( $m12f$ : 188,577,  $m12i$ : 231,599,  $m12m$ : 218,891).

The three Latte galaxies, while all Milky Way mass, have varying amounts of halo substructure in the form of tidal streams and shells and a fairly broad range of satellite properties. Their halos are also in different states of dynamical equilibrium, have a range of realistic morphologies, have experienced a varied assembly history (Garrison-Kimmel et al. 2018), and have realistic satellite populations (Wetzel et al. 2016). In addition, these galaxies have halo stars with moderate amounts of tangential streaming motions (azimuthal  $\overline{v_{\phi}}$  and polar  $\overline{v_{\theta}}$ ). Figure 1 shows heatmaps of stars in these three galaxies from the  $z = 0$  snapshot color coded by  $\overline{v_{\theta}}$  (upper row) and  $\overline{v_{\phi}}$  (lower row). In addition to substructure in the form of streams, shells and over-densities, there is recent observational evidence that the Milky Way disk is moving with respect to the outer stellar halo as a result of on-going tidal interaction with the Large Magellanic Cloud (LMC) (Petersen & Peñarrubia 2021) and that the stellar halo

density distribution shows a distinct bi-symmetry that is also likely to be due to the LMC (Conroy et al. 2021). Recent studies that account for the disequilibrium arising from the LMC suggest that dynamical models that do not account for it correctly can overestimate the total mass of the Galaxy (Erkal et al. 2020; Deason et al. 2021). In this paper we do not consider the effects of a significant perturbations, such as that of the LMC.

### 3.3 Adding observationally motivated selection functions and errors

One must choose a solar position or a Local Standard of Rest (LSR) before imposing observationally motivated selection functions or magnitude dependent errors on mock halo stars from simulated galaxies. As part of the construction of the *Ananke* synthetic Gaia catalog (Sanderson et al. 2020), the Latte collaboration provided three LSRs with precomputed solar velocities for each of the three galaxies  $m12f$ ,  $m12i$ , and  $m12m$ . The first LSR position, LSR0, places the sun at  $(x, y, z) = (0, 8.2, 0)$  kpc and the other two, LSR1 and LSR2, are located at the same Galactocentric radius, in the galactic midplane, and rotated  $120^\circ$  in either direction from LSR0. We transform the coordinates of the star particles to ICRS and Galactic coordinates frame with Astropy (Astropy Collaboration et al. 2013, 2018) when applying the following observational selection functions and errors.

We start imposing the magnitude selection function by calculating each star particle’s Gaia  $G$ -band apparent magnitude. We assume that each star particle is a single RR-Lyrae star with  $G$ -band absolute magnitude  $M_G = 0.58$ , as quoted for metal-poor RRL in Iorio & Belokurov (2019). We then calculate the  $G$ -band apparent magnitudes with the assumed  $M_G$ , each star particle’s distance modulus  $\mathcal{D}$ , and heliocentric distance  $d$ .

$$\mathcal{D} = 5 * \log(d/\text{kpc}) + 10 \quad (5)$$

$$G = M_G + \mathcal{D} \quad (6)$$

Once we have  $G$  for each mock RRL, we select only those with magnitudes  $G < 20.7$ <sup>2</sup>

We then impose a sky-selection function similar to the fiducial DESI footprint with a declination cut  $-35^\circ \leq \delta \leq 90^\circ$  and a galactic latitude cut  $30^\circ \leq |b| \leq 90^\circ$ . Because the resulting sample is dependent on the LSR, imposing these selection functions produces 3 datasets for each Latte galaxy or 9 total error-free datasets with observational selection functions and roughly 30,000–40,000 remaining particles each.

We impose realistic Gaia- and DESI-like errors on the phase space coordinates of the 9 datasets with observational selection functions to assess how their presence biases the Jeans estimate. The Gaia-like errors are present on the proper motions ( $\mu_{\alpha*}$ ,  $\mu_\delta$ ) and heliocentric distances ( $d$ ), and the DESI-like errors are present on the line-of-sight velocities ( $v_{\text{los}}$ ). Lastly, we assume the right ascension ( $\alpha^*$ ) and the declination ( $\delta$ ) have no error.

The errors are randomly generated from independent Gaussian distributions on each quantity for each star particle with standard deviations determined by characteristic or expected errors for Gaia and DESI observations. The Gaia heliocentric distance error is the same for all star particles at  $\sigma_d = 0.1107$  (Hattori et al. 2020). The Gaia proper motion errors ( $\sigma_{\mu\alpha*}$ ,  $\sigma_{\mu\delta}$ ) are a function of  $G$

<sup>2</sup> Gaia Expected Science Performance <https://www.cosmos.esa.int/web/gaia/science-performance>

– and also  $d$  by Eq. 5 and 6 – and come from PyGaia<sup>3</sup>. PyGaia allows for calculations of proper motion errors from Gaia EDR3, and expected proper motion errors from DR4 (containing the first 5 1/2 years of observations) and DR5 (containing up to 10 years of observations). Figure 2 shows the 3D velocities squared of tracers in  $m12f$  after imposing the observational selection functions and errors at LSR0 over the different Gaia data releases. Because the Gaia proper motion errors scale up with  $d$ , stars at large radii tend to be given very large velocities, often above a visual escape velocity curve. These extremely high velocity stars will cause Jeans modeling to blow up at large radii. This is discussed further in Section 4.3 and Figure 5.

