Universiteit
Leiden
The Netherlands

The Milky Way's plane of satellites is consistent with LambdaCDM

Sawala, T.; Cautun, M.; Frenk, C.; Helly, J.; Jasche, J.; Jenkins, A.; ... ; Schaller, M.

Citation
Sawala, T., Cautun, M., Frenk, C., Helly, J., Jasche, J., Jenkins, A., ... Schaller, M. (2022). The Milky Way's plane of satellites is consistent with LambdaCDM. *Nature Astronomy*. doi:10.1038/s41550-022-01856-z

Version:	Publisher's Version
License:	Creative Commons CC BY 4.0 license
Downloaded from:	https://hdl.handle.net/1887/3515389

Note: To cite this publication please use the final published version (if applicable).

nature astronomy

Article

https://doi.org/10.1038/s41550-022-01856-z

The Milky Way's plane of satellites is consistent with ΛCDM

Received: 5 May 2022

Accepted: 3 November 2022

Published online: 19 December 2022

Check for updates

Till Sawala[1,2] ✉, Marius Cautun[3], Carlos Frenk[2], John Helly[2], Jens Jasche[4], Adrian Jenkins[2], Peter H. Johansson[1], Guilhem Lavaux[5], Stuart McAlpine[1,4] & Matthieu Schaller[3,6]

The Milky Way is surrounded by 11 'classical' satellite galaxies in a remarkable configuration: a thin plane that is possibly rotationally supported. Such a structure is thought to be highly unlikely to arise in the standard (ΛCDM) cosmological model (Λ cold dark matter model, where Λ is the cosmological constant). While other apparent discrepancies between predictions and observations of Milky Way satellite galaxies may be explained either through baryonic effects or by invoking alternative forms of dark matter particles, there is no known mechanism for making rotating satellite planes within the dispersion-supported dark matter haloes predicted to surround galaxies such as the Milky Way. This is the so-called 'plane of satellites problem', which challenges not only the ΛCDM model but the entire concept of dark matter. Here we show that the reportedly exceptional anisotropy of the Milky Way satellites is explained, in large part, by their lopsided radial distribution combined with the temporary conjunction of the two most distant satellites, Leo I and Leo II. Using Gaia proper motions, we show that the orbital pole alignment is much more common than previously reported, and reveal the plane of satellites to be transient rather than rotationally supported. Comparing with new simulations, where such short-lived planes are common, we find the Milky Way satellites to be compatible with standard model expectations.

The original discovery of the Milky Way's 'plane of satellites' (then just five galaxies)[1] preceded the advent of the Λ cold dark matter (ΛCDM) model, where Λ is the cosmological constant, as the paradigm for galaxy and structure formation[2]. A key ΛCDM prediction is that galaxies such as the Milky Way (MW) are surrounded by a dispersion-supported dark matter halo and by satellite galaxies formed within its substructures. However, while several ΛCDM predictions, including the discovery of dozens of additional MW satellites, have since been borne out[3], the 'plane of satellites problem'[4,5] has emerged as its most persistent challenge[6–11].

That the 'plane of satellites' problem has so far eluded resolution is not for lack of trying. Planes of satellites form with the same (low) frequency in collisionless and hydrodynamic cosmological simulations[12–14] and in MW analogues in isolation or in pairs[15,16], with no significant correlation with other properties of the host halo[16]. There is evidence that filamentary accretion[17,18], a compact satellite system[19] or the presence of massive satellites[20] can generate some anisotropy, but systems as thin as the Milky Way's are still very rare[12]. Moreover, any planes that do form in ΛCDM are transient, chance alignments of substructures[14,20–22] rather than long-lived, rotationally supported disks.

[1]Department of Physics, University of Helsinki, Helsinki, Finland. [2]Institute for Computational Cosmology, Durham University, Durham, UK. [3]Leiden Observatory, Leiden University, Leiden, the Netherlands. [4]The Oskar Klein Centre, Department of Physics, Stockholm University, Stockholm, Sweden. [5]Institut d'Astrophysique de Paris, CNRS & Sorbonne Université, Paris, France. [6]Lorentz Institute for Theoretical Physics, Leiden University, Leiden, the Netherlands. ✉e-mail: till.sawala@helsinki.fi

https://doi.org/10.1038/s41550-022-01856-z

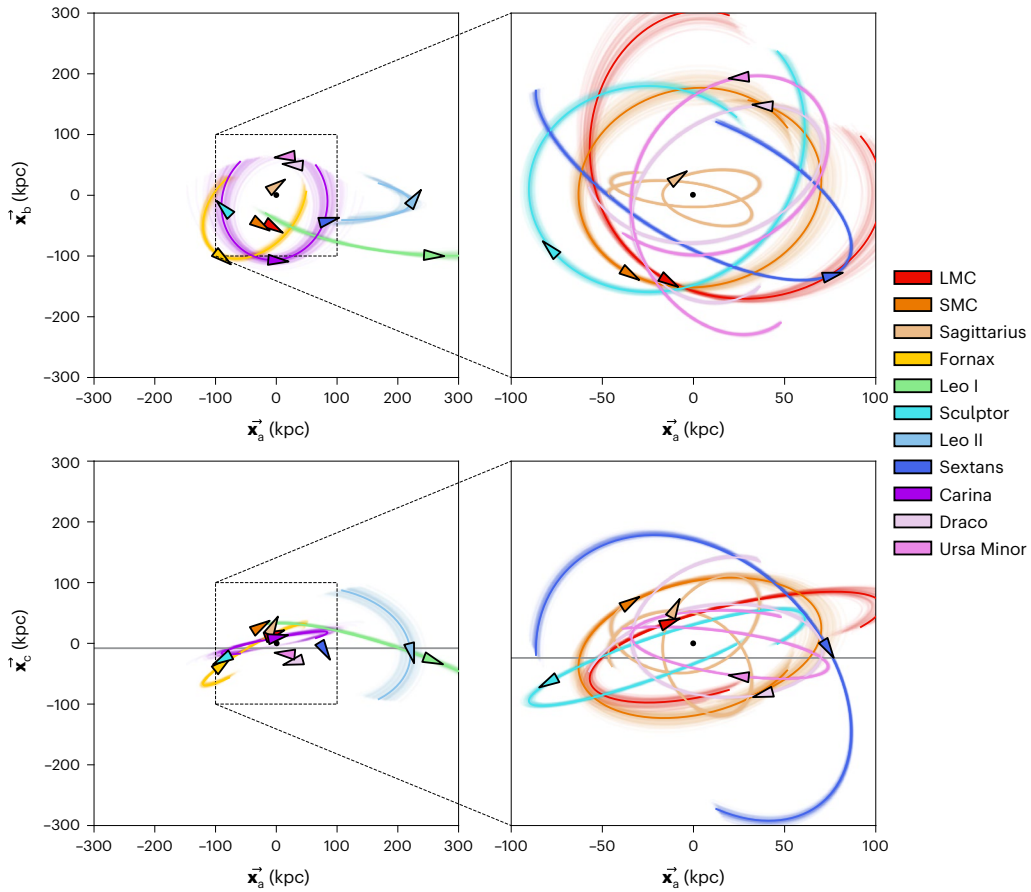


Fig. 1 | Observed positions and orbits of the Milky Way's classical satellites. Arrowheads show maximum likelihood positions, projected face-on (top panels) and edge-on (bottom panels) according to the eigenvectors of the full inertia tensor. Lines show orbits integrated for 1 Gyr into the past and future in a fixed halo potential of mass $10^{12}\,M_\odot$. Faint lines show 200 Monte Carlo samples of the observations and bold lines show the maximum likelihood orbit. The left panels show the orbits of the 4 satellites with Galactocentric distances beyond 100 kpc and the right panels show the remaining 7 orbits in an inset of ±100 kpc around the Galactic Centre. With the Gaia EDR3 measurements, the proper motions are very well constrained, with the exception of the LMC and SMC. It can also be seen that the MW satellites are highly concentrated, with 7 out of 11 within 100 kpc and only 2, Leo I and Leo II, at $r > 200$ kpc. Several galaxies, including Leo I and II, are presently crossing the 'plane' (indicated by the grey horizontal lines in the bottom panels), which soon disperses as a result.

We examine here the contention that the MW contains an exceptional plane of satellites, explain the origin of the observed anisotropy and study its time evolution in light of proper motion measurements by the Gaia space telescope[23].

The present MW plane of satellites

The Milky Way's 'plane of satellites' canonically consists of the 11 'classical' satellites, the brightest within a radius of $r = 300$ kpc of the Galactic Centre, believed to constitute a complete sample. To characterize the spatial anisotropy of a system of N satellites, it is customary to consider the inertia tensor, defined as

$$I_{ij} = \sum_{n=1}^{N} x_{n,i} x_{n,j}, \qquad (1)$$

where x_n are the coordinates of the nth satellite relative to the centre of positions, and i and j index the three spatial dimensions. We label the square roots of its eigenvalues as a, b and c, corresponding to the dispersions in position along the unit eigenvectors $\mathbf{x}_a$, $\mathbf{x}_b$ and $\mathbf{x}_c$. A related metric is the 'reduced' inertia tensor[24], defined after projection of the positions onto a unit sphere. We label the square roots of its eigenvalues as $a_{\rm red}$, $b_{\rm red}$ and $c_{\rm red}$. Both c/a and $(c/a)_{\rm red} \equiv c_{\rm red}/a_{\rm red}$ parametrize the spatial anisotropy but differ in the weight attached to each galaxy. c/a measures the full spatial anisotropy, whereas $(c/a)_{\rm red}$ considers only its angular component, independently of the radial distribution. For both c/a and $(c/a)_{\rm red}$, smaller values imply greater anisotropy; for example, a sphere has $c/a = 1$ while a perfectly thin disk has $c/a = 0$. Note that for small N, the expectation values of c/a and $(c/a)_{\rm red}$ decrease, regardless of the underlying anisotropy[25].

Figure 1 shows the present positions and estimated orbits of the 11 brightest MW satellites projected along the principal axes. We measure $c/a = 0.183 \pm 0.004$ and $(c/a)_{\rm red} = 0.3676 \pm 0.0004$, mean ± s.d.

The effect of the radial distribution

Earlier work[26] has found that only 0.7% of ΛCDM simulations produce systems as anisotropic as the MW. However, we find this likely to be a severe underestimate caused by the artificial disruption of satellites in numerical simulations, which can cause artificially extended radial profiles[27–30]. To identify analogues of the classical MW satellites in simulations, it is common to select at $z = 0$ the 11 satellites with the highest value of $v_{\rm peak}$, the peak value of the maximum circular velocity, $v_{\rm max}$, across a halo's history[17]. In our own analysis of 202 MW analogues resimulated in the SIBELIUS constrained simulations intended to reproduce the structures observed in the Local Universe[31], including only surviving subhaloes, we find only 2 (1%) systems with c/a as low as the MW, and none that reproduce the radial distribution. However, this is an 'incomplete' sample. Accounting for artificial disruption[32] and selecting the 11 satellites among a 'complete' sample containing both the

https://doi.org/10.1038/s41550-022-01856-z

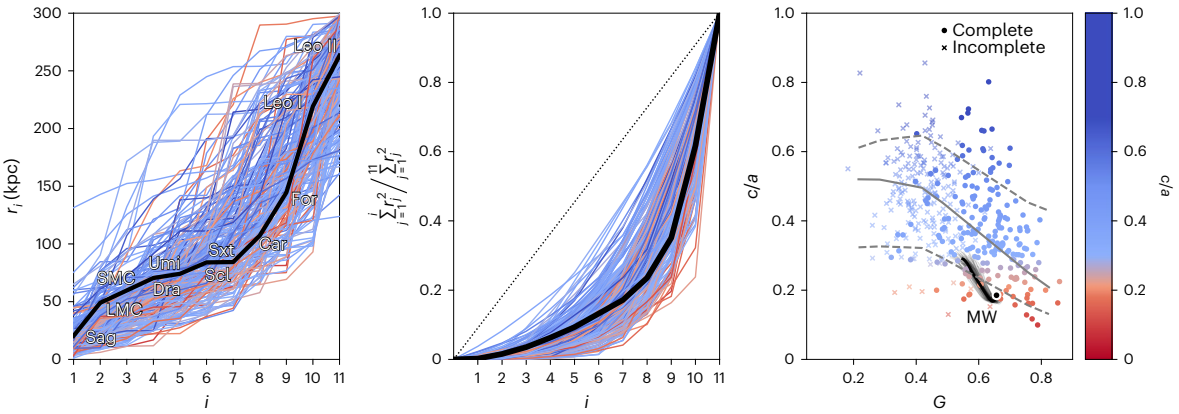


Fig. 2 | Comparison of the radial distribution and anisotropy of the classical MW satellites with those of simulated ΛCDM counterparts. On the first two panels, the thick black lines show data for the MW satellites and the thin lines show data from the simulations. Sag: Sagittarius; Umi: Ursa Minor; Dra: Draco; Sxt: Sextans; Scl: Sculptor; Car: Carina; For: Fornax. On all panels, lines and symbols for the simulations are coloured by c/a as shown on the right to identify highly anisotropic (red) and more isotropic (blue) systems. The left panel shows the radius, r_i, of the ith closest satellite. The centre panel shows the sum of the squares of the radii of the closest i satellites, normalized by the sum of all 11 satellites. The square of the radius determines each satellite's contribution to the inertia tensor; the Gini coefficient of inertia, G, quantifies the inequality of these contributions and corresponds to the distance of each line from the dotted line. The right panel shows the correlation between the Gini coefficient and anisotropy, c/a, for 'complete' sets of satellites (circles), corrected for artificial disruption, coloured by c/a. The solid line shows the estimated median and the dashed lines the 10th and 90th percentiles. The black circle denotes the Milky Way's present values of G and c/a and the lines around it show its most likely (bold) and Monte Carlo sampled (thin) evolution over the past 0.5 Gyr. The anisotropy correlates with the Gini coefficient. Without accounting for artificial disruption of satellites in ΛCDM simulations ('incomplete', faint crosses), the Milky Way's satellites are much more centrally concentrated and much more anisotropic. However, when this is taken into account, the MW lies within the scatter.

