

The effect of environment on the mass assembly history of the Milky Way and M31

Ewoud Wempe^{1,*}, Amina Helmi¹, Simon D.M. White², Jens Jasche³, and Guilhem Lavaux⁴

¹ Kapteyn Astronomical Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

² Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85748 Garching, Germany

³ The Oskar Klein Centre, Department of Physics, Stockholm University, Albanova University Center, SE 106 91 Stockholm, Sweden

⁴ CNRS & Sorbonne Université, UMR 7095, Institut d'Astrophysique de Paris, 98 bis boulevard Arago, F-75014 Paris, France

Received 13 January 2025 / Accepted 25 June 2025

ABSTRACT

We study the mass growth histories of the haloes of Milky Way and M31 analogues formed in constrained cosmological simulations of the Local Group. These gravity-only simulations constitute a fair and representative set of Λ cold dark matter (Λ CDM) realisations conditioned on the masses, positions, and relative velocities of the two big haloes and on the observed recession velocities at the positions of nearby isolated galaxies. Our M31 haloes have similar mass growth histories as the isolated analogues in the TNG dark-matter-only simulations, while our Milky Ways typically form earlier, with suppressed growth at late times. On average, our Milky Ways assemble half their halo mass by $\langle z_{50} \rangle = 1.4$ and our M31s by $\langle z_{50} \rangle = 1.2$, whereas $\langle z_{50} \rangle = 1.1$ for their isolated analogues. Mass growth associated with major and minor mergers is also biased early for the Milky Way in comparison to M31. Most accretion occurs 1–4 Gyr after the Big Bang; growth at later times is relatively quiescent. Based on the mass ratio and time of infall, we find that 32% of our Milky Ways experienced a Gaia-Enceladus-Sausage-like merger, 13% host a massive Large Magellanic Cloud-like satellite at the present day, and 5% have both. In one case, a Small Magellanic Cloud analogue and a Sagittarius analogue are also present, showing that the most important mergers of the Milky Way can be reproduced in its Local Group environment in Λ CDM. We find that the material that makes up the Milky Way and M31 haloes at the present day initially collapsed onto a plane roughly aligned with the Local Sheet and super-galactic plane; after $z \sim 2$, accretion occurred mostly within this plane, with the tidal effects of the heavier companion, M31, significantly impacting the late growth history of the Milky Way.

Key words. galaxies: evolution – galaxies: formation – Local Group – dark matter

1. Introduction

In recent years, tremendous progress has been made on constraining the assembly history of the Milky Way (MW) using *Gaia* data (Gaia Collaboration 2018, 2023) in combination with ground-based spectroscopic surveys such as APOGEE (Majewski et al. 2017). We now know that the Galaxy experienced its last major merger (LMM) approximately 10 Gyr ago, with a system dubbed Gaia-Enceladus (Helmi et al. 2018); this created a kinematic structure known as the Gaia-Sausage (Belokurov et al. 2018), and hence the event is often referred to as the Gaia-Enceladus-Sausage (GES). We also know that since then accretion has been primarily driven by minor mergers (Helmi 2020; Naidu et al. 2020; Deason & Belokurov 2024).

This relatively quiescent recent merger history has implications for key properties of the MW. For instance, it is consistent, and perhaps even necessary for the preservation of its thin, kinematically cold disc, which could be disrupted or thickened by a significant late merger (Toth & Ostriker 1992; Stewart et al. 2008). That said, recent hydrodynamical simulations suggest that thin discs can reform quickly even after a major merger if sufficient gas is available (Sotillo-Ramos et al. 2022). The LMM may also have influenced the chemo-dynamical properties of the Galaxy (e.g. Buck 2020; Buck et al. 2023; Renaud et al. 2021; Ciucă et al. 2024). Such a quiescent assembly also raises

questions about how typical the MW is within the standard Λ cold dark matter (Λ CDM) model, with some studies suggesting the MW is an outlier with respect to other local spirals (Hammer et al. 2007).

The idea that the MW is a typical L_* galaxy may also be challenged by the presence of two bright satellites, the Large and Small Magellanic Clouds (LMC and SMC, respectively). Observationally, such systems are uncommon; for example, Liu et al. (2011) found only 11% of their SDSS sample of MW-like galaxies to have a satellite of similar luminosity to the LMC, with only 3.5% having analogues of both clouds. More recently, in the SAGA survey (Mao et al. 2024) a third of the MW-like galaxies contain an LMC-mass satellite, and these systems often have richer satellite populations than the MW. Cosmological simulations also produce relatively few MW-mass haloes with sub-haloes massive enough to host the LMC and SMC (Boylan-Kolchin et al. 2011; Busha et al. 2011; Boylan-Kolchin et al. 2012).

