

Quantifying the Milky Way, LMC and their interaction using all-sky kinematics of outer halo stars

Richard A. N. Brooks¹, Jason L. Sanders¹, Adam M. Dillamore¹, Nicolás Garavito-Camargo², Vedant Chandra³, Adrian M. Price-Whelan⁴, Phillip Cargile³

¹Department of Physics and Astronomy, University College London, London, WC1E 6BT, UK

²Department of Astronomy, University of Maryland, College Park, MD 20742, USA

³Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

⁴Center for Computational Astrophysics, Flatiron Institute, Simons Foundation, 162 Fifth Avenue, New York, NY 10010, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The recent pericentric passage of the Large Magellanic Cloud (LMC) has dislodged the Milky Way’s (MW) centre of mass, inducing dynamical disequilibrium, the *reflex motion*, in the kinematics of outer stellar halo stars. Using data out to 160 kpc from the combined H3+SEGUE+MagE outer halo survey, we constrain the mass of the MW and LMC, as well as the resulting reflex motion and the stellar halo velocity anisotropy. Using a suite of 32,000 rigid MW–LMC simulations, each with a MW stellar halo evolved to the present day in the combined MW–LMC potential, we perform Simulation Based Inference by training a neural posterior estimator on the means and dispersions of the radial and tangential velocities of stars from the combined H3+SEGUE+MagE outer halo sample. Relative to halo stars at 100 kpc, we find the magnitude of the reflex velocity to be $v_{\text{travel}} = 38.6^{+8.3}_{-7.8} \text{ km s}^{-1}$. Simultaneously, we determine the enclosed MW mass, $M_{\text{MW}}(< 50 \text{ kpc}) = 3.36 \pm 0.15 \times 10^{11} M_{\odot}$ and the enclosed LMC mass, $M_{\text{LMC}}(< 50 \text{ kpc}) = 8.76^{+1.94}_{-1.77} \times 10^{10} M_{\odot}$. Our results suggest that the total LMC mass must be at least $\sim 20\%$ that of the MW. The velocity anisotropy prior to the LMC’s infall is constrained to be $\beta_0 = 0.68 \pm 0.02$. Finally, we demonstrate that neglecting the LMC in models biases the estimated MW mass to prefer more massive values.

Key words: Galaxy: kinematics and dynamics – Galaxy: halo – Galaxy: evolution – Magellanic Clouds – software: machine learning – software: simulations

1 INTRODUCTION

The mass of the Milky Way (MW) is one of its most fundamental properties, with its value crucial to constructing dynamical models for our Galaxy. The MW mass has been estimated through a variety of techniques and visible tracers such as: stellar streams (e.g., Gibbons et al. 2014; Küpper et al. 2015; Malhan & Ibata 2019; Erkal et al. 2019a; Vasiliev et al. 2021; Dillamore et al. 2025), satellite galaxies and globular clusters (e.g., Zaritsky et al. 1989; Vasiliev 2019b; Correa Magnus & Vasiliev 2022), and MW halo stars (e.g., Deason et al. 2012; Williams & Evans 2015; Williams et al. 2017; McMillan 2017; Bird et al. 2022; Shen et al. 2022; Medina et al. 2025), as well as via dynamics-free approaches (e.g., Zaritsky & Courtois 2017). Most studies converge around constrained values for the virial mass of the MW being $M_{\text{MW}} \sim 1 \times 10^{12} M_{\odot}$ with an uncertainty of approximately 20 – 30% (see, Wang et al. 2020; Bobylev & Baykova 2023; Hunt & Vasiliev 2025, for compilations). Many of these previous estimates have relied on the MW being well described by equilibrium models (e.g., Jeans models, Jeans 1915), i.e., not undergoing a significant merger episode.

Yet, the MW is currently experiencing a merger with the Large Magellanic Cloud (LMC, see Vasiliev 2023, for a comprehensive review of the effect of the LMC on the MW) satellite galaxy. The LMC is thought to be on its first pericentric passage and to have a dark matter mass $M_{\text{LMC}} \sim 10^{11} M_{\odot}$ (e.g., Besla et al. 2007, 2010; Boylan-Kolchin et al. 2011; Peñarrubia et al. 2016; Erkal et al. 2019a; Koposov et al. 2023; Kravtsov & Winney 2024; Benisty 2024; Watkins et al. 2024; Brooks et al. 2025). As the LMC has just completed its most recent pericentric passage at a relative velocity $> 300 \text{ km s}^{-1}$ and has a significant mass ratio relative to the MW ($\sim 1 : 4 - 10$), the interaction of these two galaxies has induced a state of dynamical disequilibrium throughout the entire MW (Deason & Belokurov 2024; Hunt & Vasiliev 2025). This is also expected and seen within MW–LMC analogues identified within cosmological simulations (e.g., Darragh-Ford et al. 2025; Arora et al. 2025; Mansfield et al. 2026). The infall of the LMC into the MW has generated a density wake in the MW’s dark matter halo (Weinberg 1998; Garavito-Camargo et al. 2019, 2021a; Tamfal et al. 2021; Rozier et al. 2022; Foote et al. 2023), as well as leaving observable signatures in the density and kinematics of stars in the stellar halo (e.g., Belokurov et al. 2019; Conroy et al. 2021; Cavieres et al. 2025; Chandra et al. 2025b; Yaaqib et al. 2024; Byström et al. 2025; Fushimi et al. 2024; Amarante et al. 2024; Sheng et al. 2025). On large scales, the inner and outer parts of the MW

* E-mail: richard.brooks.22@ucl.ac.uk

† NASA Hubble Fellowship Program, Einstein Fellow

halo have experienced different strengths of acceleration towards the LMC (e.g., Gómez et al. 2015). To a Galactocentric observer, the Galactic northern sky appears to be red-shifted and the Galactic southern sky blue-shifted because the halo moves preferentially ‘up’, towards the Galactic north. This ‘reflex’ displacement manifests as a dipole signal in density (Garavito-Camargo et al. 2021b; Conroy et al. 2021; Amarante et al. 2024) that is higher in the Galactic North, and similarly in stellar radial velocities (Erkal et al. 2019b; Petersen & Peñarrubia 2020, 2021; Erkal et al. 2021; Yaaqib et al. 2024; Chandra et al. 2025b; Byström et al. 2025; Brooks et al. 2025). The magnitude of the reflex velocity dipole is called the *travel velocity* and its on-sky orientation is called the *apex* direction of the reflex motion. If models of the MW ignore the reflex motion induced by the LMC, then the inference for the MW mass can be overestimated by up to 50% depending on the observed regions of the sky (Erkal et al. 2020).

Simulation Based Inference (SBI, see Cranmer et al. 2020, for a conceptual overview) is a statistical framework for inference in complex modelling scenarios where traditional analytic methods are impractical. Rather than defining a likelihood function, SBI uses many forward simulations to learn the statistical relationship between model parameters and observational data. Within the astrophysics community, SBI has been increasingly applied to a variety of astrophysics problems (e.g., Weyant et al. 2013; Alsing et al. 2019; Jeffrey et al. 2021; Lemos et al. 2021; Hermans et al. 2021; von Wietersheim-Kramsta et al. 2024; Lovell et al. 2024; Ho et al. 2024; Widmark & Johnston 2025; Sante et al. 2025; Xiangyuan Ma et al. 2025; Jeffrey et al. 2025; Saoulis et al. 2025; Nguyen et al. 2025; Brooks et al. 2025, 2026). One of the main probes of the dynamical MW–LMC interaction is the direction and magnitude of the reflex motion. Many previous studies have used a likelihood-based approach to constrain this interaction (e.g., Petersen & Peñarrubia 2021; Yaaqib et al. 2024; Byström et al. 2025; Chandra et al. 2025b), which may lead to biases being introduced due to, e.g., survey selection criteria. As a result, SBI offers a novel and flexible method to determine precise measurements of this induced motion, while simultaneously constraining other properties of the MW and LMC (Brooks et al. 2025, 2026).

In Brooks et al. (2026), we presented and validated a novel SBI architecture for exploring large MW–LMC model parameter spaces to enable rapid and reliable parameter inference of model parameters using the dynamics of outer halo stars. We showed that our SBI framework trained on many rigid MW–LMC simulations retains enough of the relevant physics that more complex simulations capture (e.g., N -body simulations, Garavito-Camargo et al. 2019) to avoid model misspecification and vastly improve the computational efficiency of the inference. Subsequently in Brooks et al. (2025), we demonstrated the first application of our SBI architecture to data from the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al. 2025) survey and the H3+SEGUE+MagE (Chandra et al. 2025b) combined outer halo survey to constrain the LMC mass, the reflex motion and the strength of dynamical friction. One limitation of these studies is that only mean velocities of outer halo stars were used for inference, limiting the ability to constrain the MW mass, which is sensitive to their velocity dispersions (see the spherical Jeans equations, Jeans 1915; Binney & Tremaine 2008). Expanding this SBI framework to incorporate the velocity dispersions of outer halo stars will enable simultaneous constraints to be placed on the MW and LMC masses as well as their reflex interaction.

In this work, we present an enhanced SBI architecture to constrain the MW–LMC interaction. We use both the mean *and* dispersions of outer halo star velocities to place constraints on the MW and LMC

masses, the velocity anisotropy parameter, and the reflex motion. Additionally, we improve on the previous distance-independent reflex motion parametrisation by implementing a linear continuity model to capture any radial variation of the reflex motion. Using all-sky kinematics data from the H3+SEGUE+MagE (Chandra et al. 2025b) combined with these model enhancements, our SBI framework is able to simultaneously produce precise constraints on the MW and LMC masses, the reflex motion and velocity anisotropy. Furthermore, with the release of upcoming datasets from Gaia Data Release 4 (DR4) and the Sloan Digital Sky Survey V (SDSS-V, Chandra et al. 2025a; SDSS Collaboration et al. 2025), our flexible SBI framework will be uniquely placed to properly account for variation across surveys when performing inference.

The layout of this paper is as follows. In Sec. 2 we describe the data used throughout this work and define the summary statistics used for the inference. In Sec. 3, we give a description of the MW–LMC simulations used to train the SBI architecture used for parameter inference. In Sec. 4, we describe the SBI architecture, Bayesian statistics and the machine learning models used for inference. In Sec. 5, we give the constraints on the MW and LMC enclosed and total masses, the reflex motion and the velocity anisotropy model parameters. In Sec. 6, we discuss our results in context of previous literature, assess model limitations and investigate the relative importance of including the LMC in models. Finally, we conclude and provide an outlook for upcoming surveys in Sec. 7.

2 DATA

2.1 H3, SEGUE & MagE Outer Halo Surveys

2.1.1 H3

The H3 Spectroscopic Survey (Conroy et al. 2019) has conducted a spectroscopic survey of halo stars with the Hectochelle instrument (Szentgyorgyi et al. 2011) on the 6.5m MMT telescope at the Whipple Observatory in Arizona. We use the sample of H3 stars observed up to January 2024 that have reliable stellar parameters from MINESweeper (Cargile et al. 2020) and are not associated with known MW substructures.

2.1.2 SEGUE

The Sloan Extension for Galactic Understanding and Exploration (SEGUE, Yanny et al. 2009) survey observed $\sim 250,000$ stars with the low-resolution BOSS spectrograph as a part of the Sloan Digital Sky Survey (SDSS, York et al. 2000). These spectra have been fitted using the MINESweeper routine to provide reliable stellar parameters (Chandra et al. 2025a; Cargile & Chandra 2025).

2.1.3 MagE

Over the past two years, a tailored spectroscopic survey of luminous RGB stars in the outer halo has been conducted with The Magellan Echelle Spectrograph (MagE, Marshall et al. 2008) on the 6.5m Magellan Baade Telescope at Las Campanas Observatory. The selection procedure for the target sample of RGB stars is described in Chandra et al. (2023a), and the details of the spectroscopic survey are further described in Chandra et al. (2023b) and Chandra et al. (2025b). Stellar parameters are estimated with the MINESweeper code (Cargile et al. 2020) including parallax measurements from Gaia (Gaia Collaboration et al. 2023). As of May 2024, a total of 400

stars have been observed, of which ~ 300 are spectroscopically confirmed to be at a heliocentric distance beyond 50 kpc, and ~ 100 are beyond 100 kpc making it the largest dataset of outer halo stars beyond 50 kpc. For a much more detailed account of the *MagE* survey, we direct the reader to Chandra et al. (2023a,b, 2025b).

2.1.4 Combined H3+SEGUE+MagE sample

The above surveys combine to produce a pure sample of stars with homogeneous stellar parameters and their uncertainties derived with the MINESweeper pipeline, e.g., distance estimates with an uncertainty floor of 10%. Subsequently, we adopt the same selection procedure to exactly reproduce the high-fidelity subset in sec. 2.3, Chandra et al. (2025b), e.g., fractional distance uncertainties $< 25\%$. This selection procedure ensures that known substructures and unphysical quantities are removed from the sample. In particular, the Sagittarius stream (e.g., Majewski et al. 2003; Vasiliev et al. 2021) is excised using the angular momentum cuts proposed by Johnson et al. (2020). We note that specific choices for removing Sagittarius stream stars may alter results, e.g. see appendix A in Chandra et al. (2025b) as well as sec. 2.6 and appendix D in Byström et al. (2025). The all-sky sample of H3+SEGUE+MagE stars used in this work contains 1,296 field stars between $r_{\text{gal}} \in [30 - 160]$ kpc. The same sample of stars was used in Brooks et al. (2025) when considering the full depth of the H3+SEGUE+MagE combined outer halo survey. Although that study also considered other data selections to match DESI survey selection criteria. Note that some figures in this study will use the shorthand ‘H3+’ to refer to this combined sample.

2.2 Observational summary statistics

SBI requires a set of summary statistics to be used for the inference process. For this application of an SBI framework to the dynamics of the outer MW stars, we adopt summary statistics based on their velocity field distributions. Specifically, we focus on the mean and dispersions of the radial and tangential velocities as a function of distance. This choice of velocity summary statistics is insensitive to differential selection between the individual H3, SEGUE and MagE surveys.

2.2.1 Radial velocities

We calculate the stellar radial velocities in the *Galactic Standard of Rest* (GSR) frame, v_{GSR} , which accounts for the solar motion with respect to the Galactic centre. For a non-rotating and/or axisymmetric galaxy in equilibrium, the mean radial velocity, $\langle v_{\text{GSR}} \rangle$, is expected to be zero throughout the entire MW. However, our Galaxy is in disequilibrium due to the merger with the LMC and hence $\langle v_{\text{GSR}} \rangle \neq 0$ ubiquitously across the MW.