Finally, we use the nominal line-of-sight velocity error expected from the DESI Milky Way survey  $\sigma_{v_{\text{los}}} = 2$  km/s (Allende Prieto et al. 2020).

## 4 RESULTS

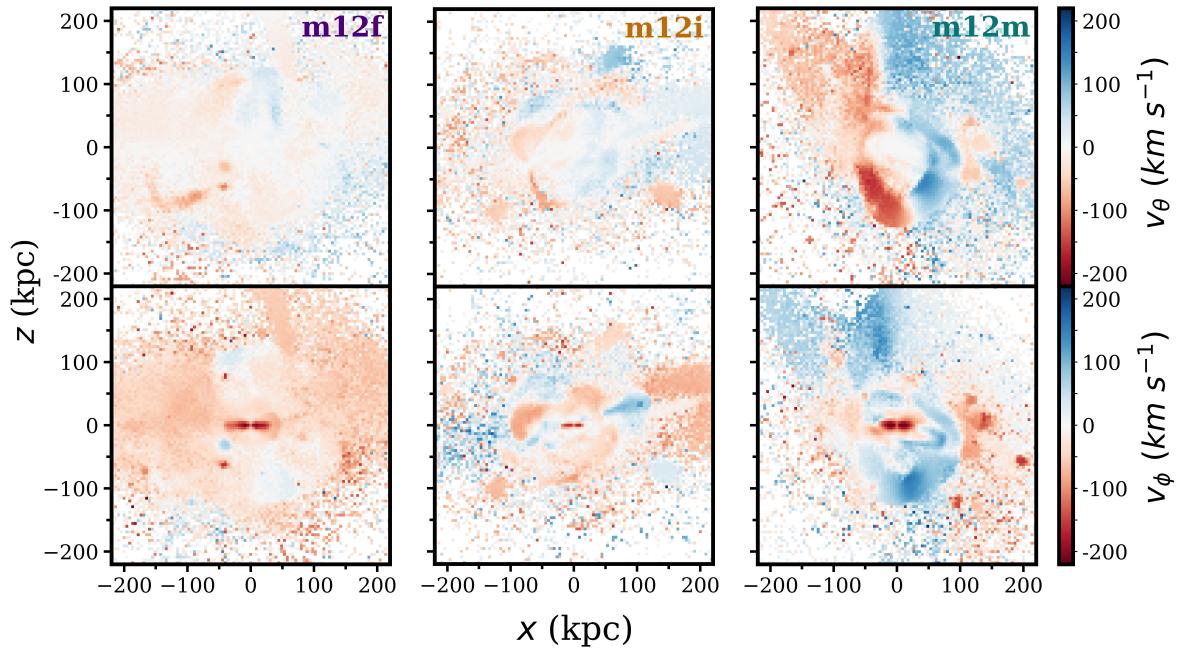
The estimated mass profile from each mock is judged for accuracy by computing the percentage error relative to the corresponding true mass profile. Figures 3–6 show the estimated (solid line) and true mass profiles (dashed line) for a set of mock datasets and the error as a function of radius. The mocks described in each subsection of Section 3 correspond to a matching subsection in Section 4.

### 4.1 Initial tests with self-consistent equilibrium mocks

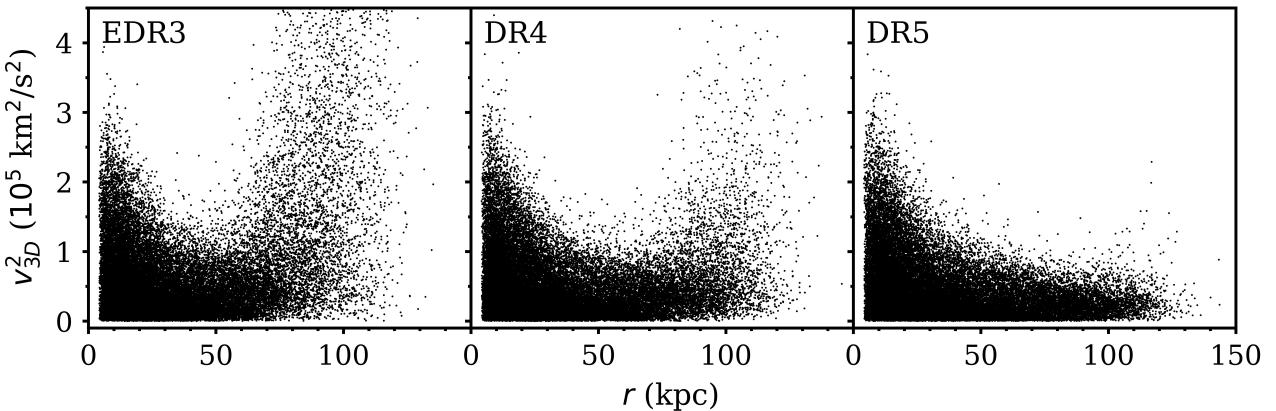
In this section we describe the results from tests with the self-consistent equilibrium mock datasets described in Section 3.1. Figure 3 shows the results for the three HA-models, whose B-spline fits are each run with  $N_{\text{knots}} = 6$  knots logarithmically spaced between  $r_{\min} = 1$  kpc and  $r_{\max} = 70$  kpc. The blip in error near the beginning of the curve is merely an artifact of small number statistics at small radii. Aside from this, the progression of flattening from  $q = 1.0$  to  $q = 0.6$  only has a very small effect on the accuracy of the estimation. The  $q = 0.6$  mock oscillates around the true profile and is clearly less stable than the  $q = 1.0$  and  $q = 0.8$  mocks. For all three, the error is less than 4% beyond  $r = 5$  kpc. We also run tests on HA models with  $q = 0.9$  and  $q = 0.7$ , which perform in-line with the three models shown here.