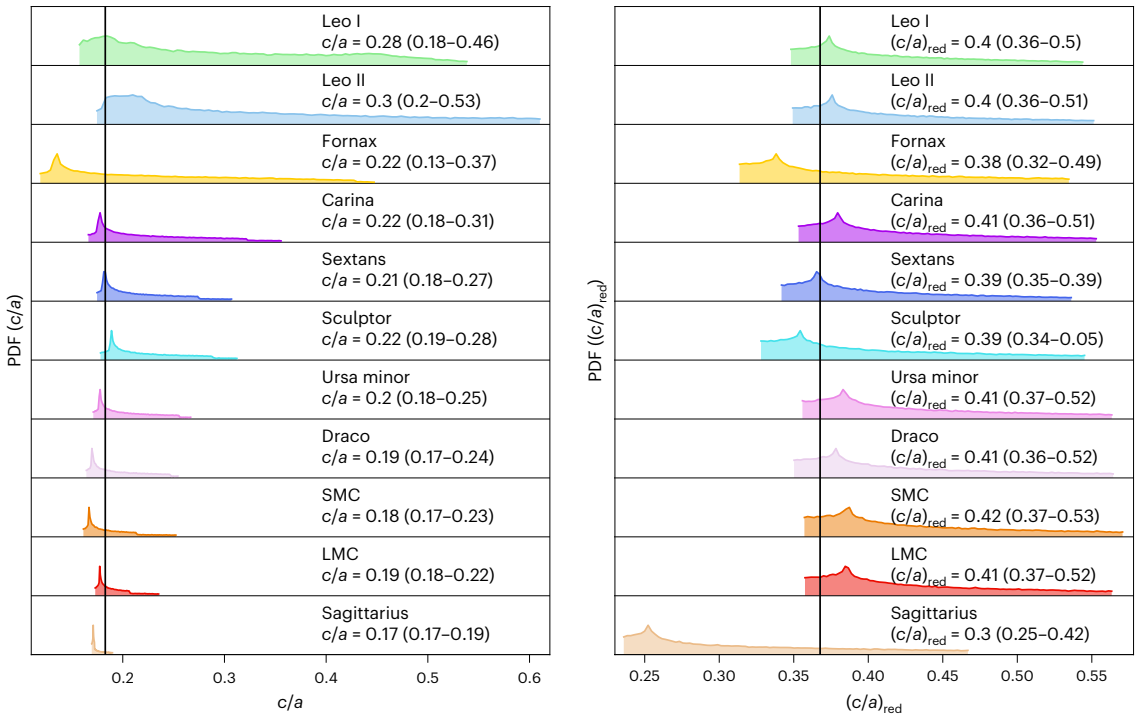


Fig. 3 | Sensitivity of the anisotropy to the positions of single satellites. Probability density functions (PDFs) of c/a (left) and $(c/a)_{red}$ (right) for the 11 brightest satellites when the angular coordinates of each galaxy are randomized in turn, with the distance set to the observed value and the coordinates of all other galaxies kept fixed. Numbers show the median values of c/a and $(c/a)_{red}$; those in parentheses show the 10th and 90th percentiles. Galaxies are sorted from top to bottom in decreasing order of radius. Black horizontal lines indicate the vertical offset and the black vertical lines show the values with all 11 galaxies at their observed positions. Each galaxy impacts the distribution of c/a differently, and the range of possible c/a values correlates with the radius of the satellite. Just placing either Leo I or Leo II at different angular coordinates at their respective radius could result in a completely different anisotropy, including cases that are more isotropic than the majority of ΛCDM systems.

https://doi.org/10.1038/s41550-022-01856-z

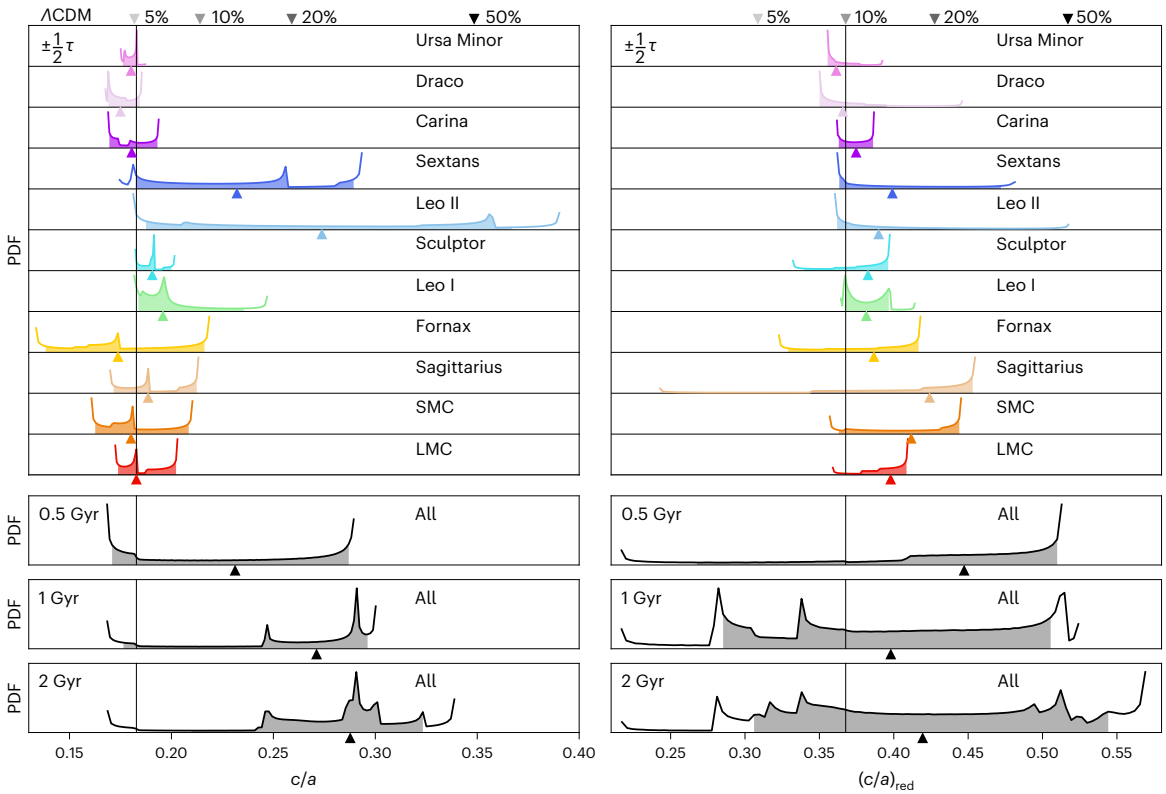


Fig. 4 | Time-averaged distributions of the anisotropy of the Milky Way satellite system. The top panel (coloured plots) shows time-averaged PDFs for the anisotropy, c/a (left) and $(c/a)_{red}$ (right), when the position of each satellite is moved along its most likely orbit for one orbital period, τ, while the other 10 satellites remain fixed at their present positions. The bottom panels (greyscale plots) show PDFs over lookback times of 0.5, 1 and 2 Gyr, evolving all orbits simultaneously. Triangles below each graph indicate the time-averaged medians and filled areas extend from the 10th to 90th percentiles. Black vertical lines show the anisotropy of the Milky Way's satellites with all 11 galaxies at their present positions, $c/a = 0.183$ and $(c/a)_{red} = 0.364$. Downward triangles at the top indicate percentiles in the ΛCDM simulations. Given the orbits of the satellites, over the past 1 Gyr c/a could have been as low as 0.17 or as high as 0.31 and $(c/a)_{red}$ could have been as low as 0.22 or as high as 0.52. The present value of c/a, contingent on the close but fleeting proximity of Leo I and Leo II, is significantly below the time-averaged median, even for a short period of 0.5 Gyr. Note that the sharp limits and features are not numerical artefacts but reflect the orbital velocity gradients and the locations of satellites relative to each other.

surviving and the recovered, artificially disrupted satellites, we find radial distributions resembling the MW's, as shown in the left panel of Fig. 2. This reproduces the results of very-high-resolution simulations[30,33]. As each satellite contributes to the inertia (equation (1)) in proportion to r_i^2, c/a is sensitive to the radial profile. To quantify this relationship, we introduce the Gini coefficient formalism. The central panel of Fig. 2 shows the summed weights of the closest i satellites from the centre, $\sum_{j=1}^{i} r_j^2$, normalized by the total weight of all 11 satellites, $\sum_{j=1}^{11} r_j^2$. The area between each curve and the diagonal measures the inequality of the satellites' contributions to the inertia, or the sample Gini coefficient of inertia,

$$G = \frac{1}{N-1}\sum_{i=1}^{N}(2i - N - 1)r_i^2 / \sum_{i=1}^{N} r_i^2. \quad (2)$$

Compared to a more equal distribution, the Milky Way's centrally concentrated radial profile is equivalent to sampling a system with fewer points, lowering the expectation value of c/a.

The right panel of Fig. 2 shows the relationship between G and c/a. Systems with higher central concentration (higher G) tend to be more anisotropic (lower c/a). Accounting for artificial disruption (filled circles), 58% of ΛCDM systems have G above the MW and 11 (~5%) have $c/a < 0.183$. Neglecting this effect (faint crosses) produces no systems with G as high the MW and only two (1%) with as low c/a.

The Milky Way's anisotropy additionally results from the fact that its two outermost satellites, Leo I and Leo II, which contribute two thirds of the total inertia, are currently in close proximity. However, as is already apparent from Fig. 1, this constellation is short-lived.