We explore the full depth and on-sky coverage of the H3+SEGUE+MagE dataset by dividing the sky into quadrants and measuring a set of mean radial and tangential velocity summary statistics. Throughout this work, we define the quadrants as follows: Quadrant 1 (Q1) as $l \in [+180^\circ, 0^\circ]$, $b \in [0^\circ, +90^\circ]$, Quadrant 2 (Q2) as $l \in [0^\circ, -180^\circ]$, $b \in [0^\circ, +90^\circ]$, Quadrant 3 (Q3) as $l \in [+180^\circ, 0^\circ]$, $b \in [-90^\circ, 0^\circ]$ and Quadrant 4 (Q4) as $l \in [0^\circ, -180^\circ]$, $b \in [-90^\circ, 0^\circ]$. A visual representation can be seen in the Mollweide figure above Fig. 1 and Fig. 2. For the all-sky H3+SEGUE+MagE dataset, we calculate the 3σ -clipped mean, $\langle v_{\text{GSR}} \rangle$, and dispersions, $\langle \sigma_{v_{\text{GSR}}} \rangle$, values for the Galactocentric distance range $r_{\text{gal}} \in [30 - 160]$ kpc in 5 equally spaced distance bins

in the northern quadrants (Q1 and Q2) and the southern quadrants (Q3 and Q4); see the upper and middle panels of Fig. 1 & Fig. 2. The distance bins are larger than any associated distance uncertainties to minimise correlations between adjacent bins. Note, that neither the data nor the simulations account for uncertainty in the solar motion with respect to the Galactic centre. In these figures, data points with increased transparency, and distinct markers, are those which are poorly represented by simulated data; see Sec. 5 for more detail.

2.2.2 Tangential velocities

The tangential velocity of a star, particularly in the Galactic latitude direction, $v_{t,b}$, traces the LMC’s perturbation in the MW (Erkal et al. 2021; Sheng et al. 2024; Chandra et al. 2025b). This component of the tangential velocity captures the apparent ‘upward’ reflex motion of outer halo stars as the MW’s centre of mass is dragged ‘downward’. We calculate the 3σ -clipped mean, $\langle v_{t,b} \rangle$, and dispersions, $\langle \sigma_{v_{t,b}} \rangle$, values for the Galactocentric distance range $r_{\text{gal}} \in [30 - 160]$ kpc in 6 equally spaced distance bins; see the lower panels of Fig. 1 & Fig. 2. These tangential velocities have been corrected for the solar motion and uncertainties determined via bootstrap resampling. In this study, we do not consider the tangential velocities in the Galactic longitudinal direction, $v_{t,l}$, though future work can utilise them to constrain the LMC’s present-day phase space coordinates and its past orbit (Sheng et al. 2025).

3 SIMULATIONS

3.1 The Milky Way – LMC potentials

3.1.1 The Milky Way

To model the MW dark matter halo we use a Navarro-Frenk-White (NFW, Navarro et al. 1996, 1997) dark matter halo density profile described by M_{200} and c_{200} ; see Table. 1 for the mass prior range considered. These quantities are defined by a sphere enclosing an overdensity that is 200 times the critical density of the Universe, $\rho_{\text{crit}} = 3H_0^2/8\pi G$, as denoted by the ‘200’ subscript, where the Hubble parameter, H_0 , is taken to be $67.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$ using the default cosmology in *astropy* (Astropy Collaboration et al. 2013). We constrain the normalisation of the halo mass profile such that the circular velocity at the solar position is approximately 235 km s^{-1} (e.g., matching constraints from McMillan 2017, within the associated uncertainty).¹ The MW stellar components are modelled using a spherical bulge with a total mass $1.2 \times 10^{10} M_\odot$, and an exponential stellar disc with a total mass $5 \times 10^{10} M_\odot$. The stellar distributions remain fixed taking the values suggested by McMillan (2017).

3.1.2 The LMC

We model the LMC as a Hernquist dark matter halo (Hernquist 1990). We normalise the profile such that the derived rotation curve peaks at $(91.7 \pm 18.8) \text{ km s}^{-1}$ at a distance of 8.7 kpc from its centre. This corresponds to an enclosed dynamical mass of $M_{\text{LMC}}(r < 8.7 \text{ kpc}) = (1.7 \pm 0.7) \times 10^{10} M_\odot$ (van der Marel & Kallivayalil 2014). See Table. 1 for the chosen total LMC mass prior distribution.

¹ Also see Horta et al. (2026) for a recent empirical measurement using element abundance gradients.

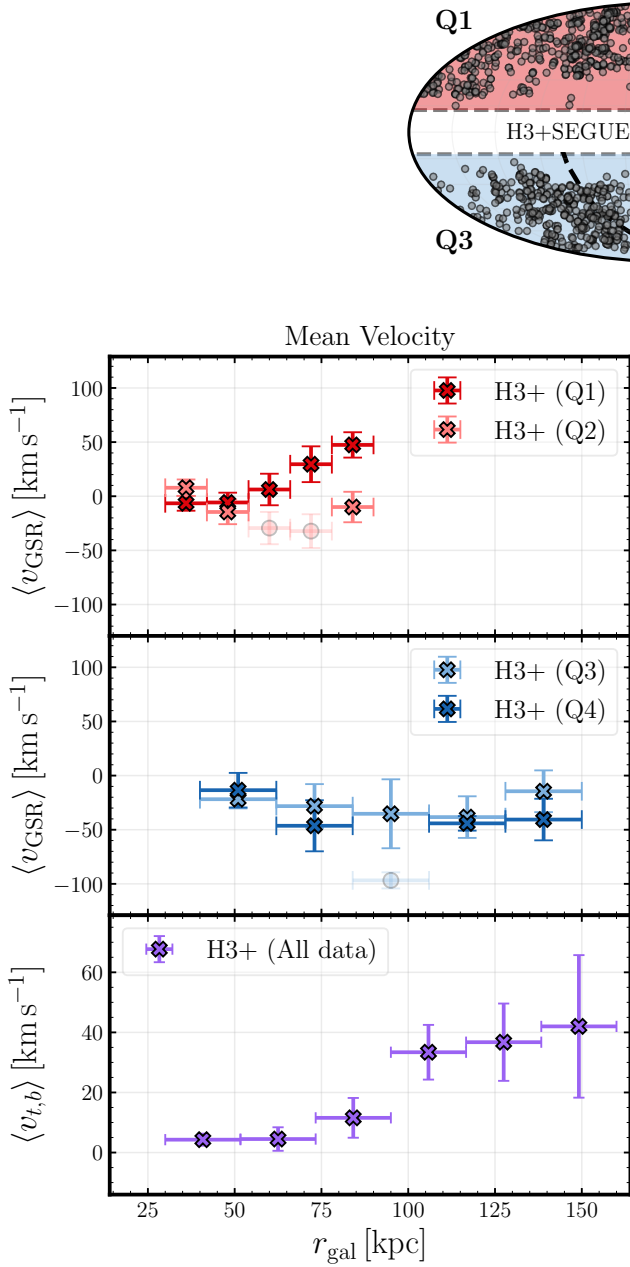


Figure 1. Mean velocity distributions in the radial and tangential components, $\langle v_{\text{GSR}} \rangle$ and $\langle v_{t,b} \rangle$, as a function of Galactocentric distance for the H3+SEGUE+MagE data divided into on-sky quadrant footprints. Q1/2 are in the Galactic North, Q3/4 are in the Galactic South. Vertical error bars represent the 1σ uncertainties determined via bootstrap resampling. Horizontal error bars display the bin width. Data points with increased transparency, and distinct markers, are poorly represented by simulated data counterparts.

3.1.3 The MW–LMC interaction

The trajectories of the MW and LMC under their mutual gravitational attraction are numerically integrated (see equ. 3-6 in [Brooks et al. 2026](#)). We account for acceleration due to Chandrasekhar dynamical friction, $\mathbf{a}_{\text{DF}}$, on the trajectory of the LMC ([Chandrasekhar 1943](#); [Binney & Tremaine 2008](#); [Jethwa et al. 2016](#)) as:

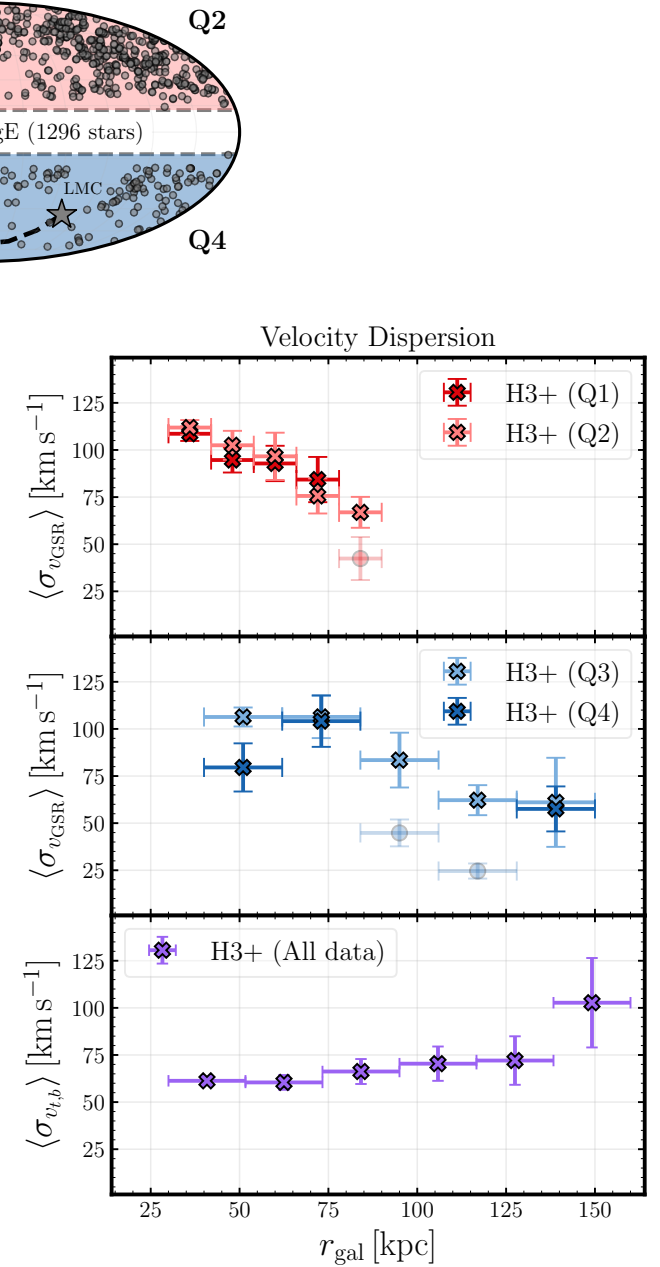


Figure 2. Velocity dispersion distributions in the radial and tangential components, $\langle \sigma_{v_{\text{GSR}}} \rangle$ and $\langle \sigma_{v_{t,b}} \rangle$, as a function of Galactocentric distance for the H3+SEGUE+MagE data divided into on-sky quadrant footprints. Q1/2 are in the Galactic North, Q3/4 are in the Galactic South. Vertical error bars represent the 1σ uncertainties determined via bootstrap resampling. Horizontal error bars display the bin width. Data points with increased transparency, and distinct markers, are poorly represented by simulated data counterparts.

$$\mathbf{a}_{\text{DF}} = -\frac{4\pi G^2 M_{\text{LMC}} \rho_{\text{MW}} \ln \Lambda}{v_{\text{LMC}}^3} \left[\text{erf}(X) - \frac{2X}{\sqrt{\pi}} e^{-X^2} \right] \lambda_{\text{DF}} \mathbf{v}_{\text{LMC}}, \quad (1)$$

where $X = v_{\text{LMC}}/\sqrt{2}\sigma_{\text{MW}}$ and σ_{MW} and ρ_{MW} are the velocity dispersion and total density field of the MW. Following [Vasiliev et al. \(2021\)](#), we take a fixed value of $\sigma_{\text{MW}} = 120 \text{ km s}^{-1}$ for the velocity dispersion as the dynamical friction is insensitive to the precise

value. For the Coulomb logarithm we adopt $\ln \Lambda = \ln(100 \text{ kpc}/\epsilon)$. The softening length, ϵ , depends on the satellite's density profile (White 1976). We adopt $\epsilon = 1.6 a_{\text{LMC}}$, where a_{LMC} is the LMC scale radius, as this has been used previously when modelling the LMC as a Plummer sphere (e.g., Hashimoto et al. 2003; Besla et al. 2007; van der Marel et al. 2012; Sohn et al. 2013; Kallivayalil et al. 2013; Patel et al. 2020). The numerator in the Coulomb logarithm expression is an arbitrarily chosen value that loosely describes the average separation of the MW and LMC. This value could be updated through the integration of the LMC orbit, though it would have a small effect. We use a dimensionless parameter, λ_{DF} , to modulate the strength of the dynamical friction that the LMC experiences. In principle, this will take into account changes to the fixed Coulomb logarithm value per simulation.

3.2 The Milky Way Stellar Halo

To generate a mock MW stellar halo for each simulation, we draw phase-space samples from radially-biased distribution functions as implemented in AGAMA (Vasiliev 2019a). This requires instances of a tracer density profile, a potential, and a prescription for the radial velocity anisotropy. We use a spheroidal tracer density profile,

$$\rho \propto \left(\frac{\tilde{r}}{r_{\text{dens}}} \right)^{-1} \left[1 + \left(\frac{\tilde{r}}{r_{\text{dens}}} \right) \right]^{1-\gamma}, \quad (2)$$

allowing the logarithmic slope of the outer halo to vary between $\gamma \in [3, 5]^2$ to account for the range of previous literature values (e.g., Han et al. 2022; Amarante et al. 2024; Cavieres et al. 2025; Li et al. 2025). We note that the inner slope is fixed, $\rho \propto r^{-1}$, making it formally inconsistent with very radially biased distributions via the *cusp slope – anisotropy theorem* (An & Evans 2006). In this study, we use stars well beyond the scale radius to perform parameter inference, making this inconsistency negligible here. For the potential, we use the combined rigid MW profile, consisting of an NFW dark matter halo, a stellar disc and bulge (Navarro et al. 1996, 1997) with relevant parameters adopted from each unique MW–LMC simulation; see, Sec. 3.1.1. We implement the radial bias of stellar velocities (Osipkov 1979; Merritt 1985; Binney & Tremaine 2008) using a constant velocity anisotropy profile such that the anisotropy remains constant and equal to β_0 over all radial scales.

