



Detection of the Milky Way reflex motion due to the Large Magellanic Cloud infall

Michael S. Petersen^{1,3} and Jorge Peñarrubia^{1,2,3}

The Large Magellanic Cloud is the most massive satellite galaxy of the Milky Way, with an estimated mass exceeding a tenth of the mass of the Milky Way^{1–5}. Just past its closest approach of about 50 kpc, and flying past the Milky Way at an astonishing speed of 327 km s⁻¹ (ref. ⁶), the Large Magellanic Cloud can affect our Galaxy in a number of ways, including dislodging the Milky Way disk from the Galactic centre of mass^{7–9}. Here, we report evidence that the Milky Way disk is moving with respect to stellar tracers in the outer halo in a direction that points at an earlier location on the Large Magellanic Cloud trajectory. The resulting reflex motion is detected in the kinematics of outer halo stars and Milky Way satellite galaxies with accurate distances, proper motions and line-of-sight velocities. Our results indicate that dynamical models of our Galaxy cannot neglect gravitational perturbations induced by the Large Magellanic Cloud infall, nor can observations of the stellar halo be treated in a reference frame that does not correct for disk reflex motion. Future spectroscopic surveys of the stellar halo combined with Gaia astrometry will allow for sophisticated modelling of the Large Magellanic Cloud trajectory across the Milky Way, constraining the dark matter distribution in both galaxies with unprecedented detail.

Numerical simulations that follow the accretion of a massive Large Magellanic Cloud (LMC) into our Galaxy predict that the Milky Way (MW) disk will be displaced with a speed of $v_{\text{travel}} \approx v_{\text{LMC}}(M_{\text{LMC}}/M_{\text{MW}})$ relative to the initial Galactic barycentre^{7–9}. If true, the disk motion should be detectable in the apparent kinematics of halo stars with long dynamical times as a simple dipole⁹. The peak of the blue-shifted velocities, or apex, is the direction towards which the MW disk is currently travelling. Here, we measure the direction (ℓ_{apex} , b_{apex}) and the magnitude (v_{travel}) of the disk motion by applying a Bayesian analysis of proper motions and radial velocities of hundreds of bright, distant stars in the smooth MW stellar halo: K giants, blue horizontal branch (BHB) stars and the MW satellites, each with individual 6-dimensional position–velocity information^{10–13}. Distant tracers have long dynamical times and are thus slow to react to the gravitational pull of the LMC⁹. Because the MW disk moves in response to the LMC, but the outer halo does not, a reflex motion arises as a global kinematic pattern on the sky: the distant halo appears to move relative to the MW disk, when in fact it is the MW disk moving relative to the distant halo stars, an effect akin to the apparent motion of rain drops seen from a car driving through a downpour.

In Fig. 1, we show the on-sky probability density function for the direction of the MW disk travel imprinted in the kinematics of each of the three halo tracers (ℓ_{apex} , b_{apex}), against a backdrop of RR Lyrae stars^{14–16}. All three tracers return consistent values at the 67% confidence level (Table 1). Strikingly, the apex direction is

not pointing towards the present-day location of the LMC on the sky, but rather along its historical trajectory across the MW. This implies that, as the LMC approached its current location with great speed, the MW disk was not able to keep up. As a result, the MW disk appears to be travelling in the direction of the LMC at some earlier point on its trajectory, as indicated by the plane of the LMC orbit (dashed curve). A similar effect is observed in the numerical models that follow the accretion of the LMC onto the MW (Supplementary Information).

In Fig. 2, we show the probability density function of the magnitude of the disk motion v_{travel} derived from each of the studied tracers. The strongest constraint is provided by the K giant stars, with $v_{\text{travel,Kgiants}} = 35^{+5}_{-5}$ km s⁻¹. Being the most numerous bright tracers in our data sample and having accurate proper motions out to large Galactocentric distances (Extended Data Fig. 1 and Supplementary Information), these stars currently provide the best kinematic targets of the outermost halo. The full phase-space information also allows for removal of major substructure in the halo (Extended Data Fig. 2). BHB stars and satellites return slightly less constraining measurements of the reflex motion, $v_{\text{travel,BHB}} = 26^{+9}_{-9}$ km s⁻¹ and $v_{\text{travel,satellites}} = 55^{+22}_{-23}$ km s⁻¹, respectively. All three datasets are consistent at the 67% confidence level (Table 1 and Extended Data Fig. 3). Fitting the combined sample of BHB and K giant stars yields $v_{\text{travel,combined}} = 32^{+4}_{-4}$ km s⁻¹. The substantial reflex motion and location of the apex is consistent with a massive LMC falling in for the first time. To test our Bayesian fitting technique we use mock datasets of the stellar halo drawn from numerical models of the MW-LMC interaction (Extended Data Fig. 4 and Supplementary Information). The magnitude of the reflex motion observed in the stellar halo of the MW is consistent with numerical models where the infall mass of the LMC is greater than $10^{11} M_{\odot}$ (where M_{\odot} is the solar mass; see Extended Data Figs. 5, 6 and 7), which is much larger than the mass enclosed within its luminous radius, $M_{\text{LMC}}(<8.7 \text{ kpc}) \simeq 0.17 \times 10^{11} M_{\odot}$ (ref. ¹⁷), suggesting that the LMC fell in surrounded by an extended dark matter halo.

Our results complicate theoretical studies of the MW stellar halo and firmly indicate that our Galaxy cannot be treated in dynamical equilibrium. In addition, Earth observers must correct for non-inertial effects introduced by the MW disk motion when translating velocities measured in a heliocentric frame into a coordinate system whose origin is at the Galactic centre of mass. Although corrections from the solar reflex are commonly applied, the reflex motion of the MW disk has so far been neglected in dynamical studies of distant objects. Failure to properly treat the LMC infall may bias measurements of various quantities, such as the total mass of the MW¹⁸. The measurements presented here, constraining both the amplitude and direction of travel of the MW disk, will help to

¹Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, UK. ²Centre for Statistics, University of Edinburgh, School of Mathematics, Edinburgh, UK. ³These authors contributed equally: Michael S. Petersen, Jorge Peñarrubia. e-mail: michael.petersen@roe.ac.uk

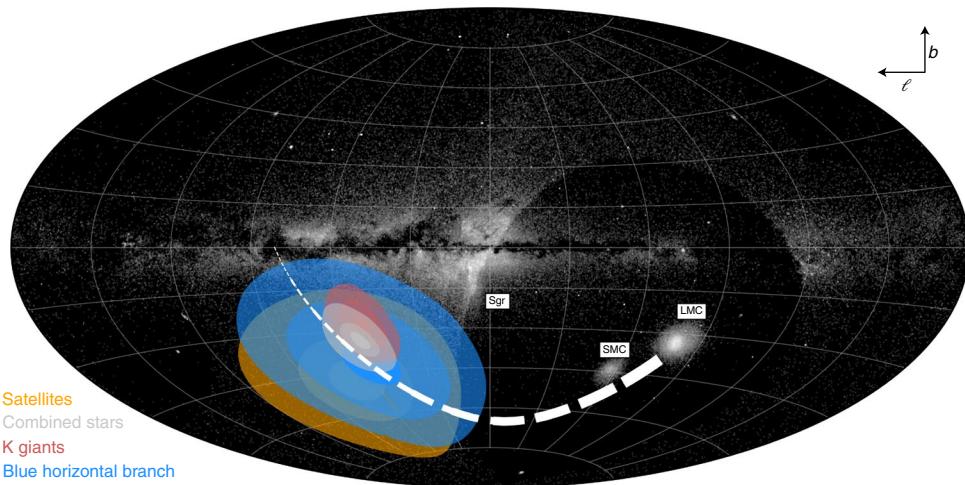


Fig. 1 | Posteriors on the direction of the MW disk motion relative to halo stars located at Galactocentric distances $r > 40$ kpc, shown in Aitoff projection. Shaded contours denote the 67%, 90% and 95% probabilities and are colour-coded according to each sample considered in this work. The black-and-white background image is the observed projected density of RR Lyrae stars identified in the union of the Pan-STARRS DR1 and Gaia DR2 catalogues^{14–16}. The LMC, as well as two less-massive satellites, the Small Magellanic Cloud (SMC) and the Sagittarius dwarf galaxy (Sgr), are labelled. Comparison with the plane of the LMC orbit derived from its proper motions (dashed white curve) indicates that the MW disk is roughly moving towards a point along the past trajectory of the LMC across our Galaxy. The Pan-STARRS survey covers only the Northern sky, and thus the region around the LMC is covered only by Gaia and is thus comparatively undersampled in RR Lyrae stars.