The clustering of orbital poles

Supporting the notion that the satellite plane constitutes a spinning disk[34], the orbital poles of 7 of the 11 classical satellites (the Large Magellanic Cloud (LMC), the Small Magellanic Cloud (SMC), Fornax, Leo II, Carina, Draco and Ursa Minor) are reported to cluster with a standard deviation in direction of only 16.0° (ref. [26]), found in only 0.04% of ΛCDM systems. Using the more precise proper motions from Gaia Early Data Release 3 (EDR3) for the same satellites, we find that this angle increases to $23.2^{\circ}\,^{+3.5}_{-2.8}$. We also repeated the analysis, and find that a different subset (that includes Leo I instead of Leo II) yields a smaller dispersion of $18.9^{\circ}\,^{+1.9}_{-1.4}$. Among our sample of 202 simulated systems, adopting either 18.9° or 23.2° and accounting for a minimum look-elsewhere effect, we find three or five systems with subsets of satellites with a smaller probability to arise from isotropic distributions. That is, we find that ~2% of ΛCDM systems contain satellites whose orbital poles are even more anisotropic than the most clustered subset of the Milky Way, an ~50-fold increase over previous results. The orbital clustering of a subset of the Milky Way satellites is unusual in ΛCDM, but not astronomically improbable.

https://doi.org/10.1038/s41550-022-01856-z

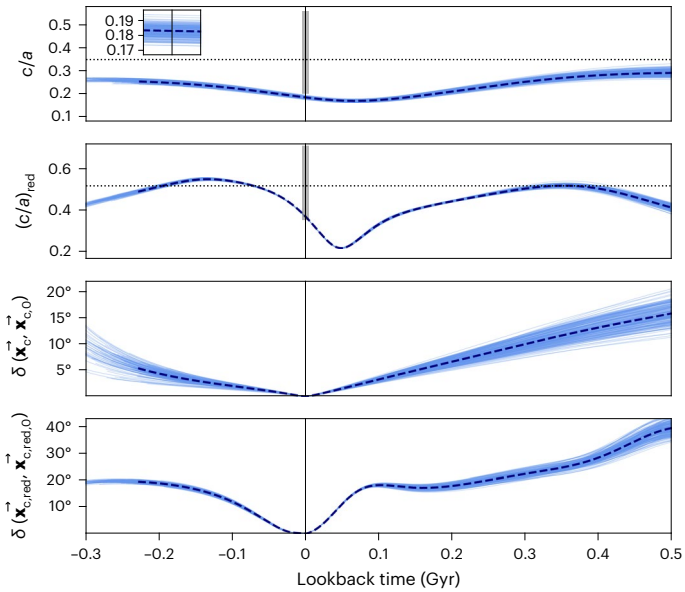


Fig. 5 | Evolution of the anisotropy of the MW satellite system and of the orientation of the plane. From top to bottom, the panels show c/a, $(c/a)_{red}$, the direction of $\mathbf{x}_c$ and the direction of $\mathbf{x}_{c,red}$; that is, the normals to the plane defined by c/a and $(c/a)_{red}$, relative to the directions at the present day ($\delta(\mathbf{x}_c, \mathbf{x}_{c,0})$ and $\delta(\mathbf{x}_{c,red}, \mathbf{x}_{c,red,0})$, respectively). Dashed dark blue lines show results based on the most likely values of the observables and light blue lines show 200 Monte Carlo samples. Lines end when one satellite lies beyond 300 kpc from the Galactic Centre. The small inset shows $t = 0 \pm 1$ Myr. Grey vertical bars at $t = 0$ show the 10th and 90th percentiles for the simulated MW analogues at $z = 0$ and dotted horizontal lines show the medians of these distributions. c/a evolves towards the median of the ΛCDM systems, while $(c/a)_{red}$ varies substantially, exceeding the median both in the near past and near future. The evolution of $\mathbf{x}_c$ and $\mathbf{x}_{c,red}$ reveal the plane to be tilting by either definition. Observational uncertainties result in greater uncertainties in the distant past and future, but the plane evolves substantially before these become an important factor.

Notably, one of our systems has both a c/a value below the MW value (0.166 compared to 0.183 for MW) and an orbital pole clustering more unlikely to arise out of isotropy than the MW's (probability of 0.005 compared to 0.009 for the MW). With no strong correlation between orbital pole clustering and anisotropy, and estimated probabilities below 10^{-3} and 10^{-2}, respectively[26], the combination of the two had previously been considered extremely unlikely.

Importantly, while the 'plane of satellites' includes all 11 classical satellites, the orbital anisotropy only concerns a subset, which is in fact more spatially isotropic than the system as a whole. The orbital pole clustering does not drive the spatial anisotropy.

The plane of satellites is transient

Another defining feature of a rotationally supported disk would be a significantly higher velocity dispersion parallel to the plane ($\sigma_{v\parallel}$) than perpendicular to it ($\sigma_{v\perp}$). However, for the MW's classical satellites, we measure $\sigma_{v\parallel} = 165.1 \pm 1.2$ km s^{-1} and $\sigma_{v\perp} = 121.6 \pm 0.4$ km s^{-1}. The ratio, $\sigma_{v\parallel}/\sigma_{v\perp} = 1.36$, is indistinguishable from the purely geometrical factor of $\sqrt{2}$. By this basic measure, the plane is not rotation supported.

The longevity of the plane can also be tested directly via orbital integration. This method was first, but inconclusively, applied using pre-Gaia data[35]. In this work, we benefit from significantly more precise observations, including Gaia EDR3 proper motions and more accurate distances.

In Fig. 1, we saw that several satellites are presently crossing the plane, while Leo I and II, which dominate the inertia, are moving apart. To elucidate the impact of such 'fortuitous alignments' on the anisotropy, we show in Fig. 3 the effect on c/a and $(c/a)_{red}$ of randomizing the position of each galaxy on the sky at the observed radius, keeping all other galaxies at their observed positions. For Sagittarius, the satellite with the smallest distance, the c/a distribution is extremely narrow: Sagittarius contributes less than 1% to the inertia tensor. However, randomizing the angular position of either of the two outer satellites, Leo I or Leo II, raises the median value of c/a to 0.28 and 0.31, respectively, with maxima of 0.53 and 0.63, far above the ΛCDM median.

While placing satellites at random sky coordinates highlights the sensitivity of c/a to individual galaxies, it is not a physical process. However, we also show at the top of Fig. 4 the anisotropy distributions when each satellite simply moves along its orbit. The time-averaged anisotropy of the system is then calculated over one full orbital period centred on the present time. Depending on the orbital phase of Leo II alone, c/a could be as high as 0.39, more isotropic than most ΛCDM systems. In other words, almost the entire anisotropy of the classical MW satellites is due to the orbital phases of Leo I and Leo II. Simply omitting both Leo I and Leo II from the analysis would yield $c/a = 0.279$, also more isotropic than 37% of ΛCDM systems of 9 satellites.

The bottom panels of Fig. 4 show time-averaged probability densities of c/a and $(c/a)_{red}$ when all satellites evolve simultaneously. The current value of c/a is significantly lower than in the recent past: over the past 0.5 and 1 Gyr, the time-averaged medians of c/a are 0.23 and 0.27 respectively, greater than 13% and 23% of ΛCDM systems. $(c/a)_{red}$ has varied widely. Neither metric is an invariant of the satellite system, both are sensitive to the orbital phases.

The four panels of Fig. 5 show the evolution of c/a, $(c/a)_{red}$ and of the orientations of the planes defined by the full and reduced inertia tensors, which we parametrize by the angles between the vectors normal to the planes, $\mathbf{x}_c$ and $\mathbf{x}_{c,red}$, and their present day equivalents, $\mathbf{x}_{c,0}$ and $\mathbf{x}_{c,0,red}$. The value of c/a approaches the ΛCDM median within a lookback time of 0.5 Gyr. $(c/a)_{red}$ evolves more rapidly, twice exceeding the ΛCDM median. That c/a and $(c/a)_{red}$ vary on such different timescales is a further consequence of the radial distribution and of the fact that c/a and $(c/a)_{red}$ are determined by satellites at different radii and on orbits with different periods: while Leo I and Leo II, which largely determine c/a, have orbital periods of 3.4 ± 0.2 and 7.2 ± 0.3 Gyr, respectively, the eight inner satellites, which largely determine $(c/a)_{red}$, all have orbital periods under 2 Gyr.

We further see that the orientation of the Milky Way's plane of satellites is not stable, but has tilted by ~17° over the past 0.5 Gyr (and ~40° for the reduced definition). This is readily understood considering that the plane is largely defined by only two outlying satellites. This is demonstrated clearly in Fig. 6, where we show the orbits and positions of the 11 classical satellites projected edge-on, in the frame defined by the minor and major axes of the inertia tensor at $z = 0$. The central panel, which shows the positions of the satellites at $z = 0$, shows the plane aligned with the major axis, analogous to the bottom left panel of Fig. 1. The other four panels show the evolution up to 1 Gyr in the past and the future. The orientation of the plane of satellites is changing, so that at each moment, it points toward the current position of the outermost satellites. Rather than satellites orbiting inside a stable plane, the plane tilts as it tracks the positions of its most distant members.

Discussion

The high reported anisotropy of the MW satellite system can largely be attributed to its high central concentration, not previously reproduced in simulations, combined with the close but fleeting contiguity of its two most distant members. Accounting for the radial distribution reveals the MW satellites to be consistent with ΛCDM expectations. Compared with previous work, we also find a much higher likelihood of subsets whose orbital poles are as clustered as the MW. Although the Milky Way contains such a subset, the plane of satellites does not constitute a rotationally supported disk. Instead, it evolves on

https://doi.org/10.1038/s41550-022-01856-z

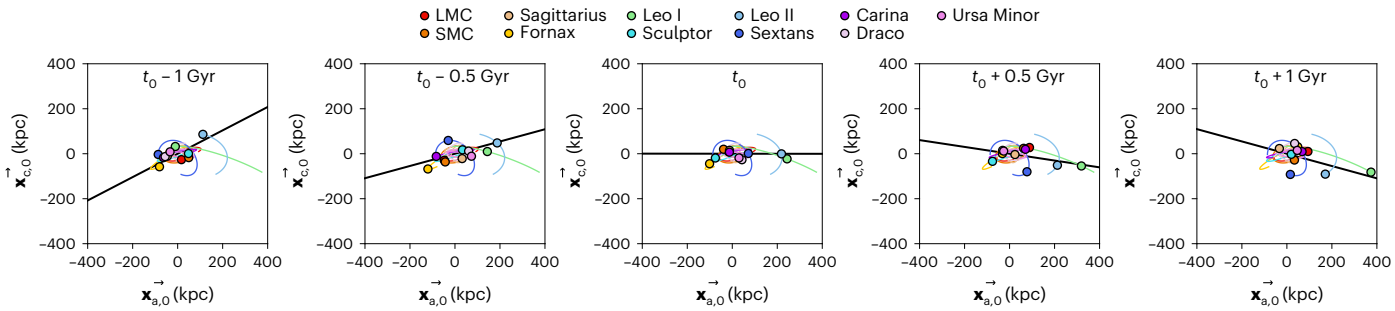


Fig. 6 | Tilt of the plane of satellites over time. Shown on all panels are the most likely orbits of the MW satellites for ±1 Gyr from the present time, analogous to Fig. 1. From left to right, panels show the position of satellites at 1 Gyr and 0.5 Gyr in the past, today (denoted with t_0) and 0.5 and 1 Gyr into the future. All panels are plotted in the frame of the major and minor eigenvectors of the inertia tensor at time t_0, $\mathbf{x}_{a,0}$ and $\mathbf{x}_{c,0}$. Thick black lines show the plane at each time edge-on. In the centre panel, by construction the plane axes align with the coordinate axes. In the other panels, the tilt of the plane can be observed. The orientation of the plane closely follows the locations of the most distant satellites.

timescales similar to the 'transient' alignments previously found in ΛCDM simulations.

Our orbital integration assumes a static MW potential with satellites as massless test particles. We Monte Carlo sample all sources of observational error and also vary the components of the potential within the uncertainties (Methods). We find our results to be robust, and while the real potential is more complex (for example, due to the presence of the LMC), these simplifications are valid within a dynamical time (~2 Gyr) of the halo, particularly for the important outer satellites[36,37]. Furthermore, a more complex potential would only accelerate the dissolution of a planar configuration[38]. Our simulations do not include the potential disruption of satellites by the disk of the MW. This might slightly extend the radial distributions, but not enough to change our conclusions. Importantly, among systems that have radial distributions similar to the MW, flattened 'planes of satellites' are quite common.

This work only directly addresses the archetypal 'plane of satellites' around the MW. Anisotropic satellite distributions have also been reported around several other galaxies[14,39] with different criteria, which can exacerbate the look-elsewhere effect. Assessing their significance requires a careful statistical analysis[12,40]. While not all criteria are equally sensitive to the radial distribution, we also expect that the much higher anisotropy we report here for simulated MW analogues will apply to ΛCDM analogues of other, similarly defined systems.

After centuries of observations, the Milky Way and its satellites are the best studied system of galaxies in the Universe. Viewed in sufficient detail, every system inevitably reveals some remarkable features. However, based on the best currently available data, there is no evidence for a plane of satellites incompatible with, or even particularly remarkable in, ΛCDM. On the contrary, as judged by the spatial anisotropy of the brightest satellite galaxies, we appear to live in a fairly regular ΛCDM system.

Methods

Observations

We adopt the sky positions and radial velocities from the McConnachie[41] catalogue, and combine these, where available, with McConnachie and Venn[23] proper motion measurements based on Gaia EDR3 (ref. [42]). The systemic proper motions were measured within a Bayesian framework that combines information from stars with full astrometric data with information from stars with only photometric and proper motion data. The method is a mixture model that associates a probability for each candidate star to be associated with the target galaxy taking into account foreground and background contaminants. It constitutes the best technique currently available in the literature to determine proper motions and provides the most precise measurements so far[37]. For the three innermost satellites (Sagittarius dSph, the LMC and the SMC), where the McConnachie & Venn catalogue does not include proper motions, we use the Gaia Data Release 2 (DR2) proper motions of ref. [43]. We further compiled the most recent estimates of the distance moduli of each satellite. The sky coordinates[41], proper motions[42,44] and distance moduli[45–50] used in this study, including their sources, are listed in Supplementary Table 2.

