

The Milky Way – Large Magellanic Cloud Interaction with Simulation Based Inference

Richard A. N. Brooks^{1,2,★}, Jason L. Sanders¹, Vedant Chandra³, Nicolás Garavito-Camargo^{4,†}, Adam M. Dillamore¹, Adrian M. Price-Whelan², Yuan-Sen Ting^{5,6}

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

²Center for Computational Astrophysics, Flatiron Institute, Simons Foundation, 162 Fifth Avenue, New York, NY 10010, USA

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

⁴Department of Astronomy, University of Maryland, College Park, MD 20742, USA

⁵Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA

⁶Center for Cosmology and AstroParticle Physics (CCAPP), The Ohio State University, Columbus, OH 43210, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The infall of the Large Magellanic Cloud (LMC) into the Milky Way (MW) has displaced the MW’s centre of mass, manifesting as an observed reflex motion in the velocities of outer halo stars. We use a Simulation Based Inference framework to constrain properties of the MW, LMC and the induced reflex motion using the dynamics of outer MW halo stars. Specifically, we use the mean radial and tangential velocities of outer halo stars calculated in a set of distance and on-sky bins. We train neural networks to estimate parameter posterior distributions using a set of 128,000 rigid MW–LMC simulations conditioned upon velocity data from the Dark Energy Spectroscopic Instrument (DESI) and the combined H3+SEGUE+MagE outer halo surveys. We constrain the reflex motion velocity and the enclosed LMC mass within 50 kpc using the DESI or H3+SEGUE+MagE dataset while varying the survey sky coverage and depth. Using the radial and tangential velocity data from the H3+SEGUE+MagE survey and on-sky quadrants, we report a distance-averaged reflex motion velocity for the outer halo samples, the speed at which the MW lurches towards the LMC, of $v_{\text{travel}} = 26.4^{+5.5}_{-4.4}$ km s^{−1}, while simultaneously finding an enclosed LMC mass of $M_{\text{LMC}}(< 50 \text{ kpc}) = 9.2^{+1.9}_{-2.3} \times 10^{10} M_{\odot}$. Quoted uncertainties are statistical. Our results suggest that the LMC’s total mass is at least $\approx 10 - 15\%$ of that of the MW. This inference framework is flexible such that it can provide rapid constraints when applied to any future survey measuring the velocities of outer halo stars.

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

1 INTRODUCTION

The Milky Way (MW) is undergoing a merger with the Large Magellanic Cloud (LMC, see Vasiliev 2023, for a comprehensive review of the effect of the LMC on the MW). The LMC is thought to be on its first pericentric passage and to have a dark matter mass $M_{\text{LMC}} \sim 10^{11} M_{\odot}$ (Besla et al. 2007, 2010; Boylan-Kolchin et al. 2011; Peñarrubia et al. 2016; Kravtsov & Winney 2024). An alternative scenario proposes the LMC is on its second pericentric passage; however, most observable features of this earlier passage scenario are superseded by the more recent, closer, pericentric passage (Vasiliev 2024). Such a large mass for the LMC is required to explain a plethora of Local Group phenomena: for example, the kinematics of its globular clusters (Watkins et al. 2024), the kinematics of MW satellites (Patel et al. 2020; Correa Magnus & Vasiliev 2022; Kravtsov & Winney 2024); dynamical models of MW stellar streams

(Erkal et al. 2019a; Koposov et al. 2019; Shipp et al. 2021; Vasiliev et al. 2021; Koposov et al. 2023; Warren et al. 2025); and the timing argument (Peñarrubia et al. 2016, but see also Benisty et al. 2022; Chamberlain et al. 2023; Benisty 2024) all require an LMC mass $M_{\text{LMC}} \sim 1 - 2 \times 10^{11} M_{\odot}$ (see fig. 1 of Vasiliev 2023, for a summary of LMC mass estimates). At present-day the LMC is at a distance of $d = (49.6 \pm 0.5)$ kpc (Pietrzyński et al. 2019), and heliocentric line of sight velocity of $v_{\text{los}} = (262.2 \pm 3.4)$ km s^{−1} (van der Marel & Kallivayalil 2014). An open question remains on the exact position, and hence velocity, of the LMC centre. A compilation of the recent reported LMC centre position and proper motion measurements is given in table 2 of Vasiliev (2023). The orbit of the LMC is sensitive to the assumed MW potential (see fig. 3 of Vasiliev 2023) and, because the LMC has a mass comparable to that of the MW ($M_{\text{MW}} \sim 10^{12} M_{\odot}$ Wang et al. 2020), it is also subject to dynamical friction from the MW dark matter halo (Chandrasekhar 1943).

The recent infall of the LMC into the MW generates a density wake in the MW dark matter halo (Chandrasekhar 1943). This occurs as the infalling LMC has a range of orbital frequencies which resonate

★ E-mail: richard.brooks.22@ucl.ac.uk

† NASA Hubble Fellowship Program, Einstein Fellow

with the orbits of dark matter particles of the MW’s dark matter halo (Mulder 1983; Weinberg 1986). The classical ‘conic’ wake trailing the LMC is typically labelled as the *transient response*, whereas the response elsewhere in the MW halo is labelled the *collective response* mainly caused by the shift in the systems’ barycentre as well as containing some large scale resonance features (Weinberg 1989; Garavito-Camargo et al. 2019, 2021; Tamfal et al. 2021; Rozier et al. 2022; Foote et al. 2023).

The large mass ratio of the LMC to the MW, $\sim 10 - 20\%$, combined with the fact that it has just completed its most recent pericentre passage at a relative velocity $> 300 \text{ km s}^{-1}$, has caused significant dynamical disequilibrium throughout the Galaxy (Hunt & Vasiliev 2025). In particular, the inner and outer parts of the MW halo have experienced different strengths of acceleration towards the LMC. 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. 2021; Conroy et al. 2021; Amarante et al. 2024) that is higher in the Galactic North, and a dipole in stellar radial velocities (Erkal et al. 2019a; Petersen & Peñarrubia 2020a, 2021; Erkal et al. 2021; Yaaqib et al. 2024; Chandra et al. 2025; Byström et al. 2025). Moreover, the density wake of the LMC is predicted to leave an observable signature in the density and kinematics of MW halo stars (e.g., Belokurov et al. 2019; Conroy et al. 2021; Cavieres et al. 2025; Chandra et al. 2025; Yaaqib et al. 2024; Byström et al. 2025; Fushimi et al. 2024; Amarante et al. 2024; Sheng et al. 2025). The magnitude of the reflex velocity dipole is called the *travel velocity* and its orientation is called the *apex* direction of the reflex motion. Recent studies have shown that the direction of the travel velocity vector points towards a point along the past orbit of the LMC; however, there is variance between studies on the exact preferred direction (e.g., Byström et al. 2025, fig. 9). This is likely due to each study using different stellar tracers over varying radial ranges and sky coverages. A flexible and reliable technique which can account for differences between surveys will be powerful when performing inference for the reflex velocity, especially given the upcoming influx of outer halo datasets.

The MW–LMC system can be described by various levels of simulation fidelity. For any type of simulation, there exists a trade-off between the simulation fidelity and the ability to explore a large model parameter space. The simplest prescriptions are *rigid* models of the MW and LMC galaxies. Rigid models describe the MW and LMC as analytic potentials that have fixed functional forms, i.e., they are time-invariant, although they are allowed to move in response to each other. The fidelity of the MW–LMC system can be increased by using *N*-body simulations in combination with Basis Function Expansions (BFEs, e.g., Lilley et al. 2018a,b; Sanders et al. 2020; Petersen & Peñarrubia 2020a; Garavito-Camargo et al. 2019, 2021; Lilleengen et al. 2023; Vasiliev 2024, often called deforming MW–LMC simulations). They aim to match the present-day conditions of the MW and LMC while also accounting for the dark matter halo responses of both galaxies during the infall of the LMC. Furthermore, MW–LMC analogues have been identified in state-of-the-art cosmological hydrodynamical zoom-in simulations. For example, there are many examples within the *Feedback In Realistic Environments* Latte (FIRE, Samuel et al. 2021; Wetzel et al. 2023; Garavito-Camargo et al. 2024; Arora et al. 2024), the *Milky Way-est* (Buch et al. 2024), the *APOSTLE* (Santos-Santos et al. 2021), the *Auriga* (Grand et al. 2017; Smith-Orlik et al. 2023; Grand et al. 2024) and the *DREAMS* (Rose et al. 2025) simulation suites. Although cosmological analogues are inherently higher fidelity than *N*-body equivalents, they

do not aim to exactly match present-day condition of the MW and LMC.

Simulation Based Inference (SBI, see, Cranmer et al. 2020, for a conceptual overview) offers a medium to explore large and complex parameter spaces by producing many forward models of the system of interest. SBI is a powerful statistical framework for performing inference in difficult modelling scenarios where traditional analytic methods are impractical or impossible. It is particularly well-suited for estimating parameter posterior distributions when the likelihood function cannot be explicitly defined and/or model parameter spaces are large and complex. Instead of defining the likelihood, many forward simulations are used to generate samples of the data from a set of model parameters. 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. 2025; Lovell et al. 2025; Ho et al. 2024; Widmark & Johnston 2025; Sante et al. 2025; Xiangyuan Ma et al. 2025; Jeffrey et al. 2025; Saoulis et al. 2025a) and more generally across many fields of research e.g., seismology (e.g., Saoulis et al. 2025b). In Brooks et al. (2026b), we have previously shown for the MW–LMC system that inference on model parameters through an SBI framework trained on many, $O(10^5)$, rigid simulations retains enough of the relevant physics that more complex simulations capture (e.g., deforming *N*-body simulations, Garavito-Camargo et al. 2019) to avoid model misspecification and vastly improve the computational efficiency of the inference.

We present an SBI architecture for model parameter inference of the MW, LMC and the induced reflex motion using the measured velocities of outer MW halo stars. This builds upon the results of Brooks et al. (2026b) where we demonstrated that an SBI framework trained on many rigid MW–LMC simulations can infer model parameters from dynamical measurements of outer halo tracers. The presented SBI architecture allows the exploration of large MW–LMC model parameter spaces, while incorporating time dependence, to enable the rapid and reliable inference of model parameters. For the first time, we demonstrate the application of our SBI architecture (Brooks et al. 2026b) to the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al. 2026) Blue Horizontal Branch (BHB) and H3+SEGUE+MagE (Chandra et al. 2025) Red Giant Branch (RGB) datasets to provide constraints on the enclosed MW and LMC masses within 50 kpc, the reflex motion and the strength of dynamical friction. Additionally, our framework allows us to explore the effects of survey sky coverage and availability of velocity information for stars in the outer halo.

The plan of the paper is as follows. In Sec. 2 we describe the datasets used throughout this work and define the summary statistics used for model parameter inference. In Sec. 3, we give a concise description of the MW–LMC simulations used to train the SBI architecture used for parameter inference. In Sec. 4, we describe the SBI architecture, detailing the use of Bayesian statistics and the machine learning models used for parameter inference. In Sec. 5, we present and compare the constraints on the MW and LMC masses, the reflex motion and the dynamical friction model parameters. In Sec. 6, we perform a series of diagnostic tests on the estimated posterior distributions. In Sec. 7, we discuss our results and assess any model limitations. Finally, we conclude and provide an outlook for upcoming surveys in Sec. 8.

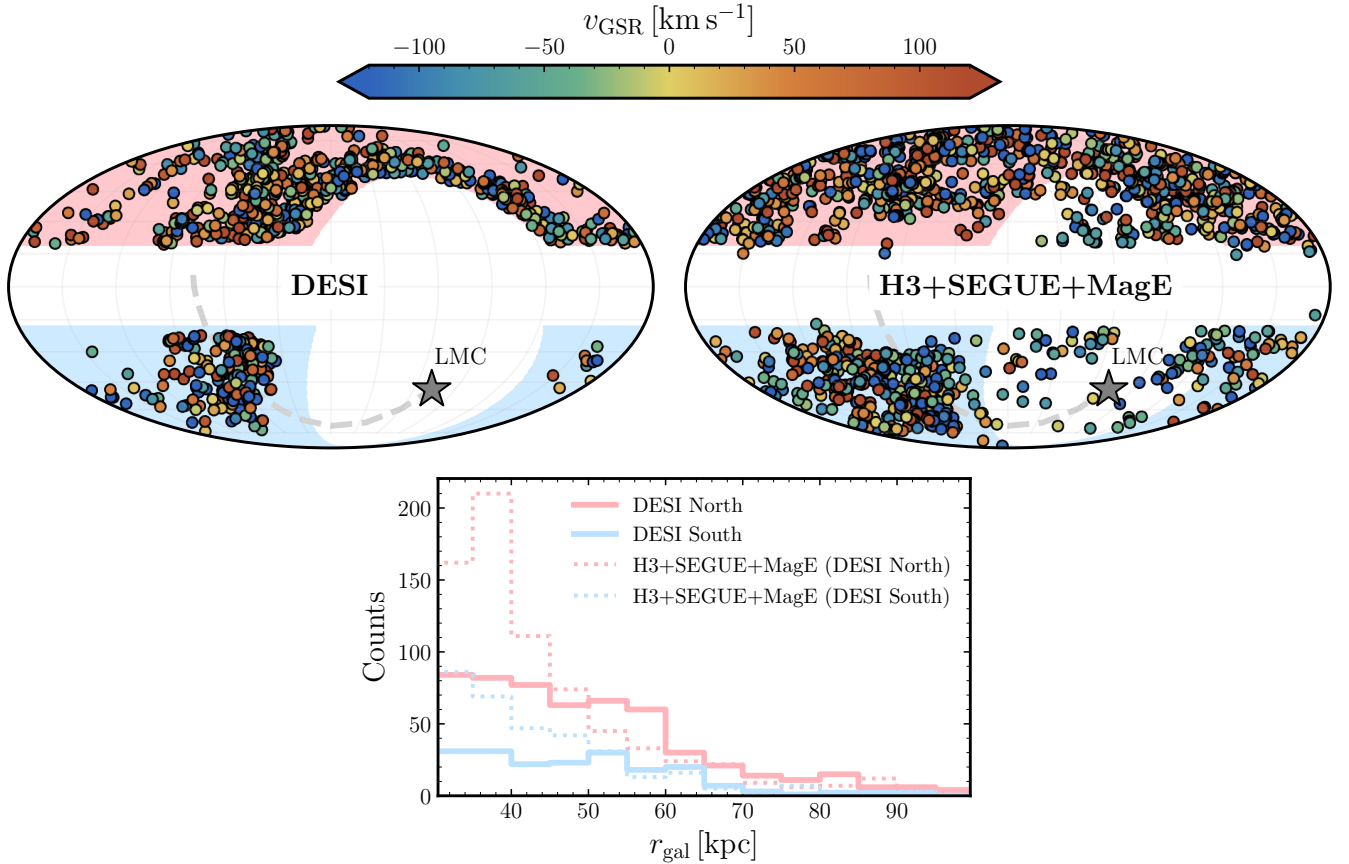


Figure 1. Top panels: The distribution of all sources beyond a Galactocentric distance of 30 kpc for the Dark Energy Spectroscopic Survey (DESI, left) and the all-sky H3+SEGUE+MagE outer halo surveys (right) in Galactic coordinates. The colour of individual sources reflects their solar corrected radial velocity, v_{GSR} . The present-day position of the LMC is shown as the grey star along with an illustrative past orbit as the grey dashed line. The DESI northern (pink) and southern (blue) Galactic fields are shown as shaded regions. **Bottom panel:** The Galactocentric radial distribution of all sources between 30 – 100 kpc for DESI (solid lines) and H3+SEGUE+MagE (dotted lines) contained within the DESI northern and southern Galactic fields.