Table 1 | Posteriors on model parameters

Description	Parameter	Units	K giant + BHB stars	K giant stars	BHB stars	Satellites
Disk barycentre velocity	v_{travel}	km s^{-1}	32^{+4}_{-4}	35^{+5}_{-5}	26^{+9}_{-9}	55^{+22}_{-23}
Apex longitude	ℓ_{apex}	deg	56^{+9}_{-9}	53^{+9}_{-9}	59^{+25}_{-24}	74^{+46}_{-34}
Apex latitude	b_{apex}	deg	-34^{+10}_{-9}	-28^{+10}_{-9}	-42^{+19}_{-16}	-49^{+20}_{-17}
Mean halo radial velocity	$\langle v_r \rangle$	km s^{-1}	-18^{+4}_{-4}	-17^{+5}_{-5}	-22^{+8}_{-8}	-35^{+16}_{-16}
Mean halo azimuthal velocity	$\langle v_\phi \rangle$	km s^{-1}	-20^{+5}_{-4}	-22^{+5}_{-5}	-9^{+9}_{-9}	-25^{+20}_{-20}
Mean halo polar velocity	$\langle v_\theta \rangle$	km s^{-1}	14^{+5}_{-5}	18^{+6}_{-6}	-2^{+10}_{-9}	-19^{+25}_{-24}
Line-of-sight velocity hyperparameter	$\sigma_{h,\text{los}}$	km s^{-1}	93^{+2}_{-2}	94^{+3}_{-3}	91^{+3}_{-4}	92^{+13}_{-10}
Galactic longitude velocity hyperparameter	$\sigma_{h,\ell}$	km s^{-1}	81^{+4}_{-4}	81^{+4}_{-4}	76^{+8}_{-8}	102^{+15}_{-12}
Galactic latitude velocity hyperparameter	$\sigma_{h,b}$	km s^{-1}	76^{+3}_{-3}	71^{+3}_{-3}	96^{+8}_{-7}	108^{+16}_{-13}

remove non-inertial perspective effects from the kinematics of the outer stellar halo and the satellite galaxy population.

Previous studies have used similar datasets to those considered here to measure the rotation pattern and velocity dispersion profile of the stellar halo^{19,20}, but did not account for the reflex motion of the MW disk. A recent work¹⁸ finds a net positive v_z motion in a sample of globular clusters and dwarf galaxies with 6-dimensional observations. As the LMC is expected to induce the disk to move downwards, one expects an upward motion in the outer halo. However, the magnitude and direction of the reflex motion signal has remained unconstrained so far. In particular, the localization of the apex signal is only enabled by combining full phase-space information of distant halo tracers (Extended Data Fig. 5).

Our results conclusively show that the kinematics of halo stars at $r \geq 40$ kpc are affected by the displacement of the MW disk from the Galactic barycentre. At smaller radii, the magnitude of reflex motion diminishes and the kinematic pattern becomes more dynamically convoluted than a simple dipole as the inner stellar halo reacts in response to the displacement of the disk potential⁹. Tidal streams may also be affected by the dislodgement of the MW disk

from the Galactic centre of mass. Intriguingly, current constraints on the shape of the dark matter halo using tidal streams have so far reported contradictory values that cover all possible geometries: from spherical to oblate to prolate to even triaxial shapes^{21–25}. Given that the resulting reflex motion varies across the sky (Extended Data Fig. 8), and that tidal streams can wrap large regions of the Galaxy, this effect may be an important factor with which to explain the discrepant measurements.

Current state-of-the-art dynamical-modelling techniques often compute orbits in gravitational potentials centred at the MW disk in order to constrain the distribution of matter in our Galaxy. Our results indicate that an essential next theoretical step is to generate time-varying MW-LMC potentials where both galaxies are allowed to deform as a result of their mutual gravitational attraction. Once self-consistent MW-LMC potentials are developed, a number of questions arise. What was the structure of the stellar and dark matter haloes prior to the infall of the LMC? What is the exact trajectory of the LMC since it fell into the MW, and how did the dynamics of dark matter alter its path? Does the LMC lose its dark matter halo envelope to tides, and where does the stripped dark matter end up

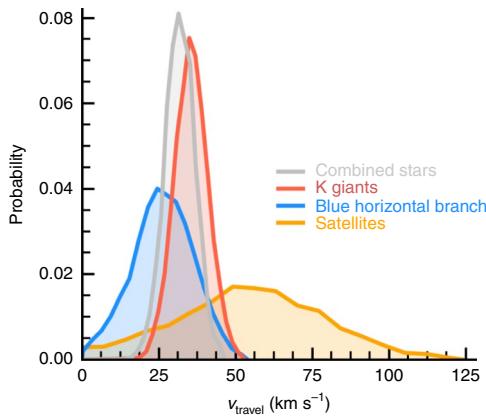


Fig. 2 | Measured velocity of the MW disk with respect to halo stars located at Galactocentric distances $r > 40$ kpc (red and blue curves), the combined sample (silver curve) and satellites (orange curve). We find that the three distinct samples of halo tracers (K giant stars, BHB stars and satellites) return consistent values within statistical uncertainties.

in our Galaxy? In this context, the detection of the reflex motion of the MW disk calls for the development of more sophisticated, self-consistent models that capture non-equilibrium features in the stellar halo. For example, the LMC is predicted to induce strong perturbations on the spatial configuration and kinematics of stellar tracers, such as the formation of a pronounced ‘wake’ in the smooth stellar halo⁸, and to perturb tidal streams out of their orbital planes⁵.

Astronomers must also turn their attention to the dynamics of the Local Group, a larger volume of galaxies that includes our nearest massive neighbour, Messier 31 (M31, the Andromeda galaxy). The quadrant that hosts M31 ($\ell < 180^\circ$, $b < 0^\circ$) is one of the regions of the sky most affected by reflex motion, a finding that suggests that previous studies treating the orbit of M31 will need to be revised in light of the localized direction of the MW disk displacement with respect to the Galaxy barycentre.

Our understanding of the MW stellar halo is on the verge of a golden age owing to data availability and analysis. Upcoming data releases from the Gaia mission will provide robust proper-motion measurements out to the most distant regions of the halo. When coupled with ongoing and future radial velocity surveys (LAMOST²⁶, 4MOST²⁷) and new instruments (MOONS²⁸), the number of halo stars with accurate 6-dimensional phase-space measurements could be potentially increased by a factor of about 100 in the next decade (Supplementary Information). Such datasets will allow us to measure the radial variation of reflex motion, the infall trajectory of the LMC and the location of its tidal debris with great detail, which in turn will constrain the dark matter distribution in both galaxies with unprecedented accuracy.

Methods

Bayesian inference method. In this section we describe the fundamentals of a Bayesian method to search for signatures of reflex motion in the outer stellar halo. We analyse stars with full phase-space information, that is, heliocentric distances D , Galactocentric coordinates (l, b) , line-of-sight velocities v_{los} and proper motions (μ_ℓ, μ_b) . The technique is applied to observational datasets as well as to mock catalogues of halo stars created from n -body realizations of the LMC–MW system (Supplementary Information).

Our models contain nine free parameters that are fitted simultaneously to the data. These are the magnitude of the disk velocity, v_{travel} , its apex direction in Galactocentric coordinates $(\ell_{\text{apex}}, b_{\text{apex}})$, the mean velocity of halo stars in spherical coordinates with origin at the Galactic centre $\langle v_r \rangle$, $\langle v_\phi \rangle$ and $\langle v_\theta \rangle$, plus three hyperparameters $\sigma_{\text{h,los}}$, $\sigma_{\text{h},\ell}$ and $\sigma_{\text{h},b}$ with units of km s^{-1} , which are added in quadrature to the observational measurement variance in order to minimize covariance with the rest of model parameters. Hyperparameters provide a useful tool with which to assign weights to datasets beyond those derived from

statistical errors²⁹. Their chief role is to minimize biases arising from (unknown) model imperfections by increasing the uncertainty associated with the posterior distributions. These statistical objects effectively act as ‘nuisance’ parameters with no direct physical meaning, which are fitted to the data to be subsequently marginalized over the posterior distributions on the quantities of interest.

Our analysis allows for non-zero mean velocities in each degree of freedom, $\langle v_r \rangle$, $\langle v_\phi \rangle$ and $\langle v_\theta \rangle$, as well as a reflex motion of the MW disk owing to the gravitational pull of the LMC, Δv_{travel} . The motion of the MW disk with respect to the stellar halo is modelled as a simple dipole. Comparison against live n -body models shows that this analytical model describes the reflex motion of halo stars at large Galactocentric distances very well¹. To this aim, we rotate the Galactocentric reference system in such a way that the new z -axis is aligned with the disk motion, which points in the direction $(\ell_{\text{apex}}, b_{\text{apex}})$. The spherical azimuthal and polar coordinates of any given star are transformed to the rotated frame as $(\phi, \theta) \rightarrow (\phi_1, \phi_2)$. In this frame, the dipole $\Delta v_{\text{travel}} = (\Delta v_r, \Delta v_{\phi_1}, \Delta v_{\phi_2})$ can be simply computed as

$$\begin{aligned}\Delta v_r &= -v_{\text{travel}} \cos(\phi_2) \\ \Delta v_{\phi_2} &= +v_{\text{travel}} \sin(\phi_2) \\ \Delta v_{\phi_1} &= 0.\end{aligned}\quad (1)$$

The rotation of the Galactocentric axes is done using an Euler–XYZ transformation with angles $(\phi, \theta, \psi)_{\text{rot}} = (l_{\text{apex}}, \pi/2 - b_{\text{apex}}, 0)$. Conversion between heliocentric and Galactocentric frames is done by adopting a right-handed Cartesian coordinate system where the centre of the MW is at $(x, y, z) = (0, 0, 0)$, and the Sun is at $\mathbf{r}_{\odot \rightarrow \text{MW}} = (-8.30, 0.00, 0.02) \text{ kpc}$ (refs. ^{30,31}). In these coordinates the Sun moves with a velocity $\mathbf{v}_{\odot \rightarrow \text{MW}} = (11.10, 244.24, 7.25) \text{ km s}^{-1}$ (refs. ^{32,33}). Changes of up to a few per cent in these quantities do not affect our analysis³⁴; however, v_{travel} is modestly dependent on the choice of the rotation velocity at the solar circle such that increasing the rotation velocity at the solar circle decreases the magnitude of the reflex motion, and vice versa.