There have been several recent attempts to estimate the likelihood of a MW-like halo experiencing a GES-like merger and simultaneously hosting massive satellites like the Magellanic Clouds. These studies have been carried out on large suites of dark-matter-only (DMO) and hydrodynamical simulations. For example, Evans et al. (2020) find that 5% of MW-mass haloes in the EAGLE simulations (Schaye et al. 2015) have a GES-like event and then no merger with an equally massive object after $z = 1$, while 16% have an LMC analogue at the present day; only

* Corresponding author: ewoudwempe@gmail.com

0.65% satisfy the two constraints simultaneously. The halo progenitors of these 0.65% of haloes appear to be much less massive than the average at early times (before the merger with the GES). In more recent work, Buch et al. (2024) searched cosmological DMO simulations for MW-like haloes that have both a GES-like event at early times and an LMC analogue at late times. Depending on the exact requirements, as few as 0.75% of all MW candidates were acceptable, of which only a small subset were in a group-like environment.

An important factor, often neglected in searches for MW analogues, may be the specific cosmic environment. While many papers consider relatively isolated systems, the real MW is part of the Local Group (LG) and has a massive companion (M31). The immediate surroundings of the LG may also influence galaxy assembly in a significant way. For instance, Santos-Santos et al. (2021) find that LMC-like satellites are about twice as likely around LG analogues in the APOSTLE simulations as around isolated galaxies (Sawala et al. 2016; Fattahi et al. 2016). Earlier assembly was also found in the constrained simulations of LG formation of Carlesi et al. (2020) and in the three LG analogues found by Santistevan et al. (2020) in their FIRE-2 hydrodynamical simulations. These studies all suggest that a proper understanding of the MW’s assembly history, and whether it is unusual, requires accounting for the LG context.

We studied the effect of environment on the mass assembly history of the MW (in particular, its embedding in the LG, of which it is not the most massive member) using the constrained simulations presented in Wempe et al. (2024, hereafter W24). By using a suite of simulations that match the detailed properties of the LG, we aim to understand whether the MW is expected to assemble early and how its environment has affected its growth history.

2. Methodology

2.1. W24 Local Group simulations

The constrained LG simulations presented in W24 constitute a fair and representative set of Λ CDM realisations conditional on properties of the LG; specifically, they are constrained to match the masses, positions and 3D velocities of the MW-M31 pair, and are also constrained by the observed velocity field within ~ 4 Mpc. Hence, if we compare these simulations with MW-mass or M31-mass galaxies identified in large cosmological simulations of random volumes, we can quantify the effect of their specific environment on the MW and M31 mass assembly histories.

For this work, we extracted a set of 224 LG realisations¹ from the chains² run in W24. For each sample, we took the initial condition field, added random small-scale perturbations at scales below those included in the chain simulations, and generated initial conditions at redshift $z = 63$ ³. We describe the exact procedure for the up-sampling of the initial conditions in Appendix A. We then integrated these initial conditions

¹ This selection was made from a homogeneous subsample of the chains, after excluding 15 cases (out of 239 total), where either the MW or M31 was part of a galaxy pair (defined as having a satellite with mass ratio > 0.3).

² We ran the chains from W24 twice as long to increase the effective sample size. In Appendix B we discuss in more detail the degree to which the realisations can be considered to be (quasi-)independent.

³ Although this is relatively low given that first-order LPT was used, the differences in mass growth history compared to simulations started at $z = 127$ are negligible.

Table 1. Numerical parameters used in the LG Gadget-4 simulations presented in this paper.

	HR region	LR region
Particle mass	$1.88 \times 10^7 M_{\odot}$	$9.65 \times 10^9 M_{\odot}$
Softening length	1.5 kpc/h	24 kpc/h
Box size	10 Mpc	40 Mpc

to $z = 0$ using Gadget-4 (Springel et al. 2021). The numerical parameters used in the gravity-only simulations are summarised in Table 1. Although the resolution of our re-simulations is relatively low, it is comparable to that of the Millennium-II simulation by Boylan-Kolchin et al. (2009, 2010); who demonstrated convergence for MW-mass halo growth histories by comparing to the higher resolution Aquarius level-2 simulation suite from Springel et al. (2008).

We used the friends-of-friends (FoF) group-finder, SUBFIND-HBT, and merger tree implementations bundled with Gadget-4 to obtain the group catalogues, sub-halo catalogues, and the merger trees, respectively. In this paper, we denote the virial mass as M_{200c} and define it as the mass within a sphere of average density $200\rho_c$, where ρ_c is the critical density of the Universe⁴. The virial masses for the Milky Way and M31 DM haloes are $M_{200c, MW} = 1.23^{+0.20}_{-0.22} \times 10^{12} M_{\odot}$ and $M_{200c, M31} = 2.08^{+0.40}_{-0.39} \times 10^{12} M_{\odot}$ (mean and 16th and 84th percentiles). As stated earlier, our MW and M31 haloes are constructed to match the observational constraints, with their masses reproduced within observational uncertainties in our posterior simulations with Gadget. The distances between the main haloes are 806^{+124}_{-133} kpc, they are infalling with radial velocities of $84^{+32}_{-31} \text{ km s}^{-1}$, and have tangential velocities of $27^{+17}_{-16} \text{ km s}^{-1}$. The spread in these quantities, especially the radial velocity and distance, is larger than the scatter in the low-resolution runs presented in W24. This is caused by the (significant) increase in resolution of the re-simulations, as well as the added entropy resulting from the augmenting of initial conditions with small-scale random fluctuations.