We also repeated our analysis using the Gaia EDR3 proper motions of Battaglia et al.[37] and the Gaia DR2 proper motions described in Riley et al.[43]. A comparison of the evolution of c/a and $(c/a)_{\rm red}$, analogous to Fig. 5, is shown in Extended Data Fig. 1. Unsurprisingly, the evolution based on the two EDR3 datasets are in excellent agreement. The main difference when using the DR2 data is the larger uncertainty (the astrometry errors of EDR3 are reduced by approximately a factor of two compared with DR2), but the evolution of both c/a and $(c/a)_{\rm red}$ is essentially the same in all three datasets, and the results are consistent within the respective errors.

Monte Carlo samples. We account for measurement errors by generating Monte Carlo samples of the satellites in the space of observed quantities: distance modulus, radial velocity and proper motions, as well as the position of the Sun relative to the Galactic Centre. We model each observable as a Gaussian distribution with the mean value and standard deviation given by the measurements and their quoted errors. For the Sun's distance from the Galactic Centre we assume $R_\odot = (8.178 \pm 0.022)$ kpc (ref. [51]), for the circular velocity at the Sun's position $V_{\rm circ} = (234.7 \pm 1.7)$ km s^{-1} (ref. [52]) and for Sun's motion with respect to the local standard of rest $(U, V, W) = (11.10 \pm 0.72, 12.24 \pm 0.47, 7.25 \pm 0.37)$ km s^{-1} (ref. [53]), where U is defined positive toward the Galactic center, V is positive in the direction of Galactic rotation and W is positive toward the North Galactic Pole.

The clustering of orbital poles. To characterize the clustering of orbital poles, we adopt the orbital pole dispersion for a subset of $N_{\rm s}$ satellites, $\Delta_{\rm std}$, defined by Pawlowski and Kroupa[26] as:

$$\Delta_{\rm std}(N_{\rm s}) = \sqrt{\frac{1}{N_{\rm s}} \sum_{i=1}^{N_{\rm s}} \theta_i^2}, \qquad (3)$$

where θ_i is the angle between the orbital pole of the ith satellite and the mean orbital pole of the satellites in the subset. To compute the clustering of an observed system relative to expectation, the same analysis is performed on the observations and simulations.

Based on earlier Gaia DR2 and Hubble Space Telescope proper motions, Pawlowski and Kroupa[26] calculated orbital pole dispersions for all possible satellite subsets in the MW and in simulations with

N_s = 3...11, and discovered that N_s = 7 yielded the most unusual configuration. However, there is no a priori reason to specifically consider N_s = 7. When considering only a proper subset of satellites, the interpretation of $\Delta_{std}(N_s)$ as evidence for unusual clustering is subject to the 'look-elsewhere effect'. To account for this, we follow here the method of ref. [12], which involves performing the same analysis for the simulated systems. As in the observations, we consider all subsets of size N_s = 3...11 in each simulated system, and identify the most unlikely to arise by chance from an isotropic distribution, which we calculate based on 10^5 isotropic distributions of N = 11 points, and the probability distributions of $\Delta_{std}(N_s)$ for all N_s = 3...11 possible subsets.

In Extended Data Fig. 2, we show a Hammer equal area projection of the most likely orbital poles and their Monte Carlo sampled uncertainty. The 3 black circles identify the clustering of 7 of the 11 satellites. The dotted line corresponds to $\Delta_{std}(7) = 16.0°$, the dispersion calculated by Pawlowski and Kroupa using pre-Gaia EDR3 data. The solid line shows our results, $\Delta_{std}(7) = 23.2°$, for the same satellites using Gaia EDR3 data. The dashed line shows our result for the most clustered subset $\Delta_{std}(7) = 18.9°$, exchanging Leo II for Leo I.

Time evolution and time integration

To infer the time evolution of the Milky Way satellite system, the orbits of the satellites are integrated numerically as massless test particles in a static potential using the Gala package[54]. The potential consists of a disk, stellar nucleus and bulge, and a dark matter halo. The disk is modelled as an axisymmetric Miyamoto–Nagai disk[55], which, for our default model, has disk mass $5.17 \times 10^{10}\,M_\odot$, a = 3 kpc and b = 0.028 kpc (ref. [56]). The nucleus and stellar bulge are both modelled as spherically symmetric Hernquist profiles[57]. For the nucleus we assume a mass of $1.71 \times 10^9\,M_\odot$ and a scale radius a = 0.07 kpc, and for the bulge we assume a mass of $5.0 \times 10^9\,M_\odot$ and a = 1.0 kpc. For the dark matter halo, we assume a spherically symmetric Navarro, Frenk and White (NFW)[58] potential.

Until recently, the Milky Way halo mass may have been a prohibitive source of uncertainty for calculating the orbital evolution of the satellites, as its value was known only to within a factor of two[59]. However, the Galactic halo mass has now been estimated with an uncertainty of only about 20% using Gaia data. Multiple dynamical probes, such as the stellar rotation curve, the escape velocity, the motions of halo stars, globular clusters and satellite galaxies[60–64], consistently imply a dark matter halo mass for the MW of $M_{200} = (1.0 \pm 0.2) \times 10^{12}\,M_\odot$ and NFW concentration, $c_{200} = 11 \pm 2$.

Based on these results, we adopt a reference MW halo of mass $1.0 \times 10^{12}\,M_\odot$ and a concentration parameter c_{200} = 11, corresponding to an NFW scale radius of r_s = 19.2 kpc. The positions and velocities relative to the plane of satellites, and the orbital periods and apocentre distances for the default potential, are listed in Supplementary Table 1, where the quoted uncertainties reflect 68% confidence intervals for all quantities based on Monte Carlo sampling. Varying the MW potential within the observational uncertainties does not significantly affect the conclusions of our study, as we show in Extended Data Fig. 3 where we vary the halo mass between 0.8 and $1.2 \times 10^{12}\,M_\odot$. The inferred anisotropies, as measured by c/a and $(c/a)_{red}$, begin to diverge for lookback times beyond ~500 Myr, but the behaviour is qualitatively similar across the halo mass range. As shown in Extended Data Fig. 4, we also tested the impact of varying the other three mass components and the concentration parameter within the observational uncertainties, and found no significant effects. The true Milky Way potential evolves with time, but the dynamical time of the halo (~2 Gyr at z = 0) is much longer than the timescale for the reported dissolution of the plane of satellites (several hundred Myr). A further possible source of uncertainty may be the mass of the LMC and its perturbation of the potential. However, at a distance of 49 kpc from the centre, even a massive LMC does not significantly perturb the acceleration field at distances above 150 kpc (ref. [36]) and would not significantly affect the orbits of the two most distant satellites, Leo I and Leo II.

Numerical simulations

Initial conditions. The simulations used in this work are cosmological zoom-in-constrained simulations, based on initial conditions created for the SIBELIUS project[31] and designed to reproduce Local Group (LG) analogues within the observed large-scale structure. The simulations assume a ΛCDM cosmology with density parameters $\Omega_0 = 0.307$ and $\Omega_\Lambda = 0.693$ for matter and dark energy, respectively, an r.m.s. density variation on a scale of $8h^{-1}$ Mpc of $\sigma_8 = 0.8288$ and a Hubble parameter of $h = 0.6777$. We use physical units throughout this work. Building on an octree representation of the phase information[65], we used the methods described in ref. [66] to supplement the observationally constrained scales in the initial density field[67,68] with independent random information below 3.2 comoving Mpc (cMpc). In total, we generated 60,000 simulations, resulting in several thousand loosely defined Local Group analogues. From these, we selected 112 for the high-resolution resimulations used in this work. All initial conditions refine a Lagrangian region extending to at least r = 3 Mpc around the centre of the LG at z = 0, with the remainder of a $1{,}000^3$ cMpc^3 volume populated with progressively more massive particles. The simulations start at z = 127. More details about the initial conditions may be found in ref. [31].

Simulations. All simulations used in this work were performed with the public version of the Gadget-4 code[69], on the Cosmology Machine at the University of Durham and at the Finnish IT Center for Science. Extended Data Fig. 5 shows a comparison of one LG analogue at two different resolutions. The left panel corresponds to the resolution of the set of 60,000 simulations of the SIBELIUS project (particle mass $2.0 \times 10^9\,M_\odot$). At this resolution, a MW analogue halo contains approximately 500 particles, and only the largest substructures are identifiable. The right panel shows a simulation with a particle mass of $1.0 \times 10^6\,M_\odot$. At this resolution, a MW analogue halo contains approximately 10^6 particles, and an average of ~200 subhaloes down to $2 \times 10^7\,M_\odot$ can be identified within 300 kpc from the centre. All results presented in this paper are based on simulations at this resolution.

Structure finding. Structures and self-bound substructures were identified using the Friends-of-Friends and Subfind algorithms implemented in Gadget-4 at 60 snapshots equally spaced in time, from z = 4 until a lookback time of 1 Gyr, and a further 40 snapshots equally spaced over the final 1 Gyr up to z = 0. Given their mass and separation, the two most massive self-bound substructures of the LG analogues can either belong to the same or to separate Friends-of-Friends structures. Throughout this work, we refer to the two principal self-bound substructures of each LG analogue at z = 0 simply as 'haloes' and to the lower mass substructures within 300 kpc of the centre of potential of each halo as 'satellites'. We select Local Group analogues as pairs of haloes with individual masses in the range $0.6–2.2 \times 10^{12}\,M_\odot$, separated by 500–1,000 kpc, with radial velocity $-150 < v_r < -50$ km s^{-1} and transverse velocity $v_t < 70$ km s^{-1}. In total, our set of high-resolution simulations contains 101 LG analogues, and, for the purposes of this work, we consider both haloes as a MW analogue.

We use Gadget's on-the-fly merger tree construction to find the progenitors of these subhaloes at previous snapshots. We cut the chain of links when a subhalo's progenitor is no longer found, or when a clear discontinuity in mass and position indicates that a satellite's progenitor has been erroneously identified as the main halo. At each snapshot, we record the maximum circular velocity of each subhalo, $\mathbf{x}_{c,red}$ and define v_{peak} as the highest value of v_{max} of a subhalo and its progenitors over time. Following ref. [17], we use the standard procedure to rank satellites by v_{peak}, and identify the top 11 within 300 kpc of each MW analogue at z = 0 as analogues to the classical MW satellites.

Obtaining a complete satellite sample. As noted above, the radial distribution of satellites is important for the anisotropy. Numerical simulations suffer from the artificial disruption of substructures, that

can affect subhaloes far beyond the particle number limit at which they can theoretically be identified[28,70]. According to van den Bosch and Ogiya[71], the main cause of this artificial tidal disruption is inaccurate force softening, which can cause force errors that, for particles within substructures, do not cancel out: once a particle is lost to the main halo, it cannot be recovered. Additionally, the amplification of discrete noise in the presence of a strong tidal field near the centre of a halo can lead to a runaway instability that can lead to the complete, but purely numerical, disruption of a subhalo.

These effects can, however, be mitigated using semi-analytical models (which populate merger trees constructed from simulated dark matter subhaloes with galaxies). These models include so-called 'orphan' galaxies, that is, galaxies whose dark matter subhalo has been numerically disrupted. After the subhalo is disrupted numerically, its subsequent evolution is followed by tracing the positions of its most bound particle[32]. Our 'complete' sample includes these 'orphan' subhaloes.

One important result of this work is that the 'incomplete' and 'complete' samples of satellite haloes have different radial distributions. Even though our high-resolution simulations resolve, on average, 200 surviving satellite haloes inside 300 kpc of each MW analogue at $z = 0$ and we rank the satellites by v_{peak} (v_{max} being more strongly affected by tidal stripping), we find that the radial distribution of the top 11 surviving satellites in the 'incomplete' samples are systematically and significantly less centrally concentrated than the MW's brightest satellites.