In a heliocentric coordinate system, the average velocity vector of the stellar halo can be calculated from the mean velocity components fitted in our analysis and the reflex motion of the disk as

$$\langle \mathbf{v} \rangle = \Delta \mathbf{v}_{\text{travel}} + \langle v_r \rangle + \langle v_\phi \rangle + \langle v_\theta \rangle - \mathbf{v}_{\odot \rightarrow \text{MW}}. \quad (2)$$

The observed components of the vector (2) correspond to the following projections

$$\begin{aligned}\langle v_{\text{los}} \rangle &= \langle \mathbf{v} \rangle \cdot \hat{\mathbf{u}}_{\text{los}} \\ \langle \mu_l \rangle &= \frac{\langle \mathbf{v} \rangle \cdot \hat{\mathbf{u}}_l}{kD} \\ \langle \mu_b \rangle &= \frac{\langle \mathbf{v} \rangle \cdot \hat{\mathbf{u}}_b}{kD},\end{aligned}\quad (3)$$

where $\hat{\mathbf{u}}_{\text{los}}$, $\hat{\mathbf{u}}_l$ and $\hat{\mathbf{u}}_b$ are unit vectors aligned with the line-of-sight direction and the Galactocentric coordinates (ℓ, b) , respectively. We use the conversion from velocity to proper motion

$$v_i = kD\mu_i, \quad (4)$$

with a sub-index $i = \ell, b$ and a geometrical factor $k = 4.74057 \text{ km s}^{-1} \text{ kpc}^{-1} (\text{milliarcsec/yr})^{-1}$. Here, D is the heliocentric distance measured in kiloparsecs.

To derive posteriors on the model parameters we adopt a Gaussian likelihood

$$\mathcal{L}(\{D_i, \ell_i, b_i, v_{\text{los},i}, \mu_{\ell,i}, \mu_{b,i}\}_{i=1}^{N_{\text{sample}}} | \mathbf{S}) = \prod_{i=1}^{N_{\text{sample}}} p_{\text{los}}(v_{\text{los}}) p_{\text{pm}}(\mu_\ell, \mu_b); \quad (5)$$

where $\mathbf{S} = (v_{\text{travel}}, \ell_{\text{apex}}, b_{\text{apex}}, \langle v_r \rangle, \langle v_\phi \rangle, \langle v_\theta \rangle, \sigma_{\text{h,los}}, \sigma_{\text{h},\ell}, \sigma_{\text{h},b})$ is a vector that comprises the model parameters. p_{los} is the one-dimensional normal probability function associated with line-of-sight velocities

$$p_{\text{los}}(v_{\text{los}}) = \frac{1}{\sqrt{2\pi\sigma_{\text{los}}^2}} \exp\left[-\frac{(v_{\text{los}} - \langle v_{\text{los}} \rangle)^2}{2\sigma_{\text{los}}^2}\right]; \quad (6)$$

with a one-dimensional variance along the line-of-sight direction

$$\sigma_{\text{los}}^2 = e_{\text{los}}^2 + \sigma_{\text{h,los}}^2, \quad (7)$$

where e_{los} is the error associated with the line-of-sight velocity, plus the additional ‘freedom’ provided by the hyperparameter $\sigma_{\text{h,los}}$.

The likelihood associated with proper motions is modelled as a bivariate Gaussian,

$$p_{\text{pm}}(\mu_\ell, \mu_b) = \frac{1}{2\pi\sqrt{\det(C)}} \exp\left[-(\chi - \langle \chi \rangle)^T C^{-1}(\chi - \langle \chi \rangle)\right]; \quad (8)$$

where $\chi = (\mu_\ell, \mu_b)$ is the data vector, and $\langle \chi \rangle = (\langle \mu_\ell \rangle, \langle \mu_b \rangle)$ is the vector containing the systemic proper motions. The covariance matrix, C , contains the correlation between the proper motion errors

$$C = \begin{bmatrix} \epsilon_{\ell}^{\prime 2} & \epsilon_{\ell} \epsilon_b \rho \\ \epsilon_{\ell} \epsilon_b \rho & \epsilon_b^{\prime 2} \end{bmatrix} \quad (9)$$

where ρ and $\epsilon_{\ell,b}$ are correlation coefficients and measurement uncertainties, respectively. The two-dimensional variances in proper motion space ($\epsilon_{\ell,b}'$) can be estimated as

$$\begin{aligned} \epsilon_{\ell}'^2 &= \epsilon_{\ell}^2 + \epsilon_D^2 |\frac{d\mu_{\ell}}{dD}|_D^2 + \frac{\sigma_{h,\ell}^2}{(kD)^2} = \epsilon_{\ell}^2 + \epsilon_D^2 \frac{|\mu_{\ell}|^2}{D^2} + \frac{\sigma_{h,\ell}^2}{(kD)^2} \\ \epsilon_b'^2 &= \epsilon_b^2 + \epsilon_D^2 |\frac{d\mu_b}{dD}|_D^2 + \frac{\sigma_{h,b}^2}{(kD)^2} = \epsilon_b^2 + \epsilon_D^2 \frac{|\mu_b|^2}{D^2} + \frac{\sigma_{h,b}^2}{(kD)^2} \end{aligned} \quad (10)$$

which include proper motion ($\epsilon_{\ell,b}'$) as well as distance (ϵ_D) errors³⁵. In practice, we find that the the variances $\epsilon_{\ell}'^2$ and $\epsilon_b'^2$ are typically dominated by the nuisance parameters $\sigma_{h,\ell}^2$ and $\sigma_{h,b}^2$, respectively. These statistical parameters absorb random motions in the stellar halo, which in general are much higher than the velocities associated with measurement errors.

We adopt flat priors for v_{travel} , ℓ_{apex} , $\cos(b_{\text{apex}})$, $\langle v_r \rangle$, $\langle v_{\phi} \rangle$ and $\langle v_{\theta} \rangle$, and Jeffreys priors for $\sigma_{h,\ell}$, $\sigma_{h,\ell}$ and $\sigma_{h,b}$, with ranges that include reasonable parameter values.

See Extended Data Fig. 3 for full covariances between the nine fitting parameters.

Inference code. We apply a nested-sampling technique³⁶ to calculate posterior distributions for our parameters and the evidence of the model. Here, we use the code MultiNest, a Bayesian inference tool which also produces posterior samplings and returns error estimates of the evidence^{37,38}.

Datasets. We use available datasets of two luminous stellar tracers at $r > 40$ kpc: BHB stars^{10,39}, and K giant stars^{11,40} identified using SDSS/SEGUE photometry and spectroscopy⁴¹. We accept the positions, distances, and uncertainties in the published tables. We apply a colour-cleaning technique to reduce Blue Straggler contamination in the BHB sample, eliminating stars that fall in the region of colour space suggested in ref.⁴². We cross-match both the BHB and K giant star catalogs at 1" tolerance with Gaia DR2^{12,43} to obtain proper motions in right ascension and declination. We filter the samples by the reduced unit weighted error—a collected measure of the data quality—accepting only stars for which it is < 1.4 (ref.⁴⁴). We convert the observed proper motions to Galactocentric coordinates, including proper motion uncertainty covariance.

While Gaia DR2 has known systematic errors in proper motion for bright stars ($G < 12$), the systematics for faint stars do not appear to show any large-scale structure that would bias our results and can be assumed to be in a non-rotating frame⁴⁴. We accept the reported mean and root-mean-square all-sky uncertainties (in milliarcseconds per year: $\langle \mu_{RA} \rangle = 0$; $\sigma_{RA} = 0.039$; $\langle \mu_{dec} \rangle = 0.011$; $\sigma_{dec} = 0.037$) and emphasize that the root-mean-square uncertainties are an order of magnitude smaller than the typical errors on individual sources.

Although the catalogues do not distinguish between disk and halo stars, we do not expect any disk stars at distances $r > 40$ kpc. Independently of published distances, we have also checked that BHB and K giant stars in our catalogue have unresolved Gaia parallaxes, which confirms their location beyond the disk volume. We also exclude stars with latitudes $|b| < 10^\circ$ to minimize contamination from faint disk stars misidentified as BHB or K giant stars in the published catalogues.