Figure 4 shows the results of tests with the original HDB-model and its variants. For these models, the B-spline fits are all performed with  $N_{\text{knots}} = 6$  knots logarithmically spaced between  $r_{\min} = 15$  kpc and  $r_{\max} = 100$  kpc. The error in the estimation on the HDB-models is within 10% at all radii. The error is extremely low ( $\leq 1\%$ ) between  $25 < r \text{ kpc} < 60$  for all HDB-models. Notably, the accuracy of the HDB variants is excellent despite the significant modifications to the original HDB dataset. We also run an additional test for the disk contamination variant, setting  $r_{\min} = 5$  kpc to include more of the injected disk particles in the B-spline fitting. This yields similarly excellent results with error  $\leq 10\%$  at all radii, but we show only the  $r_{\min} = 15$  kpc test for consistency with other tests.

<sup>3</sup> PyGaia: <https://github.com/agabrown/PyGaia> written by A.G. Brown



**Figure 1.** 2-dimensional heatmaps colored by  $\langle v_\theta \rangle$  and  $\langle v_\phi \rangle$  of three Latte galaxies  $m12f$  (left column),  $m12i$  (center column), and  $m12m$  (right column). Each of the three show significant streaming motion and substructure as deviations from dynamical equilibrium - breaking the assumptions of the spherical Jeans equation. Of the three,  $m12m$  has the most dramatic streaming motion, so we expect our mass profile estimation to perform the worst on the mock dataset derived from it.

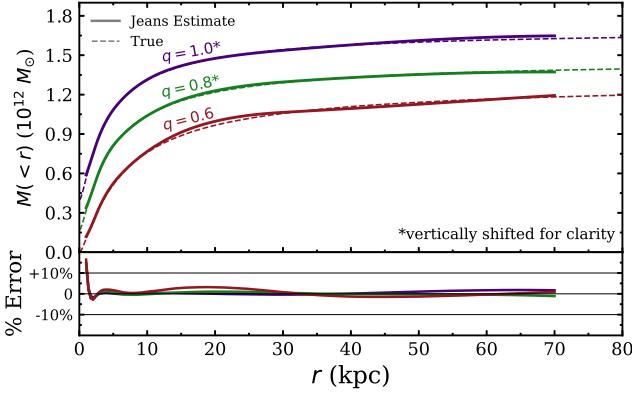


**Figure 2.** 3D velocities squared ( $v_{3D}^2$ ) of tracers in the observational mock of  $m12f$  at LSRO with errors on proper motion ( $\mu\alpha^*, \mu\delta$ ), heliocentric distance ( $d$ ), and line-of-sight velocity ( $v_{los}$ ). From left to right:  $\sigma_{\mu\alpha^*}$  and  $\sigma_{\mu\delta}$  characteristic of Gaia data release EDR3, DR4, and DR5 respectively.  $\sigma_{\mu\alpha^*}$  and  $\sigma_{\mu\delta}$  increase with dimmer  $G$ -band apparent magnitude and thus with larger  $d$ , so stars at large galactocentric radius  $r$  tend to have larger  $\sigma_{\mu\alpha^*}$ , and  $\sigma_{\mu\delta}$ . When errors are imposed, their  $v_{3D}^2$  increase dramatically, making them unbound and unfit for use in Jeans modeling. The amount of unbound stars in each dramatically decreases with progressive Gaia data releases as the expected  $\sigma_{\mu\alpha^*}$  and  $\sigma_{\mu\delta}$  decrease.