2 DATA

To transform between the Heliocentric Cartesian and Galactocentric Cartesian coordinate frame in this work, we adopt a right-handed Cartesian coordinate system with the Sun positioned at $\vec{r}_{\odot} = (-8.3, 0.0, 0.02)$ kpc (GRAVITY Collaboration et al. 2019; Bennett & Bovy 2019), with velocity $\vec{v}_{\odot} = (11.1, 244.24, 7.24)$ km s⁻¹ (Schönrich et al. 2010; Eilers et al. 2019).

2.1 Dark Energy Spectroscopic Survey

DESI is a multi-object spectrograph designed for ground-based wide-field surveys that operates on the Mayall 4-meter telescope at Kitt Peak National Observatory. The DESI spectroscopic survey has a large sky coverage footprint of 14,000 deg². The instruments consist of 5,000 fibres and cover a wavelength range of 360 – 960 nm with a resolution between 2,000 and 5,500 depending on the wavelength (DESI Collaboration et al. 2022). The main operating purpose of the survey is to obtain spectra for ~ 40 million galaxies and quasars to probe the nature of dark energy (DESI Collaboration et al. 2016). Nevertheless, the DESI Milky Way Survey (MWS) working group also publishes the data of millions of individual stars in our Galaxy. In March 2025, the DESI Data Release 1 (DR1, DESI Collaboration

et al. 2026)¹ was made available and included a catalogue containing over 6 million MW sources with radial velocities and stellar parameters (Koposov et al. 2026). In this work, we use the Blue Horizontal Branch (BHB) star catalogue² concurrently released by the MWS working group as a part of DESI DR1. The MWS provides radial velocities to all its targets via the radial velocity and stellar parameter fitting code RVSpecfit (Koposov et al. 2011, 2019).

In DR1, there are a total of 10,695 BHB targets, with the targeting procedure described in Cooper et al. (2023) and the precise distances to each BHB star is provided in Byström et al. (2025). We use the public BHB DESI DR1 catalogue, which has already been cleaned to ensure that known substructures, contaminating stars & unphysical stellar quantities have been removed from the sample (see, Byström et al. 2025, sec. 2.4-2.6). We use the PRIMARY = True flag to remove duplicate observations³, and make a further selection to keep only stars between $r_{\text{gal}} \in [30 - 100]$ kpc to produce a final sample of 853 DESI BHB stars for our analysis. In the upper left panel of Fig. 1,

¹ DESI Public Data Release 1: <https://data.desi.lbl.gov/doc/releases/dr1/>

² The DESI BHB catalogue: <https://data.desi.lbl.gov/doc/releases/dr1/vac/mws-bhb/>

³ Application of this flag exactly reproduces the BHB catalogue used in Byström et al. (2025).

we show the on-sky distribution of these sources. Furthermore, in the lower panel of Fig. 1 we show the number counts of sources as a function of Galactocentric distance, divided into the northern (pink) and southern (blue) DESI fields. The observed on-sky and distance density distribution of DESI sources is non-uniform/linear. Although these effects may be small, from a simulation perspective, properly forward modelling MW stellar haloes should account for these effects. This is because SBI methods are sensitive to systematic biases if observational selection effects are not exactly modelled in the same way for the simulations used for inference.

2.2 H3, SEGUE & MagE Outer Halo Surveys

2.2.1 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.

2.2.2 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. 2026; Cargile & Chandra 2025).

2.2.3 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, 2025). Stellar parameters are estimated with the MINESweeper code (Cargile et al. 2020) including parallaxes 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, 2025).

2.2.4 Combined H3+SEGUE+MagE sample

The above surveys combine to produce a pure sample of stars with homogeneous stellar parameters derived with the MINESweeper pipeline. Subsequently, we adopt the same selection procedure to exactly reproduce the high-fidelity subset in sec. 2.3, Chandra et al. (2025). This selection procedure ensures that known substructures (e.g., the Sagittarius stream, Majewski et al. 2003; Vasiliev et al. 2021) and unphysical quantities are removed from the sample. The all-sky sample of H3+SEGUE+MagE stars used in this work contains 1296 field stars between $r_{\text{gal}} \in [30-100]$ kpc. In the upper right panel of Fig. 1, we show the on-sky distribution of these sources. To

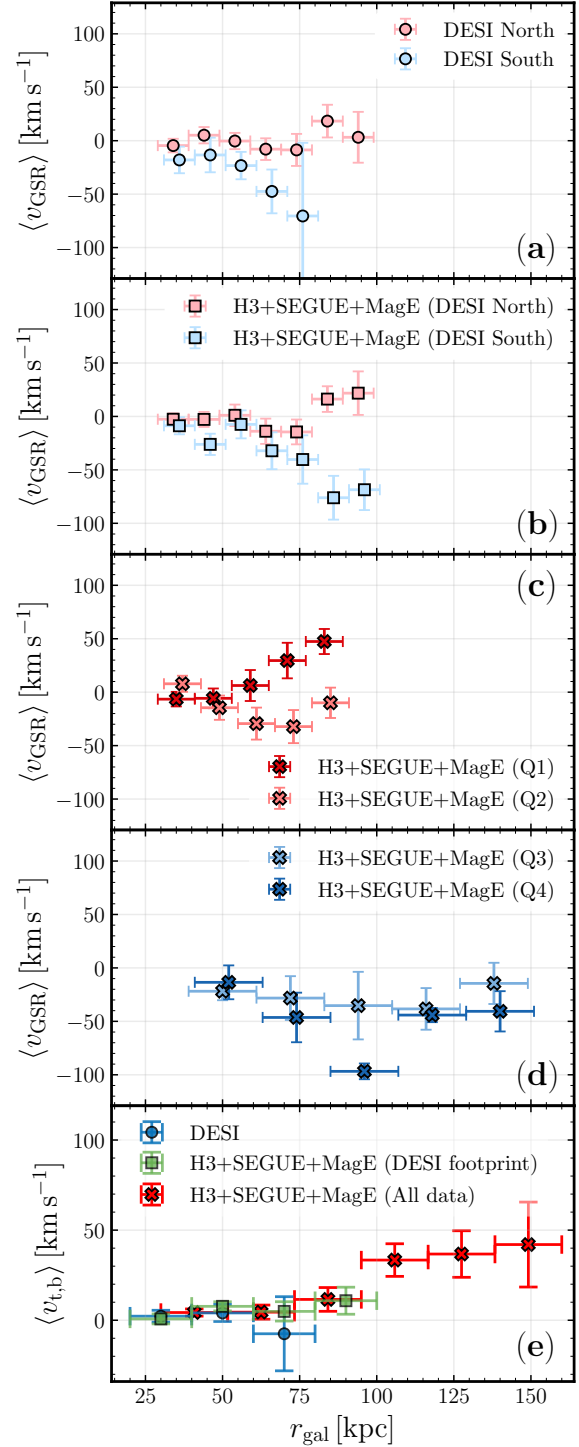


Figure 2. Panel (a): Mean radial velocity distributions, $\langle v_{\text{GSR}} \rangle$, as a function of Galactocentric distance for the DESI sources in its northern (pink) and southern (blue) observing footprints. Panel (b): Same as the first panel for H3+SEGUE+MagE sources within the DESI survey footprints. Panel (c/d): The H3+SEGUE+MagE data divided into on-sky quadrant footprints. Panel (e): All-sky mean tangential velocity, $\langle v_{t,b} \rangle$, distributions as a function of Galactocentric distance for DESI (blue), H3+SEGUE+MagE within 100 kpc (green) and the full H3+SEGUE+MagE datasets (red). Points are offset in distance for improved readability. All 1σ uncertainties are determined via bootstrap resampling. There is a clear increase in $\langle v_{t,b} \rangle$ with distance out to 160 kpc for the H3+SEGUE+MagE data.

ensure a fair comparison to the DESI survey, we apply a DESI sky coverage selection on this sample. This DESI sky coverage sample contains 1049 field stars. We show their Galactocentric distance distribution in the lower panel of Fig. 1, divided into the northern (pink dotted) and southern (blue dotted) DESI fields.

2.3 Observational summary statistics

SBI requires a set of summary statistics to be used for the inference process. For this first 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 radial and tangential velocity distance distributions.

2.3.1 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 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 $\langle v_{\text{GSR}} \rangle \neq 0 \text{ km s}^{-1}$ ubiquitously across the MW.

For both datasets, we calculate the 3σ -clipped mean $\langle v_{\text{GSR}} \rangle$ values for the Galactocentric distance range $r_{\text{gal}} \in [30 - 100] \text{ kpc}$ in bins of 10 kpc width for the northern (pink) and southern (blue) DESI footprints; see panel (a, for DESI data) and panel (b, for H3+SEGUE+MagE data) in Fig. 2. We choose this sky coverage and distance range so that the DESI and H3+SEGUE+MagE samples can be consistently compared, and to avoid any contamination in the inner halo, $r_{\text{gal}} \lesssim 30 \text{ kpc}$. We note that both surveys reach more distant parts of the halo beyond 100 kpc; however at these extreme distances the number of sources are limited. For this selection criteria, the remaining differences are intrinsic to the datasets themselves, e.g., measurement uncertainties and on-sky/distance density distributions.

In addition to comparing DESI and H3+SEGUE+MagE in this way, we will also 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 $l \in [0^\circ, -180^\circ]$, $b \in [-90^\circ, 0^\circ]$; see the all-sky inset in Fig. A3 for a visual representation of these quadrants. For the all-sky H3+SEGUE+MagE dataset, we calculate the 3σ -clipped mean $\langle v_{\text{GSR}} \rangle$ values for the Galactocentric distance range $r_{\text{gal}} \in [30 - 160] \text{ kpc}$ in 5 bins of equal width in the northern quadrants (Q1 and Q2) and the southern quadrants (Q3 and Q4); see panels (c/d) in Fig. 2.

Comparing the measured mean $\langle v_{\text{GSR}} \rangle$ for the DESI and H3+SEGUE+MagE surveys using the consistent selection criteria (panels a/b in Fig. 2), the general trends in the northern and southern DESI footprints are similar. The northern footprint displays behaviour almost consistent with expectations of dynamical equilibrium, whereas the southern footprint demonstrates an increasingly negative signal with increasing Galactocentric distance. For the DESI dataset, this is consistent with previous studies (fig. 10, [Byström et al. 2025](#)). For the H3+SEGUE+MagE data, we also see that the northern radial velocity distribution is consistent with expectations of dynamical equilibrium. This is a subtle difference to [Chandra et al. \(2025\)](#)

who showed an increasingly positive radial velocity amplitude in the northern Galactic hemisphere with Galactocentric distance. This can be explained by the choice of sky coverage. In [Chandra et al. \(2025\)](#), they selected the northern quadrant directly opposite the LMC (Q1) and a southern quadrant including the LMC (Q4) to measure the mean $\langle v_{\text{GSR}} \rangle$ values. This choice of sky coverage, particularly in the northern Galactic hemisphere, is what drives the flattening of the mean $\langle v_{\text{GSR}} \rangle$ distribution seen in the panels (c/d) of Fig. 2. Although not the main result in this work, this insight nonetheless highlights the importance of survey sky coverage, particularly when using mean velocity values as summary statistics, as signal can be averaged out across larger areas.

2.3.2 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. 2025](#)). From the perspective of a Galactocentric observer, 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’. For both datasets, we calculate the 3σ -clipped mean $\langle v_{t,b} \rangle$ values for the Galactocentric distance range $r_{\text{gal}} \in [30 - 100] \text{ kpc}$ in distance bins of 20 kpc width. Plus, to exploit the full depth of the H3+SEGUE+MagE dataset we will also use the 3σ -clipped mean $\langle v_{t,b} \rangle$ values for the Galactocentric distance range $r_{\text{gal}} \in [30 - 160] \text{ kpc}$ in 6 equally spaced distance bins. In the panel (e) of Fig. 2 we show the all-sky mean tangential velocity distribution as a function of Galactocentric radius. These tangential velocities have been corrected for the solar motion. Both surveys demonstrate an increasing trend as a function of Galactocentric distance, albeit with DESI having very large uncertainties beyond 60 kpc. Indeed, we do not show the outermost mean tangential velocity data point for DESI as the error bar is of the order 100 km s^{-1} .