The resulting catalogue still contains stars whose distances may be highly overestimated. These objects stand out as stars with very large velocities. In order to remove contaminants we performed iterative sigma clipping following a standard procedure: for all stars in the sample, we compute the difference between model prediction and measurement ($(\langle \text{data} - \text{model} \rangle)/\sigma$), where brackets denote averages over posterior distributions, and the subindex i runs over the three velocity dimensions: radial velocities and the two proper motion components. Here, σ is the fitted dispersion, which combines individual measurement errors and the hyperparameters (equations (7) and (10)). Stars with $|(\langle \text{data} - \text{model} \rangle)/\sigma| > 3$ in any of those dimensions are considered outliers. After removing them, we re-fit the new sample. The above steps are iterated until convergence is achieved.

Given that the BHB and K giant star samples are drawn from SEGUE⁴¹, the sky coverage is not complete. We test the impact of incomplete sky coverage on our Bayesian inference method using mock catalogues drawn from live n -body simulations within the SEGUE footprint (Extended Data Figs. 5 and 6 as well as Supplementary Information). The results indicate that the limited sky area does not introduce any meaningful bias. Additionally, we measure the magnitude of proper motions at $r > 40$ kpc from the K giant star sample, finding that the proper motions are less than 1 milliarcsecond per year. This provides an estimate of the precision in proper motions required to studying an ensemble of stars in the MW halo at $r > 40$ kpc. The canonical velocity of a star in the MW halo, $v = 250$ km s⁻¹, corresponds to a proper motions of $\mu \approx 1$ milliarcsecond per year at 50 kpc (Extended Data Fig. 1).

Since we model the stellar halo as a smooth object, the presence of kinematic substructures, such as tidal streams, could potentially bias our results. The Sagittarius stream is the largest known substructure in the stellar halo of our Galaxy⁴⁵. We exclude Sagittarius stream stars from our sample on the basis of their angular momentum coordinates (see Supplementary Information for additional details). In short, first we identify stars located in the plane of the Sagittarius

stream by rotating our reference frame such that the polar angle of stream stars measured from the Galactic centre is located at $B_{\text{sgt}} = 0$ on average⁴⁶. Subsequently, we identify stars in the orbital plane of the stream by choosing those that fall within $|B_{\text{sgt}}| < 20^\circ$, and compute their angular momentum components, $\mathbf{L} = (L_x, L_y, L_z)$. Next, we draw a sphere with radius 3,000 kpc km s⁻¹ around the location of the Sagittarius dwarf $\mathbf{L}_{\text{sgt}} = (605, -4515, -1267)$ kpc km s⁻¹. Stars that fall in this region are tagged as Sagittarius members and removed from our sample (Extended Data Fig. 2). We have checked that the result does not strongly depend on the size of the sphere. We do not find that any of the stars in our sample belong to known globular clusters or satellite galaxies, apart from Sagittarius.

We also use the positions, distances, radial velocities, and proper motions of MW satellites compiled from literature sources^{13,47–52}. To identify possible LMC satellites we follow a procedure similar to that above to identify Sagittarius stream members. First, we locate the LMC and the Small Magellanic Cloud in angular momentum space at $\mathbf{L}_{\text{LMC}} = (-16044, 1236, -1520)$ kpc km s⁻¹ and $\mathbf{L}_{\text{SMC}} = (-15169, -3155, -2827)$ kpc km s⁻¹ and define the separation $|\Delta\mathbf{L}_{\text{MC}}| \equiv |\mathbf{L}_{\text{SMC}} - \mathbf{L}_{\text{LMC}}| = 4,664$ kpc km s⁻¹ as a threshold in angular momentum space. Subsequently, we compute the angular momentum of the rest of the MW satellites and rank-order them according to their separation with respect to the LMC, $|\Delta\mathbf{L}| = |\mathbf{L} - \mathbf{L}_{\text{LMC}}|$. The first two galaxies in the ranking (Horologium I and Carina I) have an angular momentum separation $|\Delta\mathbf{L}| < |\Delta\mathbf{L}_{\text{MC}}|$, which establishes a probable association. Consequently, Horologium I and Carina I are removed from the sample. A third galaxy, Phoenix II, has been recently identified as a possible satellite of the LMC⁵⁰, and indeed appears as the third galaxy in the ranking with $|\Delta\mathbf{L}| \approx |\Delta\mathbf{L}_{\text{MC}}|$. Accordingly, Phoenix II is also excluded from the analysis. Following suggestions from the literature of satellites that may be strongly affected by the LMC, we also removed Carina II, Carina III, Hydrus I, Hydra II, Draco II, Reticulum II, Sculptor, Tucana III and Segue 1^{5,53,54}.

Our final datasets consist of $N_{\text{giants}} = 543$, $N_{\text{BHB}} = 292$ and $N_{\text{satellites}} = 33$. Other stellar tracers, such as RR Lyrae stars, do not have accurate 6-dimensional information, including those with available radial velocities^{14–16,55}. Distances for RR Lyrae stars are estimated using an absolute magnitude calibration tuned for Gaia⁵⁶. Another common halo tracer, the MW globular cluster system, has 6-dimensional information available⁵⁷, but only a handful of clusters are currently located at $r > 40$ kpc (Supplementary Information). We collect these datasets in the hope that future data-taking projects will expand the usable 6-dimensional stellar sample with an all-sky coverage. In particular, an all-sky sample of RR Lyrae with spectroscopic follow-up would provide a particularly rich dataset, for which Gaia DR3 may return proper motions usable for our analysis beyond about 40 kpc in the near future.

Data availability

All data used in this study is publicly available. This work made use of data from the European Space Agency (ESA) mission Gaia (<https://www.cosmos.esa.int/gaia>), processed by the Gaia Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). The data that support the plots within this paper and other findings of this study are available from <https://github.com/michael-petersen/ReflexMotion> or from the corresponding author upon reasonable request.

Received: 4 June 2020; Accepted: 14 October 2020;

Published online: 23 November 2020

References

1. Besla, G. et al. Simulations of the Magellanic stream in a first infall scenario. *Astrophys. J. Lett.* **721**, L97–L101 (2010).
2. Boylan-Kolchin, M., Besla, G. & Hernquist, L. Dynamics of the Magellanic Clouds in a Lambda cold dark matter Universe. *Mon. Not. R. Astron. Soc.* **414**, 1560–1572 (2011).
3. Peñarrubia, J., Gómez, F. A., Besla, G., Erkal, D. & Ma, Y.-Z. A timing constraint on the (total) mass of the Large Magellanic Cloud. *Mon. Not. R. Astron. Soc.* **456**, L54–L58 (2016).
4. Shao, S., Cautun, M., Deason, A. J., Frenk, C. S. & Theuns, T. Evolution of LMC/M33-mass dwarf galaxies in the EAGLE simulation. *Mon. Not. R. Astron. Soc.* **479**, 284–296 (2018).
5. Erkal, D. et al. The total mass of the Large Magellanic Cloud from its perturbation on the Orphan stream. *Mon. Not. R. Astron. Soc.* **487**, 2685–2700 (2019).
6. Kallivayalil, N., van der Marel, R. P., Besla, G., Anderson, J. & Alcock, C. Third-epoch Magellanic Cloud Proper Motions. I. Hubble Space Telescope/WFC3 data and orbit implications. *Astrophys. J.* **764**, 161 (2013).
7. Gómez, F. A. et al. And yet it moves: the dangers of artificially fixing the Milky Way center of mass in the presence of a massive Large Magellanic Cloud. *Astrophys. J.* **802**, 128 (2015).
8. Garavito-Camargo, N. et al. Hunting for the dark matter wake Induced by the Large Magellanic Cloud. *Astrophys. J.* **884**, 51 (2019).