We sample 20,000 phase-space coordinates per stellar halo to minimise Poisson noise, ensuring observational uncertainty dominates. This precision allows us to convolve simulated data with survey-specific observational uncertainties to produce observational-like quantities used for the inference. In this study, we apply the survey-specific uncertainties to the binned radial and tangential velocity measurements from the H3+SEGUE+MagE dataset; see Sec. 2.2. This approach allows us to forward model all the stellar haloes to match a specific survey of interest, in this case the combined H3+SEGUE+MagE outer halo survey, and perform the subsequent inference. For a given MW–LMC potential with reflex motion, we integrate all particles in the stellar halo as test particles to present-day over the last 2.2 Gyr (motivated by the epoch when the LMC crosses the MW virial radius in Garavito-Camargo et al. 2019).

² Although β is often used to denote the outer density slope, we adopt γ here to avoid confusion with the velocity anisotropy parameter β .

3.2.1 The reflex motion

From the final distribution of stellar halo particles, we measure the reflex motion of the MW in response to the infalling LMC (as used in Petersen & Peñarrubia 2021; Chandra et al. 2025b; Yaqib et al. 2024, 2025; Brooks et al. 2025, 2026). This method fits an on-sky velocity model which contains nine free parameters. We model the dipole reflex motion using Galactocentric Cartesian velocities $\{v_x, v_y, v_z\}$, which can be transformed to spherical coordinates $\{v_r, v_\phi, v_\theta\}$ to combine with the mean motions. We choose to sample in Cartesian coordinates rather than spherical coordinates to avoid inefficient convergence during maximum likelihood estimation if the preferred apex direction is close to the Galactic poles. We account for non-zero mean motion in the halo's Galactocentric velocity via the mean motion parameters $\vec{v}_{\text{mean}} = (\langle v_r \rangle, \langle v_\phi \rangle, \langle v_\theta \rangle)$. This allows for any departures in the bulk halo motion from the travel velocity. Finally, we account for the intrinsic velocity dispersion in each component using the set of parameters $\{\sigma_{v_r}, \sigma_{v_\phi}, \sigma_{v_\theta}\}$. We note that within our SBI framework, these parameters act as nuisance parameters, since measurement uncertainties are incorporated into the velocity summary statistics used for inference; see Sec. 2.2. This parametrisation of the intrinsic velocity dispersion is only used to obtain a maximum likelihood estimate of the ‘true’ reflex motion parameters for each simulation. The reflex motion model is represented by the sum of the travel velocity and mean velocity parameters:

$$\langle \vec{v} \rangle = \vec{v}_{\text{travel}} + \vec{v}_{\text{mean}}, \quad (3)$$

where $\langle \vec{v} \rangle$ is the mean Galactocentric halo velocity vector. To find the maximum likelihood estimates for these parameters given each set of mock stellar halo data, we minimise a Gaussian log-likelihood for the 1-dimensional line of sight velocities and 2-dimensional proper motions using `scipy.optimize` (see equations 6 & 8 in Petersen & Peñarrubia 2021). We return the maximum likelihood estimates for all of the reflex motion model parameters. For this study, we will comment on the magnitude, v_{travel} or Galactocentric components of the reflex travel velocity, $\{v_x, v_y, v_z\}$ and the apex direction, $(l_{\text{apex}}, b_{\text{apex}})$.

We improve upon our distance-independent model in Brooks et al. (2026, 2025) by employing a reflex motion model that is allowed to vary linearly as a function of Galactocentric distance (see a similar approach in Chandra et al. 2025b). To achieve this, the reflex motion model now takes unique parameter values at $r_{\text{gal}} = 40 \text{ kpc}$ and $r_{\text{gal}} = 120 \text{ kpc}$, with a linear fit implemented between those two anchoring values. Overall, the number of reflex motion model parameters is doubled. This simple linear expansion allows us to find simulation prior distributions for reflex motion model parameters at any distance between $r_{\text{gal}} \in [40, 120] \text{ kpc}$, and allows for tomographic inference of the reflex motion; see Sec. 5.1.

3.3 Simulation Priors

We update some of the parameter prior probability distributions from Brooks et al. (2025, 2026). We show all of our priors in Table. 1. The first two parameters are the MW mass and LMC total mass. Next is λ_{DF} , the scalar multiple that modulates the strength of dynamical friction relative to classic Chandrasekhar values; see Equ. 1. The next set of parameters describe the LMC present-day position and velocity with their distributions inspired by the values in sec. 3.1 and table 2 of Vasiliev (2023). The final three parameters are the anisotropy parameter, β_0 , the scale length, r_{dens} , and outer halo slope, γ , for the tracer density profile that initialise the mock MW stellar haloes. In total, we run 32,000 MW–LMC simulations each with unique

Table 1. Simulation/model parameter prior distributions.

Model Parameter	Prior probability distribution
$\log_{10}(M_{200,\text{MW}})$	$\mathcal{U}(10.5, 12.5) M_{\odot}$
M_{LMC}	$\mathcal{U}(2.5, 50) \times 10^{10} M_{\odot}$
$\log_{10}(\lambda_{\text{DF}})$	$\mathcal{U}(-3, 1)$
α_{LMC}	$\mathcal{U}(72.5^{\circ}, 87.5^{\circ})$
δ_{LMC}	$\mathcal{U}(-75^{\circ}, -65^{\circ})$
d_{LMC}	$\mathcal{N}(49.6, 2) \text{ kpc}$
v_{los}	$\mathcal{N}(262.2, 5) \text{ km s}^{-1}$
$\mu_{\alpha_{\text{LMC}}}$	$\mathcal{N}(1.9, 0.25) \text{ mas yr}^{-1}$
$\mu_{\delta_{\text{LMC}}}$	$\mathcal{N}(0.33, 0.25) \text{ mas yr}^{-1}$
β_0	$\mathcal{U}(0, 0.85)$
r_{dens}	$\mathcal{U}(10, 20) \text{ kpc}$
γ	$\mathcal{U}(3, 5)$

parameter values and 20,000 particles to represent the MW stellar halo.

4 SIMULATION BASED INFERENCE

In the Bayesian approach, a problem is often posed as calculating the probability of the model parameters, θ (e.g., $M_{200,\text{MW}}$), given some observed data, D_{obs} (e.g., $\langle v_{\text{GSR}} \rangle$), and a theoretical model, I (e.g., rigid MW–LMC simulations). In other words, we want to find the posterior probability distribution, $\mathcal{P} = p(\theta|D_{\text{obs}}, I)$. This is possible using Bayes’ Theorem:

$$p(\theta|D_{\text{obs}}, I) = \frac{p(D_{\text{obs}}|\theta, I)p(\theta|I)}{p(D_{\text{obs}}|I)} \iff \mathcal{P} = \frac{\mathcal{L} \times \Pi}{\mathcal{Z}} \quad (4)$$

where $\mathcal{L} = p(D_{\text{obs}}|\theta, I)$ is the likelihood, $\Pi = p(\theta|I)$ is the prior, and $\mathcal{Z} = p(D_{\text{obs}}|I)$ is the Bayesian evidence. The Bayesian evidence acts as a normalisation in parameter estimation and can be ignored for our application. Given a choice of prior distribution for parameters and a likelihood function, we can find the posterior distribution. SBI allows an estimate for the likelihood, and hence the posterior, distributions to be found without the need for assumptions about the analytic form for the likelihood.

The roots of SBI lie within the Approximate Bayesian Computation framework (ABC, e.g., Rubin 1984; Pritchard et al. 1999; Fearnhead & Prangle 2010). An ABC framework selects forward simulations that are the most similar to the observed data based on some distance measure involving summary statistics of the simulation. Another way to compute the posterior is via Density Estimation Likelihood Free Inference (DELFI). In this approach, forward simulations are used to learn a conditional density distribution of the data D_{obs} , given the simulation parameters θ , using a density estimation algorithm, e.g., normalising flows, that utilise a series of bijective transformations to convert a simple base distribution e.g., a Gaussian, into the target probability distribution (Jimenez Rezende & Mohamed 2015).

We adopt the DELFI approach as implemented in the `sbi`³ Python package (Tejero-Cantero et al. 2020) to directly estimate the posterior

³ Also available via the `1tu-ili` Python package which unifies several of the available SBI codes (Ho et al. 2024)

distribution; often called *neural posterior estimation*. The specific choice of neural network architecture used can influence the estimated posterior. Firstly, the choice of modelling hyperparameters, e.g., the number of layers and nodes defining the neural network, can affect the quality of the trained SBI model. We find that a Masked Autoregressive Flows (MAF, Papamakarios et al. 2017, 2019) architecture produces well calibrated posteriors using 5 transformation layers in the neural network, each with a width of 50 nodes. Secondly, we account for the model (epistemic) uncertainties, i.e., those stemming from the uncertainty in the choice of network weights and biases, by model ensemble averaging.

DELFI is advantageous over the simpler ABC approach as it does not rely on a choice of a distance measure and it uses all available forward simulations to build the posterior distribution, making it far more efficient (Alsing et al. 2019). Once a neural posterior estimator has been trained on a precomputed simulation dataset, a posterior can be returned for many unique observations without having to retrain the flow; this is known as *amortisation* (Mittal et al. 2025).

The simulations used for SBI in this work are described in Sec. 3. We use a MAF density estimator to directly obtain the posterior distribution that can be evaluated at any observed data point for any data realisation, i.e., $p(\theta|D_{\text{obs}}, I)$. In general, directly estimating the posterior is preferred over estimating the likelihood when a neural network maps high-dimensional inputs to low-dimensional outputs (Ho et al. 2024). We ensure the reliability of the estimated posteriors through some diagnostic checks including coverage probabilities and predictive posterior checks in Appendix A.

5 RESULTS

In this section we show posterior distributions for MW–LMC model parameters conditioned on different subsets of the outer halo H3+SEGUE+MagE survey. All figures use 10,000 samples drawn from their respective posterior density distributions.

5.1 Inference set-up

We will show the posterior distributions for MW and LMC masses enclosed within 50 kpc, the Galactocentric Cartesian components of the reflex motion and the velocity anisotropy parameter. Additionally, using these posterior samples, we will show the magnitude and on-sky apex direction of the reflex motion, as well as the radial variation in these reflex motion parameters.

For the input data that is used to condition the posterior estimation on, we use the mean and dispersions of the radial and tangential velocities from the outer halo H3+SEGUE+MagE survey; see Fig. 1 and Fig. 2. To investigate the effect of varying levels of velocity data provided we use various subsets of this velocity data.

Yellow contours, H3+ $\langle v \rangle$: The entire mean velocity data is used to condition the posterior estimation, i.e., all data points in Fig. 1.

Blue contours, H3+ $\langle v \rangle$, $\langle \sigma_v \rangle$: The entire mean and dispersion velocity data is used to condition the posterior estimation, i.e., all data points in Fig. 1 and Fig. 2.

Red contours, H3+ $\langle v \rangle^*$, $\langle \sigma_v \rangle^*$: Only the mean and velocity dispersion data points which are well represented by our simulations are used to condition the posterior estimation. Specifically, we remove all data points that lie 2σ or more away from the corresponding simulated data mean. These ‘discrepant’ data points are indicated by the

transparent markers in Fig. 1 and Fig. 2. Overall, 6/52 data points are removed.

The removal of discrepant data allows us to investigate any bias introduced by including data points which are not well represented by the current simulations. This effect can be two-fold. On one hand, the observational dataset itself could still contain undiagnosed substructure producing outlier values. On the other hand, the rigid simulation set-up may not contain enough of the relevant physics (e.g., Yaaqib et al. 2025) to truly represent these data points; see Sec. 6 for further discussion on the origin of discrepancies between the simulations and data. This is especially important to consider when using an SBI approach to perform parameter inference as the simulations must be able to properly represent the observed data. As a reminder, observable uncertainties are accounted for in our simulated stellar halo measurements; see Sec. 3.2 for details on accounting for survey specific uncertainties. Finally, we adopt the present-day Galactocentric position and velocity of the LMC to be $\vec{x}_{\text{LMC}} = \{-0.6, -41.3, -27.1\}$ kpc, $\vec{v}_{\text{LMC}} = \{-63.9, -213.8, 206.6\}$ km s⁻¹ (Vasiliev et al. 2021, and references therein) and the outer slope of the stellar halo to be $\gamma = 4.55$ (global stellar halo fit, Amarante et al. 2024). We note that both the LMC present-day coordinates and the outer halo slope are model parameters explored in their respective prior spaces. We choose to condition our inference on these values as they are well motivated and are not the main focus of this work. Future studies will use this SBI framework to constrain these model parameters.

5.2 The mass of the Milky Way and Large Magellanic Cloud

Previously, in Brooks et al. (2025) we used the H3+SEGUE+MagE dataset and SBI to constrain the enclosed mass of the MW and LMC. In that work, only the means of the radial and tangential velocities were used for the inference. In this study, we use the velocity means *and* dispersions to perform the inference. By conditioning the posteriors on both of these velocity moments we are able to constrain the MW mass to a much greater precision (e.g., Jeans 1915; Binney & Tremaine 2008; Correa Magnus & Vasiliev 2022; Sheng et al. 2025). In Fig. 3 we show the posterior distributions for the MW and LMC masses enclosed within 50 kpc from this study. The yellow and blue contours represent the posteriors conditioned using the means, or means *and* dispersions, of the stellar velocities, respectively. The red contours represent the posterior conditioned on the mean and dispersion data after removing any discrepant data points. The contours delineate the 1 σ and 2 σ confidence intervals. For the 1D posterior panels we show the 16th – 84th percentiles as shaded regions. The prior distributions are shown as the filled grey contours. Note that these priors differ from those used in the previous analysis of this dataset (Brooks et al. 2025). We choose here to implement broader and less informative mass priors to reduce the influence of the choice of an informative prior on our results.