3 SIMULATIONS

In this section, we provide a concise description of the rigid MW–LMC simulations presented in [Brooks et al. \(2026b, sec. 2\)](#). We detail only the key information of these simulations and highlight any significant changes to the modelling. This rich set of low fidelity simulations, spanning a large model parameter space, is used to train the neural networks for parameter inference.

3.1 The Milky Way – LMC potentials

We use the galaxy dynamics C++/Python package agama ([Vasiliev 2019](#)) to generate 128,000 rigid MW–LMC simulations, each with a unique combination of model parameters.

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} . 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 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). 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](#)).

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. 2026b](#)). 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:

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

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 $\epsilon = 1.6 a_{\text{LMC}}$ 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](#)). The numerator in the Coulomb logarithm expression an arbitrarily chosen value that loosely describes the average separation of the MW and LMC. In principle this value could be updated through the integration of the LMC orbit, however this will 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. The final MW–LMC potential includes the acceleration of the MW’s centre of mass towards the LMC.

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 2019](#)). This requires instances of a tracer density profile, a potential, and a prescription for the radial velocity anisotropy. We use a Dehnen tracer density profile ([Dehnen 1993](#)) and an NFW profile for the potential ([Navarro et al. 1996, 1997](#)) with relevant parameters adopted from each unique MW–LMC simulation. We implement the radial bias of stellar velocities ([Osipkov 1979](#); [Merritt 1985](#); [Binney & Tremaine 2008](#)) using the velocity anisotropy profile:

$$\beta(r) \equiv 1 - \frac{\sigma_t^2}{2\sigma_r^2} = \frac{\beta_0 + (r/r_a)^2}{1 + (r/r_a)^2}. \quad (2)$$

where β_0 is the limiting value of anisotropy in the centre, and if $r_a < \infty$, the anisotropy coefficient tends to 1 at large r , otherwise it is constant and equal to β_0 over all scales which we adopt.

A key improvement from our previous stellar halo simulations in [Brooks et al. \(2026b\)](#) is that we sample ~ 4.5 times more phase-space coordinates to initialise the MW stellar haloes before any LMC-induced disequilibrium. This ~ 4.5 -fold increase in the number of particles, 20,000 for each stellar halo, used to represent the stellar halo ensures that Poisson noise does not dominate the uncertainty on measurable quantities. Thus the main source of uncertainty is from the observations themselves. With this increased measurement precision from our simulations, we can convolve any measured value with the observational error from any given survey. This allows us to correctly forward model the simulations to produce observational-like quantities which are subsequently used for the evaluation of the posterior. Hence, prior to training the inference framework, we apply the survey-specific uncertainties to the binned radial and tangential velocity measurements from the DESI and H3+SEGUE+MagE datasets; see Sec. 2.3. This approach more closely aligned with SBI ideologies, allowing us to better forward model all the stellar haloes to match a specific survey of interest and perform the subsequent inference. For a given MW–LMC potential with reflex motion, we integrate all particles in the stellar halo to present-day over the last 2.2 Gyr.

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. 2025](#); [Yaaqib et al. 2024, 2025](#); [Brooks et al. 2026b](#)). To do this, we use all stars beyond 50 kpc. Although, it is worth noting that this calculation from the dynamics of outer halo stars is not necessarily exactly the same as the induced velocity of the MW centre relative to the initial inertial frame prior to the LMC’s infall. This method fits an on-sky velocity model which contains nine free parameters. We model the dipole reflex motion using Galactocentric Cartesian velocities $\{v_x, v_y, v_z\}$. 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 hyperparameters, $\{\sigma_{v_r}, \sigma_{v_\phi}, \sigma_{v_\theta}\}$, which account for measurement uncertainties. The reflex motion model is represented by the sum of the dipole and mean motion parameters:

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

where $\vec{v}$ 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 equ. 6 & 8 in [Petersen & Peñarrubia 2021](#)). We return the maximum likelihood estimates for all of the reflex motion model parameters. However, in the context of this work, we will only comment on the magnitude, v_{travel} , and Galactocentric components of the reflex travel velocity, $\{v_x, v_y, v_z\}$.

3.3 Simulation Priors

We use the same parameter prior probability distributions as in [Brooks et al. \(2026b\)](#). For clarity, we repeat that information here in Table. 1. The first two parameters are the MW total mass and enclosed mass within 50 kpc. The next two parameters are the infall LMC total mass and its enclosed mass within 50 kpc. Next is the scalar multiple that modulates the strength of dynamical friction

Table 1. Simulation model parameter prior distributions. The mass enclosed priors (grey) are derived using the priors for the total masses.

Model Parameter	Prior probability distribution
$M_{200, \text{MW}}$	$\mathcal{N}(15, 5) \times 10^{11} M_{\odot}$
$M_{\text{MW}} (< 50 \text{ kpc})$	$\mathcal{N}(4.8, 0.8) \times 10^{11} M_{\odot}$
M_{LMC}	$\mathcal{N}(15, 10) \times 10^{10} M_{\odot}$
$M_{\text{LMC}} (< 50 \text{ kpc})$	$\mathcal{N}(8.6, 2.9) \times 10^{10} M_{\odot}$
$\log_{10}(\lambda_{\text{DF}})$	$\mathcal{U}(-3, 1)$
α_{LMC}	$\mathcal{U}(60^{\circ}, 90^{\circ})$
δ_{LMC}	$\mathcal{U}(-80^{\circ}, -50^{\circ})$
d_{LMC}	$\mathcal{N}(49.6, 5) \text{ kpc}$
v_{los}	$\mathcal{N}(262.2, 10) \text{ km s}^{-1}$
$\mu_{\alpha_{\text{LMC}}}$	$\mathcal{N}(1.9, 0.25) \text{ mas yr}^{-1}$
$\mu_{\delta_{\text{LMC}}}$	$\mathcal{N}(0.33, 0.25) \text{ mas yr}^{-1}$
β_0	$\mathcal{U}(0, 0.9)$
r_{Dehnen}	$\mathcal{U}(10, 15) \text{ kpc}$

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 two parameters are the anisotropy parameter, β_0 , and the Dehnen tracer density profile scale length, r_{Dehnen} , that initialise the mock MW stellar haloes. In total, we run 128,000 MW–LMC simulations each with unique 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 , given some observed data D_{obs} , and a theoretical model I . 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. In the case where a likelihood function need not, or cannot be explicitly defined, it is possible to instead use SBI to estimate the posterior distribution.

The simplest form of SBI is known as Approximate Bayesian Computation (ABC, e.g., Rubin 1984; Pritchard et al. 1999; Fearnhead & Prangle 2010). The ABC framework selects forward simulations that are the most similar to the observed data based on some distance measure involving the 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., Gaussian, into the target probability distribution (Jimenez Rezende & Mohamed 2015). We adopt a DELFI approach using the sbi Python package (Tejero-Cantero et al. 2020), and estimate the posterior distribution from the forward simulations using Masked Autoregressive Flows (MAF, Papamakarios et al. 2017, 2019) with 5 transformation layers in the neural network, each with a width of 50 nodes. The exact neural network architecture used can influence the estimated posterior. However, we have found that varying the number of layers and nodes has very little effect on our results. 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 normalising flow has been trained on a precomputed simulation dataset, the posterior can be returned for many observations without having to retrain the flow; this is known as *amortisation* (Mittal et al. 2025).

Often, some form of data compression is required (e.g., Alsing et al. 2018; Alsing & Wandelt 2019; Heavens et al. 2020; Jeffrey et al. 2021; Widmark & Johnston 2025; Jeffrey et al. 2025). However, the application to our problem is relatively low in dimensionality, i.e., the number of parameters of interest and data points, hence no data compression is required. Generally, for SBI, the more simulations that are available to use, the better. Within a cosmological context, the estimated number of simulations that are required for reliable SBI analysis is $\sim 10^4$ (Bairagi et al. 2025).

The simulations used in this work are described in Sec. 3. We use a MAF density estimator from the sbi package (Tejero-Cantero et al. 2020) to directly obtain the posterior distribution that can be evaluated at any observed data point for any data realisation, i.e., $p(\theta|D_{\text{obs}}, I)$. We ensure the reliability of the estimated posteriors through some diagnostic checks including coverage probabilities and predictive posterior checks in Sec. 6.

5 RESULTS

We show posterior distributions for MW, LMC and reflex motion model parameters conditioned on different data subsets. All figures use 10,000 samples drawn from their respective posterior density distributions.

5.1 Inference set-up

Throughout this section, we return the posterior distributions for MW and LMC masses enclosed within 50 kpc, the Cartesian components of the travel velocity and the dynamical friction strength. Following the Bayesian notation in Equ. 4, this reads as:

$$\theta = \{M_{\text{MW}}(< 50 \text{ kpc}), M_{\text{LMC}}(< 50 \text{ kpc}), v_x, v_y, v_z, \log \lambda_{\text{DF}}\}$$

For the input stellar velocity data used to condition the posterior on, we will make it clear in each sub-section which on-sky selection and dataset is used. In Sec. 5.2, for sources that lie within the DESI survey footprint boundaries, we present the constraints obtained when using either the DESI or H3+SEGUE+MagE datasets. In Sec. 5.3, we use an on-sky selection defined by four quadrant footprints, and explore the full depth of the H3+SEGUE+MagE dataset. We note that this exploration of using quadrants is only possible for all-sky datasets like H3+SEGUE+MagE. Therefore the same analysis cannot be carried out using a limited sky coverage survey like DESI. As a reminder, observable uncertainties are accounted for in our simulated stellar halo velocity measurements; see

Sec. 3.2 for details on accounting for survey specific uncertainties. Finally, the only observed data points that are consistently used for each posterior estimation are the present-day Galactocentric position and velocity of the LMC as $\vec{x}_{\text{LMC}} = \{-0.6, -41.3, -27.1\}$ kpc, $\vec{v}_{\text{LMC}} = \{-63.9, -213.8, 206.6\}$ km s⁻¹ (Vasiliev et al. 2021, and references therein). Uncertainty in the LMC position and velocity is accounted for by convolving each simulated value with the uncertainties given in sec. 3.1 and table 2 of Vasiliev (2023). In our models, the true present-day LMC phase space coordinates are treated as model parameters, with the observed phase space coordinates, i.e. $(\vec{x}_{\text{LMC}}, \vec{v}_{\text{LMC}})$ from Vasiliev et al. (2021), providing noisy measurements. One could infer a posterior over these parameters, though they are well constrained and not the focus of this work, so we choose to condition our inference on their observed values within the explored prior space.

5.2 DESI footprint

This section only uses the data from sources that are contained within the DESI northern and southern survey footprints. This is described in detail in Sec. 2.3, and the measurements for the binned velocity field summary statistic that are used as input data are shown in Fig. 2. Following the Bayesian notation in Equ. 4, the input data reads as:

$$D_{\text{obs}} = \{(\vec{x}, \vec{v})_{\text{LMC}}, \langle v_{\text{GSR},N} \rangle, \langle v_{\text{GSR},S} \rangle\}, \text{ or} \\ D_{\text{obs}} = \{(\vec{x}, \vec{v})_{\text{LMC}}, \langle v_{\text{GSR},N} \rangle, \langle v_{\text{GSR},S} \rangle, \langle v_{t,b} \rangle\},$$

as we vary whether or not the tangential velocity data points, $\langle v_{t,b} \rangle$, are provided. The subscripts N and S signify the binned radial velocity data, $\langle v_{\text{GSR}} \rangle$, in the northern and southern DESI survey fields, respectively.

5.2.1 DESI data

The measured data for the binned radial velocities is shown in panel (a) of Fig. 2 and the binned tangential velocities are shown as the blue points in the panel (e) of Fig. 2. We show the posteriors when providing information on the LMC present-day position, velocity and the radial and tangential velocities as the open blue contours in Fig. 3, for the MW and LMC enclosed masses, and in Fig. 4 for the Galactocentric cartesian components of the travel velocity. In Appendix A, we show the full posterior distribution for the model parameter constraints using the DESI dataset in Fig. A1. This includes the results when we only provide the radial velocities information for the outer halo tracers. We show the model priors, or simulation values, as the filled grey contours. The joint posterior distributions show the 1 σ and 2 σ confidence intervals. The individual posterior distributions show the 16th and 84th percentiles as filled bands. For all parameters shown, the inclusion of DESI tangential velocity information does not greatly improve the precision of the inferred parameters. This is not unexpected as the uncertainties on the measured values are very large; see panel (e) of Fig. 2. We summarise these results in the top two rows of Table. 2 and compare them to previous results in Sec. 5.4.