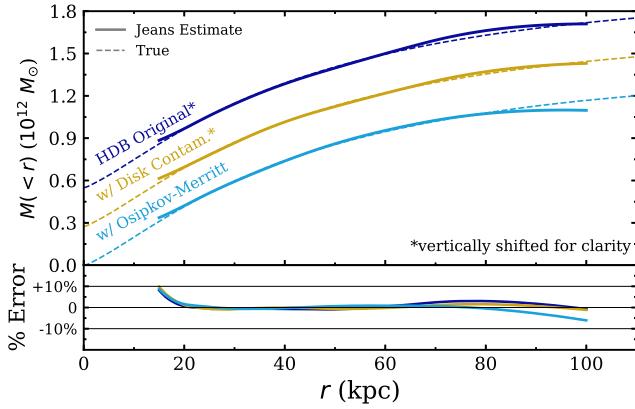
## 4.2 Early tests with metal-poor particles from cosmological simulations

In this section we describe tests of our Jeans modeling routine with mock halo stars drawn from the Latte (FIRE-2) cosmological simulations,  $m12f$ ,  $m12i$ , and  $m12m$ . The datasets we generate using these three galaxies select a population of star particles with  $[M/H] < -1.5$  meant to represent a halo RR Lyrae population. The B-spline fits on these datasets are each run with  $N_{\text{knots}} = 4$  knots logarithmically spaced between  $r_{\min} = 15$  kpc and  $r_{\max} = 100$  kpc. We decrease the number of knots from the tests with equilibrium mocks (Section 4.1) to increase the smoothing in the fits. Figure 5

shows the results of the tests for these three galaxies when all metal poor stars within the radial region above are selected. The accuracy with which our Jeans modeling routine recovers the input mass distribution is, of course, worse for these mocks than for the idealized mocks in the previous section but over most of the radial range the error remains  $\leq 10\%$ , increasing to a maximum error of  $\sim 14\%$  at 100 kpc. This is despite the fact that all the Latte galaxies we use have triaxial dark matter halos, show significant halo substructure (shells and tidal streams), show varying amounts of tangential streaming (see Fig. 1 of  $\langle v_\phi \rangle$ ), and have satellites within 100 kpc.



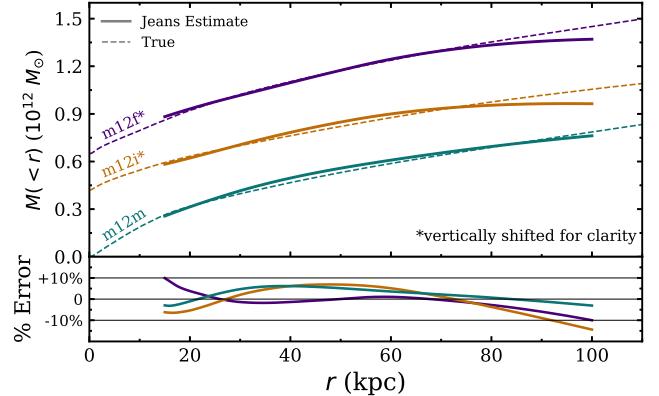
**Figure 3.** Mass profiles of three “Halo-alone” mock datasets with increasing flattening quantified by the axis ratio  $q = 1.0$  (spherical), 0.8, and 0.6 (oblate axisymmetric). The flattened models are to test for accuracy when breaking the assumption of sphericity made by the Jeans equation. The error due to flattening appears as instability in the  $q = 0.6$  mock, which yields less than 4% error at most radii. The  $q = 1.0$  and  $q = 0.8$  mocks give even smaller errors at most radii. The deviation at small radii in all three mocks is inconsequential and most likely due to noise in the simulation.



**Figure 4.** Mass profiles of the “Halo-Disk-Bulge” mock datasets including variants with contamination from disk particles and a radially varying anisotropy profile. All variants perform well ( $\leq 10\%$  error) and very consistently. The inclusion of a disk and bulge into the system’s potential to break Jeans modeling’s assumption of sphericity has little effect on the accuracy. Similarly, there is little effect on the accuracy due to contamination from disk particles with large tangential velocities or a radially varying anisotropy profile.

#### 4.3 Tests with observational errors and selection functions on cosmological simulations

All our previous tests have assumed that the phase space coordinates available to our model have no errors, a simplification that allows us to assess the systematic error introduced by the code itself. In reality, all observations come with errors. In this section, we describe tests on the observationally motivated mock datasets described in Section 3.3. These mocks best capture the data we expect to obtain from Gaia and DESI by imposing the observational selection functions and errors of these surveys onto the halo star particles from the Latte galaxies  $m12f$ ,  $m12i$ , and  $m12m$ . The B-spline fits on these datasets are run with  $N_{\text{knots}} = 4$  knots logarithmically spaced



**Figure 5.** Mass profiles of the mock datasets generated from star particles with  $[\text{M}/\text{H}] < -1.5$  in the Latte  $m12f$ ,  $m12i$ , and  $m12m$  cosmological simulations. All three resulting mass profiles are within 15% of the truth at all radii. These early tests with cosmological simulations illustrate our B-spline routine’s ability to smooth out disequilibrium like tidal streams, halo substructure, and tangential streaming motion to accurately recover the underlying mass distribution.