9. Petersen, M. S. & Peñarrubia, J. Reflex motion in the Milky Way stellar halo resulting from the Large Magellanic Cloud infall. *Mon. Not. R. Astron. Soc.* **494**, L11–L16 (2020).
10. Xue, X.-X. et al. Quantifying kinematic substructure in the Milky Way's Stellar Halo. *Astrophys. J.* **738**, 79 (2011).
11. Xue, X.-X. et al. The Radial Profile and Flattening of the Milky Way's Stellar Halo to 80 kpc from the SEGUE K-giant Survey. *Astrophys. J.* **809**, 144 (2015).
12. Gaia Collaboration. Gaia Data Release 2. Summary of the contents and survey properties. *Astron. Astrophys.* **616**, A1 (2018).
13. McConnachie, A. W. & Venn, K. A. Revised and new proper motions for confirmed and candidate Milky Way dwarf galaxies. *Astron. J.* **160**, 124 (2020).
14. Sesar, B. et al. Machine-learned identification of RR Lyrae stars from sparse, multi-band data: the PS1 sample. *Astron. J.* **153**, 204 (2017).
15. Holl, B. et al. Gaia Data Release 2. Summary of the variability processing and analysis results. *Astron. Astrophys.* **618**, A30 (2018).
16. Rimoldini, L. et al. Gaia Data Release 2. All-sky classification of high-amplitude pulsating stars. *Astron. Astrophys.* **625**, A97 (2019).
17. van der Marel, R. P. & Kallivayalil, N. Third-epoch Magellanic Cloud proper motions. II. The Large Magellanic Cloud rotation field in three dimensions. *Astrophys. J.* **781**, 121 (2014).
18. Erkal, D., Belokurov, V. & Parkin, D. L. Equilibrium models of the Milky Way mass are biased high by the LMC. *Mon. Not. R. Astron. Soc.* **498**, 5574–5580 (2020).
19. Deason, A. J. et al. The slight spin of the old stellar halo. *Mon. Not. R. Astron. Soc.* **470**, 1259–1273 (2017).
20. Bird, S. A. et al. Constraints on the assembly history of the Milky Way's smooth, diffuse stellar halo from the metallicity-dependent, radially-dominated velocity anisotropy profiles probed with K giants and BHB stars using LAMOST, SDSS/SEGUE, and Gaia. Preprint at <https://arxiv.org/abs/2005.05980> (2020).
21. Johnston, K. V., Law, D. R. & Majewski, S. R. A Two Micron All Sky Survey view of the Sagittarius dwarf galaxy. III. Constraints on the flattening of the galactic halo. *Astrophys. J.* **619**, 800–806 (2005).
22. Law, D. R. & Majewski, S. R. The Sagittarius dwarf galaxy: a model for evolution in a triaxial Milky Way halo. *Astrophys. J.* **714**, 229–254 (2010).
23. Bovy, J., Bahmanyar, A., Fritz, T. K. & Kallivayalil, N. The shape of the inner Milky Way halo from observations of the Pal 5 and GD-1 stellar streams. *Astrophys. J.* **833**, 31 (2016).
24. Fardal, M. A. et al. Connecting the Milky Way potential profile to the orbital time-scales and spatial structure of the Sagittarius stream. *Mon. Not. R. Astron. Soc.* **483**, 4724–4741 (2019).
25. Malhan, K. & Ibata, R. A. Constraining the Milky Way halo potential with the GD-1 stellar stream. *Mon. Not. R. Astron. Soc.* **486**, 2995–3005 (2019).
26. Deng, L.-C. et al. LAMOST Experiment for Galactic Understanding and Exploration (LEGUE)—The survey's science plan. *Res. Astron. Astrophys.* **12**, 735–754 (2012).
27. de Jong, R. S. et al. 4MOST: Project overview and information for the first call for proposals. *Messenger* **175**, 3–11 (2019).
28. Cirasuolo, M. et al. MOONS: the Multi-Object Optical and Near-infrared Spectrograph for the VLT. *Proc. SPIE* **9147**, 91470N (2014).
29. Hobson, M. P., Bridle, S. L. & Lahav, O. Combining cosmological data sets: hyperparameters and Bayesian evidence. *Mon. Not. R. Astron. Soc.* **335**, 377–388 (2002).
30. Gravity Collaboration. A geometric distance measurement to the Galactic center black hole with 0.3% uncertainty. *Astron. Astrophys.* **625**, L10 (2019).
31. Bennett, M. & Bovy, J. Vertical waves in the solar neighbourhood in Gaia DR2. *Mon. Not. R. Astron. Soc.* **482**, 1417–1425 (2019).
32. Schönrich, R., Binney, J. & Dehnen, W. Local kinematics and the local standard of rest. *Mon. Not. R. Astron. Soc.* **403**, 1829–1833 (2010).
33. McMillan, P. J. The mass distribution and gravitational potential of the Milky Way. *Mon. Not. R. Astron. Soc.* **465**, 76–94 (2017).
34. Drimmel, R. & Poggio, E. On the solar velocity. *Res. Not. Am. Astron. Soc.* **2**, 210 (2018).
35. Ma, Y.-Z., Hinshaw, G. & Scott, D. WMAP observations of Planck ESZ clusters. *Astrophys. J.* **771**, 137 (2013).
36. Skilling, J. Nested sampling. *AIP Conf. Proc.* **735**, 395 (2004).
37. Feroz, F. & Hobson, M. P. Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. *Mon. Not. R. Astron. Soc.* **384**, 449–463 (2008).
38. Feroz, F., Hobson, M. P. & Bridges, M. MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Not. R. Astron. Soc.* **398**, 1601–1614 (2009).
39. Xue, X. X. et al. The Milky Way's circular velocity curve to 60 kpc and an estimate of the dark matter halo mass from the kinematics of ~2400 SDSS blue horizontal-branch stars. *Astrophys. J.* **684**, 1143–1158 (2008).
40. Xue, X.-X. et al. The SEGUE K giant survey. II. A catalog of distance determinations for the SEGUE K giants in the galactic halo. *Astrophys. J.* **784**, 170 (2014).
41. Yanny, B. et al. SEGUE: a spectroscopic survey of 240,000 stars with $g = 14\text{--}20$. *Astron. J.* **137**, 4377–4399 (2009).
42. Lancaster, L., Koposov, S. E., Belokurov, V., Evans, N. W. & Deason, A. J. The halo's ancient metal-rich progenitor revealed with BHB stars. *Mon. Not. R. Astron. Soc.* **486**, 378–389 (2019).
43. Gaia Collaboration. The Gaia mission. *Astron. Astrophys.* **595**, A1 (2016).
44. Lindegren, L. et al. Gaia Data Release 2. The astrometric solution. *Astron. Astrophys.* **616**, A2 (2018).
45. Niederste-Ostholt, M., Belokurov, V., Evans, N. W. & Peñarrubia, J. Re-assembling the Sagittarius dwarf galaxy. *Astrophys. J.* **712**, 516–526 (2010).
46. Majewski, S. R., Skrutskie, M. F., Weinberg, M. D. & Ostheimer, J. C. A Two Micron All Sky Survey view of the Sagittarius dwarf galaxy. I. Morphology of the Sagittarius core and tidal arms. *Astrophys. J.* **599**, 1082–1115 (2003).
47. Fritz, T. K. et al. Gaia DR2 proper motions of dwarf galaxies within 420 kpc. Orbita, Milky Way mass, tidal influences, planar alignments, and group infall. *Astron. Astrophys.* **619**, A103 (2018).
48. Riley, A. H. et al. The velocity anisotropy of the Milky Way satellite system. *Mon. Not. R. Astron. Soc.* **486**, 2679–2694 (2019).
49. Pace, A. B. & Li, T. S. Proper motions of Milky Way ultra-faint satellites with Gaia DR2 x DES DR1. *Astrophys. J.* **875**, 77 (2019).
50. Fritz, T. K., Carrera, R., Battaglia, G. & Taibi, S. Gaia DR 2 and VLT/FLAMES search for new satellites of the LMC. *Astron. Astrophys.* **623**, A129 (2019).
51. Torrealba, G. et al. The hidden giant: discovery of an enormous Galactic dwarf satellite in Gaia DR2. *Mon. Not. R. Astron. Soc.* **488**, 2743–2766 (2019).
52. Simon, J. D. et al. Birds of a feather? Magellan/IMACS spectroscopy of the ultra-faint satellites Grus II, Tucana IV, and Tucana V. *Astrophys. J.* **892**, 137 (2020).
53. Kallivayalil, N. et al. The missing satellites of the Magellanic clouds? Gaia proper motions of the recently discovered ultra-faint galaxies. *Astrophys. J.* **867**, 19 (2018).
54. Patel, E. et al. The orbital histories of Magellanic satellites using Gaia DR2 proper motions. *Astrophys. J.* **893**, 121 (2020).
55. Cohen, J. G. et al. The outer halo of the Milky Way as probed by RR Lyr variables from the Palomar Transient Facility. *Astrophys. J.* **849**, 150 (2017).
56. Iorio, G. & Belokurov, V. The shape of the Galactic halo with Gaia DR2 RR Lyr. Anatomy of an ancient major merger. *Mon. Not. R. Astron. Soc.* **482**, 3868–3879 (2019).
57. Vasiliev, E. Proper motions and dynamics of the Milky Way globular cluster system from Gaia DR2. *Mon. Not. R. Astron. Soc.* **484**, 2832–2850 (2019).

Acknowledgements

M.S.P. acknowledges funding from the UK Science and Technology Facilities Council (STFC) Consolidated Grant and support from M. Weinberg for use of the EXP code. Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

Author contributions

Both authors assisted in the interpretation of the results and writing of the paper.

Competing interests

The authors declare no competing financial interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41550-020-01254-3>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41550-020-01254-3>.

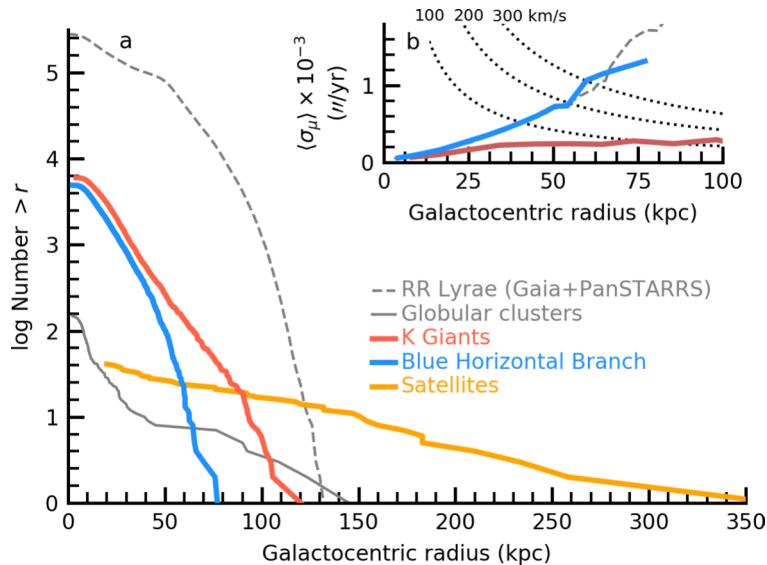
Correspondence and requests for materials should be addressed to M.S.P.