5.2.2 H3+SEGUE+MagE data

The measured data for the binned radial velocities is shown in panel (b) of Fig. 2 and the binned tangential velocities are shown as the green squares in panel (e) of Fig. 2. We show the posteriors when providing information on the LMC present-day position, velocity and the radial and tangential velocities as the open green contours

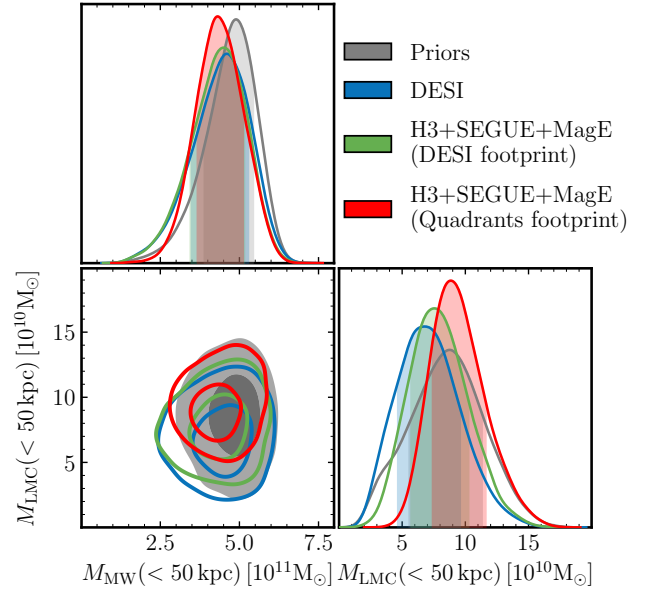


Figure 3. Posterior distributions - MW and LMC enclosed masses: The joint, and individual, posterior distributions for the MW and LMC masses enclosed within 50 kpc. We show these distributions conditioned on the LMC centre present-day position & velocity, and radial and tangential velocities as data points. The open blue, green and red contours represent the posteriors conditioned using the DESI, H3+SEGUE+MagE (DESI footprint) and H3+SEGUE+MagE (Quadrants footprint) survey data, respectively. 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.

in Fig. 3, for the MW and LMC enclosed masses, and in Fig. 4 for the Galactocentric cartesian components of the travel velocity. In Appendix A, we show the full posterior distribution for the model parameter constraints using the H3+SEGUE+MagE dataset in Fig. A2. This includes the results when we only provide the radial velocities information for the outer halo tracers. We summarise these constraints in the middle two rows of Table. 2. The constraints on all model parameters returned using DESI or H3+SEGUE+MagE data contained within the DESI survey footprints have similar values and are consistent within uncertainties. We get slightly more precise constraints on the travel velocity and enclosed LMC mass, $\sim 10\%$, using the H3+SEGUE+MagE dataset, although this is expected as there are more sources compared to DESI. We make further comparisons to previous results in Sec. 5.4.

5.3 Quadrants footprint

In this section, we explore the full H3+SEGUE+MagE dataset to extract the maximum constraining power by dividing the sky into quadrants and measuring a set of mean radial and tangential velocity summary statistics. A visual representation of these quadrants is shown in the all-sky inset of Fig. A3. We use all H3+SEGUE+MagE sources between 30 and 160 kpc. The measured data for the binned radial velocities is shown in panels (c/d) of Fig. 2 and the binned tangential velocities are shown as the red points in the panel (e) of Fig. 2. The distance bins used to measure the average radial and tangential velocity fields are the same as in Chandra et al. (2025, figures 4 & 5). Utilising an on-sky quadrant footprint is only possible

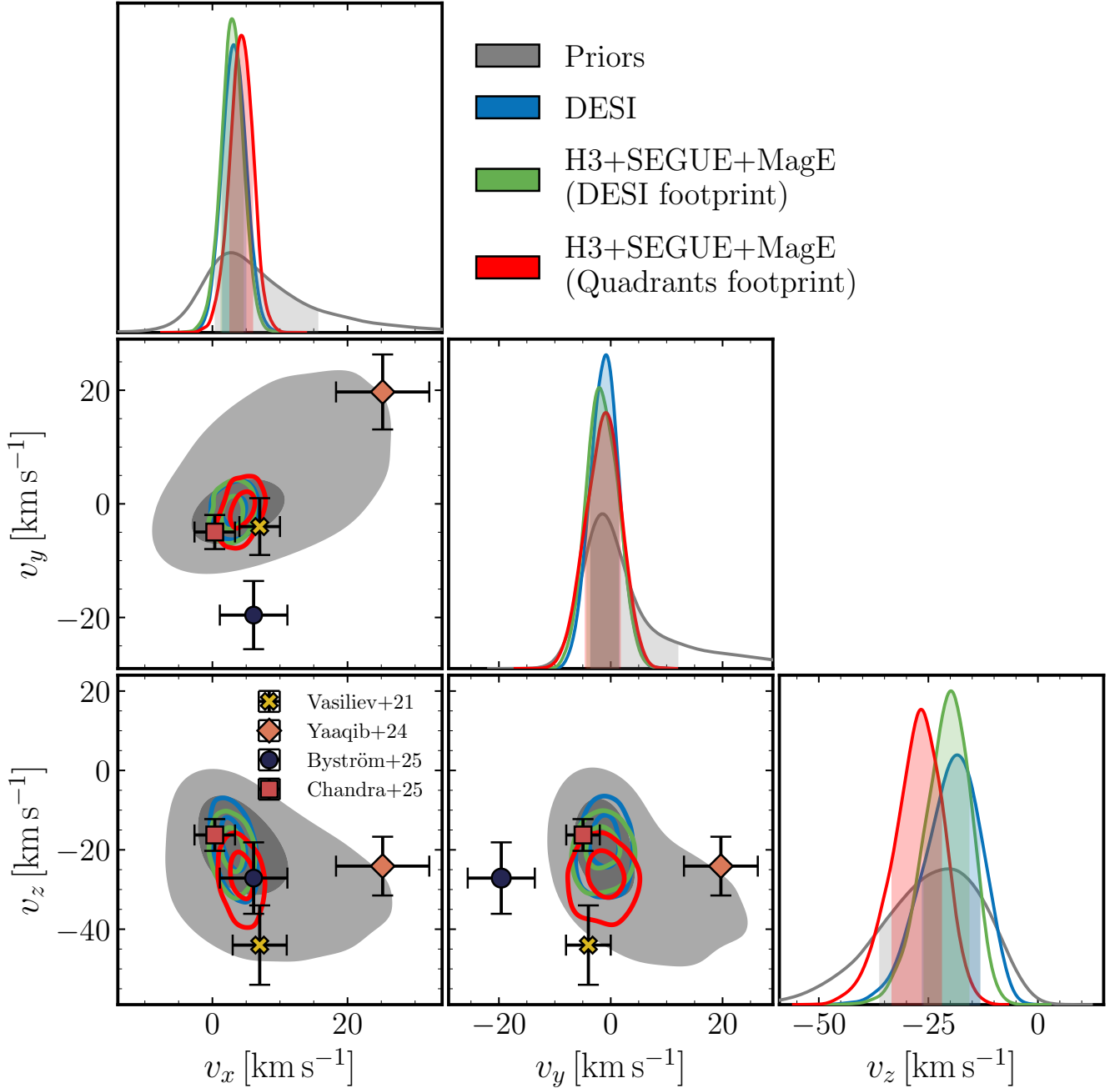


Figure 4. Posterior distributions - Reflex motion velocity: The joint, and individual, posterior distributions of the Galactocentric Cartesian travel velocity components. We show these distributions conditioned on the LMC centre present-day position & velocity, and radial and tangential velocities as data points. The open blue, green and red contours represent the posteriors conditioned using the DESI, H3+SEGUE+MagE (DESI footprint) and H3+SEGUE+MagE (Quadrants footprint) survey data, respectively. 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. (2025, red square) are shown in each panel for comparison.

for all-sky datasets. Hence we cannot carry out a similar analysis using the DESI data. In principle, as the H3+SEGUE+MagE dataset contains a larger sample of stars that have been observed deeper in the MW halo, and allows for the finer on-sky division, we expect these results to be the most precise of all those we present in this work.

Following the Bayesian notation in Equ. 4, the input data reads as:

$$D_{\text{obs}} = \{(\vec{x}, \vec{v})_{\text{LMC}}, \langle v_{\text{GSR}, Q1} \rangle, \langle v_{\text{GSR}, Q2} \rangle, \langle v_{\text{GSR}, Q3} \rangle, \langle v_{\text{GSR}, Q4} \rangle\}$$

or,

$$D_{\text{obs}} = \{(\vec{x}, \vec{v})_{\text{LMC}}, \langle v_{\text{GSR}, Q1} \rangle, \langle v_{\text{GSR}, Q2} \rangle, \langle v_{\text{GSR}, Q3} \rangle, \langle v_{\text{GSR}, Q4} \rangle, \langle v_{\text{t}, b} \rangle\},$$

as we vary whether or not the tangential velocity data points, $\langle v_{t,b} \rangle$, are provided. The numerical subscripts signify the binned radial velocity data, $\langle v_{GSR} \rangle$, in each respective on-sky quadrant.

Similar to before, we show the posteriors when providing information on the LMC present-day position, velocity and the radial and tangential velocities as the open red contours in Fig. 3, for the MW and LMC enclosed masses, and in Fig. 4 for the Galactocentric cartesian components of the travel velocity. In Appendix A, we show the full posterior distribution for the model parameter constraints using the H3+SEGUE+MagE dataset in Fig. A2. This includes the results when only provide the radial velocities information for the outer halo tracers. We show the prior distributions as the filled grey contours. The joint posterior distributions show the 1σ and 2σ confidence intervals. The individual posterior distributions show the 16th and 84th percentiles as filled bands. We see improvement in the precision of the inferred parameters, $\sim 15 - 20\%$, by including the tangential velocity information. Although, the parameters are already well constrained using the radial velocity information alone. We summarise these results in the final two rows of Table. 2 and compare them to previous results in Sec. 5.4.

5.4 Results in context

5.4.1 The reflex motion

Previously, the reflex motion has been described by the magnitude of the velocity dipole vector, namely the travel velocity, v_{travel} , and its orientation is called the *apex* direction 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. 2025). These recent studies agree that the direction of the travel velocity points towards a location on the past orbit of the LMC, though they do not converge on a consistent direction (e.g., Byström et al. 2025, fig. 9). In this work, instead of transforming to Galactic coordinates, we return the posterior constraints on the Galactocentric Cartesian components of the travel velocity vector i.e., $\vec{v}_{\text{travel}} = (v_x, v_y, v_z)$. This is because these are the parameters the MLE fits for in our model to avoid inefficient convergence, see Sec. 3.2. Although, once the posterior has been determined for $\{v_x, v_y, v_z\}$, one is free to transform to any preferred choice of coordinate system. The velocity component posterior distributions are shown alongside literature values in Fig. 4; for a tabulated comparison, see Table. 2. 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. (2025, H3+SEGUE+MagE data, distance-independent best fit values in their table. 4, red square) in each panel. We find that for both DESI and H3+SEGUE+MagE data, no matter the chosen on-sky footprint, the $v_x - v_y$ plane preferentially agrees with the measurement of Vasiliev et al. (2021) and the distance-independent measurement of Chandra et al. (2025). In fact, the prior distribution (grey contours) struggles to allow values consistent with Yaaqib et al. (2024, for positive v_y), Byström et al. (2025, for negative v_y). On the other hand, the $v_x - v_z$ plane posteriors tend to agree more with Byström et al. (2025). In that study, the components of the travel velocity are constrained as free parameters. Whereas, in Vasiliev et al. (2021) and this work, they are somewhat constrained by the simulation priors themselves. Furthermore, the prior and posteriors in the $v_y - v_z$ plane struggle to be consistent with any of the previous measured values. While the posteriors in this work show disagreement with existing measurements in some velocity projections, those measurements are also inconsistent with

each other. This further highlights the importance of the intrinsic properties of the datasets themselves e.g., survey selection functions, on the reported measurements on the reflex motion parameters.

We summarise the magnitude of the travel velocity in Fig. 5, as derived from the posteriors shown in Fig. 4. Also see Table. 2 for tabulated values. We present the median and 16th – 84th percentiles as the points with errors. The dataset and on-sky footprint used to provide each constraint are shown in the label on the left-hand side. As shaded bands, we show the measured 16th – 84th percentiles confidence intervals from Yaaqib et al. (2024, best fit values for $r_{\text{gal}} > 50$ kpc using SDSS/SEGUE data, in orange), Chandra et al. (2025, distance-independent best fit values using H3+SEGUE+MagE data, in red) and Byström et al. (2025, DESI data, in blue). We find that the constraints in this study are smaller in magnitude than those previously measured by Vasiliev et al. (2021) and Byström et al. (2025), but larger in magnitude than the distance-independent measurement of Chandra et al. (2025). This is perhaps unsurprising as we compute the travel velocity in our simulations using all stars beyond 50 kpc, and therefore any constraint on the travel velocity in this work should be interpreted as the mean value in the outer MW halo beyond 50 kpc. Hence, the average radial distance of star particles in our stellar haloes is possibly closer/further away than those quoted from the observational datasets. Moreover, the medians of the previous literature values are generally larger in magnitude than the median of the prior distribution. Hence, the SBI estimates of the travel velocity could be expected to be biased to lower values based purely on the simulation prior alone. Finally, we do not consider the effects of radial variation which may bias the constraint on our travel velocity; see Sec. 7.4. Nevertheless, we obtain our most precise constraint on the travel velocity using the radial and tangential velocity data from H3+SEGUE+MagE in combination with the quadrant sky footprints. Using this dataset, we report a measured travel velocity of $v_{\text{travel}} = 26.4^{+5.5}_{-4.4}$ km s⁻¹.

5.4.2 The enclosed mass of the LMC

A variety of techniques have been used to constrain the enclosed mass of the LMC, e.g. dynamical models of MW stellar streams (Shipp et al. 2021; Vasiliev et al. 2021; Koposov et al. 2023; Warren et al. 2025). All of the constraints in this study agree with the previous measurement within 16th – 84th percentiles uncertainties from e.g., Koposov et al. (2023). The prior of the enclosed LMC mass could be biasing our results to agree with previous measurements. Yet, as shown in Brooks et al. (2026b, sec. 6.2), adopting a wide uninformative prior instead of an informative Gaussian prior for the LMC mass did not bias our inference of the LMC mass. Therefore, we can take confidence that the choice of an informative prior is not significantly biasing the returned constraints in this work. We obtain our most precise constraint on the enclosed LMC mass using the radial and tangential velocity data from H3+SEGUE+MagE in combination with the quadrant sky footprints. Using this dataset, we report a measured LMC mass enclosed of $M_{\text{LMC}}(< 50 \text{ kpc}) = 9.2^{+1.9}_{-2.3} \times 10^{10} M_{\odot}$.