Figure 2 in the main text showed consistency between the radial distributions of the complete sample and the Milky Way satellites. It is reproduced in the top row of Extended Data Fig. 6. By contrast, the bottom row of Extended Data Fig. 6 shows how our comparison between simulations and observations would have looked had we considered only subhaloes from the incomplete samples. In the bottom left, it can be seen that their radial distribution is systematically offset from the MW data (shown as thick black line). For example, the innermost nine MW satellites are found within a distance of 140 kpc from the Galactic Centre, but none of the 202 incomplete samples has 9 satellites within this radius. The bottom-centre panel shows that the more uniform radial distributions of the incomplete samples lead to more equal contributions to the moment of inertia than in the case of the MW satellites. In fact, as shown on the bottom-right panel, none of the incomplete satellite systems have Gini coefficients as high as the MW's and only two have c/a as low or lower. This is in line with previous studies (for example, refs. [22,26]), which, presumably using incomplete subhalo populations, have found c/a values as low as the MW's to be very rare.

To demonstrate the dependence of the anisotropy on the Gini coefficient and the radial concentration independently of our simulations, we also repeat the analysis on synthetic random data. The distributions shown in the top row of Extended Data Fig. 7 use a parent radial distribution that is uniform in $r^{1/2}$, while those in the bottom row use one that is uniform in r; both parent angular distributions are isotropic. The relation between c/a and G is independent of the radial distribution, but the more centrally concentrated $r^{1/2}$ distribution attains larger values of G and higher anisotropies than the more extended one.

Baryonic effects. In hydrodynamic simulations, Milky Way mass haloes tend to be slightly more spherical than their counterparts in collisionless simulations[72], which might reduce the anisotropy of subhaloes. Additionally, the adiabatic contraction induced by baryons can increase the concentration of the dark matter halo[73], which can lead to a more central concentration of subhaloes but, conversely, could also enhance the disruption of subhaloes near the centre.

The presence of a stellar disk could also lead to additional disruption of subhaloes close to the centre of the galaxy[74–77]. To estimate its possible impact, we have calculated the fraction of subhaloes and orphans that pass through a cylinder whose radius, $r = 4.3$ kpc, is twice the scale-length and whose height, $h = 0.5$ kpc, is twice the scale-height of the MW disk during the past 1.3 Gyr. We find that in approximately half of the systems, none of the 11 satellites that we consider analogues of the classical satellites have passed through this disk, and at most one satellite has in ~83% of systems. Since satellites that passed through the disk are more likely to be found near the centre, this disruption could make the radial profiles slightly more extended. We find that this does not have a significant effect on our results: removing all satellites identified as having passed through the disk and replacing them with the next highest by v_{peak} only reduces the median Gini coefficient from 0.63 to 0.61, while it increases the median value of c/a from 0.35 to 0.36.

Some authors have argued that, in aggregate, the various baryonic effects that could individually lead to higher or lower anisotropy can increase the anisotropy of satellite systems in ΛCDM[13,35,78]. On the other hand, refs. [26] and [20] concluded that baryons only have a negligible effect. A fully realistic simulation of the Milky Way satellite system would clearly include baryons, but we believe that the inclusion of baryons would, at the very least, not decrease the anisotropy significantly. In other words, including baryons is unlikely to make the 'plane of satellite problem' worse.

Data availability

The observational data that we use are listed with references in Supplementary Table 2. Both the observational data and the simulation data necessary to recreate all the figures and tables in the paper are available at https://github.com/TillSawala/plane-of-satellites. The entirety of the raw simulation data, comprising 23 TB, are archived at the DiRAC Data Centric system at Durham University. Access may be provided by reasonable and specific request to the corresponding author.

Code availability

The analysis in this paper was performed in Python 3 and makes extensive use of open-source libraries, including Matplotlib 3.4.2, NumPy 1.21.1 (ref. [79]), SciPy 1.7.0 (ref. [80]), GalPy 1.7.0 (ref. [81]), Py-SPHViewer[82], TensorFlow[83] and Gala 1.4.1 (ref. [54]). The complete analysis code is available at https://github.com/TillSawala/plane-of-satellites.

References

1. Lynden-Bell, D. Dwarf galaxies and globular clusters in high velocity hydrogen streams. *Mon. Not. R. Astron. Soc.* **174**, 695–710 (1976).
2. Davis, M., Efstathiou, G., Frenk, C. S. & White, S. D. M. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophys. J.* **292**, 371–394 (1985).
3. Sawala, T. et al. The APOSTLE simulations: solutions to the Local Group's cosmic puzzles. *Mon. Not. R. Astron. Soc.* **457**, 1931–1943 (2016).
4. Kroupa, P., Theis, C. & Boily, C. M. The great disk of Milky-Way satellites and cosmological sub-structures. *Astron. Astrophys.* **431**, 517–521 (2005).
5. Pawlowski, M. S. The planes of satellite galaxies problem, suggested solutions, and open questions. *Mod. Phys. Lett. A* **33**, 1830004 (2018).
6. Kroupa, P. The dark matter crisis: falsification of the current standard model of cosmology. *Publ. Astron. Soc. Aust.* **29**, 395–433 (2012).
7. Bullock, J. S. & Boylan-Kolchin, M. Small-scale challenges to the ΛCDM paradigm. *Annu. Rev. Astron. Astrophys.* **55**, 343–387 (2017).
8. Perivolaropoulos, L. & Skara, F. Challenges for ΛCDM: an update. *New Astron. Rev.* **95** (2022).
9. Pawlowski, M. S. It's time for some plane speaking. *Nat. Astron.* **5**, 1185–1187 (2021).
10. Boylan-Kolchin, M. Planes of satellites are not a problem for (just) ΛCDM. *Nat. Astron.* **5**, 1188–1190 (2021).

11. Sales, L.V., Wetzel, A. & Fattahi, A. Baryonic solutions and challenges for cosmological models of dwarf galaxies. *Nat. Astron.* https://doi.org/10.1038/s41550-022-01689-w (2022).
12. Cautun, M. et al. Planes of satellite galaxies: when exceptions are the rule. *Mon. Not. R. Astron. Soc.* **452**, 3838–3852 (2015).
13. Ahmed, S. H., Brooks, A. M. & Christensen, C. R. The role of baryons in creating statistically significant planes of satellites around Milky Way-mass galaxies. *Mon. Not. R. Astron. Soc.* **466**, 3119–3132 (2017).
14. Müller, O. et al. The coherent motion of Cen A dwarf satellite galaxies remains a challenge for ΛCDM cosmology. *Astron. Astrophys.* **645**, 5 (2021).
15. Forero-Romero, J. E. & Arias, V. We are not the 99 percent: quantifying asphericity in the distribution of Local Group satellites. *Mon. Not. R. Astron. Soc.* **478**, 5533–5546 (2018).
16. Pawlowski, M. S., Bullock, J. S., Kelley, T. & Famaey, B. Do halos that form early, have high concentration, are part of a pair, or contain a central galaxy potential host more pronounced planes of satellite galaxies? *Astrophys. J.* **875**, 105 (2019).
17. Libeskind, N. I. et al. The distribution of satellite galaxies: the great pancake. *Mon. Not. R. Astron. Soc.* **363**, 146–152 (2005).
18. Shao, S. et al. The multiplicity and anisotropy of galactic satellite accretion. *Mon. Not. R. Astron. Soc.* **476**, 1796–1810 (2018).
19. Santos-Santos, I. et al. Planes of satellites around simulated disk galaxies. I. Finding high-quality planar configurations from positional information and their comparison to MW/M31 data. *Astrophys. J.* **897**, 71 (2020).
20. Samuel, J. et al. Planes of satellites around Milky Way/M31-mass galaxies in the FIRE simulations and comparisons with the Local Group. *Mon. Not. R. Astron. Soc.* **504**, 1379–1397 (2021).
21. Buck, T., Dutton, A. A. & Macciò, A. V. Simulated ΛCDM analogues of the thin plane of satellites around the Andromeda galaxy are not kinematically coherent structures. *Mon. Not. R. Astron. Soc.* **460**, 4348–4365 (2016).
22. Shao, S., Cautun, M. & Frenk, C. S. Evolution of galactic planes of satellites in the EAGLE simulation. *Mon. Not. R. Astron. Soc.* **488**, 1166–1179 (2019).
23. McConnachie, A. W. & Venn, K. A. Updated proper motions for Local Group dwarf galaxies using Gaia Early Data Release 3. *Res. Notes AAS* **4**, 229 (2020).
24. Bailin, J. & Steinmetz, M. Internal and external alignment of the shapes and angular momenta of ΛCDM halos. *Astrophys. J.* **627**, 647–665 (2005).
25. Santos-Santos, I. M., Domínguez-Tenreiro, R. & Pawlowski, M. S. An updated detailed characterization of planes of satellites in the MW and M31. *Mon. Not. R. Astron. Soc.* **499**, 3755–3774 (2020).
26. Pawlowski, M. S. & Kroupa, P. The Milky Way's disc of classical satellite galaxies in light of Gaia DR2. *Mon. Not. R. Astron. Soc.* **491**, 3042–3059 (2019).
27. Guo, Q. & White, S. Numerical resolution limits on subhalo abundance matching. *Mon. Not. R. Astron. Soc.* **437**, 3228–3235 (2014).
28. van den Bosch, F. C. & Ogiya, G. Dark matter substructure in numerical simulations: a tale of discreteness noise, runaway instabilities, and artificial disruption. *Mon. Not. R. Astron. Soc.* **475**, 4066–4087 (2018).
29. Webb, J. J. & Bovy, J. High-resolution simulations of dark matter subhalo disruption in a Milky-Way-like tidal field. *Mon. Not. R. Astron. Soc.* **499**, 116–128 (2020).
30. Grand, R. J. J. et al. Determining the full satellite population of a Milky Way-mass halo in a highly resolved cosmological hydrodynamic simulation. *Mon. Not. R. Astron. Soc.* **507**, 4953–4967 (2021).
31. Sawala, T. et al. The SIBELIUS project: e pluribus unum. *Mon. Not. R. Astron. Soc.* **509**, 1432–1446 (2022).
32. Simha, V. & Cole, S. Modelling galaxy merger time-scales and tidal destruction. *Mon. Not. R. Astron. Soc.* **472**, 1392–1400 (2017).
33. Samuel, J. et al. A profile in FIRE: resolving the radial distributions of satellite galaxies in the Local Group with simulations. *Mon. Not. R. Astron. Soc.* **491**, 1471–1490 (2020).
34. Metz, M., Kroupa, P. & Libeskind, N. I. The orbital poles of Milky Way satellite galaxies: a rotationally supported disk of satellites. *Astrophys. J.* **680**, 287–294 (2008).
35. Maji, M., Zhu, Q., Marinacci, F. & Li, Y. Is there a disk of satellites around the Milky Way? *Astrophys. J.* **843**, 62 (2017).
36. Garavito-Camargo, N. et al. Quantifying the impact of the Large Magellanic Cloud on the structure of the Milky Way's dark matter halo using basis function expansions. *Astrophys. J.* **919**, 109 (2021).
37. Battaglia, G., Taibi, S., Thomas, G. F. & Fritz, T. K. Gaia Early DR3 systemic motions of Local Group dwarf galaxies and orbital properties with a massive Large Magellanic Cloud. *Astron. Astrophys.* **657**, 54 (2022).
38. Fernando, N., Arias, V., Lewis, G. F., Ibata, R. A. & Power, C. Stability of satellite planes in M31 II: effects of the dark subhalo population. *Mon. Not. R. Astron. Soc.* **473**, 2212–2221 (2018).
39. Ibata, N. G., Ibata, R. A., Famaey, B. & Lewis, G. F. Velocity anti-correlation of diametrically opposed galaxy satellites in the low-redshift Universe. *Nature* **511**, 563–566 (2014).
40. Gillet, N. et al. Vast planes of satellites in a high-resolution simulation of the Local Group: comparison to Andromeda. *Astrophys. J.* **800**, 34 (2015).
41. McConnachie, A. W. The observed properties of dwarf galaxies in and around the Local Group. *Astron. J.* **144**, 4 (2012).
42. Fabricius, C. et al. Gaia Early Data Release 3. Catalogue validation. *Astron. Astrophys.* **649**, 5 (2021).
43. Riley, A. H. et al. The velocity anisotropy of the Milky Way satellite system. *Mon. Not. R. Astron. Soc.* **486**, 2679–2694 (2019).
44. Gaia Collaboration et al. Gaia Data Release 2. Kinematics of globular clusters and dwarf galaxies around the Milky Way. *Astron. Astrophys.* **616**, 12 (2018).
45. Pietrzyński, G. et al. A distance to the Large Magellanic Cloud that is precise to one per cent. *Nature* **567**, 200–203 (2019).
46. Graczyk, D. et al. A Distance determination to the Small Magellanic Cloud with an accuracy of better than two percent based on late-type eclipsing binary stars. *Astrophys. J.* **904**, 13 (2020).
47. Hernitschek, N. et al. Precision distances to dwarf galaxies and globular clusters from Pan-STARRS1 3π RR Lyrae. *Astrophys. J.* **871**, 49 (2019).
48. Freedman, W. L. & Madore, B. F. Astrophysical distance scale. II. Application of the JAGB method: a nearby galaxy sample. *Astrophys. J.* **899**, 67 (2020).
49. Martínez-Vázquez, C. E. et al. Variable stars in Local Group galaxies. I. Tracing the early chemical enrichment and radial gradients in the Sculptor dSph with RR Lyrae stars. *Mon. Not. R. Astron. Soc.* **454**, 1509–1516 (2015).
50. Gullieuszik, M. et al. The evolved stars of LeoII dSph galaxy from near-infrared UKIRT/WFCAM observations. *Mon. Not. R. Astron. Soc.* **388**, 1185–1197 (2008).
51. Gravity Collaboration. A geometric distance measurement to the Galactic Center black hole with 0.3% uncertainty. *Astron. Astrophys.* **625**, 10 (2019).
52. Nitschai, M. S., Eilers, A.-C., Neumayer, N., Cappellari, M. & Rix, H.-W. Dynamical model of the Milky Way using APOGEE and Gaia data. *Astrophys. J.* **916**, 112 (2021).
53. Schönrich, R., Binney, J. & Dehnen, W. Local kinematics and the local standard of rest. *Mon. Not. R. Astron. Soc.* **403**, 1829–1833 (2010).