Using only the mean velocities, i.e., the yellow contours in Fig. 3, we constrain the LMC mass enclosed to be $M_{\text{LMC}}(< 50 \text{ kpc}) = 9.31^{+2.17}_{-1.78} \times 10^{10} M_{\odot}$, in agreement with our previous study using the same dataset and a similar SBI set-up (Brooks et al. 2025). We do not report a total mass ratio here as the MW mass posterior is prior dominated.

Using the full set of mean and velocity dispersion measurements to condition the posterior estimation on, i.e., blue contours in Fig. 3, we constrain the MW enclosed mass to be $M_{\text{MW}}(< 50 \text{ kpc}) = 3.07 \pm 0.13 \times 10^{11} M_{\odot}$ and the LMC enclosed mass as $M_{\text{LMC}}(< 50 \text{ kpc}) = 9.37^{+1.94}_{-1.79} \times 10^{10} M_{\odot}$. The inclusion of this velocity dispersion data helps to constrain the MW mass (Sheng et al. 2025).

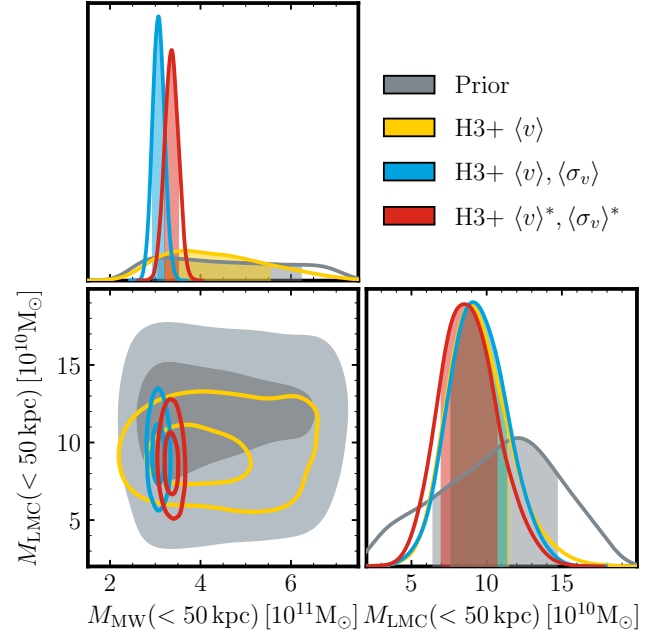


Figure 3. The MW and LMC enclosed masses: The joint, and individual, posterior distributions for the MW and LMC masses enclosed within 50 kpc using data from the H3+SEGUE+MagE survey. The yellow and blue contours represent the posteriors conditioned using the means, or means *and* dispersions, of the stellar velocities, respectively. The red contours use the mean and dispersion data, but removing discrepant data points. The prior distributions are shown as the filled grey contours. The contours delineate the 1 σ and 2 σ confidence intervals. For the 1D posterior panels we show the 16th – 84th percentiles as shaded regions.

Finally, after removing any discrepant observational data, i.e., red contours in Fig. 3, we constrain the MW enclosed mass to be $M_{\text{MW}}(< 50 \text{ kpc}) = 3.36 \pm 0.15 \times 10^{11} M_{\odot}$ and the LMC enclosed mass as $M_{\text{LMC}}(< 50 \text{ kpc}) = 8.76^{+1.94}_{-1.77} \times 10^{10} M_{\odot}$. Hence, including data which is not well captured by the current rigid MW–LMC simulations can bias the inference in the MW mass by $\sim 10\%$. As expected, the MW mass inference favours lower values when using all data points because excluded points have lower velocity dispersions, which prefer lower masses (Jeans 1915). Moreover, we find that the inference of the total masses for the MW, $M_{200, \text{MW}}$ and LMC, M_{LMC} , gives a mass ratio of $\approx 30 \pm 10\%$ for the LMC’s infall mass to the MW’s present day mass.

5.3 The reflex motion

Previously, in Brooks et al. (2025) we reported distance-independent results for the reflex motion parameters using the same H3+SEGUE+MagE dataset. In this study, we have implemented a linear continuity model, see Sec. 3.2, to capture the radial dependence in the reflex motion parameters (Yaaqib et al. 2024; Chandra et al. 2025b). As a reminder, the linear continuity model uses the information of all stars between 40 kpc and 120 kpc. We provide tabulated posterior results at 40 kpc, 100 kpc and 120 kpc in Table. 2 for the magnitude, v_{travel} , on-sky apex direction ($l_{\text{apex}}, b_{\text{apex}}$) and then Galactocentric Cartesian components, (v_x, v_y, v_z) of the reflex motion. In Fig. 4 and Fig. 5 we report the reflex motion posteriors evaluated at a reference distance of 100 kpc.

In Fig. 4 we show the posterior distribution of the Galactocen-

tric velocity components of the reflex motion. Again, we show the posterior distributions conditioned on varying subset of data as the coloured contours; see Sec. 5.1 for explicit details. The prior distributions are shown as the filled grey contours, covering a much broader region of this parameter space compared to the distance-independent approach previously used (Brooks et al. 2025). We show the measured median values from Vasiliev et al. (2021, yellow cross), Yaaqib et al. (2024, SDSS/SEGUE data, $r_{\text{gal}} > 50$ kpc distance bin, orange diamond), Byström et al. (2025, DESI data, blue circle) and Chandra et al. (2025b, H3+SEGUE+MagE data, best-fit value at 100 kpc, red square) in each panel. Although not unexpected, the addition of the velocity dispersion data does not provide any extra constraining power. In all velocity components, our result agrees with Vasiliev et al. (2021) within 1σ and with Byström et al. (2025); Chandra et al. (2025b) within 2σ . However, our result is very discrepant compared to Yaaqib et al. (2024). Interestingly, Vasiliev et al. (2021) is the only other displayed reflex motion result that is constrained by fitting a simulation of the MW-LMC system to data. However, they use a completely different data set of Sagittarius stream stars rather than outer Milky Way halo stars. In Yaaqib et al. (2024); Byström et al. (2025); Chandra et al. (2025b) the components of the travel velocity are constrained as free parameters, directly fitting the observations rather than an assumed MW-LMC model. The agreement between studies that fit simulation models to the reflex motion (i.e., this study and Vasiliev et al. 2021) should not be interpreted as a methodological success; rather, this indicates that the simulation models used share similar biases. Moreover, the MW is likely much more complex in reality. Indeed, this can be seen via the disagreement between studies directly fitting the data (i.e., Yaaqib et al. 2024; Byström et al. 2025; Chandra et al. 2025b).

In Fig. 5 we show the posterior distribution of the reflex motion described by the magnitude of the velocity dipole vector, v_{travel} , and its apex direction⁴ of the reflex motion (l_{apex} , b_{apex}) in Galactic coordinates (e.g., Vasiliev et al. 2021; Petersen & Peñarrubia 2021; Yaaqib et al. 2024; Byström et al. 2025; Chandra et al. 2025b). The coloured contours and markers represent the same information as in Fig. 4. Interestingly, the fiducial N -body simulation of the MW-LMC (see blue pentagon, Garavito-Camargo et al. 2019, MWLMC5 model) is also inconsistent with these previous observations, but agrees within the 1σ uncertainty of our posteriors. This suggests that modelling the longitudinal apex direction, l_{apex} , remains a challenge even for higher fidelity simulations of the MW-LMC system. Nevertheless, we find a good agreement between our results and the existing constraints for the magnitude of the reflex motion. The result found using the mean and dispersion velocity data after removing discrepant data points produces a travel velocity of $v_{\text{travel}} = 38.6^{+8.3}_{-7.8}$ km s⁻¹ and apex direction of $(l_{\text{apex}} [^\circ], b_{\text{apex}} [^\circ]) = (-20.0^{+58.4}_{-51.0}, -75.9^{+20.5}_{-7.0})$ at a reference distance of 100 kpc.

In Fig. 6, we explore the radial variation of the reflex motion. In the top panel, we show the distance dependence for the travel velocity posterior conditioned on the mean and dispersion velocity data, having removed discrepant data points, as the red median and shaded 1σ band. The simulation prior 1σ band is shown by the grey shaded region. There is a clear increase in the travel velocity from $v_{\text{travel}}(r_{\text{gal}} = 40 \text{ kpc}) = 17.0^{+21.1}_{-7.0}$ km s⁻¹ to $v_{\text{travel}}(r_{\text{gal}} = 120 \text{ kpc}) = 47.2^{+11.7}_{-11.0}$ km s⁻¹. This confirms the measured growth in the travel velocity as a function of distance as seen in Chandra et al. (2025b) using the same dataset and linear continuity

Table 2. The reflex motion results using the H3+SEGUE+MagE outer halo survey data. These results are from the posterior conditioned on the mean and velocity dispersion data having removed discrepant data points, i.e., the red contours in all figures. We show the results evaluated using our linear continuity model at $r_{\text{gal}} = [40, 100, 120]$ kpc.

Parameter	$r_{\text{gal}} = 40 \text{ kpc}$	$r_{\text{gal}} = 100 \text{ kpc}$	$r_{\text{gal}} = 120 \text{ kpc}$
$v_{\text{travel}} [\text{km s}^{-1}]$	$17.0^{+21.1}_{-7.0}$	$38.6^{+8.3}_{-7.8}$	$47.2^{+11.7}_{-11.0}$
$l_{\text{apex}} [^\circ]$	$2.5^{+79.5}_{-71.0}$	$-20.0^{+58.4}_{-51.0}$	$-17.1^{+65.0}_{-57.7}$
$b_{\text{apex}} [^\circ]$	$-60.4^{+32.3}_{-17.5}$	$-75.9^{+20.5}_{-7.0}$	$-76.4^{+21.3}_{-7.4}$
$v_x [\text{km s}^{-1}]$	$2.8^{+10.3}_{-4.7}$	$5.3^{+6.7}_{-5.1}$	$5.8^{+8.1}_{-6.4}$
$v_y [\text{km s}^{-1}]$	$0.1^{+10.5}_{-6.8}$	$-2.2^{+7.6}_{-6.8}$	$-1.9^{+8.8}_{-8.8}$
$v_z [\text{km s}^{-1}]$	$-12.9^{+6.9}_{-13.8}$	$-36.0^{+9.4}_{-7.7}$	$-44.1^{+13.3}_{-10.8}$

model. Moreover, Chandra et al. (2025b) demonstrated that using a linear model is sufficient to capture the overall radial dependence of the reflex motion parameters. We find that the result in this work and Chandra et al. (2025b) evaluated at a reference distance of 100 kpc agree within 1σ uncertainties. In the middle and bottom panels, we show the radial dependence of the apex parameters. Similarly, we find there is little radial variation in the apex direction, as seen previously in the data (see fig. 7 in Chandra et al. 2025b), and also within N -body simulations of the MW-LMC system (Garavito-Camargo et al. 2019).

5.4 The velocity anisotropy of the stellar halo

The velocity anisotropy parameter, β , quantifies the radial bias of stellar velocities. Previously, β has been measured as a function of distance (Cunningham et al. 2019; Bird et al. 2021; Han et al. 2024; Chandra et al. 2025b) largely suggesting that β takes values $\gtrsim 0.6$ throughout the outer MW halo. In Fig. 7 we show the posterior distributions for the average velocity anisotropy, β_0 , conditioned on the mean only, or mean and dispersion velocity data, as yellow and blue curves, respectively. The red curve uses the mean and dispersion data, but removing discrepant data points. The shaded regions show the 16th – 84th percentiles. The uniform simulation prior is shown in grey. The expectation is that the velocity dispersion data is required to produce precise constraints on β_0 . Indeed, this is what we see in Fig. 7 for the posteriors conditioned on dispersion data, returning a value $\beta_0 = 0.68 \pm 0.02$ after accounting for discrepant data. This constraint is in agreement with previous measurements of the anisotropy (Cunningham et al. 2019; Bird et al. 2021; Han et al. 2024). Interestingly, the posterior conditioned only on the mean velocity data also prefers a higher anisotropy. This slight preference can be understood as stellar haloes with different anisotropies naturally exhibiting different collective responses to a given perturbation e.g., via the interaction with the LMC (Dillamore et al. 2026). Moreover, such a consistently large value for the velocity anisotropy throughout the MW halo, combined with the dynamical influence of the LMC, can align stars on highly eccentric orbits to explain known local structures in the MW (Dillamore et al. 2026).

⁴ See this [tutorial](#) for a reproducible version of the stereographic projection showing the apex direction in Fig. 5.