5.4.3 The enclosed mass of the Milky Way

The MW mass has been constrained using a range of techniques (see, Wang et al. 2020; Medina et al. 2025, and the many references therein). The previous measurement from e.g., Erkal et al. (2019b), agrees with all of the various constraints in this work within its 16th – 84th percentiles confidence intervals. However, our constraints on the MW enclosed mass are not particularly precise with respect to

Table 2. The posterior medians and 1σ credible intervals for the magnitude of the reflex motion travel velocity, v_{travel} , its Galactocentric Cartesian vector components v_x , v_y , v_z (units, km s^{-1}), the MW (units, $10^{11} M_\odot$) and LMC (units, $10^{10} M_\odot$) masses enclosed within 50 kpc and the strength of Chandrasekhar dynamical friction, $\log \lambda_{\text{DF}}$ (units, dimensionless). The dataset and footprint used to produce each constraint are shown in the leftmost column.

Data & Footprint	v_{travel}	v_x	v_y	v_z	$M_{\text{LMC}}(< 50 \text{ kpc})$	$M_{\text{MW}}(< 50 \text{ kpc})$	$\log \lambda_{\text{DF}}$
DESI data, DESI footprint, without $v_{t,b}$	$20.4^{+7.5}_{-6.2}$	$3.2^{+2.0}_{-1.9}$	$-1.0^{+2.7}_{-2.7}$	$-20.0^{+6.3}_{-7.6}$	$7.3^{+2.8}_{-2.5}$	$4.4^{+0.8}_{-1.0}$	$0.1^{+0.6}_{-1.6}$
DESI data, DESI footprint, with $v_{t,b}$	$19.7^{+6.8}_{-5.7}$	$3.2^{+1.7}_{-1.7}$	$-1.1^{+2.5}_{-2.6}$	$-19.2^{+5.8}_{-6.8}$	$7.0^{+2.7}_{-2.4}$	$4.5^{+0.8}_{-1.0}$	$0.2^{+0.6}_{-1.7}$
H3+ data, DESI footprint, without $v_{t,b}$	$24.9^{+6.7}_{-5.7}$	$3.8^{+1.8}_{-1.7}$	$-0.5^{+3.1}_{-3.1}$	$-24.4^{+5.8}_{-6.6}$	$8.9^{+2.5}_{-2.2}$	$4.5^{+0.8}_{-0.7}$	$0.1^{+0.7}_{-1.7}$
H3+ data, DESI footprint, with $v_{t,b}$	$20.9^{+5.6}_{-4.7}$	$2.9^{+1.6}_{-1.6}$	$-1.5^{+2.9}_{-2.7}$	$-20.4^{+4.7}_{-5.6}$	$7.8^{+2.5}_{-2.2}$	$4.4^{+0.8}_{-0.9}$	$0.3^{+0.5}_{-1.8}$
H3+ data, Quadrant footprints, without $v_{t,b}$	$31.2^{+6.4}_{-5.7}$	$4.6^{+1.7}_{-1.6}$	$-0.1^{+2.6}_{-3.4}$	$-30.5^{+5.9}_{-6.6}$	$10.2^{+2.1}_{-1.9}$	$4.2^{+0.8}_{-0.7}$	$-0.2^{+0.7}_{-1.5}$
H3+ data, Quadrant footprints, with $v_{t,b}$	$26.4^{+5.5}_{-4.4}$	$4.2^{+1.4}_{-1.3}$	$-1.0^{+2.7}_{-3.6}$	$-25.3^{+4.5}_{-5.4}$	$9.2^{+2.3}_{-1.9}$	$4.4^{+0.7}_{-0.7}$	$-0.1^{+0.7}_{-1.6}$

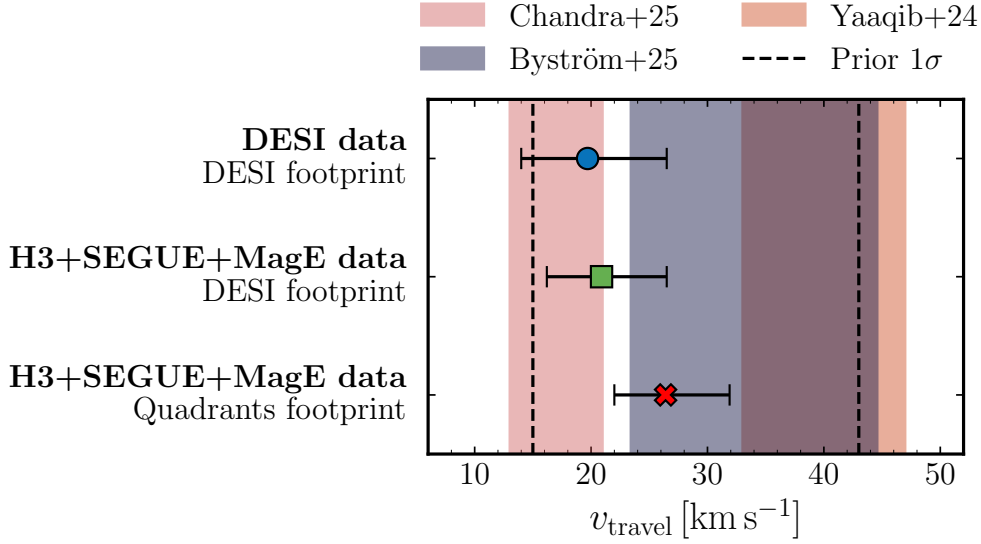


Figure 5. Comparison of reflex motion constraints: The derived medians and $16^{\text{th}} - 84^{\text{th}}$ percentiles of the magnitude of the travel velocity, v_{travel} , for the posteriors shown in Fig. 4. The labels define the dataset used, the on-sky selection footprint used to produce the constraint. The values from existing literature are shown as shaded bands (Yaaqib et al. 2024; Chandra et al. 2025; Byström et al. 2025, in orange using SDSS/SEGUE data, red using H3+SEGUE+MagE data, and blue using DESI data, respectively). The prior 1σ confidence interval is shown as the black dashed lines. The most precise constraint is produced by using the full depth of the H3+SEGUE+MagE dataset and dividing the binned velocity fields into four quadrants.⁷

Erkal et al. (2019b). For all constraints, the median and uncertainties are dominated by the enclosed MW mass prior distribution. This implies that the mean radial and/or tangential velocity data points are not particularly constraining of the MW mass in this inference set-up. Previously, the mass ratio of the MW–LMC system has been suggested to be most sensitive to the velocities of outer halo stars (Petersen & Peñarrubia 2021). The apparent model insensitivity to the MW mass implies that the mean velocities of outer halo stars are most informative of the LMC mass and not the mass ratio (see also, Sheng et al. 2025; Yaaqib et al. 2025). Indeed, future work will use the dispersions of the velocities to constrain the MW mass; see Sec. 7.4. Nevertheless, for all of the posteriors in this work conditioned on the various datasets and selection criteria, we find the total mass of the LMC is at least $\approx 10 - 15\%$ that of the MW (Consistent with Chandra et al. 2025). Although, using more generalised models of

the MW e.g., triaxiality to define the MW halo density profile, may lead to better constraints on properties of the MW.

5.4.4 Dynamical friction

In our MW–LMC models, we vary the strength of dynamical friction, λ_{DF} , that the LMC experiences. We report the median value and the $16^{\text{th}} - 84^{\text{th}}$ percentile confidence intervals from the posterior distributions in Table. 2. Regardless of the dataset or sky coverage, we cannot well constrain the dynamical friction strength. For the constraints using DESI or H3+SEGUE+MagE data within the DESI footprint, the mean of the posterior distribution is slightly greater than the standard Chandrasekhar value i.e., $\log_{10}(\lambda_{\text{DF}}) = 0$. Interestingly, this is similar to the results in Koposov et al. (2023) where they use measurements of the Orphan–Chenab stellar stream observables to constrain the dynamical friction strength. Their inference is more

precise, likely because the Orphan-Chenab stream’s trajectory constrains the closest LMC–stream separation distance which depends on the dynamical friction strength. Conversely, the constraints using the H3+SEGUE+MagE data in on-sky quadrants finds the mean to be slightly smaller than the standard Chandrasekhar values. However, the posterior 16th – 84th percentiles are large and still encompass the standard Chandrasekhar value in all cases.

6 POSTERIOR DIAGNOSTIC CHECKS

To ensure the results presented in this work can be trusted, we carry out posterior coverage probability checks and predictive posterior checks.

6.1 Coverage probability test

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. A coverage probability test is one way to assess the accuracy of such 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, 90% 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, 90% 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). 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. We use the implementation of TARP in the *sbi* Python package ([Tejero-Cantero et al. 2020](#)).

In Fig. 6, we demonstrate that the expected coverage does indeed match the credibility level for the estimated posteriors in this work conditioned using the LMC present-day position and velocity, and the radial and tangential velocities of outer halo stars. Note, we find similar results for the posteriors conditioned without using the mean tangential data as well. 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 all of the estimated posterior distributions, Figs. A1-A3, we report an 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.

6.2 Posterior Predictive 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 can provide an intuition about any bias introduced in inference e.g., determining whether or not the generated data systematically differ from the observed data used during the inference. We carry out a PPC for every posterior used throughout this work, see Figs. A1-A3. 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 average radial and tangential velocity measurements, which we can then use to compare to the observed data that was originally used to evaluate the posterior.

In Fig. 7, we demonstrate the PCC for the posteriors conditioned on DESI or H3+SEGUE+MagE data within the DESI survey footprint, i.e., open blue and green contours in Figs. 3 & 4. Using parameters sampled from these posteriors, we show the generated radial (top and middle panels for DESI and H3+SEGUE+MagE footprints, respectively) and tangential (bottom panel) velocity data using samples from these posteriors as the violin plot contours. The mean observed data points are shown as the same coloured markers; same points as in Fig. 2. Errorbars are not shown on the observed data because the generated data incorporates survey uncertainties. The sub-panels in Fig. 7 demonstrate the number of standard deviations that each observed data point lies away from the peak of the generated data distribution. Mostly, the observed data points lie within $\sim 2\sigma$ the generated data distributions, this demonstrates that the SBI posteriors are accurately representing the observations.

In Fig. 8, we demonstrate the PCC for the posteriors conditioned on the full H3+SEGUE+MagE dataset with on-sky quadrant footprints, i.e., open red contours in Figs. 3 & 4. Using parameters sampled from this posterior, we show the generated data for the average radial (top and middle panels, for quadrants 1 & 2 and 3 & 4, respectively) and tangential (bottom) velocity data as the violin plot contours. The mean observed data points are shown as the same coloured markers; same points as in Figs. 2. Again, errorbars are not shown because the generated data incorporates survey uncertainties. In most distance bins the observed data is found to be truly represented by data generated from posterior sampling, i.e. within $\sim 2\sigma$ as seen in the sub-panels. However, the observed data for the northern quadrants, Q1 and Q2, is less well represented by the generated data than the southern quadrants, Q3 and Q4. This is likely a limitation using this suite of reasonably simplistic rigid MW–LMC simulations which are able to capture global velocity perturbations, e.g. dipole signatures, but struggle to capture smaller scale perturbations e.g., quadrupole signatures. However, even higher fidelity models of the MW–LMC system, e.g., deforming simulations ([Garavito-Camargo et al. 2019](#)),

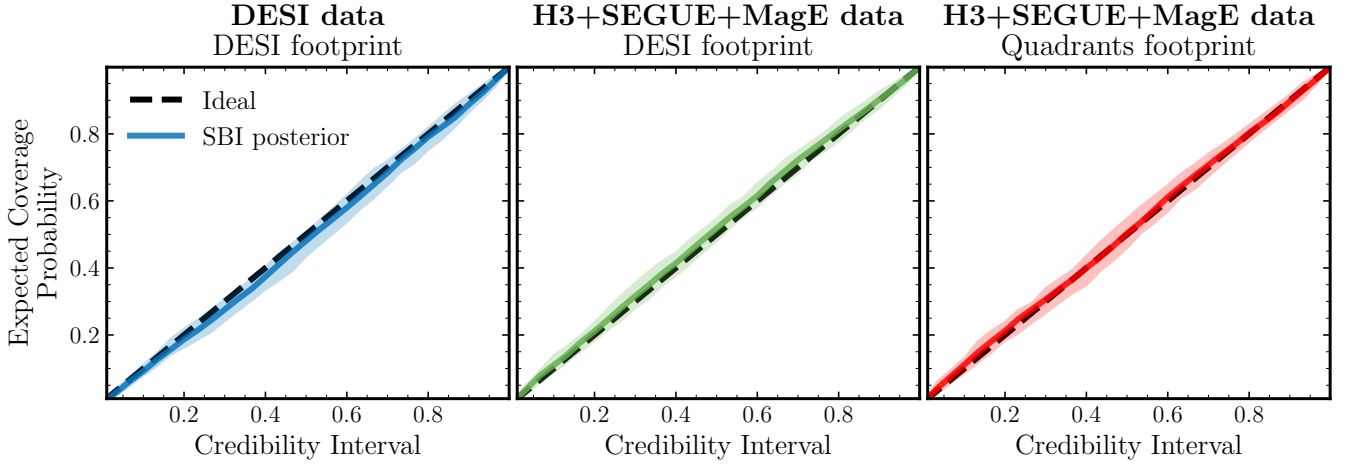


Figure 6. Coverage probability posterior check: For the posteriors estimated using SBI in this work, the probability of finding true test parameters in the appropriate credible intervals matches the expected coverage probability. A bootstrapped 2σ uncertainty is shown as the shaded region. This validates the estimated SBI posterior distributions and allows one to interpret the confidence intervals on parameter constraints as representative and reliable.

struggle to reproduce these velocity trends in the northern Galactic hemisphere (see fig. 4, Chandra et al. 2025). Additionally, the observed data point in Q4 around 80 – 100 kpc exhibits behaviour that is inconsistent with the simulations; $\sim 3\sigma$ discrepant from the peak of the generated distribution. However, this has been seen previously in Chandra et al. (2025) and been suggested to be due to unresolved substructure at these distances, for example stars stripped from the LMC. Finally, given the current precision of the observations, the rigid models and generated data in this study sufficiently represent the observations; see Fig. 7 & Fig. 8. In the future, as the observational precision increases, one must ensure this remains true otherwise more complex models and simulations may be required to operate within an SBI framework.