Peer review information *Nature Astronomy* thanks the anonymous reviewers for their contribution to the peer review of this work.

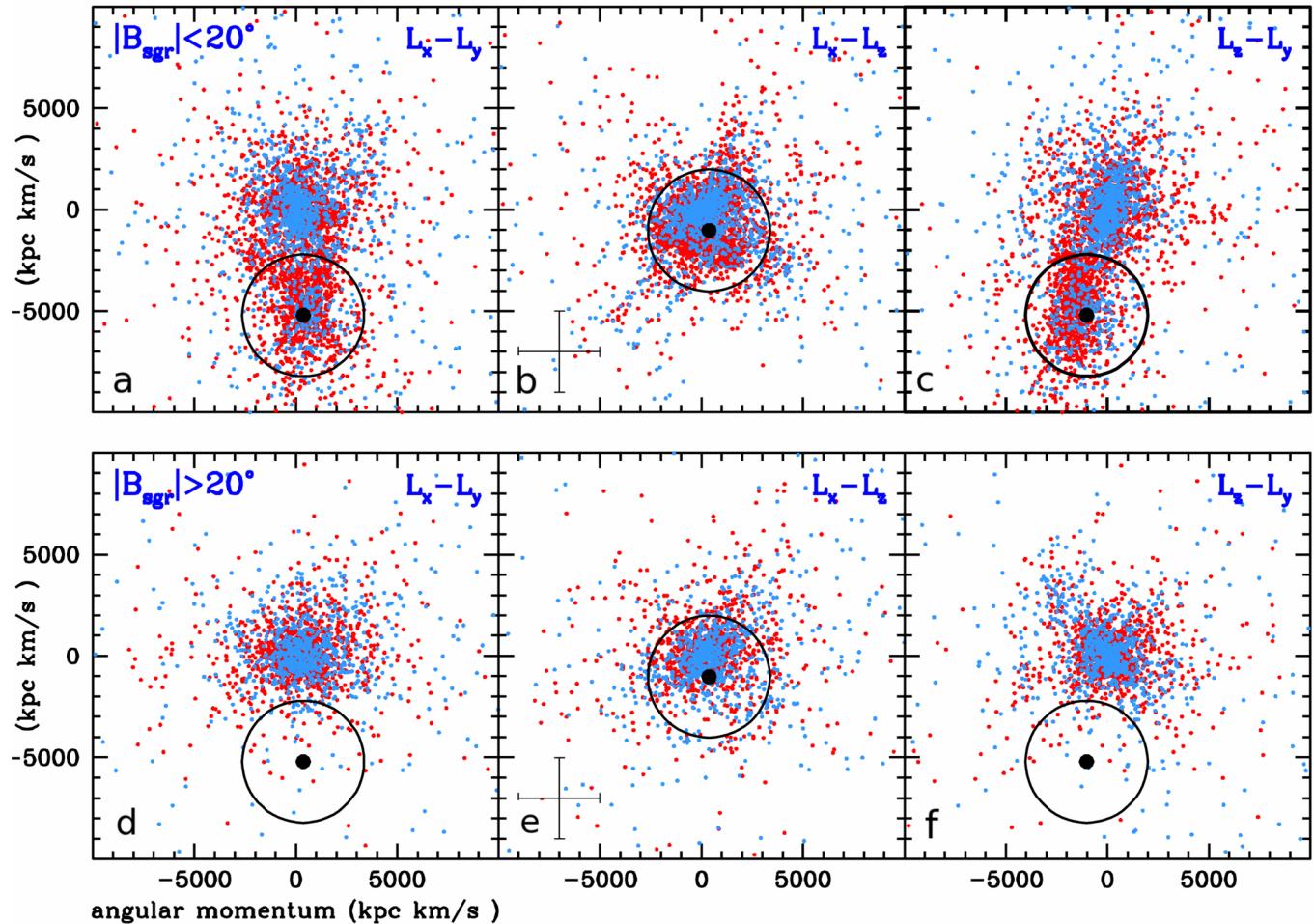
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

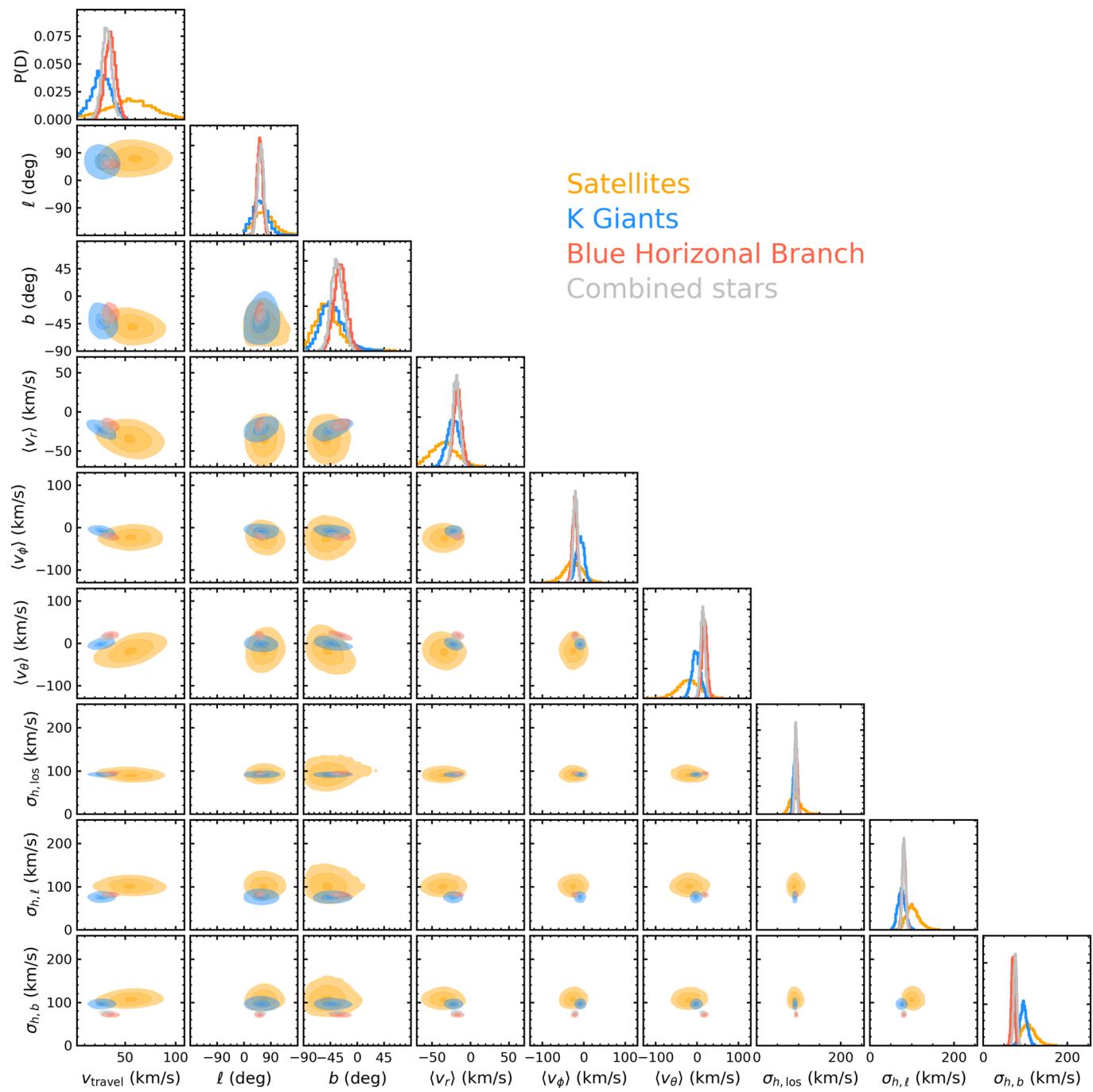
© The Author(s), under exclusive licence to Springer Nature Limited 2020