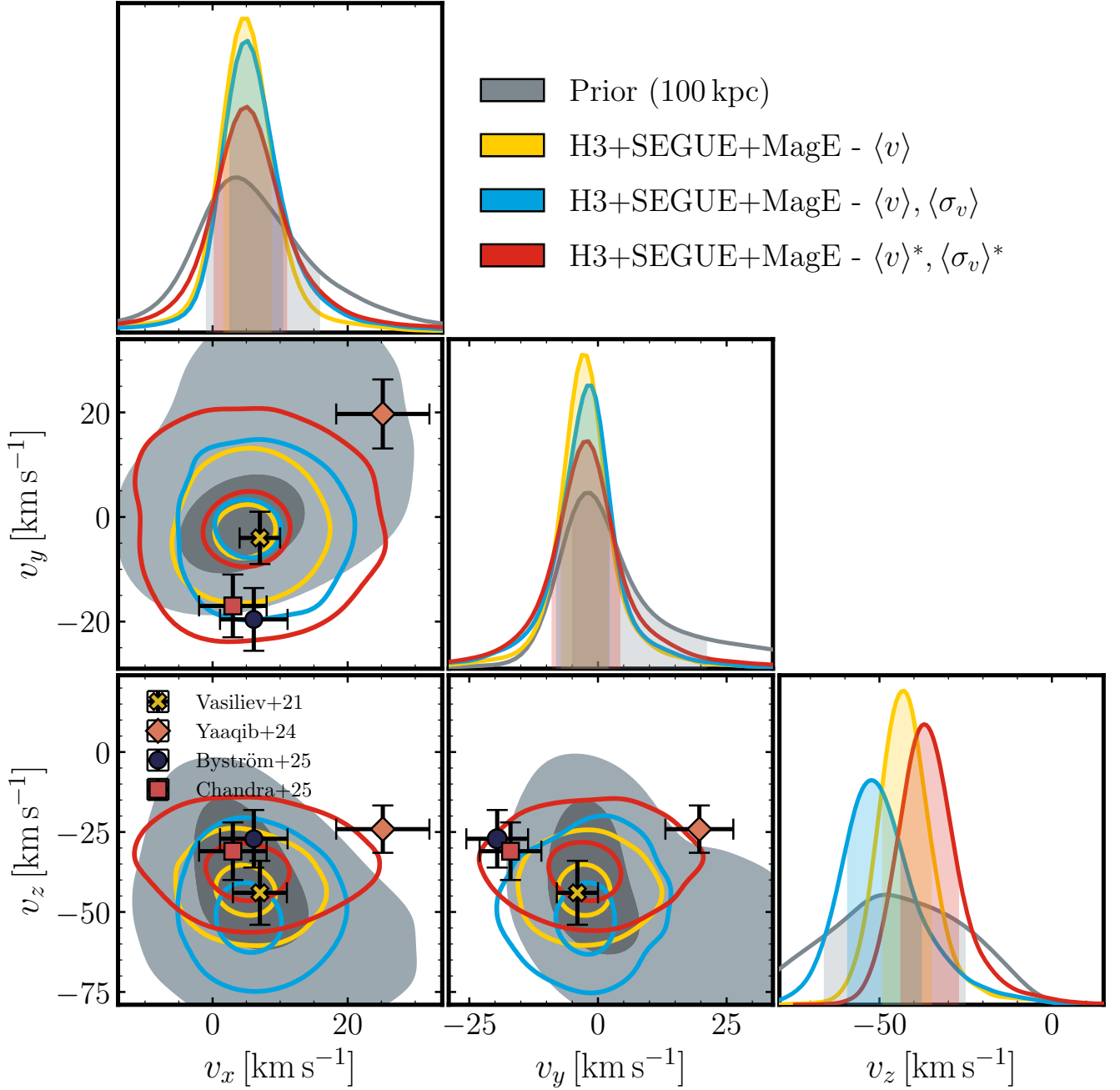


Figure 4. The reflex motion velocity: The joint, and individual, posterior distributions of the Galactocentric Cartesian travel velocity components using data from the H3+SEGUE+MagE survey. The yellow and blue contours represent the posteriors conditioned using the means, or means *and* dispersions, of the stellar velocities, respectively. The red contours use the mean and dispersion data, but removing discrepant data points. The contours delineate the 1σ and 2σ confidence intervals. For the 1D posterior panels we show the 16th – 84th percentiles as shaded regions. The measured mean and 1σ errors from Vasiliev et al. (2021, yellow cross), Yaaqib et al. (2024, orange diamond), Byström et al. (2025, blue circle) and Chandra et al. (2025b, red square) are shown in each panel for comparison.

6 DISCUSSION

6.1 The mass of the Milky Way and LMC in context

6.1.1 The Milky Way

To put the results from this study in context we compare with previous literature values using a range of techniques and dynamical tracers

to constrain the enclosed MW mass. In Fig. 8 we show the MW mass enclosed posterior conditioned on the H3+SEGUE+MagE after removing discrepant data bins. The posterior median and 1σ are represented by the solid red lines and surrounding shaded area. Each literature value is shown as a numbered square, see figure caption for references, with the colour representing the observational tracer used; blue for stellar streams, brown for Blue Horizontal Branch

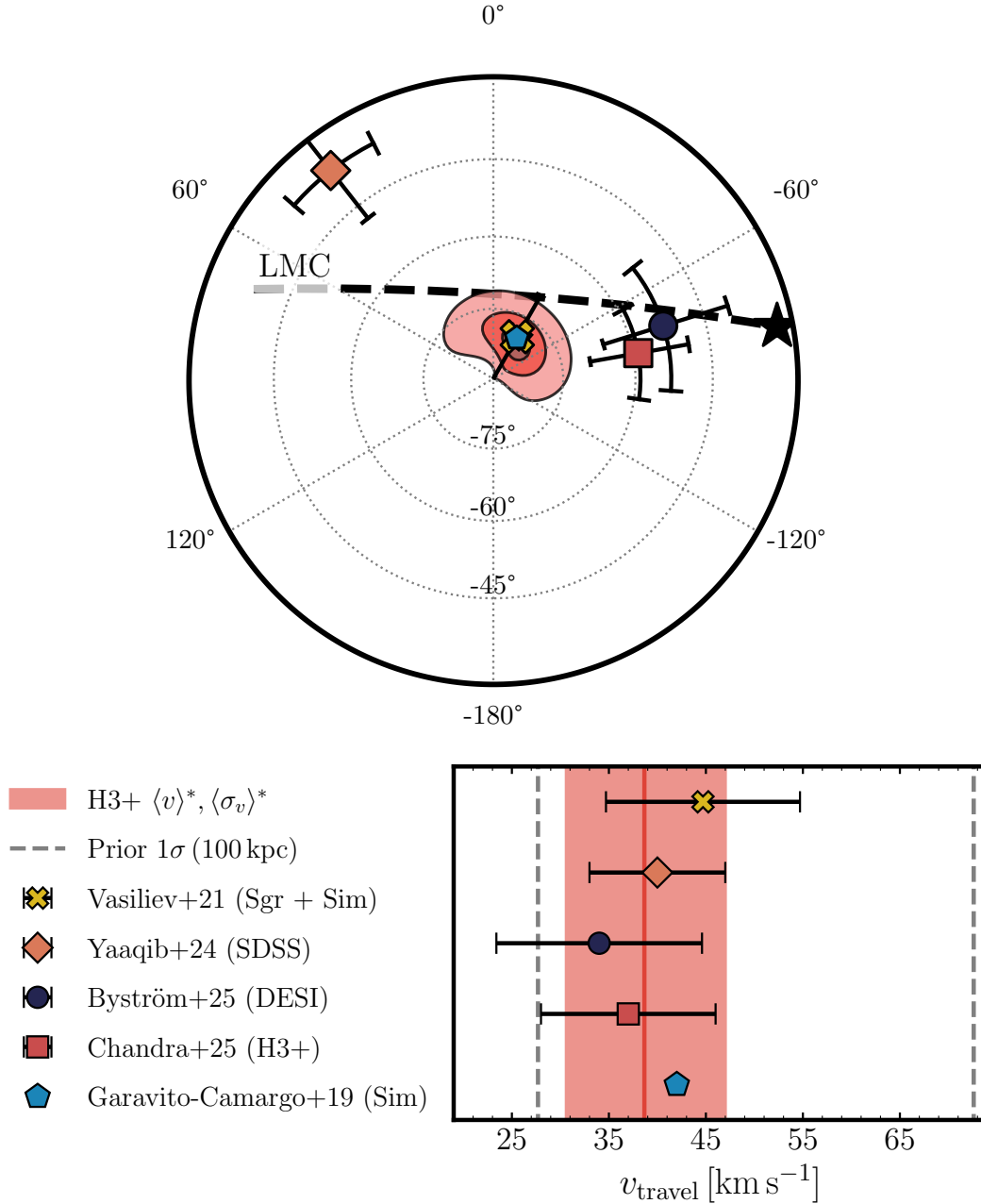


Figure 5. Apex direction and magnitude of the reflex motion – Top panel: A stereographic projection centred on the south Galactic pole showing the apex direction of the travel velocity, $(l_{\text{apex}}, b_{\text{apex}})$. The red posterior contours are placed at the 40th, 70th, and 90th quantiles, corresponding to Gaussian-equivalent levels of 0.5σ , 1σ , and 1.5σ . The posterior distribution is conditioned on the mean *and* dispersion velocity data after removing discrepant data points. The values from existing literature are shown as markers (Vasiliev et al. 2021; Yaaqib et al. 2024; Chandra et al. 2025b; Byström et al. 2025, in yellow using Sgr data, orange using SDSS data, red using H3+SEGUE+MagE data, and dark blue using DESI data, respectively). Additionally, we show the apex direction of the fiducial *N*-body simulation in Garavito-Camargo et al. (2019) as the light blue pentagon. The past LMC orbit is calculated using the best-fit values from the inference with the present-day location denoted by the black star. Lower panel: A comparison of the magnitude of the travel velocity, v_{travel} , from the same literature values and our posterior median and 1σ percentiles as the red line and shaded region. The prior 1σ confidence interval is shown as the grey dashed lines.

(BHB) stars, pink for RR Lyrae (RRL) stars, teal for satellites and globular cluster, and yellow for other stars. For the enclosed mass at 50 kpc, as shown in more detail in our Fig. 3, we find that our result is consistent with many others within uncertainty (e.g., Williams et al. 2017; Koposov et al. 2023; Medina et al. 2025). One of the most precise mass enclosed measurements at this distance is obtained

by modelling the Orphan-Chenab stream in a combined MW–LMC potential, leading to an enclosed Milky Way mass, $M_{\text{MW}}(< 50 \text{ kpc}) \approx 3.5 \pm 0.2 \times 10^{11} M_{\odot}$ (Koposov et al. 2023). The fact that we also find a value consistent within 1σ , and with a similar precision, is intriguing given the modelling scenarios are completely unique. At 100 kpc there is a large spread in previously reported values, although our

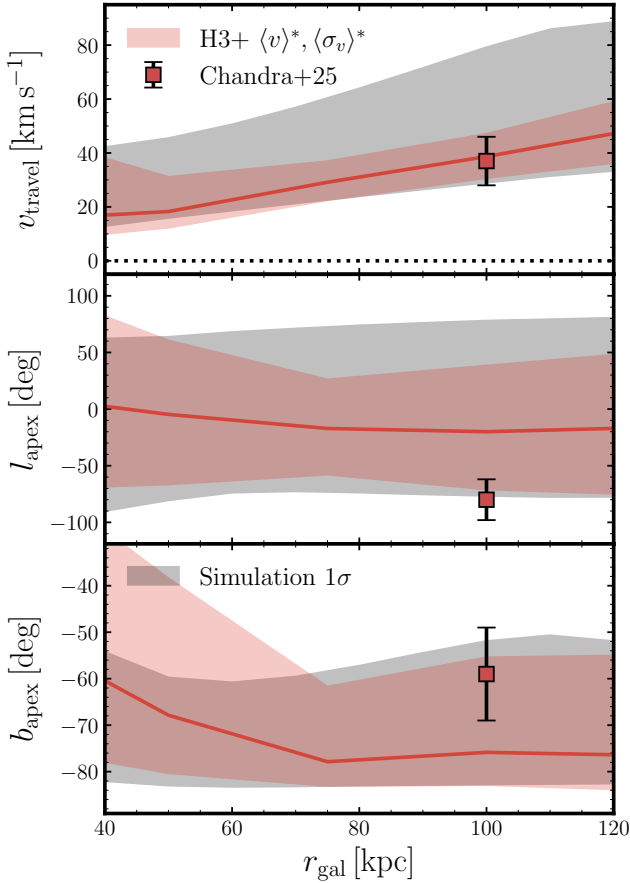


Figure 6. Radial variation of the reflex motion: The radial variation of the magnitude, v_{travel} , and the apex direction, $(l_{\text{apex}}, b_{\text{apex}})$, of the reflex motion posteriors. We show the median as the red line and the 1σ confidence interval as the shaded band. The simulation prior 1σ confidence interval is shown as the grey shaded band. The red data point represents the result from Chandra et al. (2025b) evaluated at 100 kpc using the same H3+SEGUE+MagE dataset but distinct methodology.

posterior tends to agree closely with many at $\sim 5 - 6 \times 10^{11} M_{\odot}$ (e.g., Vasiliev 2019b; Vasiliev et al. 2021). The outermost data points are all consistent with our posterior within uncertainties, although we note our posterior is much more precise. However, the H3+SEGUE+MagE sample only covers distances out to 160 kpc, hence the posterior beyond this distance range is an extrapolation. Potentially, we are subject to overconfidence in our MW mass profile posterior due to the restrictive nature of the model parametrisation i.e., an NFW profile defined by 2 parameters $M_{200, \text{MW}}$ and c_{200} .

6.1.2 The LMC

Similarly, we place our LMC mass results in context with other previous literature values using a range of techniques and dynamical tracers. In Fig. 9 we show the LMC mass enclosed posterior conditioned on the H3+SEGUE+MagE after removing for discrepant data bins. The posterior median and 1σ are represented by the solid red lines and surrounding shaded area. Each literature value is shown as a numbered square, see figure caption for references, with the colour representing the observational tracer used; blue for stellar streams, pink for Local Group, teal for satellites and yellow for LMC

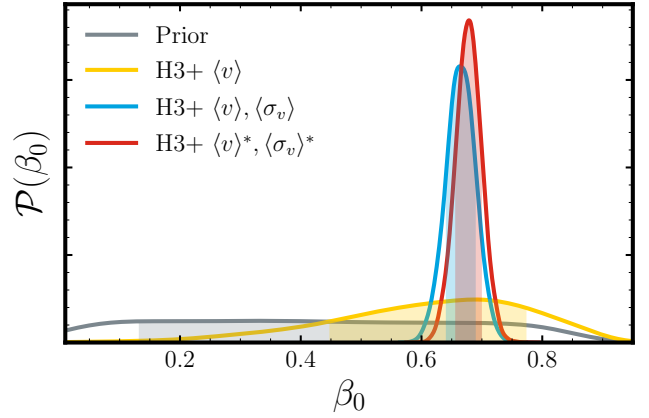


Figure 7. Stellar halo velocity anisotropy: The individual posteriors for the average velocity anisotropy of the stellar halo, β_0 . The yellow and blue contours represent the posteriors conditioned using the means, or means and dispersions, of the stellar velocities, respectively. The red contours use the mean and dispersion data, but removing discrepant data points. The prior 1σ confidence interval is shown as the grey curve.

stars. Within 50 kpc, we find that our posterior mass profile prefers slightly more massive values but remains consistent with other studies within uncertainties (e.g., Koposov et al. 2023; Warren et al. 2025). The outermost data points shown from previous studies all relate to total mass constraints for the LMC. We find that the total mass enclosed by our posterior agrees with other previous studies quoting $M_{\text{LMC}} \sim 15 \times 10^{10} M_{\odot}$.