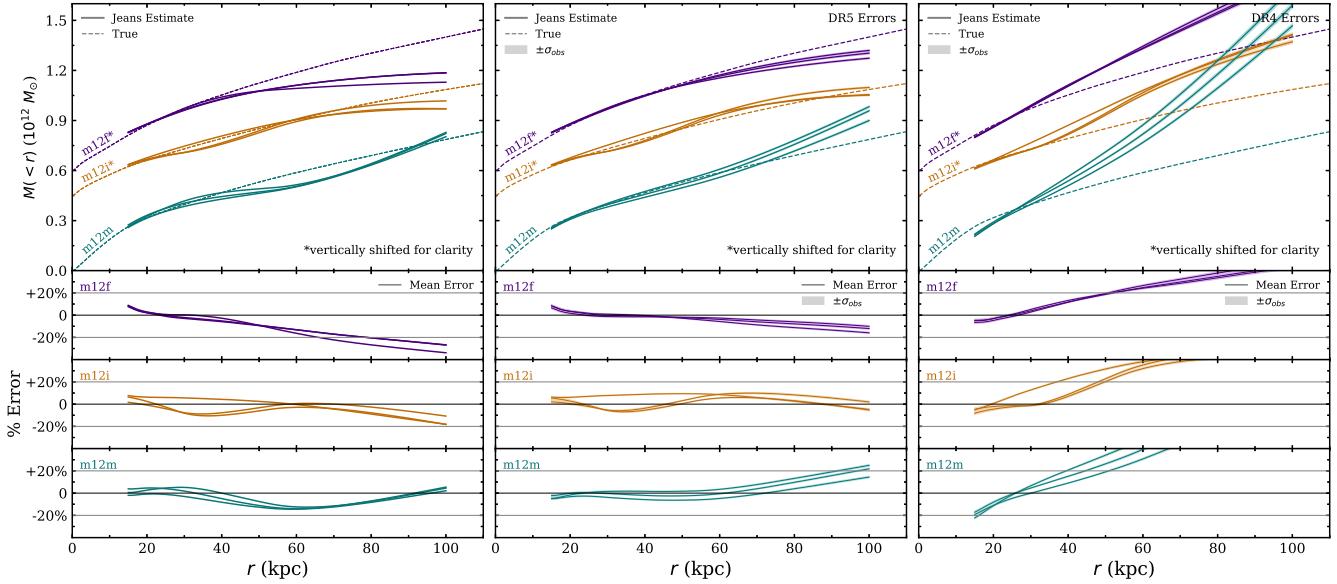
between  $r_{\min} = 15$  kpc and  $r_{\max} = 100$  kpc. Figure 6 shows the results from the error-free mock data (left panel), mock data with expected Gaia DR5 errors (middle panel), and expected Gaia DR4 errors (right panel). As described in Section 3.3 we consider three LSRs for each galaxy, shown in Figure 6 by the three solid curves of the same color. The datasets with errors imposed have 50 independent trials of random errors. The standard deviation at each radial point of the 50 trials is shown as the width of the bands above and below each solid curve. There are a few takeaways from Figure 6:

First, with the expected Gaia DR4 errors, the error in the estimated enclosed mass rises rapidly and is above 20% by  $r \sim 50$  kpc in all models and is even more significant at larger radii. As illustrated in Figure 2, the relatively large errors on proper motions in DR4 make too many very high velocity stars, which causes the Jeans estimation to blow up at large radii, especially in the case of  $m12m$ . This suggests that reducing the maximum extent of the knots to  $r_{\max} \sim 50$  kpc would include many fewer unbound stars in the B-spline fits, which would produce a more accurate estimate.

Second, all of the mocks in Figure 6 demonstrate that the spread in the mass profiles due to the position of the LSR is quite large, reaching upwards of 15% difference at some radii in the LSRs of  $m12i$ .

Third, the error-free mocks and Gaia DR5 mocks demonstrate that realistic expected errors will get small enough such that they do not introduce much deviation of the estimated mass profile from the error-free case. The Gaia DR5 mocks decrease the error on the estimated enclosed masses from the DR4 mocks to no more than 30% out to 100 kpc for all three models. It is also apparent that the spread from many trials of random error generation on the same setup (shown by the bands above and below the solid lines) is very small.

We do not show models with Gaia DR3 errors since they are worse, as would be expected, than results with Gaia DR4 errors.



**Figure 6.** Comparison of true and recovered enclosed mass estimates for the three Latte galaxies  $m12f$ ,  $m12i$ , and  $m12m$  with observational selection functions and errors. Error-free models are shown in the left column, Gaia DR5 error mocks are shown in the center and Gaia DR4 error mocks are shown in the right column. For each galaxy we consider three LSR positions, and the mass estimates for these are shown as three lines of the same color. The DR4 errors on proper motion tend to create very high velocity stars, which cause the Jeans estimate to blow up at large radii. The DR5 errors are much smaller and do not deviate significantly from the error-free case.

## 5 SUMMARY & CONCLUSION

We have presented a new non-parametric spherical Jeans modeling code that uses B-splines to fit the enclosed mass distribution within the Milky Way halo. We present tests of this code with a variety of mock datasets to assess the effects of breaking various assumptions made in deriving the spherical Jeans equations: that the underlying mass distribution is spherical, that tangential streaming motions are absent ( $\langle v_\phi \rangle = \langle v_\theta \rangle = 0$ ), and that the potential is in global dynamical equilibrium. Our main results are listed below:

- With our simplest “Halo-alone” models we show that going from a spherical mass distribution to a flattened mass distribution with  $q = 0.6$  only slightly increases the error on the estimate of enclosed mass but the error is within 4% beyond  $r = 5$  kpc (see Fig. 3).