2.2. IllustrisTNG MW- and M31-mass haloes

To assess the effect of the Large Group environment, we compared our haloes with a sample of isolated MW- and M31-mass haloes extracted from the DMO variant of TNG50 (Nelson et al. 2019a,b; Pillepich et al. 2019). We extracted two subsets of the sub-halo catalogue to represent, respectively, MW-mass and M31-mass samples. For this, we required that the M_{200c} values, centred on the position of the main sub-halo’s most bound particle, be within 1σ of the masses of the MW and M31 analogues in the W24 sample. In practice, this means $M_{200c, MW} \in (0.98-1.45) \times 10^{12} M_{\odot}$ and $M_{200c, M31} \in (1.62-2.52) \times 10^{12} M_{\odot}$. Additionally, we required there to be no sub-halo of more than half the mass of the primary halo within 1 Mpc. This resulted in 52 isolated MW-mass and 53 isolated M31-mass haloes from TNG50-DMO. Because of the isolation criterion, all sub-haloes in our sample are the main sub-halo of their respective FoF

⁴ We also considered M_{100c} , for which all results in this paper remain qualitatively unchanged. Generally, the effects reported in this paper and associated with the environment are enhanced for M_{100c} as this encompasses mass in a more extended region.

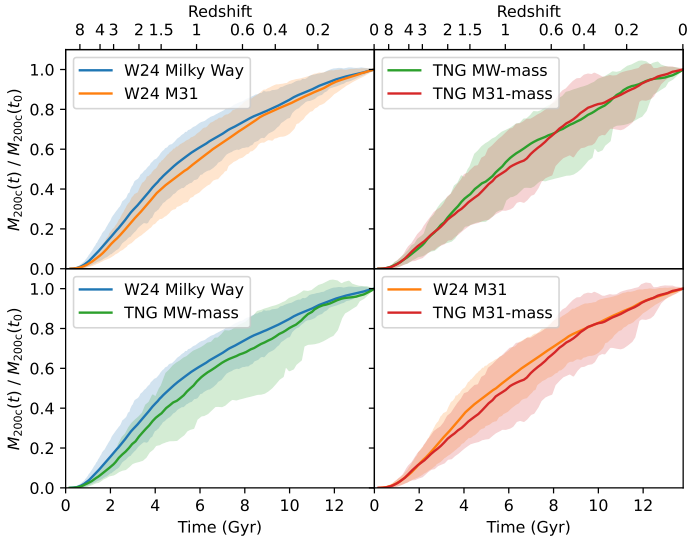


Fig. 1. Comparison of the mass growth histories of the MW and M31 galaxies from the constrained simulations of W24, compared to random MW- or M31-mass galaxies from TNG50-DMO. The central lines are the means, and shading indicates the 16–84 percentile regions. The unnormalised mass growth histories are shown in Appendix C.

group⁵. These samples are similar to, and mostly a subset of the ‘halo-based’ selection of MW/M31-mass galaxies (also from TNG50) by Pillepich et al. (2024), who selected the central sub-haloes of FoF groups with masses between $6 \times 10^{11} M_{\odot}$ and $2 \times 10^{12} M_{\odot}$. The only notable differences are the mass ranges (e.g. some of our M31-mass haloes are not present in their sample due to their $2 \times 10^{12} M_{\odot}$ upper limit), and that their sample was built on the full-physics version of TNG50 instead of TNG50-DMO.

We mostly used the Sublink merger trees part of TNG50’s public data release (Nelson et al. 2019b). This algorithm differs from the one used for the merger trees for the constrained LG simulations; however, the only quantity we used is the position of the most bound particle, which should be robust for the MW and M31 main progenitors across different sub-halo finders and linkers. Nonetheless, we found that for $\sim 10\%$ of objects, these merger trees missed a link in the main progenitor branch. We therefore decided to perform an additional correction step: if the sum of the progenitor masses of a particular sub-halo in the tree is less than half of the mass of the descendant (meaning it misses a massive progenitor), we searched in the neighbouring (comoving) 200 kpc for a sub-halo with a similar mass and, if found, manually added it to the merger tree. We verified via visual inspection of the snapshots that in all 22 such cases, the manually inserted extra link looked robust and that no wrong sub-haloes have been linked.

3. Results

3.1. Mass assembly history

We show the evolution of the virial mass M_{200c} of the MW/M31 haloes in Fig. 1. In the top-left panel, we plot the mass growth history normalised by the present-day mass in the constrained simulations. This figure shows that on average the MW assem-

⁵ In fact, if instead of the above selection we had simply required that the sub-halo be the central sub-halo of its FoF group, we would have found a very similar sample of 56 MW-mass and 56 M31-mass haloes.