6.2 Caveats and Limitations

Our approach of using rigid models for the MW–LMC systems comes with various caveats. Some of the limitations of rigid models include: assuming that the MW is in equilibrium prior to the LMC’s infall (e.g., Deason et al. 2017), assuming there is no mass lost or gained by either halo, not accounting for halo triaxiality, and using a simple Chandrasekhar prescription for dynamical friction (see discussion sections in Brooks et al. 2026, 2025). Ideally, one would use a suite of N -body MW–LMC simulations that can naturally account for most of the above limitations. Moreover, N -body simulations would allow the MW disc to move separately from its dark matter halo, and account for deformation to the potentials, better modelling the induced reflex motion from the LMC (Yaaqib et al. 2025). However, as we have shown, even state-of-the-art N -body simulations are in tension with the measured reflex motion parameters; see Fig. 5. In future studies, using a suite of N -body simulations (e.g., Sheng et al. 2025) to train an SBI architecture to estimate the posterior distribution of e.g., the reflex motion parameters, will highlight any biases in the inference of these parameters when using rigid models to train similar SBI networks.

One of the other clear limitations of our rigid MW–LMC model set-up is the inability to explore large regions of simulation parameter space in the Galactocentric y -direction of the reflex motion, v_y . Despite accounting for radial variation in the reflex motion, even the allowed simulation prior space at a reference value of 100 kpc is in tension with direct measurements from observational datasets (Yaaqib et al. 2024; Byström et al. 2025; Chandra et al. 2025b). One possible explanation is that the observational datasets contain undetected substructures which can produce larger values for v_y . For

example, [Chandra et al. \(2025b\)](#) suggest that there may still be some LMC contamination in their data sample; see 3rd bin of the Q4 data in the middle panel of Fig. 1. Currently, our MW stellar haloes are comprised purely of MW star particles and do not contain any substructures. A future interesting avenue to pursue is to investigate the level of contamination required from, e.g., LMC stars, to produce simulated data points and values for v_y which are consistent with the current H3+SEGUE+MagE dataset. Aside from undetected substructures, a rudimentary analysis of our rigid simulations shows that producing sufficiently large v_y values is predominantly sensitive to LMC properties and the strength of dynamical friction.

SBI relies on the fact that any simulations used to train the inference include a sufficient amount of physics to properly describe the observations. The Predictive Posterior Check (PPC) carried out in this work, see Appendix. A and Fig. A2 & A3, show that the simulated data from our rigid MW-LMC models capture most of the observed data points well, although clearly some are not well represented. In Sec. 5, we removed any discrepant data points, those lying 2σ away from the simulation data mean values, and trained a separate SBI posterior estimator. For the returned MW mass, removing the discrepant velocity dispersion data points is important as it biased the mass estimate low by $\sim 10\%$. It remains the scope of a future study to determine whether our rigid models are the limitation here, or whether there are persistent substructures remaining in the data to produce the observed values. Nevertheless, accounting for more of the observational selection effects e.g., better matching the exact spatial distribution of the observed data, and adopting a better set of summary statistics (e.g., using spherical harmonic expansions for the velocity fields, [Cunningham et al. 2020](#)) will help to create more realistic MW-LMC models used to perform the inference.

Finally, the choice of prior distribution for model parameters can influence the final inference result. In [Brooks et al. \(2026\)](#) we found that using either an informative or uninformative prior on the MW and LMC total masses had a negligible effect on the inference. To remove the effect of confirmation bias on our posteriors, we used uninformative priors on the MW and LMC masses in this work.

6.3 The influence of the LMC on Milky Way mass inference

In this section we investigate the importance of including an LMC within simulation models when estimating the MW mass using velocity dispersion data. Often, the MW is assumed to be in a state of dynamical equilibrium when using velocity dispersion measurements to constrain its mass (e.g., Jeans modelling, [Jeans 1915](#); [Binney & Tremaine 2008](#); [Correa Magnus & Vasiliev 2022](#); [Medina et al. 2025](#)). Using the initial distribution of stellar halo particles, one can calculate the radial and tangential velocity dispersion profiles *prior* to the LMC's infall. Therefore, we can measure the radial and tangential velocity dispersions using the same set of distance bins as before when computing the velocity dispersions from the MW stellar haloes evolved in the combined MW-LMC potential to present day; i.e., *post* the LMC's infall. Now, we are able to compare the bias introduced into the MW mass inference by using the velocity dispersion as measured either *prior* or *post* the LMC infall. This is equivalent to having *excluded* or *included* the LMC within our simulations.

To determine the level of bias in the inference, we train a separate neural network to estimate the MW mass using the H3+SEGUE+MagE velocity dispersion data, see Sec. 2, when adopting the velocity dispersions as those *prior* the LMC infall. We adopt the same outer halo slope as in the inference, $\gamma = 4.55$ ([Amarante et al. 2024](#)). All other simulation parameters remain unchanged.

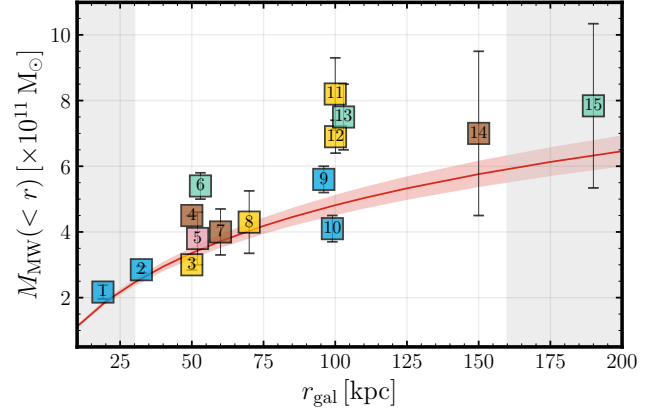


Figure 8. Milky Way mass profile: We show the posterior median and 1σ for the Milky Way enclosed mass as the red line and shaded region using the H3+SEGUE+MagE dataset. Regions outside of the radial range of the dataset, $r_{\text{gal}} \notin [30 - 160]$ kpc, are shaded out. A selection of other previous literature values are coloured by the observational tracer used; blue for stellar streams, brown for BHB stars, pink for RRL stars, teal for satellites and globular cluster, and yellow for other stars. The numerical value indicates the referenced study: 1 – [Dillamore et al. \(2025\)](#), 2 – [Koposov et al. \(2023\)](#), 3 – [Williams et al. \(2017\)](#), 4 – [Williams & Evans \(2015\)](#), 5 – [Medina et al. \(2025\)](#), 6 – [Vasiliev \(2019b\)](#), 7 – [Xue et al. \(2008\)](#), 8 – [Bird et al. \(2022\)](#), 9 – [Vasiliev et al. \(2021\)](#), 10 – [Gibbons et al. \(2014\)](#), 11 – [McMillan \(2017\)](#), 12 – [Shen et al. \(2022\)](#), 13 – [Correa Magnus & Vasiliev \(2022\)](#), 14 – [Deason et al. \(2012\)](#), 15 – [Wang et al. \(2022\)](#). Literature values with the same radial position are slightly separated for readability.

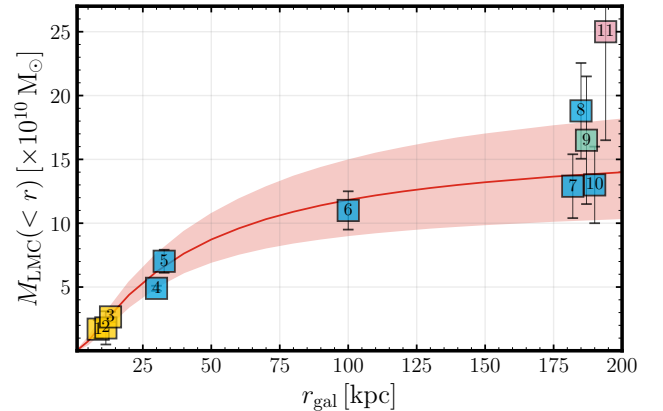


Figure 9. LMC mass profile: We show the posterior median and 1σ for the LMC enclosed mass as the red line and shaded region. Other previous literature values are shown as squares, coloured by the observational tracer used; blue for stellar streams, pink for Local Group, teal for satellites and yellow for LMC stars. The numerical value indicates the referenced study: 1 – [van der Marel & Kallivayalil \(2014\)](#), 2 – [Cullinane et al. \(2020\)](#), 3 – [Watkins et al. \(2024\)](#), 4 – [Warren et al. \(2025\)](#), 5/6/7 – [Koposov et al. \(2023\)](#), 8 – [Shipp et al. \(2021\)](#), 9 – [Correa Magnus & Vasiliev \(2022\)](#), 10 – [Vasiliev et al. \(2021\)](#) and 11 – [Peñarrubia et al. \(2016\)](#). The outermost literature values are total mass estimates, separated on the x-axis for visualisation purposes.

When *excluding* the LMC in the simulations, we find that the MW mass enclosed with 50 kpc prefers to be more massive by $\sim 5\%$ compared to the MW mass estimated when the LMC is *included* in the simulations. Furthermore, the virial mass, $M_{200, \text{MW}}$, is found to be overestimated by $\sim 10\%$ when the LMC is neglected in the simulations. Therefore, the *exclusion* of an LMC leads to the MW mass being biased to *more* massive models. This degree of bias is comparable to the measured precision of the MW mass in this work; see Sec. 5. Interestingly, our findings are similar to Erkal et al. (2020) who found that the MW mass is biased to larger values by $\sim 5\%$ for $M_{\text{MW}}(< 50 \text{ kpc})$ when accounting for the mean velocity motions of stars i.e., the reflex motion, and averaging over stars distributed across the entire sky. Additionally, this level of bias is similar to that found by Correa Magnus & Vasiliev (2022) using satellite tracer populations. Therefore, within our presented methodology, including the LMC to account for dynamical disequilibrium is essential to model the reflex motion in the mean kinematics of MW halo stars; however, the velocity dispersions remain largely insensitive, yielding a relatively unbiased MW mass inference. This is similar to Sheng et al. (2025) who advocate that the LMC should be explicitly included in dynamical models, as mean stellar velocities primarily constrain the LMC, while the velocity dispersions constrain MW properties.

Furthermore, since velocity dispersions are proportional to the total mass of a system, the theoretical expectation is that dispersions prior to the LMC's infall will be smaller in magnitude than those computed post-infall. To investigate this, we measured the dispersions prior to the LMC's infall for all of our simulations using the spherical Jeans equations. We find that the dispersions are $\sim 5 - 10 \text{ km s}^{-1}$ smaller on average prior to the LMC infall. Interestingly, we note that these differences are prominently seen in the latitudinal component of the tangential velocity, $v_{t,b}$, and in the southern quadrants radial velocities, i.e., v_{GSR} in Q3 and Q4, whereas the northern quadrants, Q1 and Q2, remain largely unaffected. This highlights the locality of the LMC's perturbative nature on the velocity dispersions of outer halo stars. Moreover, this is in agreement with Erkal et al. (2020) who also demonstrated that the stellar velocities in the northern hemisphere remain largely unaffected, allowing unbiased estimates of the MW mass when assuming equilibrium models.

7 CONCLUSIONS & OUTLOOK

We have investigated the effect of the LMC in the MW via its perturbation to the kinematics of outer halo stars. From the combined H3+SEGUE+MagE survey, we use the means and dispersions of the radial and tangential velocities of stars out to a distance of 160 kpc to constrain properties of the MW, the LMC and their interaction. We use an SBI framework to train neural posterior estimators using 32,000 rigid MW–LMC simulations. These MW–LMC simulations allow for disequilibrium to develop in the MW's stellar halo via their interaction. Our SBI framework is able to directly map between observable quantities of the MW's stellar halo and simulation parameters. After removing discrepant data relative to the simulated models, we find the following results and conclusions:

(i) The enclosed mass of the MW and LMC within 50 kpc of their respective dynamical centres are determined to be:

The enclosed MW mass: $M_{\text{MW}}(< 50 \text{ kpc}) = 3.36^{+0.15}_{-0.15} \times 10^{11} M_{\odot}$

The enclosed LMC mass: $M_{\text{LMC}}(< 50 \text{ kpc}) = 8.76^{+1.94}_{-1.77} \times 10^{10} M_{\odot}$

These new constraints are consistent within uncertainties of other previously reported values for both galaxies.

(ii) The inference of the total mass of each halo implies a ratio of the total LMC infall mass to the total present-day MW mass of $\sim 30 \pm 10\%$.

(iii) The average velocity anisotropy of the stellar halo, prior to the LMC's infall is well constrained using the velocity dispersions of outer halo stars to be $\beta_0 = 0.68 \pm 0.02$. This suggests that the MW's stellar halo was already radially anisotropic out to large distances prior to the LMC's recent passage. Moreover, this is consistent with, but generally lower than, the velocity anisotropy measured at present day in the Solar neighbourhood.

(iv) Using a linear continuity model to capture the radial variation in the reflex motion, we find the magnitude of the reflex motion, evaluated at a reference distance of 100 kpc, to be:

The reflex motion velocity: $v_{\text{travel}} = 38.6^{+8.3}_{-7.8} \text{ km s}^{-1}$

Meanwhile, the direction of the reflex motion remains challenging to constrain, further highlighting the inconsistencies with and between other previous measurements.

(v) If the influence of the LMC is not accounted for in the simulations, we find that the MW virial mass is biased to larger values by up to $\sim 10\%$. Hence, while the LMC is crucial for modelling dynamical disequilibrium in the mean halo kinematics, it has minimal impact on velocity dispersion-based constraints of the MW mass.

Our analysis has produced some of the most precise results to date on properties of the MW and LMC, as well as their relative interaction via the reflex motion. In the near future, there will be an increase in the raw number of outer halo stars from SDSS-V (SDSS Collaboration et al. 2025). Moreover, this SDSS-V catalogue can be combined with distance and proper motion measurements, which have much improved precision, from Gaia DR4. Both of these factors, increased precision and number counts of outer halo stars, will improve the inference performed using this SBI framework, provided that observational facilities remain the dominant source of uncertainty.

ACKNOWLEDGEMENTS

We thank the reviewer for a detailed and constructive report that improved this study. RANB acknowledges support from the Royal Society and would like to thank Dennis Zaritsky and Ben Johnson for comments on the manuscript as well as the entire H3 team for access to their data. JLS acknowledges the support of the Royal Society (URFR1\191555; URFR\241030). NGC acknowledges support from the Heising-Simons Foundation grant #2022-3927, through which NGC is supported by the Barbara Pichardo Future Faculty Fellowship.

DATA AVAILABILITY

The MW–LMC simulation parameters and the data for 32,000 unique stellar haloes, each with 20,000 particles, will be shared upon reasonable request. A Jupyter notebook to make the stereographic projection of the reflex motion apex direction can be found [here](#).