- Our “Halo-disk-bulge” (HDB) models are constructed to resemble the Milky Way halo, disk, and bulge but consider only a spherical halo, but with both isotropic and anisotropic velocity distributions for halo particles. One of the mock data sets also considers “contamination” from disk stars. Our Jeans modeling code performs extremely well with these HDB models with maximum errors on the enclosed mass out to 100 kpc not greater than 10% in all cases and less than 1% at most radii (see Fig. 4).

- The mock datasets with the next level of complexity are constructed from three galaxies from the Latte FIRE-2 simulations,  $m12f$ ,  $m12i$ , and  $m12m$ . These galaxies were constructed to be Milky Way analogs and contain realistic amounts of substructure and display moderate amounts of disequilibrium (see Fig. 1). With error-free data over the whole sky, the code still gives better than 15% error out to 100kpc in all three Latte galaxies (see Fig. 5).

This suggests that our Jeans modeling code is very robust and will introduce little systematic error when applied to real data in the future.

- The final datasets considered the effects of observational selec-

tion functions and errors from Gaia and the DESI survey. The Gaia component included a  $G$ -band magnitude limit and expected proper motion errors from the end of DR4 and DR5. The DESI component included its observational footprint and errors on  $v_{\text{los}} = 2 \text{ km s}^{-1}$ . These errors are imposed on the same three Latte (FIRE-2) galaxies and consider three possible positions for the LSR. The variation resulting from the LSR position alone reaches upwards of 15% at some radii in  $m12i$ . With proper motion errors expected of Gaia DR4, the error on the mass estimate is small at small radii but very quickly rises above 20% by  $r = 50$  kpc in all cases, growing to more than 40% by 100 kpc (see Fig. 6). In contrast, with proper motion errors expected of Gaia DR5, the errors on the mass estimate drop dramatically and are within 30%, often less, all the way out to 100 kpc (a radial range that includes some satellites and considerable amounts of substructure from tidal streams, see Fig. 1).

The spherical Jeans equation method is a fairly simple, but popular method for recovering the mass distribution of (nearly) spherical potentials. A fundamental feature of this method is that it depends on the radial derivatives of the tracer mass distribution as well as the radial velocity dispersion profile. Since derivatives of observational data profiles are extremely noisy (Kafle et al. 2018), most authors model the tracer density and velocity dispersion profiles with analytic functions (often power-laws) to enable convenient computation of the derivatives. In our implementation of this method, we have used B-splines to estimate the density profile and velocity dispersion profiles of the tracer particles, enabling us to avoid the need to assume a functional form but still allowing for the rapid computation<sup>4</sup> of the enclosed mass profile. Although the tests we carry out with mock data add noise with magnitudes comparable

<sup>4</sup> The routine’s run time is on the order of <10 seconds for the observational mocks.

with the expected observational errors to the line-of-sight velocities and proper motions, we do not rigorously evaluate how those errors translate to uncertainties on the galactocentric phase space coordinates, nor do we propagate those errors forward to derive uncertainties on the enclosed mass distribution. Rather, the tests with mock datasets allow us to assess the bias in the derived mass distribution, that arises from the limited size of the survey volume and proper motion errors.

Other methods such as distribution function fitting (Eadie et al. 2017, 2018; Eadie & Jurić 2019; Hattori et al. 2020; Deason et al. 2021) use Bayesian MCMC analysis to properly account for the errors on the mass estimates arising from the error on the data. These methods, while much more sophisticated, are computationally and technically challenging. Our goal in developing this non-parametric B-spline Jeans modeling code and testing it with mock data is to provide the community with a fast and easy-to-use code with well quantified systematic errors. The code is written in Python and uses the *Agama* library (Vasiliev 2019). It can be applied to a sample of observed 6D coordinates for tracer stars to construct B-spline estimates of the radial density profile, three velocity dispersion profiles and the velocity anisotropy profile, which are then used to evaluate Eq. 1.

In future, we aim to apply this code to Gaia DR4 and Gaia DR5 data with radial velocities from DESI, as well as publicly available data from WEAVE (Dalton et al. 2014) and 4MOST (de Jong et al. 2012).

## 6 ACKNOWLEDGEMENTS

We thank members of the stellar halos group at the University of Michigan for stimulating discussion. We especially thank Eric Bell and Kohei Hattori for assistance and advise. M.V. is supported in part by NASA grants NNX15AK79G and 80NSSC20K0509 and a MICDE Catalyst Grant.