7 DISCUSSION

7.1 Limitations and caveats

In Brooks et al. (2026b, sec. 6.1) we discussed many of the limitations and caveats of using rigid models for this analysis. These include: assuming that the MW is at equilibrium prior to the LMC’s infall; that the LMC is on its first pericentric passage; that there is no mass exchange of the MW and LMC; that the MW stellar haloes are non-triaxial and that dynamical friction is sufficiently well captured using a Chandrasekhar prescription. We direct the reader to that section for suggestions on how the model can be improved to account for those limitations. We also note that using rigid models fixes the MW disc at the centre of the Galaxy. Hence, the reflex motion cannot discern between the motion of the disc relative to the outer MW halo and the internal bulk motions of the MW stellar halo induced by the LMC (Yaaqib et al. 2025). Extending the MW–LMC models to include self-gravity, e.g., using the N -body simulations in Sheng et al. (2025) to train the SBI neural networks, could in theory provide a way to navigate this effect and further improve the system modelling by recovering the full dipole signal (Petersen & Peñarrubia 2020b). Nevertheless, as shown in Brooks et al. (2026b), training an SBI framework on many rigid simulations retains enough of the relevant physics that more complex simulations capture (e.g., deforming N -body simulations, Garavito-Camargo et al. 2019) to avoid model

misspecification and allow rapid exploration of large model parameter spaces at a fraction of the computational cost required to run the higher fidelity simulations.

In this study, one of the limitations is the choice of summary statistics used for the parameter inference. As a reliable and simple starting point, we implemented a binned radial and tangential velocity field approach, see Sec. 2.3. By computing average values for the radial and tangential velocities in distance bins this leads to data compression over such scales, thus potentially missing out on some constraining power. The most powerful way to constrain the MW, LMC and reflex motion parameters would be to employ a star-by-star fitting procedure. However, as there are $O(10^3)$ stars in any given dataset, this would mean the dimensionality of the inference framework would increase dramatically, potentially subjecting this methodology to the curse of dimensionality. Another approach to improve upon the current set of summary statistics would be to increase the number of distance bins and reduce the area on the sky over which the binned velocities are computed. Although, increasing the number of bins would require one to ensure the model can still reliably produce realistic observed data points. This naturally lends itself to implement a spherical harmonic expansion to the halo velocity fields in a set appropriately defined distance bins (e.g. Cunningham et al. 2020).

In any approach, the selection function of a survey should be accounted for. In this work, we consider the depth and on-sky footprint of a survey, but do not account for any non-uniformity, for example the grouping of sources near the survey footprint boundaries in DESI. As an intermediate step, a nearest neighbours algorithm could be used to select the same number density of observed sources as a function of distance and angle on-sky. This would reduce the risk of introducing any systematic biases during the inference as the accuracy of the forward modelled systems has been improved by accounting for another observational effect.

Finally the choice of prior distribution for, e.g., the MW and LMC masses, influences the resulting inference. In Brooks et al. (2026b, sec 6.2), we found that using both informative and uninformative priors produced similar MW and LMC mass posteriors. Within this study, we find the MW mass posterior is insensitive to the provided data, and the inference is prior dominated. Further to this, the max-

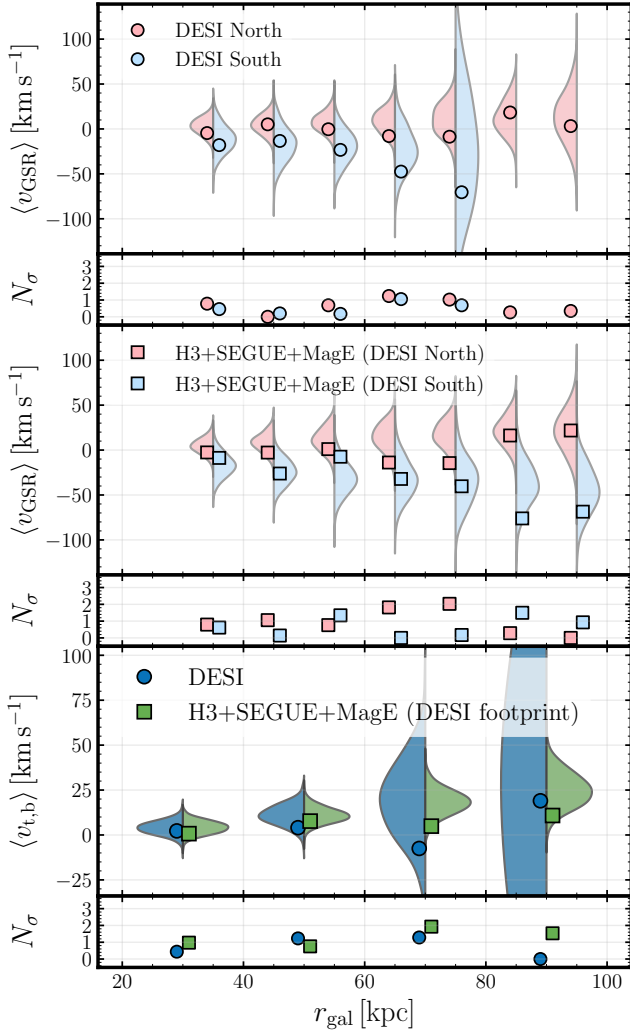


Figure 7. Predictive posterior check - DESI / H3+SEGUE+MagE data, DESI footprint: We test the quality of the estimated posterior distribution returned using the DESI and H3+SEGUE+MagE data within the DESI north and south footprints. We show the generated radial (top and middle main panels for DESI and H3+SEGUE+MagE footprints, respectively) and tangential (bottom main panel) velocity data using samples from these posteriors as the violin plot contours. The mean observed data points are shown as the same coloured markers. We do not show the data errorbars as the generated data already accounts for the survey uncertainties, allowing direct comparison to mean observed values. The sub-panels show the number of standard deviations each observed data point lies from the peak of the generated distribution. The generated and original data look sufficiently similar implying the SBI posteriors are representative of the observed data.

imum probability for the LMC mass prior and posterior are similar, although the latter is more precise; see Fig. 3. That being said, these posteriors are not in tension with the prior as the observed data is well represented by the use of rigid MW–LMC models; see the PPC in Fig. 7 & Fig. 8. Nevertheless, we note that the presented results cannot be taken as an entirely prior-independent measurements.

7.2 Sky coverage

We have varied the area over which the binned velocities were calculated by either using the DESI northern and southern survey fields

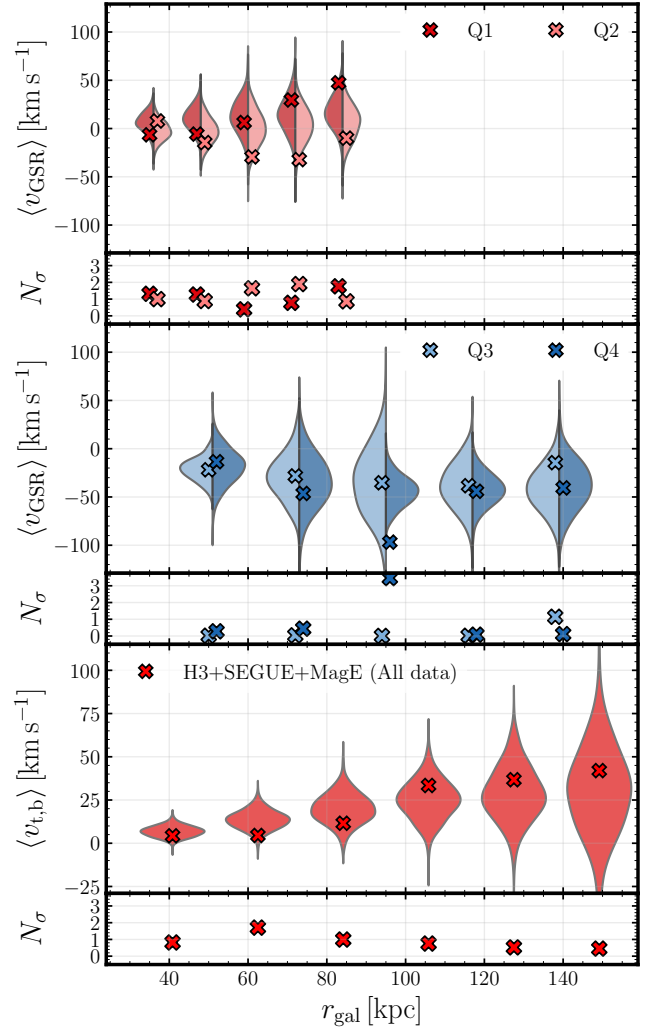


Figure 8. Predictive posterior check - H3+SEGUE+MagE data, Quadrants footprints: We test the quality of the estimated posterior distribution returned using H3+SEGUE+MagE data and on-sky quadrant footprints. We show the generated radial (top and middle main panels for northern and southern quadrants, respectively) and tangential (bottom main panel) velocity data 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, allowing direct comparison to mean observed values. The sub-panels show the number of standard deviations each observed data point lies from the peak of the generated distribution. The generated and original data look sufficiently similar implying the SBI posteriors are representative of the observed data.

or defined quadrants. In principle, as long as there is sufficient good-quality data, using the quadrants with a finer on-sky division of the data is expected to produce results that are more accurate and precise simply because of averaging over smaller sky coverages. Using the all-sky H3+SEGUE+MagE survey data, we can qualitatively assess the importance of sky coverage on returned model parameter constraints. For example, in Fig. 5, we can compare the travel velocity values found using the H3+SEGUE+MagE and choice of sky coverage. The use of the quadrants sky coverage leads to a larger value than the DESI fields, although within the uncertainties they agree with each other. Therefore, there appears to be a subtle degree of sensitivity to returned constraints given the imposed on-sky selec-

tion. Current and upcoming all-sky datasets are best placed to further investigate this effect.

7.3 Data dimensionality

We have varied whether we include information on the reflex corrected tangential velocities as the input data used to set parameter constraints. In general, including the tangential velocities slightly improves the inference precision. However, the effect is almost negligible for DESI and only just noticeable for H3+SEGUE+MagE. This can be understood from the measured uncertainties from DESI and H3+SEGUE+MagE in Sec. 2.3.2 and the panel (e) of Fig. 2. As the measured uncertainties are large, $O(10 \text{ km s}^{-1})$, when we produce survey-specific measurements from our simulations that are used to train the neural posterior estimator, this allows for a wide range of models that can explain the observed mean data points. As more precise distance and proper motion measurements become available with the release of upcoming datasets e.g., RR Lyrae stars in Gaia Data Release 4 (DR4), this will reduce the uncertainties in the measured binned tangential velocities and improve the precision on the inferred model parameters. Plus, the increase in the number of stars with available proper motion measurements e.g., Sloan Digital Sky Survey V (SDSS-V, SDSS Collaboration et al. 2025), will improve inference as well.

Throughout our inference, we have only used the mean radial and tangential velocities of outer halo stars. In future work, we can extend the current inference set-up to use information on the velocity dispersions as well; see Sec. 7.4.

We assumed knowledge of the present-day LMC position and velocity. There remains ambiguity in the definition of the LMC centre (see sec. 4.1, van der Marel & Kallivayalil 2014) which in turn corresponds to different values of the mean proper motions (Gaia Collaboration et al. 2018, 2021; Wan et al. 2020). We use well motivated LMC coordinates as input data points such that posterior probability weight is not unfairly pushed to include unphysical solutions. This helps to break model degeneracies e.g., for the dynamical friction strength. That being said, the data which provides the most constraining power on model parameters is the velocities of the outer halo stars. An interesting future application of this SBI architecture would be to constrain the present-day LMC centre’s position and velocity using the dynamics of outer halo stars as tracers.

7.4 Extending the Milky Way–LMC models

There are some improvements to the current models and methodology which will benefit future studies using this SBI approach. Firstly, we will account for radial variation in the reflex motion by implementing a linear continuity model (Chandra et al. 2025). As the amplitude of the reflex motion is expected to increase with Galactocentric distance (Yaaqib et al. 2024, 2025) this will allow the simulation priors to cover a larger area of this parameter space and allow for a more fair comparison to previous literature results. Secondly, we have only considered using the *mean* of the radial and tangential velocities of outer halo stars in this work to perform the inference. Future analysis, see Brooks et al. (2026a), will include the second-order velocity moments, i.e., the velocity *dispersions*, of outer halo stars as this allows constraints to be placed on the MW mass (Jeans 1915; Binney & Tremaine 2008; Sheng et al. 2025) and the velocity anisotropy of the stellar halo (e.g., Cunningham et al. 2019; Bird et al. 2021; Han et al. 2024; Chandra et al. 2025). Finally, we assumed the MW to be in equilibrium prior to the LMC’s infall. This model assumption is not

strictly true as there has been observed net compression (e.g., Chandra et al. 2025), and net rotation (e.g., Deason et al. 2017, though alone is not a disequilibrium state), of the MW’s stellar halo. Accounting for the pre-processed state of the MW’s stellar halo prior to the LMC’s infall is essential for future studies aiming to constrain the LMC’s past orbit (Sheng et al. 2025). Moreover, accounting for these effects may help in generating simulated data that better represents all observed data points.

8 CONCLUSIONS & OUTLOOK

We have presented an SBI architecture to infer MW–LMC parameters using the velocities of outer halo stars and provided its first application the DESI and H3+SEGUE+MagE survey datasets. The SBI framework is trained on a large set of 128,000 rigid MW–LMC simulations, with each stellar halo containing 20,000 star particles. We account for the survey-specific uncertainties and use the average radial and tangential velocities of these stars as a function of distance to constrain model parameters. We summarise our key findings as follows:

(i) We obtain our most precise MW–LMC parameter constraints using the average radial and tangential velocities of outer halo stars, divided into on-sky quadrants, from the all-sky H3+SEGUE+MagE dataset:

The reflex motion velocity, $v_{\text{travel}} = 26.4^{+5.5}_{-4.4} \text{ km s}^{-1}$

The enclosed LMC mass, $M_{\text{LMC}}(< 50 \text{ kpc}) = 9.2^{+1.9}_{-2.3} \times 10^{10} M_{\odot}$

We note the precision in the distance-averaged reflex motion velocity is slightly larger than previously found in Chandra et al. (2025) using the same dataset but unique methodology.