Software: sbi (Tejero-Cantero et al. 2020), agama (Vasiliev 2019a),

gala (Price-Whelan 2017), NumPy (Harris et al. 2020), Matplotlib (Hunter 2007), Seaborn (Waskom 2021), Astropy (Astropy Collaboration et al. 2013, 2018, 2022), SciPy (Virtanen et al. 2020).

REFERENCES

- Alsing J., Charnock T., Feeney S., Wandelt B., 2019, *MNRAS*, **488**, 4440
- Amarante J. A. S., Kposov S. E., Laporte C. F. P., 2024, *A&A*, **690**, A166
- An J. H., Evans N. W., 2006, *ApJ*, **642**, 752
- Arora A., et al., 2025, *ApJ*, **988**, 190
- Astropy Collaboration et al., 2013, *A&A*, **558**, A33
- Astropy Collaboration et al., 2018, *AJ*, **156**, 123
- Astropy Collaboration et al., 2022, *apj*, **935**, 167
- Belokurov V., Deason A. J., Erkal D., Kposov S. E., Carballo-Bello J. A., Smith M. C., Jethwa P., Navarrete C., 2019, *MNRAS*, **488**, L47
- Benisty D., 2024, *A&A*, **689**, L1
- Besla G., Kallivayalil N., Hernquist L., Robertson B., Cox T. J., van der Marel R. P., Alcock C., 2007, *ApJ*, **668**, 949
- Besla G., Kallivayalil N., Hernquist L., van der Marel R. P., Cox T. J., Kereš D., 2010, *ApJ*, **721**, L97
- Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*
- Bird S. A., Xue X.-X., Liu C., Shen J., Flynn C., Yang C., Zhao G., Tian H.-J., 2021, *ApJ*, **919**, 66
- Bird S. A., et al., 2022, *MNRAS*, **516**, 731
- Bobylev V. V., Baykova A. T., 2023, *Astronomy Reports*, **67**, 812
- Boylan-Kolchin M., Besla G., Hernquist L., 2011, *MNRAS*, **414**, 1560
- Brooks R. A. N., Sanders J. L., Chandra V., Garavito-Camargo N., Dillamore A. M., Price-Whelan A. M., Ting Y.-S., 2025, *arXiv e-prints*, p. [arXiv:2510.04735](https://arxiv.org/abs/2510.04735)
- Brooks R. A. N., Sanders J. L., Dillamore A. M., Garavito-Camargo N., Price-Whelan A. M., 2026, *MNRAS*, **545**, staf2111
- Byström A., et al., 2025, *MNRAS*, **542**, 560
- Cargile P., Chandra V., 2025, *The SEGUE-MINESweeper Catalog*, doi:10.5281/zenodo.16105186, <https://doi.org/10.5281/zenodo.16105186>
- Cargile P. A., Conroy C., Johnson B. D., Ting Y.-S., Bonaca A., Dotter A., Speagle J. S., 2020, *ApJ*, **900**, 28
- Cavieres M., et al., 2025, *ApJ*, **983**, 83
- Chandra V., et al., 2023a, *ApJ*, **951**, 26
- Chandra V., et al., 2023b, *ApJ*, **956**, 110
- Chandra V., et al., 2025a, *arXiv*, p. [arXiv:2508.00978](https://arxiv.org/abs/2508.00978)
- Chandra V., et al., 2025b, *ApJ*, **988**, 156
- Chandrasekhar S., 1943, *ApJ*, **97**, 255
- Conroy C., et al., 2019, *ApJ*, **883**, 107
- Conroy C., Naidu R. P., Garavito-Camargo N., Besla G., Zaritsky D., Bonaca A., Johnson B. D., 2021, *Nature*, **592**, 534
- Correa Magnus L., Vasiliev E., 2022, *MNRAS*, **511**, 2610
- Cranmer K., Brehmer J., Louppe G., 2020, *Proceedings of the National Academy of Science*, **117**, 30055
- Cullinane L. R., et al., 2020, *MNRAS*, **497**, 3055
- Cunningham E. C., et al., 2019, *ApJ*, **879**, 120
- Cunningham E. C., et al., 2020, *ApJ*, **898**, 4
- DESI Collaboration et al., 2025, *arXiv e-prints*, p. [arXiv:2503.14745](https://arxiv.org/abs/2503.14745)
- Darragh-Ford E., et al., 2025, *arXiv e-prints*, p. [arXiv:2511.02031](https://arxiv.org/abs/2511.02031)
- Deason A. J., Belokurov V., 2024, *New Astron. Rev.*, **99**, 101706
- Deason A. J., et al., 2012, *MNRAS*, **425**, 2840
- Deason A. J., Belokurov V., Kposov S. E., Gómez F. A., Grand R. J., Marinacci F., Pakmor R., 2017, *MNRAS*, **470**, 1259
- Dillamore A. M., Sanders J. L., Belokurov V., Zhang H., 2025, *MNRAS*, **541**, 214
- Dillamore A. M., Sanders J. L., Brooks R. A. N., 2026, *arXiv e-prints*, p. [arXiv:2603.11159](https://arxiv.org/abs/2603.11159)
- Erkal D., et al., 2019a, *MNRAS*, **487**, 2685
- Erkal D., et al., 2019b, *MNRAS*, **487**, 2685
- Erkal D., Belokurov V. A., Parkin D. L., 2020, *MNRAS*, **498**, 5574
- Erkal D., et al., 2021, *MNRAS*, **506**, 2677
- Fearnhead P., Prangle D., 2010, *arXiv e-prints*, p. [arXiv:1004.1112](https://arxiv.org/abs/1004.1112)
- Foot H. R., et al., 2023, *ApJ*, **954**, 163
- Fushimi K. J., Mosquera M. E., Dominguez M., 2024, *A&A*, **688**, A147
- Gaia Collaboration et al., 2023, *A&A*, **674**, A1
- Garavito-Camargo N., Besla G., Laporte C. F. P., Johnston K. V., Gómez F. A., Watkins L. L., 2019, *ApJ*, **884**, 51
- Garavito-Camargo N., Besla G., Laporte C. F. P., Price-Whelan A. M., Cunningham E. C., Johnston K. V., Weinberg M., Gómez F. A., 2021a, *ApJ*, **919**, 109
- Garavito-Camargo N., Besla G., Laporte C. F. P., Price-Whelan A. M., Cunningham E. C., Johnston K. V., Weinberg M., Gómez F. A., 2021b, *ApJ*, **919**, 109
- Gibbons S. L. J., Belokurov V., Evans N. W., 2014, *MNRAS*, **445**, 3788
- Gómez F. A., Besla G., Carpintero D. D., Villalobos Á., O'Shea B. W., Bell E. F., 2015, *ApJ*, **802**, 128
- Han J. J., et al., 2022, *AJ*, **164**, 249
- Han J. J., Conroy C., Zaritsky D., Bonaca A., Caldwell N., Chandra V., Ting Y.-S., 2024, *arXiv e-prints*, p. [arXiv:2406.12969](https://arxiv.org/abs/2406.12969)
- Harris C. R., et al., 2020, *Nature*, **585**, 357
- Hashimoto Y., Funato Y., Makino J., 2003, *ApJ*, **582**, 196
- Hermans J., Banik N., Weniger C., Bertone G., Louppe G., 2021, *MNRAS*, **507**, 1999
- Hernquist L., 1990, *ApJ*, **356**, 359
- Ho M., et al., 2024, *The Open Journal of Astrophysics*, **7**, 54
- Horta D., Price-Whelan A. M., Kposov S. E., Hunt J. A. S., Hogg D. W., Filion C., Daniel K. J., 2026, *arXiv e-prints*, p. [arXiv:2601.18876](https://arxiv.org/abs/2601.18876)
- Hunt J. A. S., Vasiliev E., 2025, *arXiv e-prints*, p. [arXiv:2501.04075](https://arxiv.org/abs/2501.04075)
- Hunter J. D., 2007, *Computing in Science & Engineering*, **9**, 90
- Jeans J. H., 1915, *MNRAS*, **76**, 70
- Jeffrey N., Alsing J., Lanusse F., 2021, *MNRAS*, **501**, 954
- Jeffrey N., et al., 2025, *MNRAS*, **536**, 1303
- Jethwa P., Erkal D., Belokurov V., 2016, *MNRAS*, **461**, 2212
- Jimenez Rezende D., Mohamed S., 2015, *arXiv e-prints*, p. [arXiv:1505.05770](https://arxiv.org/abs/1505.05770)
- Johnson B. D., et al., 2020, *ApJ*, **900**, 103
- Kallivayalil N., van der Marel R. P., Besla G., Anderson J., Alcock C., 2013, *ApJ*, **764**, 161
- Koposov S. E., et al., 2023, *MNRAS*, **521**, 4936
- Kravtsov A., Winney S., 2024, *arXiv e-prints*, p. [arXiv:2405.06017](https://arxiv.org/abs/2405.06017)
- Küpper A. H. W., Balbinot E., Bonaca A., Johnston K. V., Hogg D. W., Kroupa P., Santiago B. X., 2015, *ApJ*, **803**, 80
- Lemos P., Jeffrey N., Whiteway L., Lahav O., Libeskind N., Hoffman Y., 2021, *Phys. Rev. D*, **103**, 023009
- Lemos P., Coogan A., Hezaveh Y., Perreault-Levasseur L., 2023, *40th International Conference on Machine Learning*, **202**, 19256
- Li S., et al., 2025, *arXiv e-prints*, p. [arXiv:2512.01350](https://arxiv.org/abs/2512.01350)
- Lovell C. C., et al., 2024, *arXiv e-prints*, p. [arXiv:2411.13960](https://arxiv.org/abs/2411.13960)
- Lueckmann J.-M., Boelts J., Greenberg D., Gonçalves P., Macke J., 2021, in Banerjee A., Fukumizu K., eds, *Proceedings of Machine Learning Research* Vol. 130, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*. PMLR, pp 343–351, <https://proceedings.mlr.press/v130/lueckmann21a.html>
- Majewski S. R., Skrutskie M. F., Weinberg M. D., Ostheimer J. C., 2003, *ApJ*, **599**, 1082
- Malhan K., Ibata R. A., 2019, *MNRAS*, **486**, 2995
- Mansfield M., Sanderson R., Hey D., Huber D., Arora A., Cunningham E., Panithanpaisal N., 2026, *arXiv e-prints*, p. [arXiv:2602.03528](https://arxiv.org/abs/2602.03528)
- Marshall J. L., et al., 2008, in McLean I. S., Casali M. M., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* Vol. 7014, *Ground-based and Airborne Instrumentation for Astronomy II*. p. 701454 ([arXiv:0807.3774](https://arxiv.org/abs/0807.3774)), doi:10.1117/12.789972
- McMillan P. J., 2017, *MNRAS*, **465**, 76
- Medina G. E., et al., 2025, *arXiv e-prints*, p. [arXiv:2508.19351](https://arxiv.org/abs/2508.19351)
- Merritt D., 1985, *AJ*, **90**, 1027
- Mittal S., Leif Bracher N., Lajoie G., Jaini P., Brubaker M., 2025, *arXiv e-prints*, p. [arXiv:2502.06601](https://arxiv.org/abs/2502.06601)
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, **462**, 563
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, **490**, 493
- Nguyen T., et al., 2025, *arXiv e-prints*, p. [arXiv:2512.07960](https://arxiv.org/abs/2512.07960)
- Osipkov L. P., 1979, *Pisma v Astronomicheskii Zhurnal*, **5**, 77