54. Price-Whelan, A. M. Gala: a Python package for galactic dynamics. *J. Open Source Softw.* https://doi.org/10.21105/joss.00388 (2017).
55. Miyamoto, M. & Nagai, R. Three-dimensional models for the distribution of mass in galaxies. *Publ. Astron. Soc. Jpn* **27**, 533–543 (1975).
56. Licquia, T. C. & Newman, J. A. Improved estimates of the Milky Way's stellar mass and star formation rate from hierarchical Bayesian meta-analysis. *Astrophys. J.* **806**, 96 (2015).
57. Hernquist, L. An analytical model for spherical galaxies and bulges. *Astrophys. J.* **356**, 359–364 (1990).
58. Navarro, J. F., Eke, V. R. & Frenk, C. S. The cores of dwarf galaxy haloes. *Mon. Not. R. Astron. Soc.* **283**, 72–78 (1996).
59. Wang, W., Han, J., Cautun, M., Li, Z. & Ishigaki, M. N. The mass of our Milky Way. *Sci. China Phys. Mech. Astron.* **63**, 109801 (2020).
60. Monari, G. et al. The escape speed curve of the Galaxy obtained from Gaia DR2 implies a heavy Milky Way. *Astron. Astrophys.* **616**, 9 (2018).
61. Callingham, T. M. et al. The mass of the Milky Way from satellite dynamics. *Mon. Not. R. Astron. Soc.* **484**, 5453–5467 (2019).
62. Deason, A. J. et al. The local high-velocity tail and the Galactic escape speed. *Mon. Not. R. Astron. Soc.* **485**, 3514–3526 (2019).
63. Cautun, M. et al. The Milky Way total mass profile as inferred from Gaia DR2. *Mon. Not. R. Astron. Soc.* **494**, 4291–4313 (2020).
64. Koppelman, H. H. & Helmi, A. Determination of the escape velocity of the Milky Way using a halo sample selected based on proper motion. *Astron. Astrophys.* **649**, 136 (2021).
65. Jenkins, A. A new way of setting the phases for cosmological multiscale Gaussian initial conditions. *Mon. Not. R. Astron. Soc.* **434**, 2094–2120 (2013).
66. Sawala, T. et al. Setting the stage: structures from Gaussian random fields. *Mon. Not. R. Astron. Soc.* **501**, 4759–4776 (2021).
67. Lavaux, G. & Jasche, J. Unmasking the masked Universe: the 2M++ catalogue through Bayesian eyes. *Mon. Not. R. Astron. Soc.* **455**, 3169–3179 (2016).
68. Jasche, J., Leclercq, F. & Wandelt, B. D. Past and present cosmic structure in the SDSS DR7 main sample. *J. Cosmol. Astropart. Phys.* **2015**, 036 (2015).
69. Springel, V., Pakmor, R., Zier, O. & Reinecke, M. Simulating cosmic structure formation with the GADGET-4 code. *Mon. Not. R. Astron. Soc.* **506**, 2871–2949 (2021).
70. Guo, Q. et al. From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology. *Mon. Not. R. Astron. Soc.* **413**, 101–131 (2011).
71. van den Bosch, F. C. & Ogiya, G. Dark matter substructure in numerical simulations: a tale of discreteness noise, runaway instabilities, and artificial disruption. *Mon. Not. R. Astron. Soc.* **475**, 4066–4087 (2018).
72. Abadi, M. G., Navarro, J. F., Fardal, M., Babul, A. & Steinmetz, M. Galaxy-induced transformation of dark matter haloes. *Mon. Not. R. Astron. Soc.* **407**, 435–446 (2010).
73. Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A. & Nagai, D. Response of dark matter halos to condensation of baryons: cosmological simulations and improved adiabatic contraction model. *Astrophys. J.* **616**, 16–26 (2004).
74. D'Onghia, E., Springel, V., Hernquist, L. & Keres, D. Substructure depletion in the Milky Way halo by the disk. *Astrophys. J.* **709**, 1138–1147 (2010).
75. Sawala, T. et al. Shaken and stirred: the Milky Way's dark substructures. *Mon. Not. R. Astron. Soc.* **467**, 4383–4400 (2017).
76. Garrison-Kimmel, S. et al. Not so lumpy after all: modelling the depletion of dark matter subhaloes by Milky Way-like galaxies. *Mon. Not. R. Astron. Soc.* **471**, 1709–1727 (2017).
77. Richings, J. et al. Subhalo destruction in the APOSTLE and AURIGA simulations. *Mon. Not. R. Astron. Soc.* **492**, 5780–5793 (2020).
78. Zhu, Q., Hernquist, L., Marinacci, F., Springel, V. & Li, Y. Baryonic impact on the dark matter orbital properties of Milky Way-sized haloes. *Mon. Not. R. Astron. Soc.* **466**, 3876–3886 (2017).
79. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).
80. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
81. Bovy, J. galpy: a Python library for galactic dynamics. *Astrophys. J. Suppl. Ser.* **216**, 29 (2015).
82. Benitez-Llambay, A. py-sphviewer: Py-SPHViewer v1.0.0 Zenodo https://doi.org/10.5281/zenodo.21703 (2015).
83. Abadi, M. et al. TensorFlow: large-scale machine learning on heterogeneous systems. ArXiv eprints, arXiv:1603.04467 (2016).

Acknowledgements

T.S. is an Academy of Finland Research Fellow. This work was supported by Academy of Finland grant numbers 314238 and 335607. P.H.J. acknowledges the support by the European Research Council via ERC Consolidator Grant KETJU (no. 818930) and the Academy of Finland grant 339127. M.S. is supported by the Netherlands Organisation for Scientific Research (NWO) through VENI grant 639.041.749. C.F. acknowledges support by the European Research Council (ERC) through Advanced Investigator DMIDAS (GA 786910). C.F. and A.J. acknowledge consolidated grant ST/T000244/1. J.J. acknowledges support by the Swedish Research Council (VR) under the project 2020-05143, 'Deciphering the Dynamics of Cosmic Structure'. G.L. acknowledges financial support by the French National Research Agency for the project BIG4, under reference ANR-16-CE23-0002. This work was supported by the Simons Collaboration on 'Learning the Universe'. This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk), and facilities hosted by the CSC—IT Centre for Science, Finland. The DiRAC system was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grants ST/H008519/1 and ST/K00087X/1, STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National UK E-Infrastructure.

Author contributions

T.S., M.C. and C.F. developed the original ideas and T.S. and M.C. performed the analysis of the MW data. T.S. generated and analysed the simulation data based on the SIBELIUS project, which was developed jointly by T.S., C.F, J.H., J.J., A.J., P.H.J., G.L., S.M. and M.S. J.H. created the merger tree data for the complete sample. T.S. created the figures and wrote the first draft and all co-authors edited and revised the manuscript.

Funding

Open Access funding provided by University of Helsinki including Helsinki University Central Hospital.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41550-022-01856-z.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41550-022-01856-z.

Correspondence and requests for materials should be addressed to Till Sawala.

Peer review information *Nature Astronomy* thanks Moupiya Maji and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

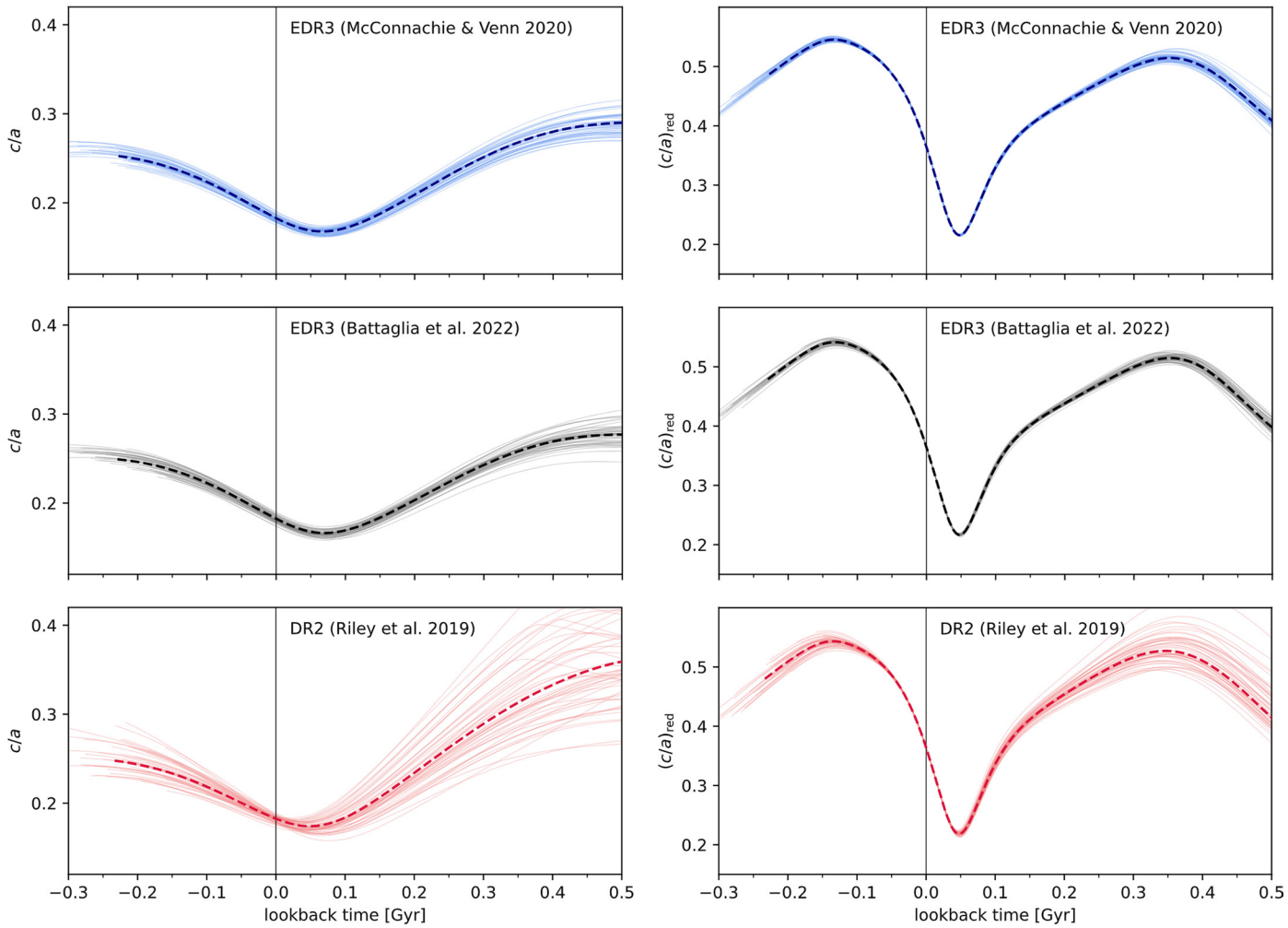
Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