(ii) We find that the infall LMC’s total mass is at least $\approx 10 - 15\%$ of the MW’s total mass. This result is consistent across the DESI and H3+SEGUE+MagE datasets, independent of the velocity information or on-sky footprint selection used.

(iii) Given the prior distributions, we find that the returned MW mass is insensitive to the mean radial and tangential velocities of outer halo stars. Future use of the dispersions of these velocity components will help to break model degeneracies (Brooks et al. 2026a).

(iv) For both the DESI and H3+SEGUE+MagE datasets, when using the DESI sky coverage, constraints on v_{travel} and M_{LMC} are systematically pushed towards lower values relative to using on-sky quadrants, but remain consistent within uncertainties.

(v) The inclusion of tangential velocities of outer halo stars as input data for the inference process provides minimal extra constraining power for MW–LMC model parameters. This is because the measured tangential velocity uncertainties remain large. Upcoming data releases e.g., Gaia DR4, will greatly improve their precision and in turn the constraints on MW–LMC model parameters.

(vi) We find little preference for the strength of dynamical friction, λ_{DF} , to deviate from the standard Chandrasekhar value.

(vii) We have developed an SBI framework that enables rapid inference across large MW–LMC parameter spaces while incorporating time dependence. This approach will be essential for exploiting upcoming outer halo datasets.

We look forward to the upcoming data releases from SDSS-V and Gaia DR4 to provide an increased number of outer halo stars with precise velocity measurements. Using this exquisite data in combination with the SBI architecture presented in this work, we will produce

the most up-to-date, rapid and reliable constraints on the masses of the MW, LMC and the induced reflex motion.

ACKNOWLEDGEMENTS

We thank the referee for their feedback that improved the quality of the manuscript. RANB acknowledges support from the Royal Society and the Flatiron Institute, Simons Foundation. RANB would like to thank all of the 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 DESI DR1 BHB catalogue is available here: <https://data.desi.lbl.gov/doc/releases/dr1/vac/mws-bhb/>. The MW–LMC simulation parameters and the data for 128,000 unique stellar haloes, each with 20,000 particles, will be shared upon reasonable request. The posteriors produced using SBI in this work are available [here](#).

Software: `sbi` (Tejero-Cantero et al. 2020), `agama` (Vasiliev 2019), `gala` (Price-Whelan 2017), `NumPy` (Harris et al. 2020), `Matplotlib` (Hunter 2007), `Seaborn` (Waskom 2021), `corner` (Foreman-Mackey 2016), `Astropy` (Astropy Collaboration et al. 2013, 2018, 2022), `SciPy` (Virtanen et al. 2020).

REFERENCES

- Alsing J., Wandelt B., 2019, *MNRAS*, **488**, 5093
- Alsing J., Wandelt B., Feeney S., 2018, *MNRAS*, **477**, 2874
- Alsing J., Charnock T., Feeney S., Wandelt B., 2019, *MNRAS*, **488**, 4440
- Amarante J. A. S., Koposov S. E., Laporte C. F. P., 2024, *A&A*, **690**, A166
- Arora A., Garavito-Camargo N., Sanderson R. E., Cunningham E. C., Wetzel A., Panithanpaisal N., Barry M., 2024, *ApJ*, **974**, 286
- 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
- Bairagi A., Wandelt B., Villares-Navarro F., 2025, *arXiv e-prints*, p. [arXiv:2503.13755](#)
- Belokurov V., Deason A. J., Erkal D., Koposov S. E., Carballo-Bello J. A., Smith M. C., Jethwa P., Navarrete C., 2019, *MNRAS*, **488**, L47
- Benisty D., 2024, *A&A*, **689**, L1
- Benisty D., Vasiliev E., Evans N. W., Davis A.-C., Hartl O. V., Strigari L. E., 2022, *ApJ*, **928**, L5
- Bennett M., Bovy J., 2019, *MNRAS*, **482**, 1417
- 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
- Boylan-Kolchin M., Besla G., Hernquist L., 2011, *MNRAS*, **414**, 1560
- Brooks R. A. N., Sanders J. L., Dillamore A. M., Garavito-Camargo N., Chandra V., Price-Whelan A. M., Cargile P., 2026a, *MNRAS*,
- Brooks R. A. N., Sanders J. L., Dillamore A. M., Garavito-Camargo N., Price-Whelan A. M., 2026b, *MNRAS*, **545**, staf2111
- Buch D., Nadler E. O., Wechsler R. H., Mao Y.-Y., 2024, *arXiv e-prints*, p. [arXiv:2404.08043](#)
- 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), <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
- Chamberlain K., Price-Whelan A. M., Besla G., Cunningham E. C., Garavito-Camargo N., Peñarrubia J., Petersen M. S., 2023, *ApJ*, **942**, 18
- Chandra V., et al., 2023a, *ApJ*, **951**, 26
- Chandra V., et al., 2023b, *ApJ*, **956**, 110
- Chandra V., et al., 2025, *ApJ*, **988**, 156
- Chandra V., et al., 2026, *ApJ*, **1000**, 283
- 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
- Cooper A. P., et al., 2023, *ApJ*, **947**, 37
- 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
- Cunningham E. C., et al., 2019, *ApJ*, **879**, 120
- Cunningham E. C., et al., 2020, *ApJ*, **898**, 4
- DESI Collaboration et al., 2016, *arXiv e-prints*, p. [arXiv:1611.00036](#)
- DESI Collaboration et al., 2022, *AJ*, **164**, 207
- DESI Collaboration et al., 2026, *AJ*, **171**, 285
- Deason A. J., Belokurov V., Koposov S. E., Gómez F. A., Grand R. J., Marinacci F., Pakmor R., 2017, *MNRAS*, **470**, 1259
- Dehnen W., 1993, *MNRAS*, **265**, 250
- Eilers A.-C., Hogg D. W., Rix H.-W., Ness M. K., 2019, *ApJ*, **871**, 120
- Erkal D., et al., 2019a, *MNRAS*, **487**, 2685
- Erkal D., et al., 2019b, *MNRAS*, **487**, 2685
- Erkal D., et al., 2021, *MNRAS*, **506**, 2677
- Fearnhead P., Prangle D., 2010, *arXiv e-prints*, p. [arXiv:1004.1112](#)
- Foot H. R., et al., 2023, *ApJ*, **954**, 163
- Foreman-Mackey D., 2016, *The Journal of Open Source Software*, **1**, 24
- Fushimi K. J., Mosquera M. E., Dominguez M., 2024, *A&A*, **688**, A147
- GRAVITY Collaboration et al., 2019, *A&A*, **625**, L10
- Gaia Collaboration et al., 2018, *A&A*, **616**, A12
- Gaia Collaboration et al., 2021, *A&A*, **649**, A7
- 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., 2021, *ApJ*, **919**, 109
- Garavito-Camargo N., et al., 2024, *ApJ*, **975**, 100
- Grand R. J. J., et al., 2017, *MNRAS*, **467**, 179
- Grand R. J. J., Fragkoudi F., Gómez F. A., Jenkins A., Marinacci F., Pakmor R., Springel V., 2024, *MNRAS*, **532**, 1814
- Han J. J., Conroy C., Zaritsky D., Bonaca A., Caldwell N., Chandra V., Ting Y.-S., 2024, *arXiv e-prints*, p. [arXiv:2406.12969](#)
- Harris C. R., et al., 2020, *Nature*, **585**, 357
- Hashimoto Y., Funato Y., Makino J., 2003, *ApJ*, **582**, 196
- Heavens A. F., Sellentin E., Jaffe A. H., 2020, *MNRAS*, **498**, 3440
- 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
- Hunt J. A. S., Vasiliev E., 2025, *New Astron. Rev.*, **100**, 101721
- 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](#)
- Kallivayalil N., van der Marel R. P., Besla G., Anderson J., Alcock C., 2013, *ApJ*, **764**, 161
- Koposov S. E., et al., 2011, *ApJ*, **736**, 146

- Koposov S. E., et al., 2019, *MNRAS*, **485**, 4726
- Koposov S. E., et al., 2023, *MNRAS*, **521**, 4936
- Koposov S., et al., 2026, *The Open Journal of Astrophysics*, **9**, 55260
- Kravtsov A., Winney S., 2024, *The Open Journal of Astrophysics*, **7**, 50
- 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
- Lilleengen S., et al., 2023, *MNRAS*, **518**, 774
- Lilley E. J., Sanders J. L., Evans N. W., Erkal D., 2018a, *MNRAS*, **476**, 2092
- Lilley E. J., Sanders J. L., Evans N. W., 2018b, *MNRAS*, **478**, 1281
- Lovell C. C., et al., 2025, *MNRAS*, **544**, 3949
- Lueckmann J.-M., Boelts J., Greenberg D., Goncalves 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., Osthheimer J. C., 2003, *ApJ*, **599**, 1082
- 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)
- Mulder W. A., 1983, *A&A*, **117**, 9
- 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
- Osipkov L. P., 1979, *Pisma v Astronomicheskii Zhurnal*, **5**, 77
- Papamakarios G., Pavlakou T., Murray I., 2017, *arXiv e-prints*, p. [arXiv:1705.07057](https://arxiv.org/abs/1705.07057)
- Papamakarios G., Nalisnick E., Jimenez Rezende D., Mohamed S., Lakshminarayanan B., 2019, *arXiv e-prints*, p. [arXiv:1912.02762](https://arxiv.org/abs/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., 2020a, *MNRAS*, **494**, L11
- Petersen M. S., Peñarrubia J., 2020b, *MNRAS*, **494**, L11
- Petersen M. S., Peñarrubia J., 2021, *Nature Astronomy*, **5**, 251
- Pietrzyński G., et al., 2019, *Nature*, **567**, 200
- 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
- Rose J. C., et al., 2025, *ApJ*, **982**, 68
- 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](https://arxiv.org/abs/2507.07093)
- Samuel J., Wetzel A., Chapman S., Tollerud E., Hopkins P. F., Boylan-Kolchin M., Bailin J., Faucher-Giguère C.-A., 2021, *MNRAS*, **504**, 1379
- Sanders J. L., Lilley E. J., Vasiliev E., Evans N. W., Erkal D., 2020, *MNRAS*, **499**, 4793
- 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)), <https://arxiv.org/abs/2502.14972>
- Santos-Santos I. M. E., Fattahi A., Sales L. V., Navarro J. F., 2021, *MNRAS*, **504**, 4551
- Saoulis A. A., Piras D., Jeffrey N., Spurio Mancini A., Ferreira A. M. G., Joachimi B., 2025a, *arXiv e-prints*, p. [arXiv:2505.21215](https://arxiv.org/abs/2505.21215)
- Saoulis A. A., Piras D., Spurio Mancini A., Joachimi B., Ferreira A. M. G., 2025b, *Geophysical Journal International*, **241**, 1741
- Schönrich R., Binney J., Dehnen W., 2010, *MNRAS*, **403**, 1829
- 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, *MNRAS*, **544**, 2434
- Shipp N., et al., 2021, *ApJ*, **923**, 149
- Smith-Orlik A., et al., 2023, *J. Cosmology Astropart. Phys.*, **2023**, 070
- 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](https://arxiv.org/abs/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., 2019, *MNRAS*, **482**, 1525
- Vasiliev E., 2023, *Galaxies*, **11**, 59
- Vasiliev E., 2024, *MNRAS*, **527**, 437
- Vasiliev E., Belokurov V., Erkal D., 2021, *MNRAS*, **501**, 2279
- Virtanen P., et al., 2020, *Nature Methods*, **17**, 261
- Wan Z., Guglielmo M., Lewis G. F., Mackey D., Ibata R. A., 2020, *MNRAS*, **492**, 782
- Wang W., Han J., Cautun M., Li Z., Ishigaki M. N., 2020, *Science China Physics, Mechanics, and Astronomy*, **63**, 109801
- Warren H. T., Newberg H. J., Guffey A. G., Lin J., Mendelsohn E. J., Roux K., 2025, *arXiv e-prints*, p. [arXiv:2508.01086](https://arxiv.org/abs/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., 1986, *ApJ*, **300**, 93
- Weinberg M. D., 1989, *MNRAS*, **239**, 549
- Wetzel A., et al., 2023, *ApJS*, **265**, 44
- 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, *ApJ*, **987**, 120
- 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](https://arxiv.org/abs/2502.03522)
- Yaaqib R., Petersen M. S., Peñarrubia J., 2024, *MNRAS*, **531**, 3524
- Yaaqib R., Petersen M. S., Peñarrubia J., 2025, *MNRAS*, **544**, 1820
- Yanny B., et al., 2009, *AJ*, **137**, 4377
- York D. G., et al., 2000, *AJ*, **120**, 1579
- 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., 2025, *A&A*, **694**, A223

APPENDIX A: FULL POSTERIOR DISTRIBUTIONS

We provide the full posterior distributions for MW and LMC enclosed masses, the strength of dynamical friction and the Galactocentric Cartesian travel velocity components conditioned on different data subsets. In Fig. A1, we show the posterior conditioned on DESI data. The dark blue and light blue open contours represent the posteriors conditioned without and with the tangential velocity information, respectively; see Sec. 5.2.1. In Fig. A2, we show the posterior conditioned on H3+SEGUE+MagE data within the DESI survey footprint. The dark green and light green open contours represent the posteriors conditioned without and with the tangential velocity information, respectively; see Sec. 5.2.2. In Fig. A3, we show the posterior conditioned on the full H3+SEGUE+MagE dataset, divided into on-sky quadrants. The dark red and light red open contours represent the posteriors conditioned without and with the tangential velocity information, respectively; see Sec. 5.3. In all figures, the prior distributions are shown as the filled grey contours. Additionally, we show the data used to produce the constraints on the all-sky projection. An illustrative LMC orbit is shown as the dashed line with its present day position denoted by the grey star marker.