- Papamakarios G., Pavlakou T., Murray I., 2017, *arXiv e-prints*, p. [arXiv:1705.07057](#)
- Papamakarios G., Nalisnick E., Jimenez Rezende D., Mohamed S., Lakshminarayanan B., 2019, *arXiv e-prints*, p. [arXiv:1912.02762](#)
- Patel E., et al., 2020, *ApJ*, **893**, 121
- Peñarrubia J., Gómez F. A., Besla G., Erkal D., Ma Y.-Z., 2016, *MNRAS*, **456**, L54
- Petersen M. S., Peñarrubia J., 2020, *MNRAS*, **494**, L11
- Petersen M. S., Peñarrubia J., 2021, *Nature Astronomy*, **5**, 251
- Price-Whelan A. M., 2017, *The Journal of Open Source Software*, **2**
- Pritchard J. K., Seielstad M. T., Perez-Lezaun A., Feldman M. W., 1999, *Molecular Biology and Evolution*, **16**, 1791
- Rozier S., Famaey B., Siebert A., Monari G., Pichon C., Ibata R., 2022, *ApJ*, **933**, 113
- Rubin D. B., 1984, *The Annals of Statistics*, **12**, 1151
- SDSS Collaboration et al., 2025, *arXiv e-prints*, p. [arXiv:2507.07093](#)
- Sante A., Kawata D., Font A. S., Grand R. J. J., 2025, GalactiKit: reconstructing mergers from $z = 0$ debris using simulation-based inference in Auriga ([arXiv:2502.14972](#)), <https://arxiv.org/abs/2502.14972>
- Saoulis A. A., Piras D., Jeffrey N., Spurio Mancini A., Ferreira A. M. G., Joachimi B., 2025, *arXiv e-prints*, p. [arXiv:2505.21215](#)
- Shen J., et al., 2022, *ApJ*, **925**, 1
- Sheng Y., Ting Y.-S., Xue X.-X., Chang J., Tian H., 2024, *MNRAS*, **534**, 2694
- Sheng Y., Ting Y.-S., Xue X.-X., 2025, *arXiv e-prints*, p. [arXiv:2507.03663](#)
- Shipp N., et al., 2021, *ApJ*, **923**, 149
- Sohn S. T., Besla G., van der Marel R. P., Boylan-Kolchin M., Majewski S. R., Bullock J. S., 2013, *ApJ*, **768**, 139
- Szentgyorgyi A., et al., 2011, *PASP*, **123**, 1188
- Talts S., Betancourt M., Simpson D., Vehtari A., Gelman A., 2018, *arXiv e-prints*, p. [arXiv:1804.06788](#)
- Tamfal T., Mayer L., Quinn T. R., Capelo P. R., Kazantzidis S., Babul A., Potter D., 2021, *ApJ*, **916**, 55
- Tejero-Cantero A., Boelts J., Deistler M., Lueckmann J.-M., Durkan C., Gonçalves P. J., Greenberg D. S., Macke J. H., 2020, *Journal of Open Source Software*, **5**, 2505
- Vasiliev E., 2019a, *MNRAS*, **482**, 1525
- Vasiliev E., 2019b, *MNRAS*, **484**, 2832
- Vasiliev E., 2023, *Galaxies*, **11**, 59
- Vasiliev E., Belokurov V., Erkal D., 2021, *MNRAS*, **501**, 2279
- Virtanen P., et al., 2020, *Nature Methods*, **17**, 261
- Wang W., Han J., Cautun M., Li Z., Ishigaki M. N., 2020, *Science China Physics, Mechanics, and Astronomy*, **63**, 109801
- Wang J., Hammer F., Yang Y., 2022, *MNRAS*, **510**, 2242
- Warren H. T., Newberg H. J., Guffey A. G., Lin J., Mendelsohn E. J., Roux K., 2025, *arXiv e-prints*, p. [arXiv:2508.01086](#)
- Waskom M. L., 2021, *Journal of Open Source Software*, **6**, 3021
- Watkins L. L., van der Marel R. P., Bennet P., 2024, *ApJ*, **963**, 84
- Weinberg M. D., 1998, *MNRAS*, **299**, 499
- Weyant A., Schafer C., Wood-Vasey W. M., 2013, *ApJ*, **764**, 116
- White S. D. M., 1976, *MNRAS*, **174**, 467
- Widmark A., Johnston K. V., 2025, *arXiv e-prints*, p. [arXiv:2501.13148](#)
- Williams A. A., Evans N. W., 2015, *MNRAS*, **454**, 698
- Williams A. A., Belokurov V., Casey A. R., Evans N. W., 2017, *MNRAS*, **468**, 2359
- Xiangyuan Ma P., Rogers K. K., Li T. S., Hložek R., Webb J., Huang R., Meunier J., 2025, *arXiv e-prints*, p. [arXiv:2502.03522](#)
- Xue X. X., et al., 2008, *ApJ*, **684**, 1143
- Yaaqib R., Petersen M. S., Peñarrubia J., 2024, *arXiv e-prints*, p. [arXiv:2402.10841](#)
- Yaaqib R., Petersen M., Peñarrubia J., 2025, *arXiv e-prints*, p. [arXiv:2508.04781](#)
- Yanny B., et al., 2009, *AJ*, **137**, 4377
- York D. G., et al., 2000, *AJ*, **120**, 1579
- Zaritsky D., Courtois H., 2017, *MNRAS*, **465**, 3724
- Zaritsky D., Olszewski E. W., Schommer R. A., Peterson R. C., Aaronson M., 1989, *ApJ*, **345**, 759
- van der Marel R. P., Kallivayalil N., 2014, *ApJ*, **781**, 121
- van der Marel R. P., Besla G., Cox T. J., Sohn S. T., Anderson J., 2012, *ApJ*, **753**, 9
- von Wietersheim-Kramsta M., Lin K., Tessore N., Joachimi B., Loureiro A., Reischke R., Wright A. H., 2024, *arXiv e-prints*, p. [arXiv:2404.15402](#)

APPENDIX A: POSTERIOR DIAGNOSTIC CHECKS

Any posterior from a generative approach should be assessed for its accuracy through a variety of diagnostic tests to gain trust that the inference has been successfully performed. To trust the results presented in this work we carry out posterior coverage probability checks and predictive posterior checks.

A1 Coverage probability test

A coverage probability test is one way to assess the accuracy of an estimated posterior. In Bayesian analysis, a coverage test checks whether credible intervals have the expected probabilities (see, [Jeffrey et al. 2025](#), sec. 2.4 for a concise explanation). In a 1-dimensional posterior setting, one can define a particular credible interval to be the narrowest interval containing, for example, 68% of the probability weight. The Bayesian inference procedure takes in some observed data, D_{obs} , and determines a posterior distribution, $p(\theta|D_{\text{obs}})$, and hence a credible interval for θ . For a coverage test, one uses a test parameter, θ_{test} , selected from the prior, $p(\theta)$, as the input to a simulation that produces the corresponding output data point, D_{test} . From this, one can derive a posterior, $p(\theta|D_{\text{test}})$, and therefore a credible interval. If the inference process has been correctly implemented, then the true test parameter value, θ_{test} , will fall in this credible interval, in this example, 68% of the time. Repeating this test for many sampled θ_{test} , and varying the credibility intervals, one can gain trust that the estimated posterior distributions are accurate and have reliable confidence intervals.

To perform a coverage test on our SBI posteriors, we use ‘Tests of Accuracy with Random Points’ (TARP, [Lemos et al. 2023](#), see their figs. 1&2 for further intuition) as implemented in the *sbi* Python package ([Tejero-Cantero et al. 2020](#)). For our application of SBI, this test is relatively straightforward as we have many pre-existing simulations with an amortised inference scheme, i.e., each data evaluation is computationally cheap without the need to retrain the neural network ([Mittal et al. 2025](#)). TARP coverage probabilities test the accuracy of estimated posteriors by only using samples from the posterior. This technique is similar to simulation based calibration ([Talts et al. 2018](#)) but extends the idea to the full-dimensional posterior space instead of being restricted to 1-dimension.

In [Fig. A1](#), we demonstrate that the expected coverage does indeed match the credibility level for the estimated posteriors in this work conditioned using both the mean *and* dispersion velocity data from H3+SEGUE+MagE. This validates our neural posterior estimation as being truly representative of the probability that each of our model parameters has some true value with truly representative uncertainties. This can be further quantified in two ways. Firstly, we can compute the area between the ideal TARP curve and our posterior TARP curves for credibility intervals greater than 0.5; namely, the Area To Curve (ATC) value. This number should be close to 0, a value $\gg 0$ indicates an estimated posterior that is too wide, conversely, a value $\ll 0$ indicates that the estimated posterior is too narrow. Secondly, we can calculate the p-value of a Kolmogorov-Smirnov test. The null hypothesis is that an exact one-to-one curve and our posterior TARP curve are identical. If this p-value is less than 0.05, then this null hypothesis is rejected. For our SBI posterior we report an

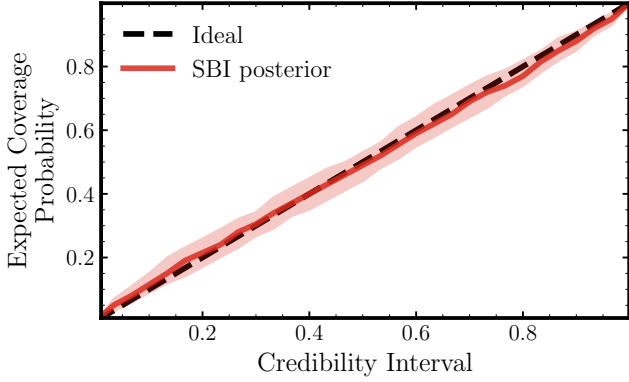


Figure A1. Coverage probability posterior check: For the SBI posterior conditioned on the mean and dispersion H3+SEGUE+MagE data, after removing discrepant data points, we show the probability of finding true test parameters in the appropriate credible intervals matches the expected coverage probability. A bootstrapped 1σ uncertainty is shown as the shaded region. This validates the estimated SBI posterior distributions and allows one to interpret the confidence intervals on parameter constraints as representative and reliable.

ATC magnitude $\lesssim 0.1$ and a Kolmogorov-Smirnov p-value of 1.0. These values suggest that we are not drastically over-/under-fitting, or biased, and are not required to reject the posterior.

A2 Predictive posterior check

We carry out a ‘Posterior Predictive Check’ (PPC) to act as a complementary diagnostic test. This test makes use of the fact that if the inference has been correctly implemented, then any generated data, D_{pp} , using simulation parameters as sampled from the posterior, θ_{pp} , should be similar to the observed data, D_{obs} (Lueckmann et al. 2021). A PPC provides intuition about any bias introduced in inference. For example, it can help to determine whether or not the generated data systematically differ from the observed data used during the inference. We carry out a PPC for the posterior conditioned on the mean and dispersion velocity data from H3+SEGUE+MagE. To do this, we first sample model parameter values from these posterior distributions. We then re-run rigid MW-LMC simulations adopting these parameter values, generating output data as the mean and dispersion of the radial and tangential velocity measurements, which we can then use to compare to the observed data that was originally used to condition the posterior.

In Fig. A2 and Fig. A3, we demonstrate the PCC for the posteriors conditioned on the full H3+SEGUE+MagE dataset i.e., open blue contours in all figures. Using parameters sampled from this posterior, we show the generated data for the radial and tangential, mean and dispersion, velocity data as the violin plot contours. The mean observed data points are shown as the same coloured markers; same points as in Fig. 1 and Fig. 2. Measurement errors are not shown because the generated data incorporates survey uncertainties. In most distance bins the observed data is found to be well represented by the generated simulation data. However, there are clearly some data points which are difficult to reconcile with the simulations. For example, the outermost bins in Q1 for the mean radial velocity, or similarly for the 3rd/4th Q4 bins for the radial velocity dispersions, are hard to explain with the current modelling. This is likely a limitation of using reasonably simple rigid MW-LMC simulations which are able to capture global velocity perturbations, but struggle to capture smaller

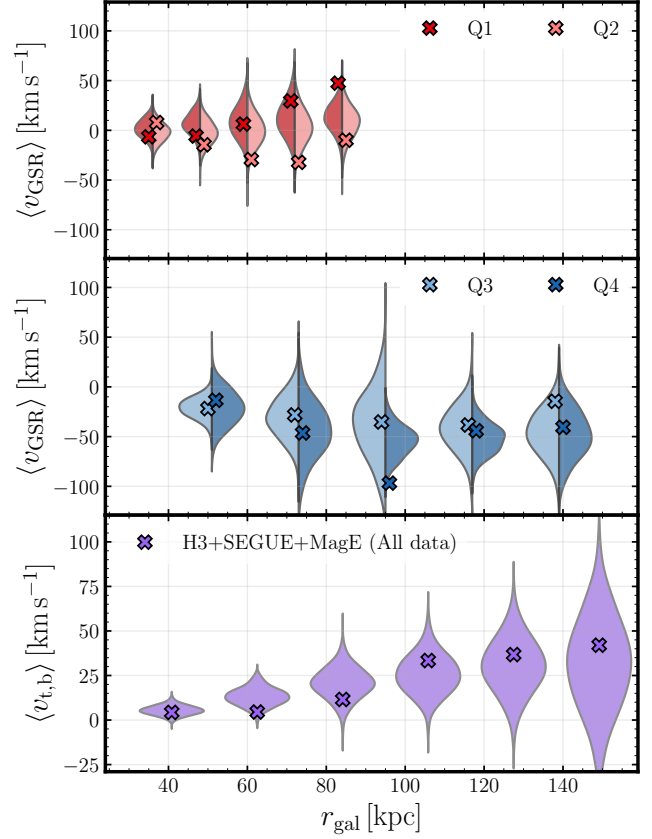


Figure A2. Predictive Posterior Check - mean velocities: We test the quality of the SBI posterior returned using the H3+SEGUE+MagE data. We show the mean radial (top and middle panels for northern and southern quadrants, respectively) and tangential (bottom panel) velocity data as generated using samples from this posterior as the violin plot contours. The observed data points are shown as the same coloured crosses. We do not show the data errorbars as the generated data already accounts for the survey uncertainties. The generated and original data look sufficiently similar implying the SBI posteriors are representative of the observed data.

scale perturbations. Although, even N -body deforming simulations (Garavito-Camargo et al. 2019), struggle to reproduce these velocity trends in the northern Galactic hemisphere (see fig. 4, Chandra et al. 2025b). Potentially, some of these discrepancies between the observed and generated data could be due to unresolved substructure across these distances, for example stars stripped from the LMC (Chandra et al. 2025b).

One of the advantages of SBI is the flexibility of choosing which data to train the inference network on. To investigate any biases introduced by including all data points in our inference, we removed data points that lie 2σ away from the mean of each simulation distribution. With this reduced set of velocity summary statistics for outer halo stars, we re-train the inference framework, resulting in the red posterior distributions shown throughout all figures in Sec. 5; see that section for more details on the degree of any bias introduced.

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.

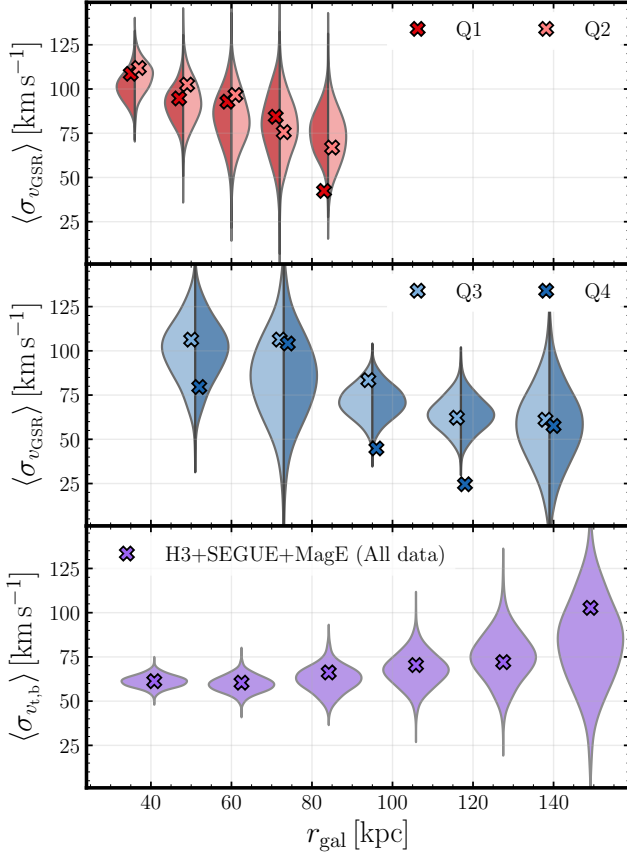


Figure A3. Predictive Posterior Check - velocity dispersions: We test the quality of the SBI posterior returned using the H3+SEGUE+MagE data. We show the dispersion of the radial (top and middle panels for northern and southern quadrants, respectively) and tangential (bottom panel) velocity data as generated using samples from this posterior as the violin plot contours. The observed data points are shown as the same coloured crosses. We do not show the data errorbars as the generated data already accounts for the survey uncertainties.