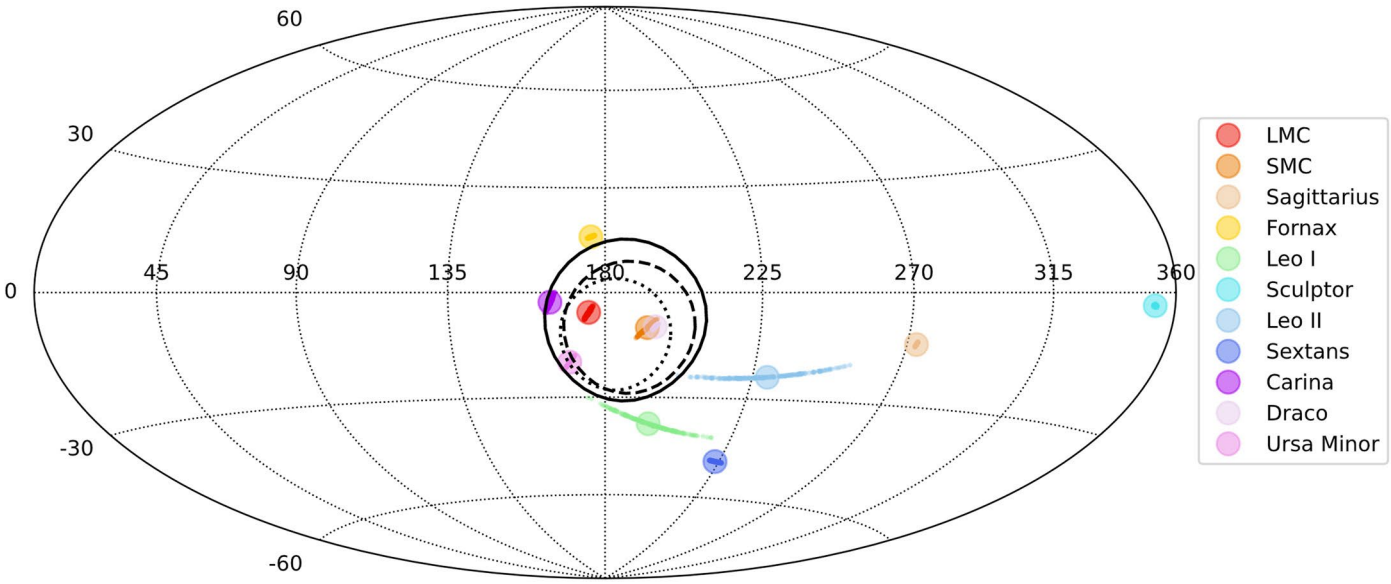
© The Author(s) 2022

https://doi.org/10.1038/s41550-022-01856-z



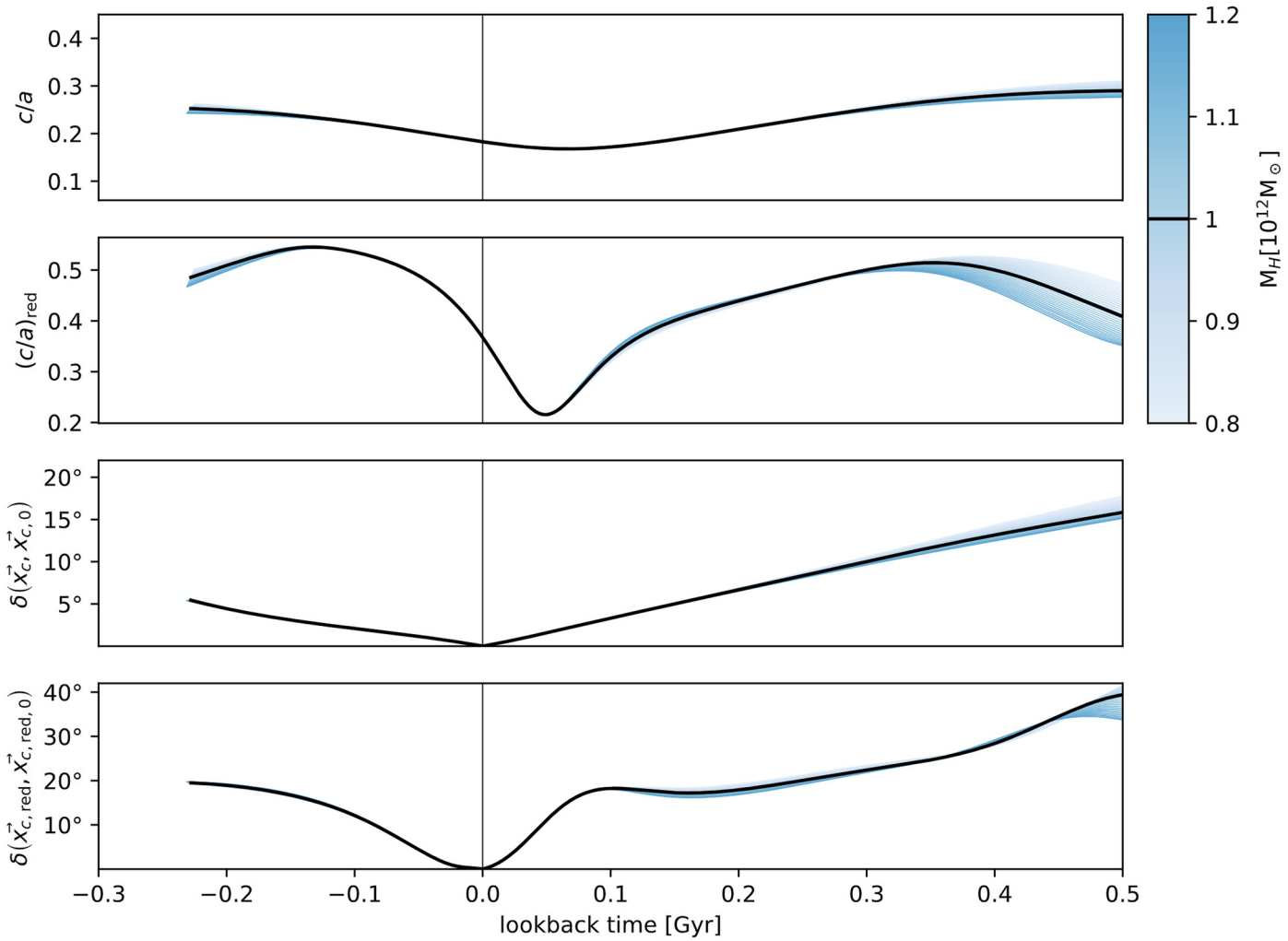
Extended Data Fig. 1 | Sensitivity of the results to the source of proper motion information. Evolution of c/a (left column) and $(c/a)_{red}$ (right column), from integrated satellite orbits, analogous to Fig. 4, showing the effect of using Gaia EDR3 proper motions by McConnachie and Venn[23] (top, blue lines), Gaia EDR3 proper motions by Battaglia et al. (2022)[27] (middle, black lines), or Gaia DR2 proper motions as described in Riley et al.[43] (bottom, red lines). For all data sets, thick dashed lines are the for the most likely observations, thin lines show 50 Monte Carlo samples. Lines extend to 0.5 Gyr into the past, and as long as all 11 satellites remain within 300 kpc of the centre into the future. The two EDR3 data sets are in excellent agreement. The main difference in the DR2 data is the larger errors, but the evolution of both c/a and $(c/a)_{red}$ is essentially the same in all three data sets.

https://doi.org/10.1038/s41550-022-01856-z



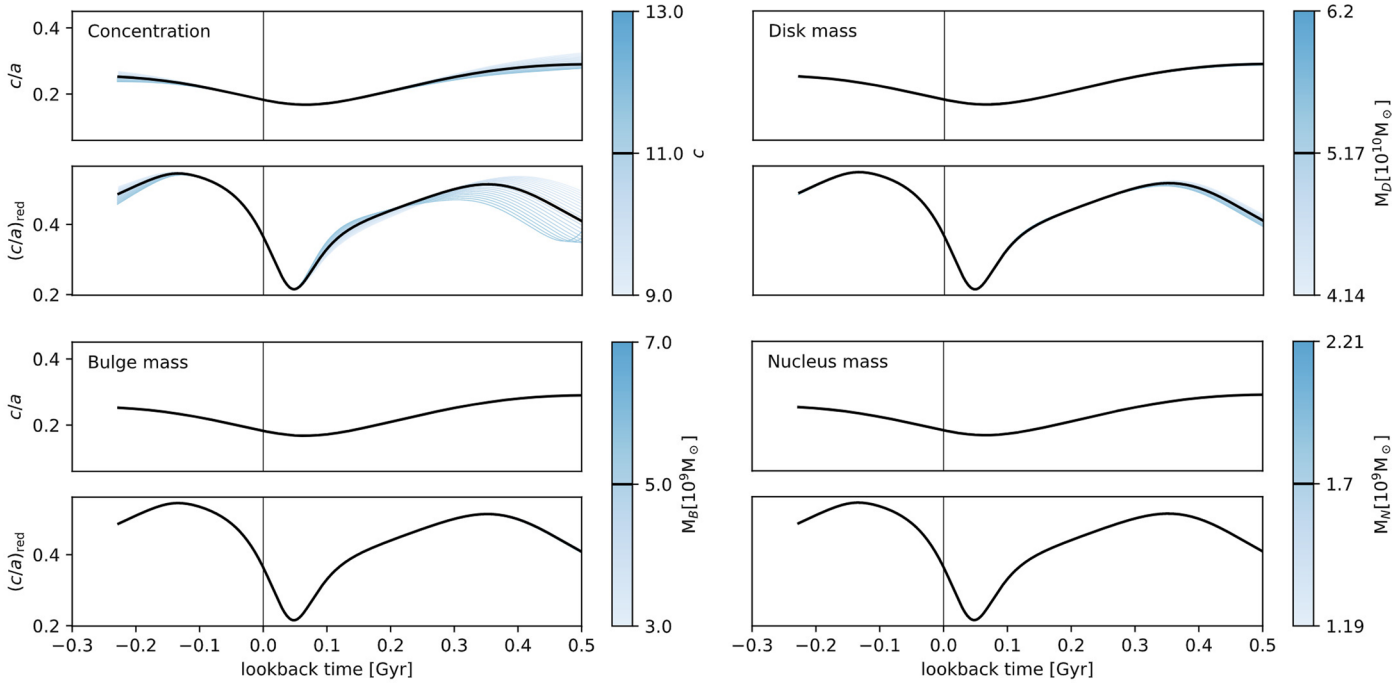
Extended Data Fig. 2 | Orbital pole clustering. Hammer projection of Milky Way satellite orbital poles, using Gaia EDR3 proper motions. Large circles show the most likely values, small circles show 200 Monte Carlo samples of the observational errors, within $\pm 1\sigma$ of the most likely values. The dotted black curve indicates the dispersion reported[26] for seven MW satellites. The solid black curve indicates the dispersion that we find for the same set based on Gaia EDR3, the dashed black curve indicates the minimum dispersion that we find for seven satellites exchanging Leo I and Leo II. The orbital poles of the MW satellites are significantly clustered, but several of our simulated ΛCDM systems contain equally or more strongly clustered satellite systems.