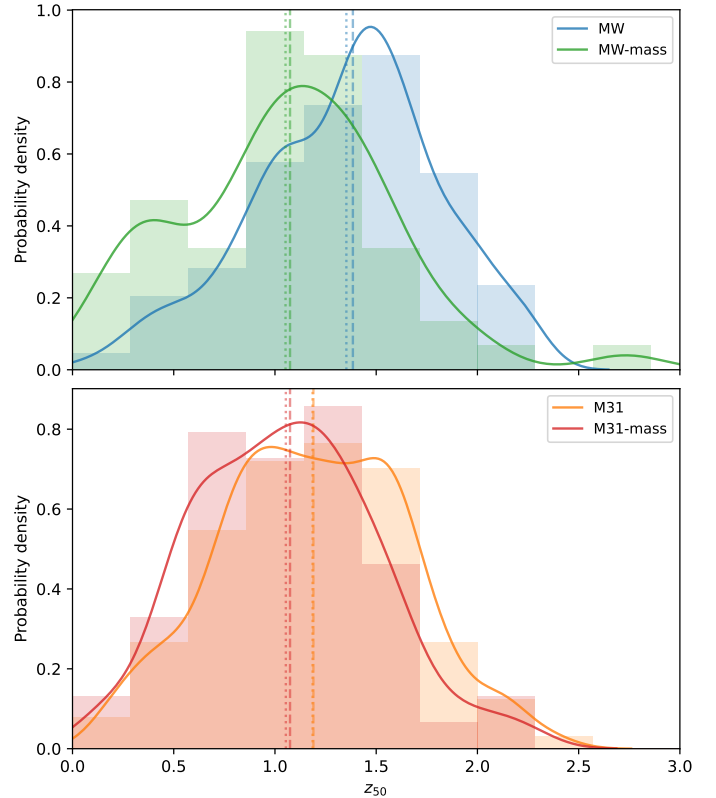


Fig. 2. Distribution of the redshifts at which half the final mass was accumulated (z_{50}) for the constrained MW and M31 haloes, compared to the random MW- or M31-mass haloes from TNG50-DMO.

bles a bit earlier than M31 in our simulations. Some of this bias towards early assembly could in principle be attributed to the MW being less massive (the ratio of the two virial masses is $M_{200c,MW}/M_{200c,M31} = 0.57^{+0.12}_{-0.16}$ in our suite) since, on average, heavier objects have a later assembly time in Λ CDM.

To analyse this finding in more depth, we considered the mass histories of MW- and M31-mass haloes from the TNG50 DMO variant, selected as described above. The top-right panel of Fig. 1 compares their mass growth histories; random MW-mass haloes indeed assemble slightly earlier on average. However, the amplitude of the effect is not enough to explain what is seen for our constrained simulations in the top-left panel. Furthermore, comparing the MW growth histories of the constrained realisation and the TNG samples (bottom-left panel), we see that the former are biased towards earlier assembly (although well within the spread), while for M31 the effect of the environment is not apparent, as shown in the bottom-right panel. In summary, it appears that the LG environment causes MW haloes to grow slightly faster at early times and more slowly at late times compared to their isolated counterparts. It also results in much less scatter in the allowed histories of the MW compared to the TNG sample of MWs. In particular, the tail of late-forming MWs that are present in the TNG sample is absent in our LG sample.

The distributions of the redshift at which 50% of the mass was attained, z_{50} , plotted in Fig. 2, support a similar view. In this figure, the top panel corresponds to the MWs and their comparison sample, and the bottom to the M31s. In 60% of cases z_{50} is larger (z_X is defined as the redshift at which $X\%$ of the present-day mass was in place), and in 69% z_{10} is larger, although the magnitude of the difference is $\lesssim 1\sigma$ of the scatter between realisations.

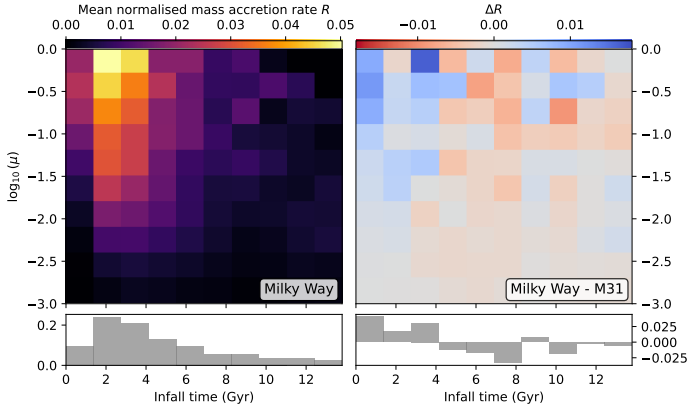


Fig. 3. Merger-associated mass accretion rates as a function of infall time and mass ratio (μ). For each bin, the peak masses of the infalling objects are summed and normalised by the total merger-associated mass accretion for the respective halo (MW or M31). This normalised mass accretion rate (R) is then averaged over simulations, with the resulting values shown by the colours. Left panel: This quantity for the MW. Right panel: Difference between the MW and M31. The histograms at the bottom represent the same quantity but marginalised over mass ratio, showing the 1D distribution as a function of infall time.