In addition to *Agama* (Vasiliev 2019) we used the following software packages: Astropy (Astropy Collaboration et al. 2013, 2018), matplotlib (Hunter 2007), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), PyGaia (<https://github.com/agabrown/PyGaia>)

## REFERENCES

- Allende Prieto C., et al., 2008, *Astronomische Nachrichten*, **329**, 1018  
 Allende Prieto C., et al., 2020, *Research Notes of the American Astronomical Society*, **4**, 188  
 Astropy Collaboration et al., 2013, *A&A*, **558**, A33  
 Astropy Collaboration et al., 2018, *AJ*, **156**, 123  
 Battaglia G., et al., 2005, *MNRAS*, **364**, 433  
 Binney J., 1980, *MNRAS*, **190**, 873  
 Binney J., Tremaine S., 2008, *Galactic Dynamics*: Second Edition. Princeton University Press  
 Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, **54**, 529  
 Bullock J. S., Johnston K. V., 2005, *ApJ*, **635**, 931  
 Conroy C., Naidu R. P., Garavito-Camargo N., Besla G., Zaritsky D., Bonaca A., Johnson B. D., 2021, arXiv e-prints, p. [arXiv:2104.09515](https://arxiv.org/abs/2104.09515)  
 Cooper A. P., et al., 2011, *ApJ*, **743**, L21  
 DESI Collaboration et al., 2016a, ArXiv e-prints, arXiv:1611.00036.  
 DESI Collaboration et al., 2016b, ArXiv e-prints, arXiv:1611.00037.  
 Dalton G., et al., 2014, Project overview and update on WEAVE: the next generation wide-field spectroscopy facility for the William Herschel Telescope. SPIE, p. 91470L, doi:[10.1117/12.2055132](https://doi.org/10.1117/12.2055132)  
 De Silva G. M., et al., 2015, *MNRAS*, **449**, 2604  
 Deason A. J., et al., 2021, *MNRAS*, **501**, 5964  
 Diakogiannis F. I., Lewis G. F., Ibata R. A., Guglielmo M., Kafle P. R., Wilkinson M. I., Power C., 2017, *MNRAS*, **470**, 2034  
 Eadie G., Jurić M., 2019, *ApJ*, **875**, 159  
 Eadie G. M., Springford A., Harris W. E., 2017, *ApJ*, **835**, 167  
 Eadie G., Keller B., Harris W. E., 2018, *ApJ*, **865**, 72  
 El-Badry K., Wetzel A. R., Geha M., Quataert E., Hopkins P. F., Kereš D., Chan T. K., Faucher-Giguère C.-A., 2017, *ApJ*, **835**, 193  
 Erkal D., Belokurov V. A., Parkin D. L., 2020, *MNRAS*, **498**, 5574  
 Evslin J., Del Popolo A., 2017, *ApJ*, **841**, 90  
 Gaia Collaboration et al., 2016, *A&A*, **595**, A2  
 Gaia Collaboration et al., 2018, *A&A*, **616**, A1  
 Garrison-Kimmel S., et al., 2018, *MNRAS*, **481**, 4133  
 Genina A., et al., 2020, *MNRAS*, **498**, 144  
 Gilmore G., et al., 2012, *The Messenger*, **147**, 25  
 Gnedin O. Y., Brown W. R., Geller M. J., Kenyon S. J., 2010, *ApJ*, **720**, L108  
 Harris C. R., et al., 2020, *Nature*, **585**, 357–362  
 Hattori K., Valluri M., Loebman S. R., Bell E. F., 2017, *ApJ*, **841**, 91  
 Hattori K., Valluri M., Vasiliev E., 2020, arXiv e-prints, p. [arXiv:2012.03908](https://arxiv.org/abs/2012.03908)  
 Hernquist L., 1990, *ApJ*, **356**, 359  
 Hopkins P. F., 2015, *MNRAS*, **450**, 53  
 Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, **445**, 581  
 Hu J., Jiang B., 2005, in J. Mikolajewska & A. Olech ed., *American Institute of Physics Conference Series* Vol. 752, *Stellar Astrophysics with the World's Largest Telescopes*. pp 84–87, doi:[10.1063/1.1893338](https://doi.org/10.1063/1.1893338)  
 Hunter J. D., 2007, *Computing in Science and Engineering*, **9**, 90  
 Iorio G., Belokurov V., 2019, *MNRAS*, **482**, 3868  
 Jeans J. H., 1915, *MNRAS*, **76**, 70  
 Johnston K. V., Bullock J. S., Sharma S., Font A., Robertson B. E., Leitner S. N., 2008, *ApJ*, **689**, 936  
 Kafle P. R., Sharma S., Robotham A. S. G., Elahi P. J., Driver S. P., 2018, *MNRAS*, **475**, 4434  
 Lindegren L., et al., 2016, *A&A*, **595**, A4  
 McMillan P. J., 2017, *MNRAS*, **465**, 76  
 Merritt D., 1985a, *AJ*, **90**, 1027  
 Merritt D., 1985b, *MNRAS*, **214**, 25P  
 Osipkov L. P., 1979, *Pisma v Astronomicheskii Zhurnal*, **5**, 77  
 Perryman M. A. C., et al., 2001, *A&A*, **369**, 339  
 Petersen M. S., Peñarrubia J., 2021, *Nature Astronomy*, **5**, 251  
 Posti L., Binney J., Nipoti C., Ciotti L., 2015, *MNRAS*, **447**, 3060  
 Read J. I., Steger P., 2017, *MNRAS*, **471**, 4541  
 Sanderson R. E., et al., 2020, *ApJS*, **246**, 6  
 Steinmetz M., et al., 2006, *AJ*, **132**, 1645  
 Vasiliev E., 2018, arXiv e-prints, p. [arXiv:1802.08255](https://arxiv.org/abs/1802.08255)  
 Vasiliev E., 2019, *MNRAS*, **482**, 1525  
 Vasiliev E., Baumgardt H., 2021, arXiv e-prints, p. [arXiv:2102.09568](https://arxiv.org/abs/2102.09568)  
 Virtanen P., et al., 2020, *Nature Methods*, **17**, 261  
 Wang W., Han J., Cole S., More S., Frenk C., Schaller M., 2018, preprint, ([arXiv:1801.07373](https://arxiv.org/abs/1801.07373))  
 Wang W., Han J., Cautun M., Li Z., Ishigaki M. N., 2020, *Science China Physics, Mechanics, and Astronomy*, **63**, 109801  
 Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, *ApJ*, **827**, L23  
 Xue X. X., et al., 2008, *ApJ*, **684**, 1143  
 Zemp M., Gnedin O. Y., Gnedin N. Y., Kravtsov A. V., 2011, *ApJS*, **197**, 30  
 de Jong R. S., et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* Vol. 8446, *Ground-based and Airborne Instrumentation for Astronomy IV*. SPIE, pp 252 – 266, doi:[10.1117/12.926239](https://doi.org/10.1117/12.926239)