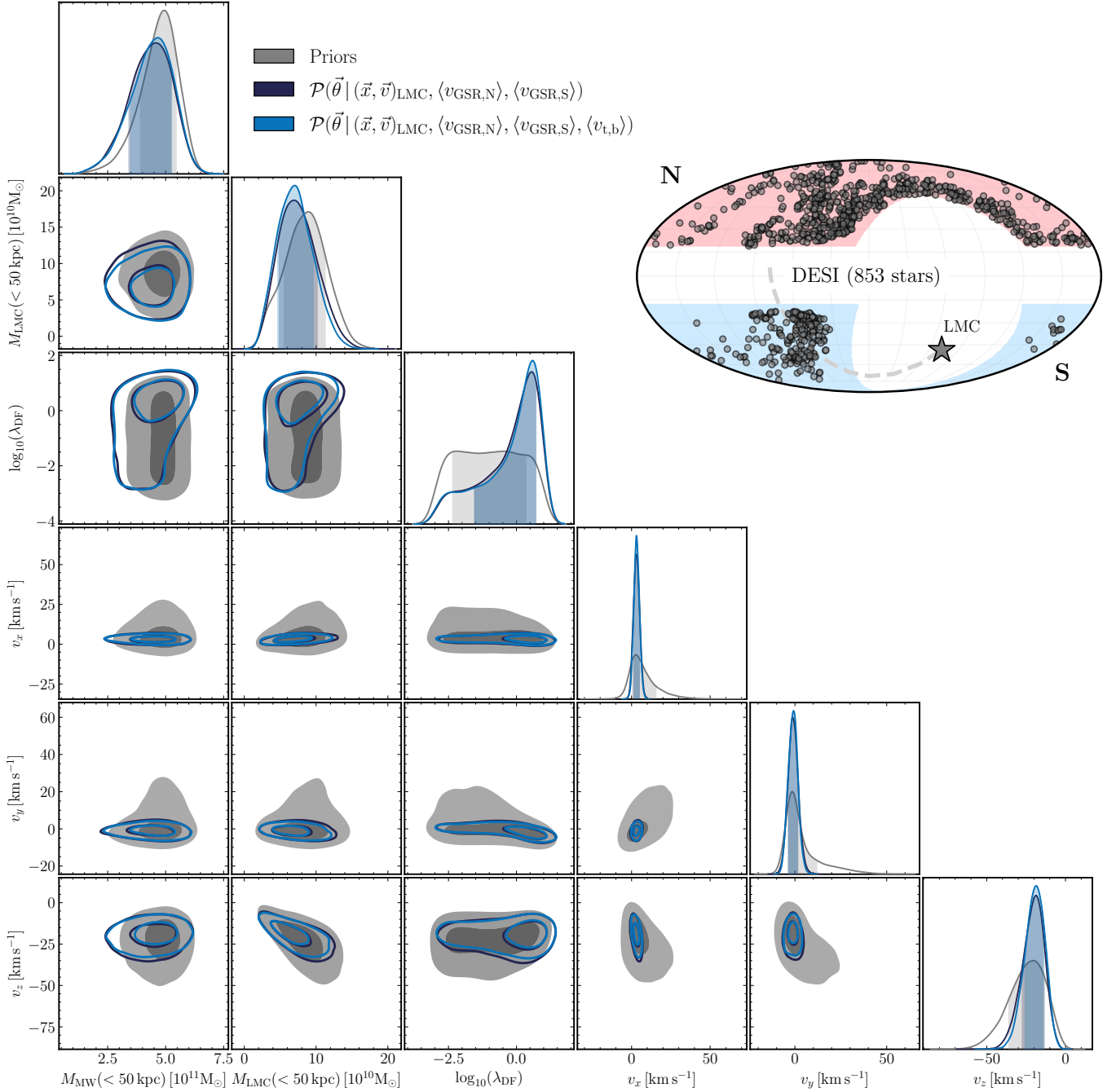


Figure A1. Posterior distributions - DESI data, DESI footprint: The joint, and individual, posterior distributions for the MW and LMC masses enclosed within 50 kpc, the dynamical friction strength, λ_{DF} , and the Cartesian components of the reflex motion travel velocity. The dark blue and light blue open contours represent the posteriors conditioned without and with the tangential velocity information, respectively. The prior distributions are shown as the filled grey contours. For the 1D posterior panels we show the 16th – 84th percentiles as shaded regions. The inclusion of DESI tangential velocity information does not greatly improve the precision of the inferred parameters. The selected DESI data between 30 – 100 kpc and contained within the northern (pink) and southern (blue) DESI survey footprints are shown in the top-right inset.

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

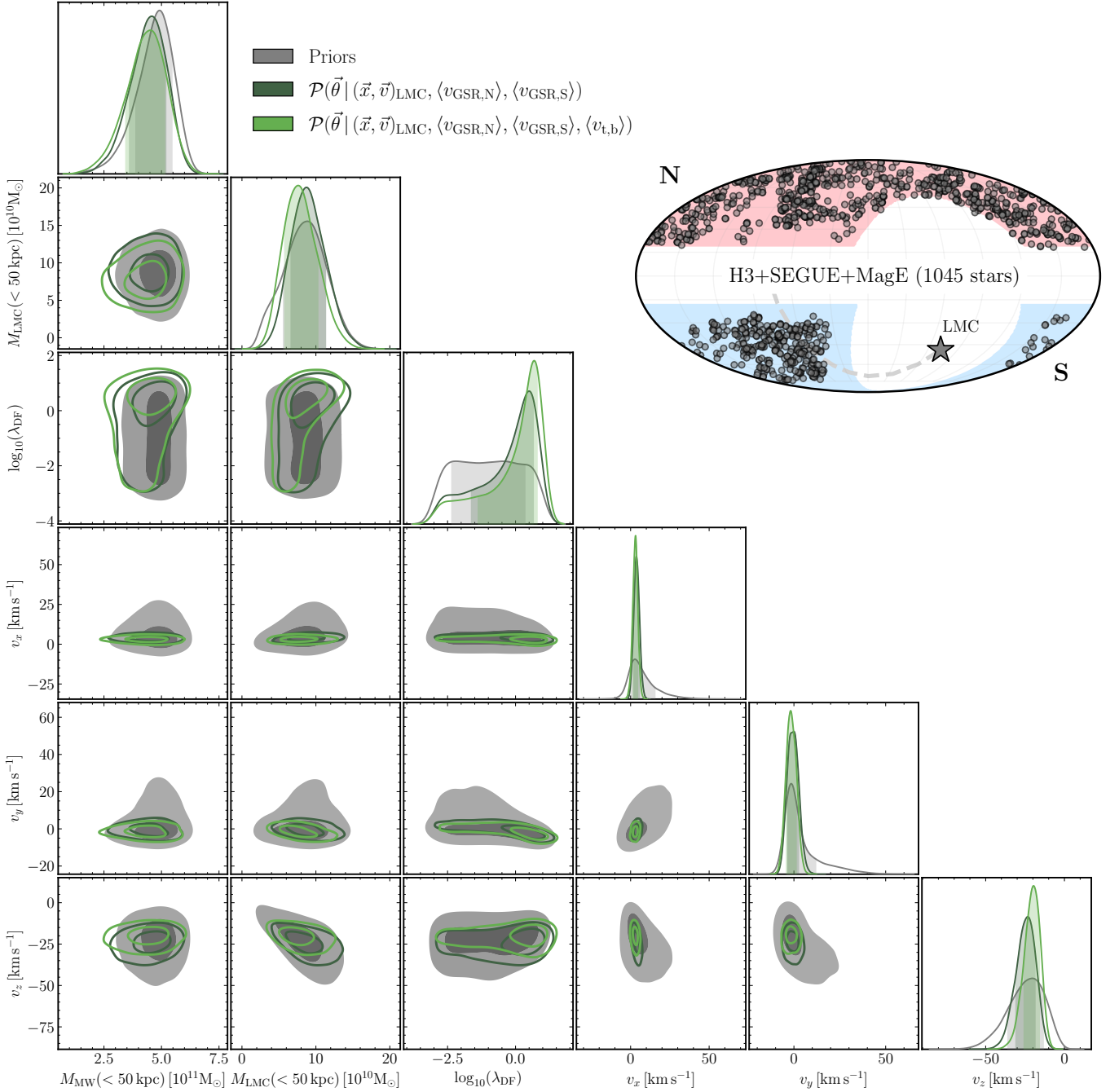


Figure A2. Posterior distributions - H3+SEGUE+MagE data, DESI footprint: The joint, and individual, posterior distributions for the MW and LMC masses enclosed within 50 kpc, the dynamical friction strength, and the Cartesian components of the reflex motion travel velocity. The dark green and light green open contours represent the posteriors conditioned without and with the tangential velocity information, respectively. The prior distributions are shown as the filled grey contours. For the 1D posterior panels we show the 16th – 84th percentiles as a shaded region. The inclusion of the tangential velocity information slightly improves the precision of the inferred parameters. The selected H3+SEGUE+MagE data between 30 – 100 kpc and contained within the northern (red) and southern (blue) DESI survey footprints are shown in the top-right inset.

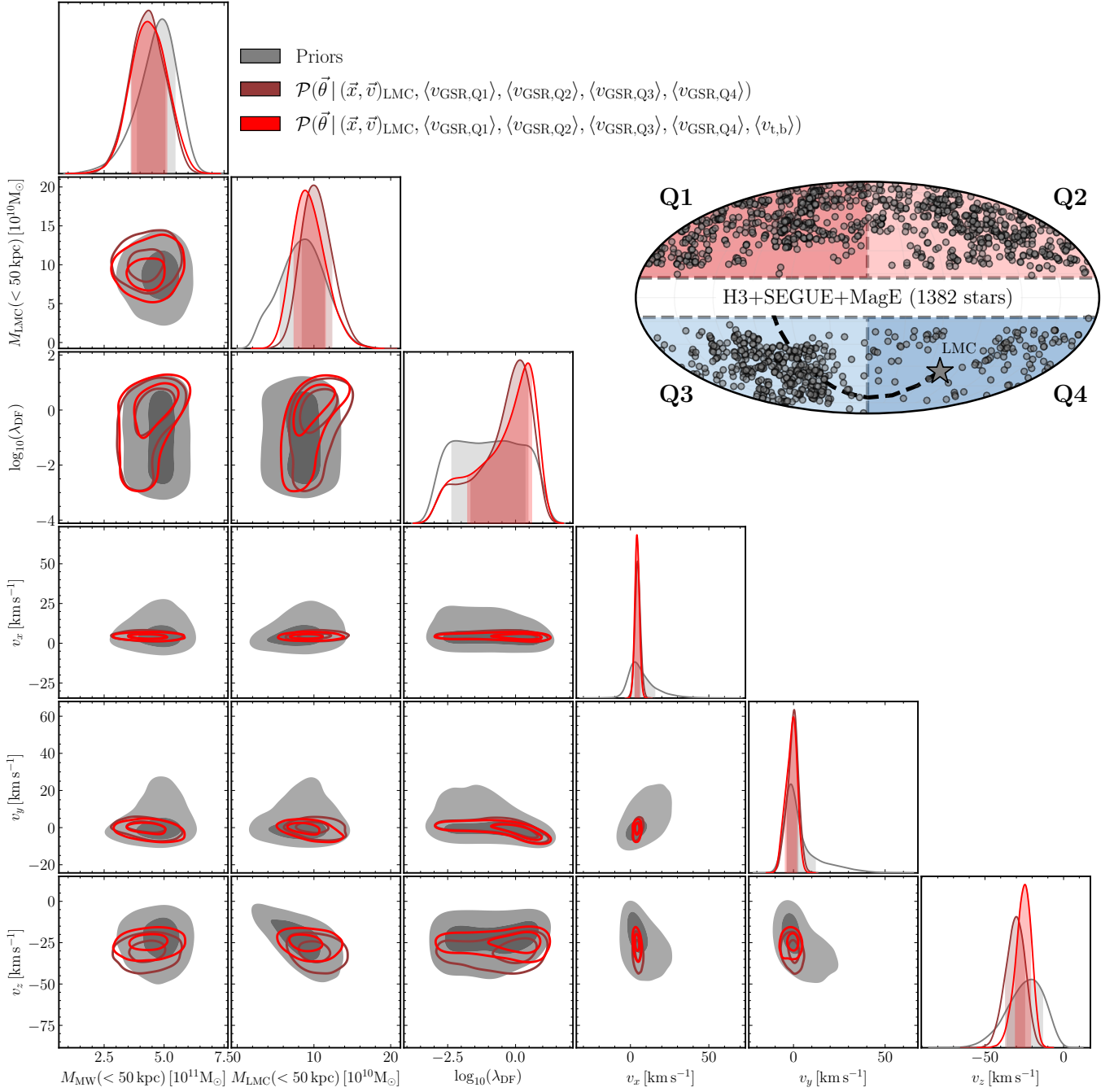


Figure A3. Posterior distributions - H3+SEGUE+MagE data, Quadrants footprint: The joint, and individual, posterior distributions for MW and LMC masses enclosed within 50 kpc, the dynamical friction strength, and the cartesian components of the reflex motion travel velocity. The dark red and light red open contours represent the posteriors conditioned without and with the tangential velocity information, respectively. The prior distributions are shown as the filled grey contours. For the 1D posterior panels we show the 16th – 84th percentiles as a shaded region. The inclusion of the tangential velocity information improves the precision of the inferred parameters. The selected H3+SEGUE+MagE data between 30 – 160 kpc and contained within the four defined quadrants, Q1-4, is shown in the top-right inset.