https://doi.org/10.1038/s41550-022-01856-z



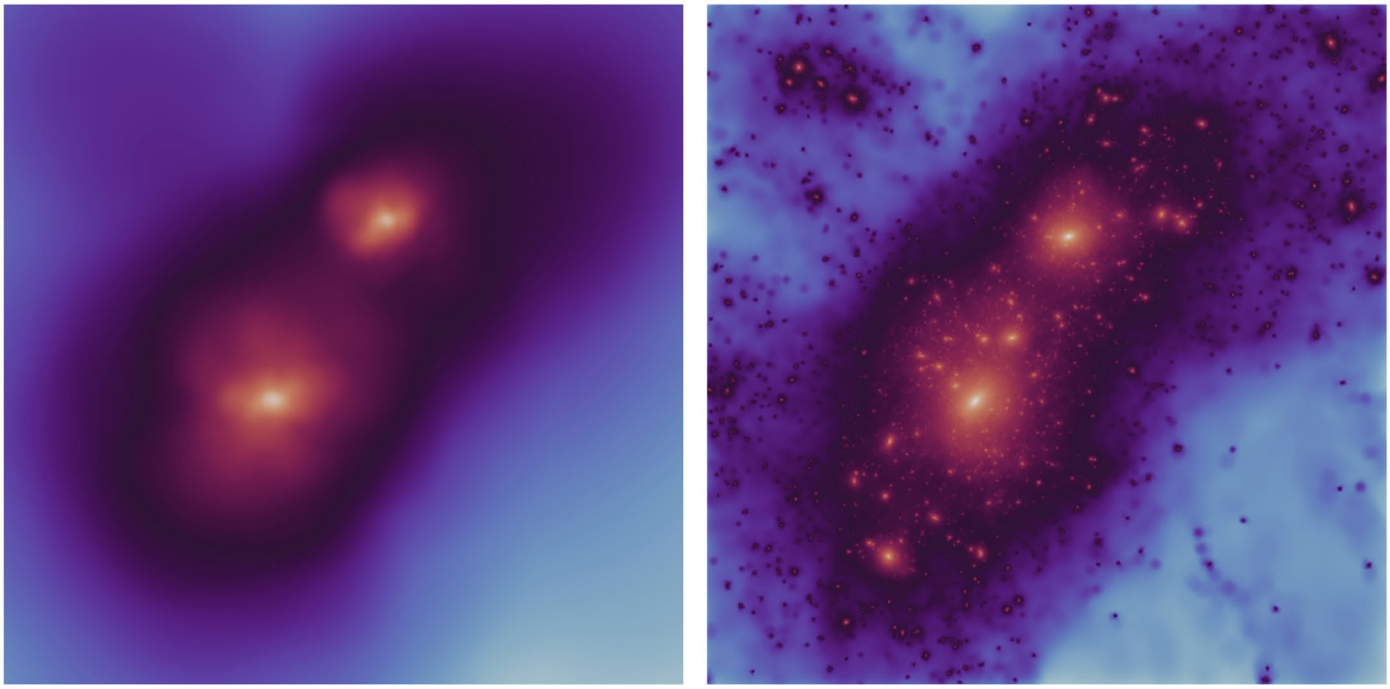
Extended Data Fig. 3 | Sensitivity of the results to the assumed halo mass. Evolution of c/a, $(c/a)_{red}$ and the direction of the vectors normal to the planes of the full and reduced inertia tensors, from integrated satellite orbits, analogous to Fig. 5 of the main part. Here, all lines are for the most likely observations, and we show the results of varying the mass of the dark matter halo. Black lines are for the default mass of $10^{12}M_{\odot}$, shades of blue show the result of varying the mass between $0.8-1.2\times10^{12}M_{\odot}$. Lines extend to 0.5 Gyr into the past, and as long as all 11 satellites remain within 300 kpc of the centre into the future. The ratio c/a reverts towards the median of the ΛCDM values and $(c/a)_{red}$ evolves rapidly for any mass. The tilt of the plane defined by either the full or reduced inertia tensor is essentially independent of the halo mass.

https://doi.org/10.1038/s41550-022-01856-z



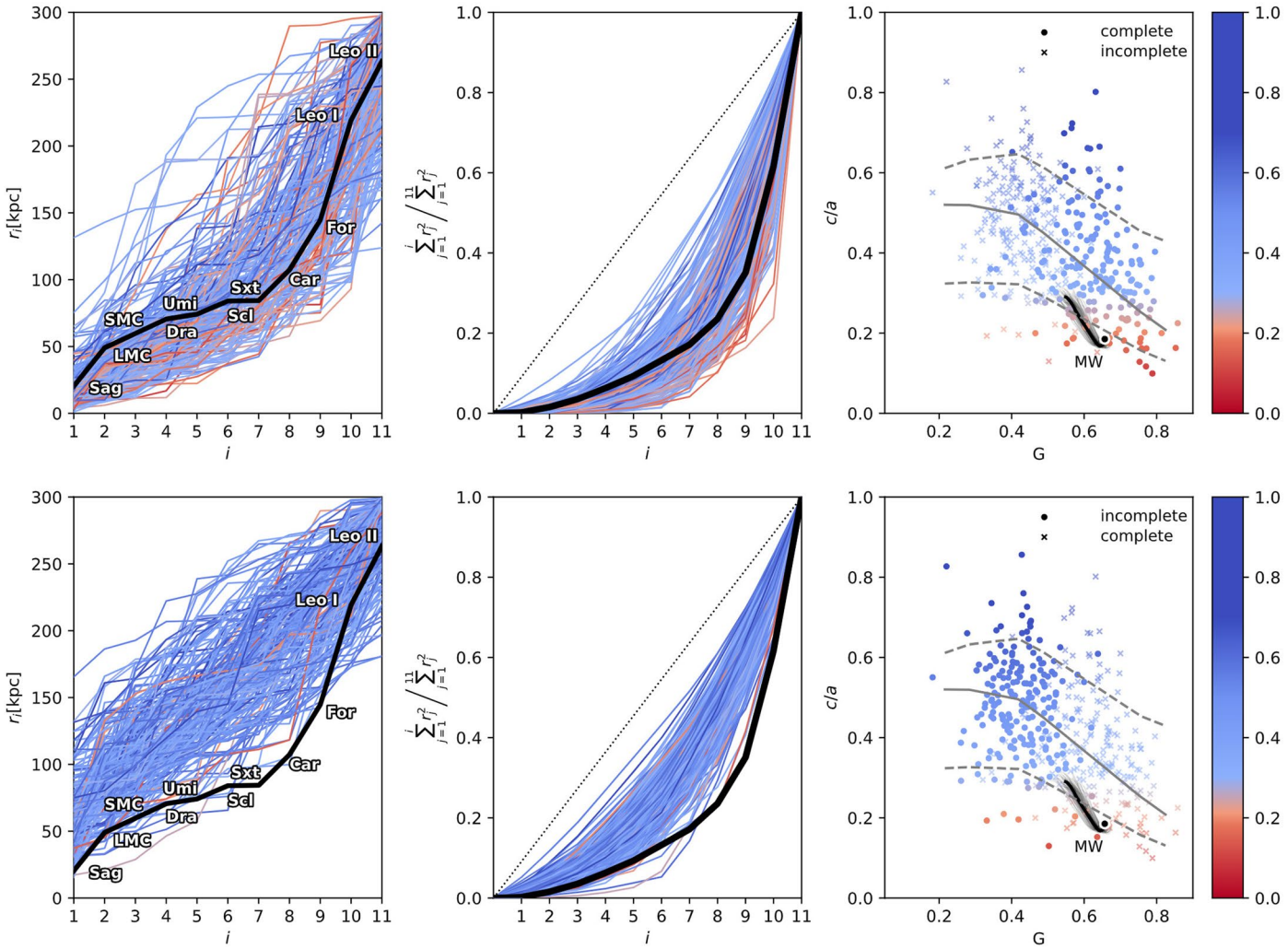
Extended Data Fig. 4 | Sensitivity of the results to the remaining parameters of the assumed Milky Way potential. Evolution of c/a and $(c/a)_{red}$ from integrated satellite orbits, analogous to Fig. 5 and Extended Data Fig. 3. All lines are for the most likely observations, and we show the results of varying the concentration of the dark matter halo (top left), disk mass (top right), bulge mass (bottom left) and nucleus mass (bottom right). Black lines are for the default parameters, shades of blue show the result of varying the parameters within the observational uncertainties. Lines extend to 0.5 Gyr into the past, and as long as all 11 satellites remain within 300 kpc of the centre into the future. Our results are not significantly affected by varying the components of the halo model within the observational uncertainties.

https://doi.org/10.1038/s41550-022-01856-z



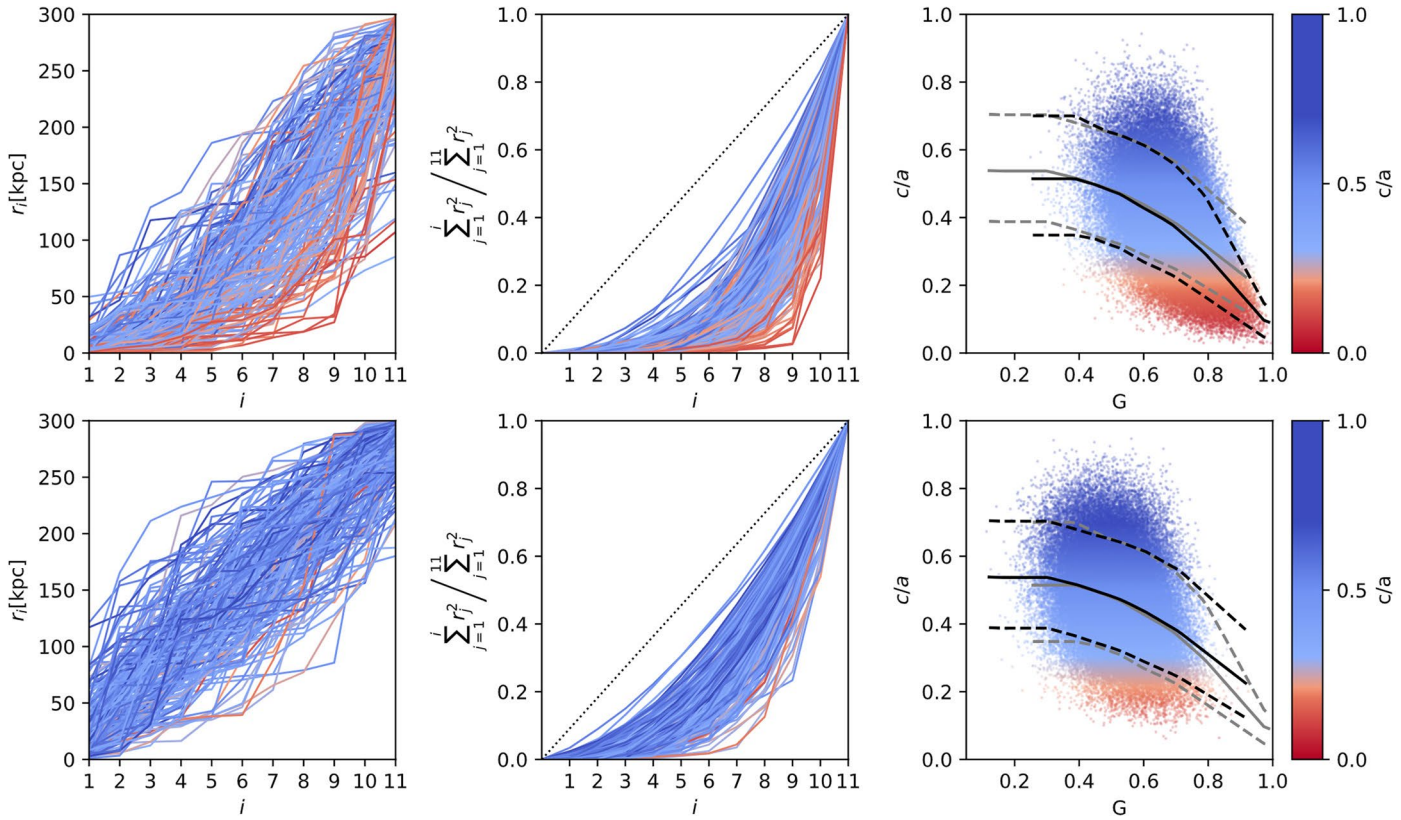
Extended Data Fig. 5 | Dark matter density in our simulations at two different resolutions. Dark matter density in one of the Local Group analogues from the Sibelius project of ΛCDM constrained simulations. Both panels show the projection of a box of side length 2 Mpc centred around the same system at z = 0. The left panel shows results from a simulation with particle mass of $1.25 \times 10^8 M_\odot$, used in ref. [31] to identify LG analogues. The right panel shows the same system resimulated with a particle mass of $10^6 M_\odot$ used in this work to identify and trace the satellite substructures.

https://doi.org/10.1038/s41550-022-01856-z



Extended Data Fig. 6 | Radial distributions, Gini coefficients, and anisotropy of complete and incomplete samples of subhaloes. Relation between Gini coefficient of inertia, G, and anisotropy, c/a, for the complete set of satellites (top, same as Fig. 2), and using only the subhaloes from the incomplete sample (bottom). While the radial distributions of the complete samples bracket the Milky Way data (left panel), those of the incomplete subhalo samples are much less centrally concentrated, resulting in much lower Gini coefficients. The complete sample contains multiple systems with G as high or higher, and c/a as low or lower than the MW, while in the incomplete sample, only two systems are as anisotropic, and none is as centrally concentrated as the MW.

https://doi.org/10.1038/s41550-022-01856-z



Extended Data Fig. 7 | Radial distributions, Gini coefficients, and anisotropy of uniformly distributed random samples. Relation between Gini coefficient of inertia, G, and anisotropy, c/a, for 10^5 random samples of 11 points, drawn from isotropic angular distributions and radial distributions uniformly distributed in r1/2 (top) or r (bottom), analogous to Fig. 2 and Extended Data Fig. 6. On the right panel, black lines denote the 90th, 50th and 10th percentile of each dataset, grey lines repeat the corresponding percentiles for the other dataset. For clarity, the left and middle panels only show the first 200 samples.