The mass growth history is at least in part driven by mergers and accretion events. In Fig. 3 we show the average mass accretion rates via mergers for our MWs as a function of mass ratio and infall time, as well as the difference between our MWs and M31s. The mass ratio is defined as the ratio of sub-halo to most massive progenitor mass when the sub-halo reaches its maximum mass, while the infall time refers to when the sub-halo first becomes part of the FoF group of the main halo’s progenitor. For the MWs, 52% of the total mass accretion can be attributed to mass bound to merging objects at some point in time, compared to 55% for M31. The remainder comes from unresolved mergers and smooth accretion. Our MWs exhibit a sharp peak in merger-associated mass accretion at early times, followed by a steep decline after the third bin at an infall time of 4 Gyr. In the difference plot on the right, this is reflected in the higher normalised mass accretion in the first three bins compared to our M31s. Conversely, at times later than 4 Gyr, the M31s display more accretion from both major and minor mergers, indicating that our MWs generally have a quieter recent merger history than our M31s. Quantitatively, when combining all infall events across our MWs and M31s, the median infall time of mass associated with major mergers (mass ratio $\mu > 0.25$) is at $t_{50,\mu>0.25} = 3.3$ Gyr, which is 0.4 Gyr earlier than for the M31s ($t_{50,\mu>0.25} = 3.7$ Gyr). Similarly, the median infall time for minor mergers ($0.1 < \mu < 0.25$) is $t_{50,0.1<\mu<0.25} = 3.6$ Gyr for the MWs, which is 0.7 Gyr earlier than for the M31 analogues ($t_{50,0.1<\mu<0.25} = 4.3$ Gyr). The bootstrap errors, estimated by resampling the merger events, are less than 0.16 Gyr, making the differences significant to 2.4σ for major mergers and 4.7σ for minor mergers. We also examined the time of the LMM (with mass ratio $\mu > 0.25$) of our MWs and M31s. The average time of the LMM ($\mu > 0.25$) shows no significant difference between the two haloes: the MWs have an average t_{LMM} of 3.13 Gyr while the M31s have 3.15 Gyr. However, t_{LMM} is a rather stochastic quantity, with a large simulation-to-simulation scatter of 2 Gyr, which could obscure any small signal.

We next identified GES-like mergers, which we defined as those satisfying the following criteria: (1) it is the last merger and had a mass ratio $\mu > 0.2$; (2) the object was fully destroyed

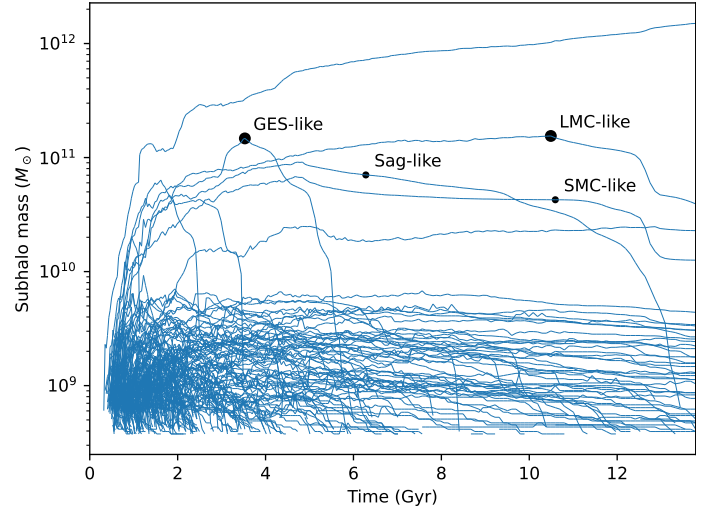


Fig. 4. Example of the mass growth history of all the sub-halos that end up in the MW’s FoF group. The top curve corresponds to the main progenitor of the MW. This particular simulation (C10 S910 in the notation of W24) has both a GES and an LMC analogue, which are indicated with the large black circles. A video of this particular simulation is available [online](#).

[2,6] Gyr after the Big Bang; and (3) the mass ratio is $0.2 < \mu < 0.5$. These criteria reflect our understanding of GES as the LMM of the MW, including its accretion time (e.g. Helmi 2020, and references therein), where we allow for substantial offset due to observational uncertainties as well as those stemming from extrapolating from gravity-only simulations. With these criteria we find that 72 out of our 224 simulations have such a merger.

Since the behaviour at late times is also interesting, we also searched for an LMC-like object, which we define as follows: (1) the LMC-sub-halo is the most massive at the present day; (2) it first crossed the MW’s r_{200c} more than 0.5 Gyr ago; (3) the mass ratio is $0.1 < \mu < 0.2$; and (4) its peak mass is higher than 5% of the mass of the main progenitor at present day. The mass criteria align with measurements of the mass of the LMC (Vasiliev 2023, and references therein). The 0.5Gyr infall criterion (2) is consistent with a wide variety of models of the LMC’s past trajectory, all of which consistently show the LMC has resided within r_{200c} for at least the past 0.5 Gyr (Vasiliev 2023, their Fig. 2). In this case, we find that 28 out of our 224 simulations have such an LMC. Furthermore, 12 out of 224 simulations have both an LMC-like object and a GES-like merger.