## APPENDIX A: FITTING DENSITY AND VELOCITY PROFILES WITH B-SPLINES

In this section we provide a brief overview of the B-spline formalism in the context of fitting and density estimation. For a more complete mathematical background see Appendix A2 in Vasiliev (2018).

B-splines are a set of piecewise-polynomial basis functions  $B_j(x)$  defined by nodes of a grid  $x_k, k = 1..K$ . B-splines of polynomial degree  $N$  are nonzero on at most  $N + 1$  consecutive grid segments, their  $N - 1$ 'th derivatives are continuous at grid nodes, and a grid with  $K$  nodes ( $K - 1$  segments) generates  $K + N - 1$  basis functions. A B-spline representation of a curve is a linear combination of basis functions with some coefficients:  $f(x) = \sum_j A_j B_j(x)$ . The most familiar case is a histogram, which is nothing else than a 0th degree B-splines basis; however, since even the curve itself is discontinuous, it is not suitable for evaluating derivatives. In this study we use 3rd degree B-splines, which are mathematically equivalent to a clamped cubic spline curve (clamped meaning that its 2nd derivatives at endpoints are not necessarily zero).

In the present context, we use B-splines to estimate the density from discrete samples. Namely, given a set of points  $x_n$  with weights  $w_n, n = 1..N_{\text{data}}$ , we seek to construct a density function  $p(x)$  such that  $\int_a^b p(x) dx \approx \sum_n w_n \Gamma(x_n; a, b)$ , where the indicator function  $\Gamma(x; a, b) = 1$  when  $a \leq x \leq b$  and zero otherwise. The density is normalized so that the integral over the entire interval containing all points equals  $\sum_n w_n$ . It turns out that B-splines are very useful for representing the logarithm of  $p$ , ensuring the positivity constraint on density. After choosing the interval and grid nodes, the coefficients of the B-spline representation  $\ln p(x) = \sum_j A_j B_j(x)$  can be found by maximizing the penalized log-likelihood of the input sample:  $\{A_j\} = \arg \max \sum_n w_n \ln p(x_n) - \lambda \int [d^2 \ln p(x)/dx^2]^2 dx$ , where the second term biases the solution towards smooth curves, and the roughness penalty  $\lambda$  is automatically determined by cross-validation. With the optimal choice of  $\lambda$ , the number of grid nodes has little effect on the resulting curve (as long as it is sufficient to resolve key features in the distribution).

For the Jeans equation, we need the density and the kinetic energy, which are estimated as follows. The distribution of points in spherical radius (the number of points in the radial interval  $r..r + dr$ ) is  $dN(< r)/dr = 4\pi r^2 \rho(r)$ , and it is convenient to use  $\chi \equiv \ln r$  instead of radius. We construct a B-spline estimate for the logarithm of the density of points in  $\chi$ , using the masses  $m_n$  of particles in the input snapshot as the weights associated with points:  $f_1(\chi) \equiv \ln [dN/d \ln r] = \ln [4\pi r^3 \rho(r)]$ ; hence the mass density is given by  $\rho(r) = (4\pi)^{-1} \exp [f_1(\chi) - 3\chi]$ . Likewise, when using the masses multiplied by squared radial and tangential velocity components  $v_r^2, v_\theta^2 + v_\phi^2$  of each particle as the weights, we obtain the estimates of the corresponding terms in the kinetic energy tensor:  $f_{v_r^2}(\chi) \equiv \ln [4\pi r^3 \rho(r) v_r^2]$ , etc. The logarithmic derivative of the radial kinetic energy  $d \ln [\rho v_r^2]/d \ln r$  appears in the Jeans equation directly, and the other terms contains the ratios of  $\exp f_1, \exp f_{v_r^2}$ , etc.