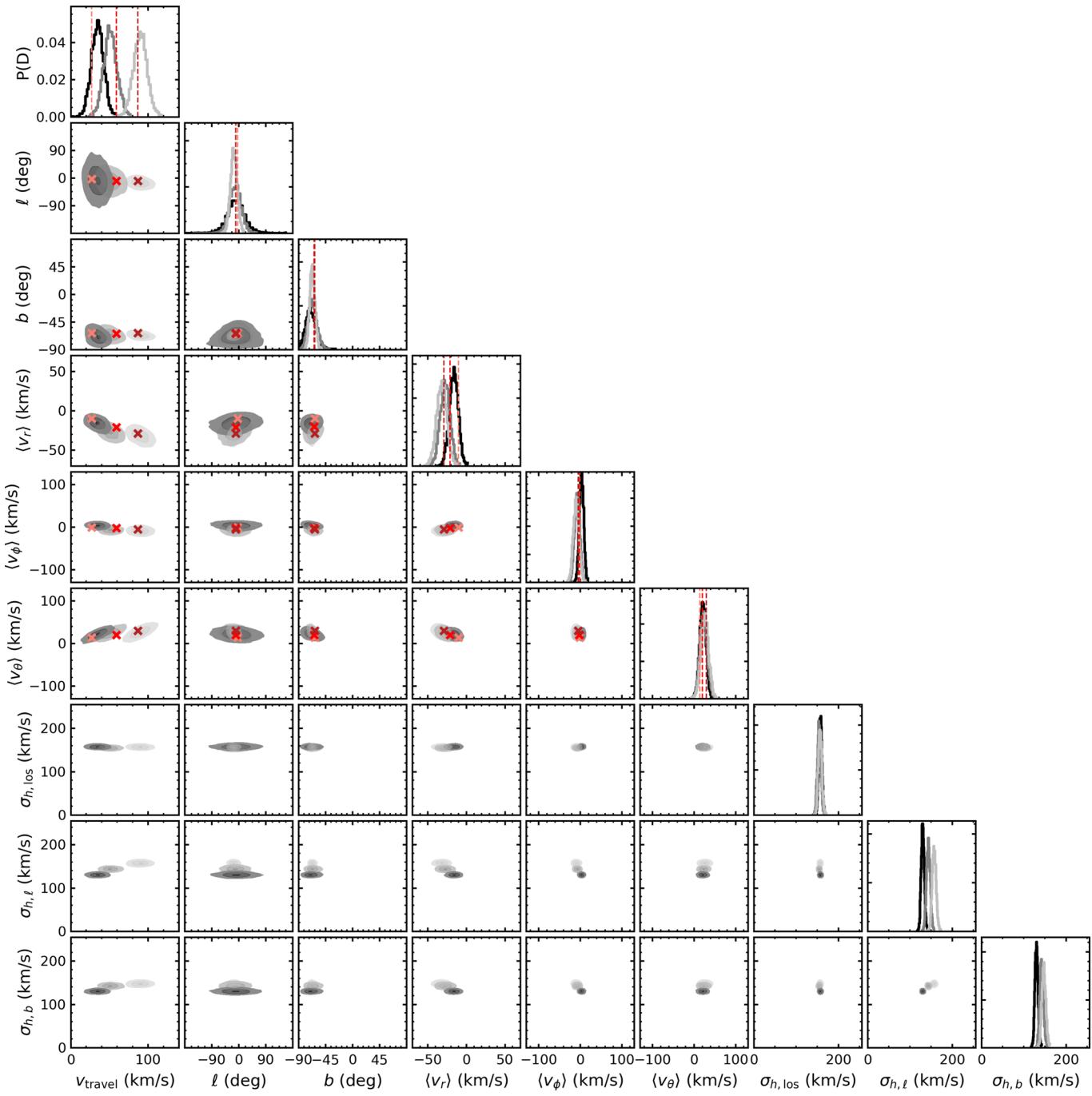
Extended Data Fig. 1 | Tracer population properties as a function of distance. **a**, Logarithm of number of sources larger than some Galactocentric radius, r , as a function of Galactocentric radius r in kiloparsecs. We show the three datasets analysed in this paper in colour (red: K Giants, blue: Blue Horizontal Branch, orange: satellites). We show two explored, but unused, datasets in grey (solid grey: globular clusters, dashed gray: RR Lyrae stars). **b**, The mean proper motion error in milliarcseconds per year for the stellar sources (K Giants, Blue Horizontal Branch, RR Lyrae), as a function of Galactocentric radius. As the brightest sample, K Giants have the smallest uncertainty. We mark 100, 200, and 300 km s^{-1} , the approximate range of halo velocities, as a dotted black curves.



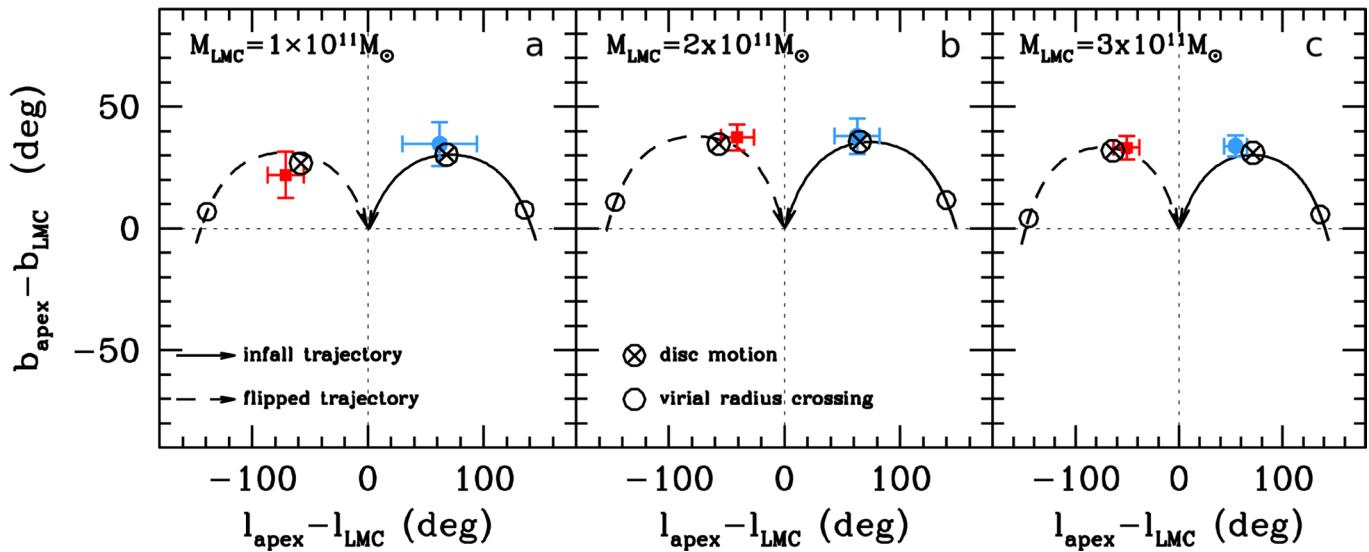
Extended Data Fig. 2 | Angular momentum components of halo stars at Galactocentric distances $r > 20$ kpc. Blue and red dots denote BHB and K giant stars, respectively. The angular momentum of the Sagittarius dwarf is marked with a black dot at \mathbf{L}_{sgr} (in units of kpc km s^{-1}) = (605, -4, 515, -1, 267). For reference, circles mark angular momentum difference $|\mathbf{L} - \mathbf{L}_{\text{sgr}}| = 3,000$ kpc km s^{-1} . Panels **a–c** and panels **d–f** correspond to stars at $|B_{\text{sgr}}| < 20^\circ$ and $|B_{\text{sgr}}| > 20^\circ$ off the orbital plane of the Sagittarius stream, respectively. We show the mean 1σ error bar at $r > 40$ kpc in the lower left of the centre column.