The mass growth history of one of these cases is shown in Fig. 4. The LMC analogue in this simulation actually comes with a companion, an SMC-analogue, and there is also a merger resembling a Sagittarius analogue and an SMC. This simulation does not, however, feature any major or minor mergers for M31 in the last 8 Gyr. The other 11 simulations with both an LMC-like and a GES-like object have less quiescent merger histories for M31, with one experiencing a minor merger 2 Gyr ago, as proposed to have taken place for M31 by D’Souza & Bell (2018). However, the MWs do not host an LMC-SMC pair in these realisations.

We quantify the effect of having a GES-like merger or LMC-like merger on the growth times of the MW halo in Table 2. A GES-like merger biases towards an earlier z_{50} (MWs with a GES-like merger have $z_{50} = 1.5^{+0.5}_{-0.5}$, while without a GES requirement we find $z_{50} = 1.4^{+0.5}_{-0.4}$), which makes sense, given that our selection criteria require no major mergers (and thus no

Table 2. Mean and 16th and 84th percentiles of mass growth histories.

Selection	N	z_{10}	z_{50}	z_{90}
Milky Way	224	$3.7^{+0.9}_{-0.9}$	$1.4^{+0.5}_{-0.4}$	$0.3^{+0.1}_{-0.2}$
w. LMC	28	$3.8^{+1.1}_{-1.0}$	$1.3^{+0.3}_{-0.4}$	$0.3^{+0.2}_{-0.1}$
w. GES	72	$4.0^{+0.9}_{-0.9}$	$1.5^{+0.5}_{-0.5}$	$0.3^{+0.1}_{-0.1}$
w. LMC+GES	12	$4.3^{+0.9}_{-0.9}$	$1.4^{+0.3}_{-0.3}$	$0.4^{+0.2}_{-0.1}$
M31	224	$3.2^{+0.7}_{-0.7}$	$1.2^{+0.4}_{-0.4}$	$0.3^{+0.2}_{-0.2}$
TNG-MW-mass	52	$3.1^{+0.6}_{-0.7}$	$1.1^{+0.5}_{-0.6}$	$0.3^{+0.2}_{-0.2}$
TNG-M31-mass	53	$3.2^{+0.9}_{-0.8}$	$1.1^{+0.4}_{-0.4}$	$0.3^{+0.2}_{-0.2}$

Notes. The number (N) of realisations (or objects) satisfying a certain selection criterion, and the redshifts at which 10%, 50%, 90% of their present-day M_{200c} values were attained.

merger-associated growth) after the GES-like merger. Requiring an LMC-like object shifts z_{50} towards lower values only slightly: although one would expect to see the effect of the LMC in the growth history, in our definition, the LMC is required to be the most massive object at the present day. Because of that, any late major mergers with mass ratio $\mu > 0.2$ are not considered in the subset, effectively disfavouring significant late growth. We find 12 simulations that have both a GES-like merger and an LMC analogue. In this case, as expected, z_{10} is biased early (and z_{90} in the opposite way), and also the scatter is somewhat reduced.

3.2. Understanding the growth in an LG environment

To further investigate how the MW and M31 have grown in our LG simulations, in the top panels of Fig. 5 we show the simulation presented in Fig. 4 and identify the positions at earlier times of the particles that end up in the MW and M31 FoF groups at the present day. In 65 out of 224 cases, the MW and M31 haloes end up in the same FoF group, and in these cases, the satellites are assigned to the closest halo. In this figure we have rotated and scaled the coordinate system such that the MW's main progenitor is positioned at $(x', y', z') = (0, 0, 0)$, and M31's main progenitor is at $(1, 0, 0)$ at all times. This leaves one free angle (the rotation around the MW–M31 axis), which we set such that the z' -axis aligns as closely as possible with the north pole of the Local Sheet (at super-galactic coordinates $(L, B) = (241.74^\circ, 82.05^\circ)$, McCall 2014), which roughly aligns with the super-galactic north pole (the separation is 7.95°). As a result, the z' -axis maintains close alignment with the Local Sheet north pole. This is not exact because of M31's slight elevation above the Local Sheet, but the misalignment is minor, averaging 23° at $z = 7$, and reducing to 12° at $z = 0$. The left column shows the distribution at redshift $z \sim 7$, which reveals a few narrow filaments already forming with low z' and a generally isotropic distribution around the two haloes. Then, at $z \sim 3.9$, a large part of the MW's material has collapsed into a sheet-like distribution, aligned with the z' direction, with a filament forming along the MW–M31 axis. At $z = 2$, the collapse along the z' direction is mostly complete, and further growth occurs mostly from within the sheet. By $z = 1$, this halo has accreted more than half of its material, and the distribution is again more axisymmetric around the MW–M31 axis.