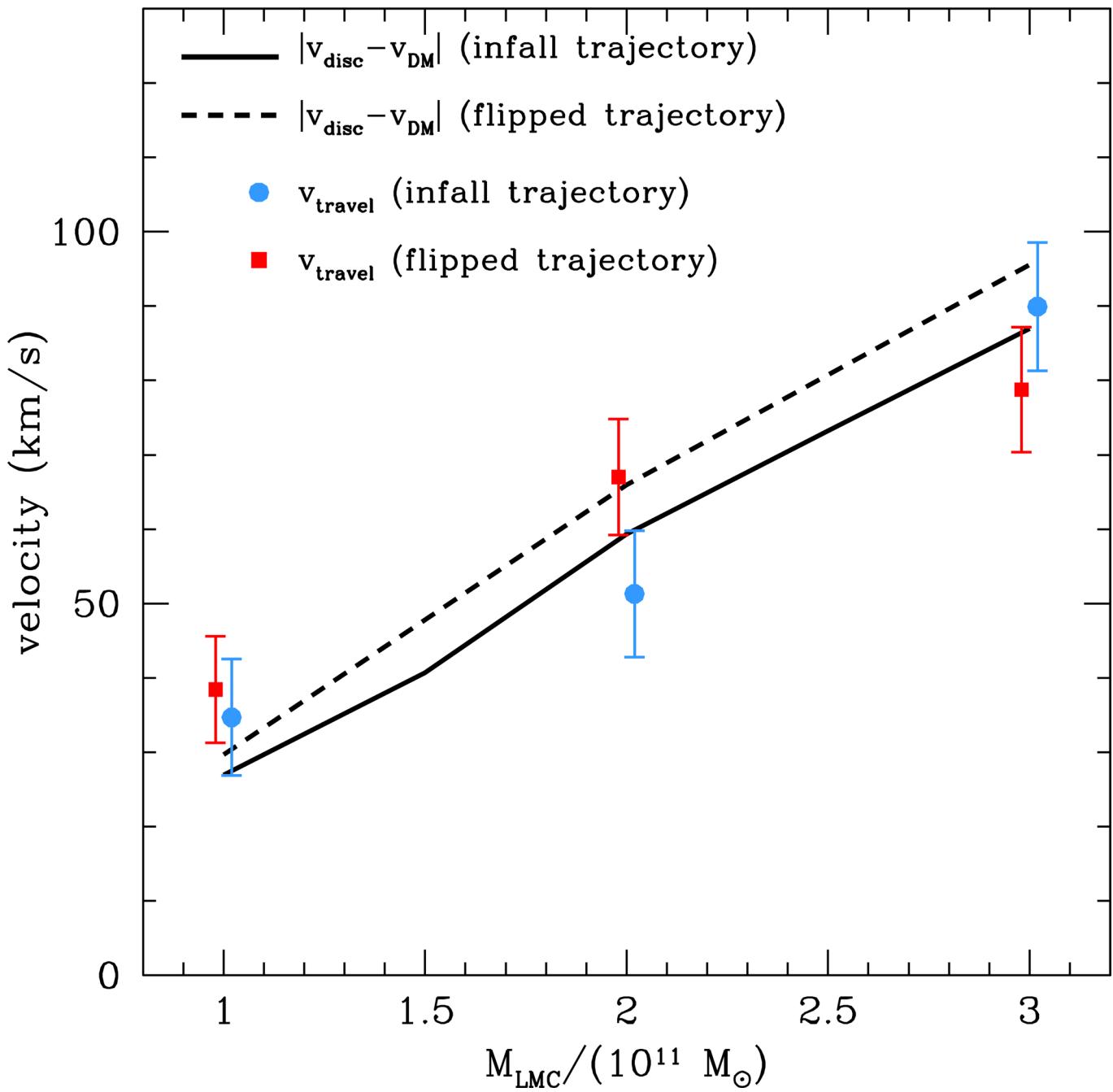
Extended Data Fig. 3 | Corner plot of covariance in the nine fitting parameters for the Blue Horizontal Branch (blue), K Giant (red), satellite (orange) and combined K Giant+Blue Horizontal Branch samples (silver). The fit parameters are disk velocity, v_{travel} ; Galactocentric longitude apex ℓ_{apex} ; Galactocentric latitude apex b_{apex} ; the mean velocity of halo stars in spherical coordinates $\langle v_r \rangle$, $\langle v_\phi \rangle$ and $\langle v_\theta \rangle$; the three hyperparameters $\sigma_{h,\text{los}}$, $\sigma_{h,\ell}$ and $\sigma_{h,b}$. A full description of the model may be found in the Supplementary Information.



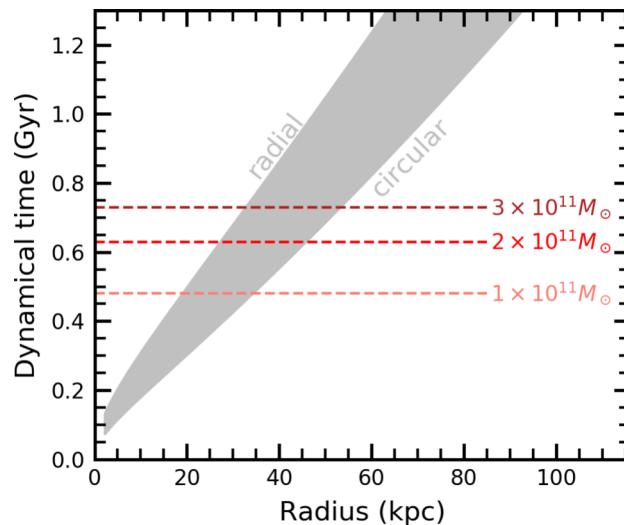
Extended Data Fig. 4 | Corner plot for mock datasets of K Giant stars at $r > 40$ kpc drawn from live n -body simulations of the MW where the LMC falls in with a mass $M_{\text{LMC}}/(10^{11}M_{\odot}) = 1, 2$ and 3 . Posteriors from the fits for the three models are shown in greyscale from dark to light to indicate increasing mass: $1 \times 10^{11}M_{\odot}$ (black), $2 \times 10^{11}M_{\odot}$ (dark grey), $3 \times 10^{11}M_{\odot}$ (light grey). See Supplementary Information for mock data details. All parameters are well constrained and show a relatively minor covariance. Note that the bounds on the apex direction and the magnitude of the reflex motion improve in proportion to the LMC mass. Values derived directly from the simulation are shown as dashed vertical lines (on histograms), or coloured 'x' markers on contour plots: light red ($1 \times 10^{11}M_{\odot}$), red ($2 \times 10^{11}M_{\odot}$), dark red ($3 \times 10^{11}M_{\odot}$).



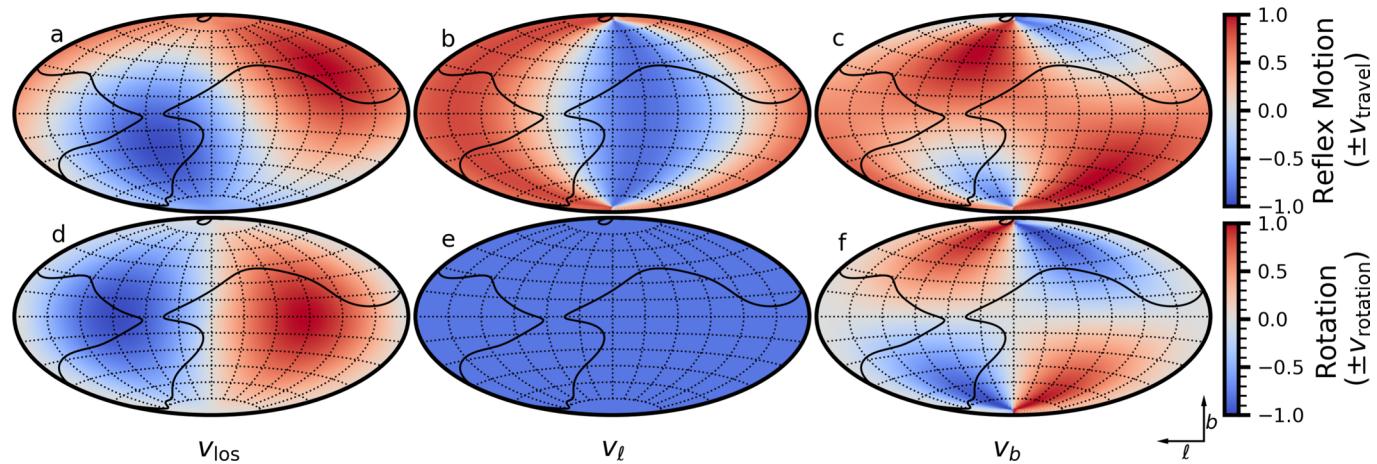
Extended Data Fig. 5 | Apex direction measured from mock samples of K Giant stars located at $r > 40$ kpc in the SDSS footprint. We use live Galaxy models that experience the infall of LMC-like galaxies with masses $M_{\text{LMC}}/10^{11} M_{\odot} = 1, 2$ and 3 (panels **a**, **b** and **c** respectively). Solid lines denote the infall trajectory of the LMC derived from backwards orbit integration of HST proper motions, while dashed lines show models where the LMC trajectory has been flipped. Red (blue) symbols denote the measurement for mock samples with an LMC-like (flipped) trajectory. Uncertainties are the standard deviation derived from the posteriors of the fit locations. Open circles show the point on the trajectory where the LMC crossed the virial radius of the Galaxy. Note that within statistical uncertainties the apex direction derived from the kinematics of distant stellar halo particles points towards the direction where the MW disk is currently moving (crossed circles). The disk component is currently travelling to an earlier point on the LMC trajectory.



Extended Data Fig. 6 | Velocity of the Galactic disk (v_{travel}) inferred from the kinematics of stellar tracers located at $r > 40$ kpc as a function of LMC mass. The solid (dashed) lines show the true speed of the disk barycentre relative to dark matter particles within a radial range $40 < r < 150$ on the infall (flipped) trajectories in the mock LMC models, where radius r is in kiloparsecs. The blue (red) symbols show the value of v_{travel} measured from the mock models in the infall (flipped) trajectory case.



Extended Data Fig. 7 | Dynamical time of particles in the model MW-LMC systems as a function of Galactocentric radius. For ease of reference, horizontal dashed lines show the lookback time for the onset of disk motion. Halo particles with a dynamical time above the horizontal lines react impulsively to the LMC infall. At a fixed radius, radial and circular orbits provide the shortest and longest periods, respectively. As a function of radius, a star's period will fall in the grey-shaded region.



Extended Data Fig. 8 | Comparison of the appearance of reflex motion versus rotation. **a–c**, Our reflex motion model, projected onto the (ℓ, b) plane in $v_{\text{los}}, v_\ell, v_b$. Velocities in each panel have been normalized to units of v_{travel} . **d–f**, A rotation model, with magnitude v_ϕ , projected onto the (ℓ, b) plane in $v_{\text{los}}, v_\ell, v_b$, where velocities in each panel have been normalized to units of v_{rotation} . The smoothed SEGUE footprint is outlined in black. Comparing the upper and lower rows shows that the line-of-sight and b velocities are similar, supporting the need for 6-dimensional data and all-sky coverage.