Although this is a single simulation, this behaviour is similar across all LG realisations, as depicted in Fig. 6. In this figure we stack all the MW and M31 progenitor particles of all simulations to get a map of the posterior mean behaviour of the positions of the progenitor particles over time (note that all simulations are

samples from the LG posterior of W24, and hence stacking them gives the posterior mean field). At $z = 7$, the progenitor particle distribution is, on average, elongated along the z' -axis (which is roughly aligned with the super-galactic z -axis). Then up until $z \sim 2$, it collapses along the z' -axis, becoming most compressed along this axis. After $z = 2$, the halo material starts to virialise and the distribution becomes roughly spherically symmetric by $z = 1$. The separatrix between the MW and M31 material is shown with black lines. It lies at $x' = 0.4$ along the MW–M31 axis at $z = 7.0$; particles with $x' < 0.4$, are more likely to become part of the MW, and those with $x' > 0.4$ are more likely to become part of M31. Its location at $x' = 0.4$ stays roughly constant throughout cosmic time, with some scatter ranging from 0.37 to 0.45. Interestingly, this range is expected from a simple computation of the location of the Lagrange point L1 in a system of two point masses (without rotation, as the tangential velocities are negligible) and with a mass ratio of 0.57 (which is the median value for our simulations).

Using Fig. 6 we can point out the effect that suppresses the MW's mass growth at late times: due to the higher mass of M31, the particles that are right of the separatrix will go to M31 instead of the MW, and this especially suppresses late-time growth, when particles from larger initial distances dominate the mass growth. This is also seen in the un-normalised mass growth histories, shown in Appendix C, which at early times largely coincide and start to deviate at later times.

4. Discussion and conclusions

The MW haloes in our constrained LG simulations generally form earlier and grow more slowly at late times than their isolated counterparts. This is principally due to the presence of M31 (which is more massive by a factor of approximately 1.75 in the median) but may also be due in part to the quiet surrounding Hubble flow. The growth of both galaxy haloes is also quite anisotropic in our constrained simulations. Proto-halo material first collapses into a planar distribution, which is roughly aligned with the Local Sheet, and then, after $z \sim 2$, starts to virialise, ending up in a significantly less flattened configuration by $z \sim 1$.

Quantitatively, our MWs reach half their final mass at $z_{50, \text{MW}} = 1.4^{+0.5}_{-0.4}$, the M31s at $z_{50, \text{M31}} = 1.2^{+0.4}_{-0.4}$, and the MW-mass and M31-mass isolated haloes from TNG50-DMO at $(z_{50, \text{MW-mass}}, z_{50, \text{M31-mass}}) = (1.1^{+0.5}_{-0.6}, 1.1^{+0.4}_{-0.4})$. Requiring the LMM to be GES-like biases the MW to $z_{50, \text{MW} | \text{GES}} = 1.5^{+0.5}_{-0.5}$. The early- and late-time behaviour can be quantified as $(z_{10, \text{MW}}, z_{10, \text{M31}}) = (3.7^{+0.9}_{-0.9}, 3.2^{+0.7}_{-0.7})$ and $(z_{90, \text{MW}}, z_{90, \text{M31}}) = (0.3^{+0.1}_{-0.2}, 0.3^{+0.2}_{-0.2})$.

These values are not very different from those of Buch et al. (2024) for MW-like isolated haloes. They find $a_{50} = 0.49$ ($z_{50} = 1.05$) for a similar sample of MW-mass objects without the requirement of an LMC (in which case $a_{50} = 0.57$, $z_{50} = 0.75$), although on average, our MWs reach their half-mass slightly earlier. In the ELVIS suite of LG analogues, Garrison-Kimmel et al. (2019) and Santistevan et al. (2020) find that for isolated haloes $z_{10} \sim 3$ and for LG-like hosts $z_{10} \sim 4$, which is in agreement with our results. These authors argue that at later times, $z \lesssim 2$, the mass growth histories of isolated and LG-like hosts are more similar, which might indicate that the initial overdensity associated with the LG in comparison to a random MW or M31 progenitor plays a role.

Similar findings are reported for the HESTIA simulations of LG-like pairs inside constrained simulations of the large-scale structure (with constraints on scales larger than ~ 4 Mpc;

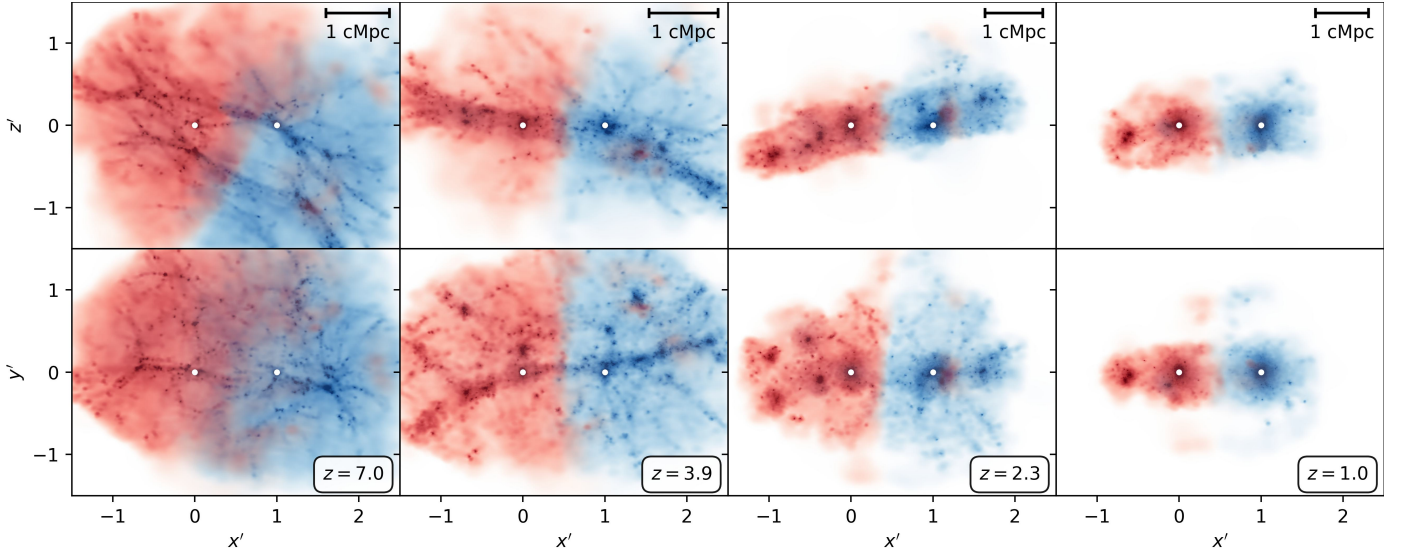


Fig. 5. Particles that end up in the MW sub-halo (in red) and in the M31 (in blue) by the present day, traced back to various redshifts for one of our realisations. Each column shows a different redshift. The coordinate system has been rotated and scaled such that the MW and M31 main progenitors are at (0,0,0) and (1,0,0), respectively, at all times (indicated with white dots). Top row: Approximately edge-on view of the Local Sheet. Bottom row: Face-on view.

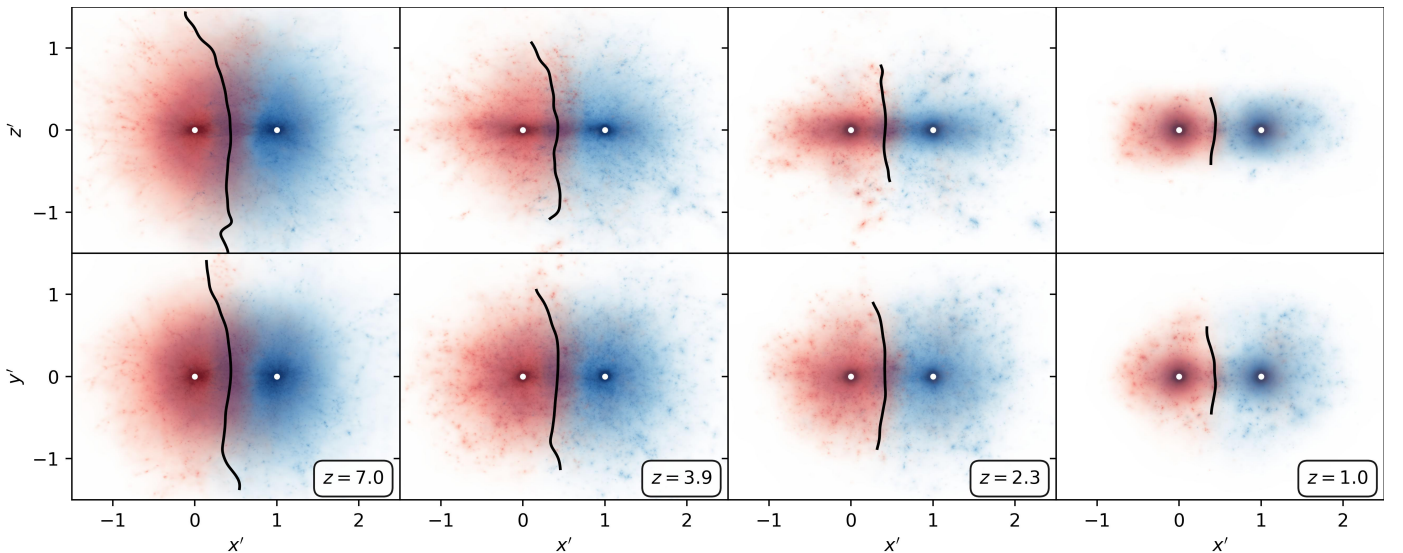


Fig. 6. Like Fig. 5 but now for the posterior mean of our simulations, a ‘stack’ of all LG re-simulations. The black lines are the isocontours at which a particle is equally likely to end up in the MW or M31. A video version of this figure is available [online](#).

Libeskind et al. 2020). In these simulations, on average $z_{50} = 1.4$, indicating that these haloes form roughly one billion years earlier than in unconstrained similarly massed MW or M31 analogues (whose z_{50} is ~ 1.1). Note, however, that these authors make no distinction between MW and M31 haloes in their estimates of z_{50} , so these numbers reflect the effect of a group environment and not differences between the two systems. In a sample of LGs extracted from CLUES (Carlesi et al. 2016), Carlesi et al. (2020) find that the median half-mass time of galaxy haloes in LGs in constrained realisations is also earlier than in LGs without constraints, although the difference is small, of the order of 0.5 Gyr both for the MWs and M31s. Carlesi et al. (2020) notice that their MWs reach their half-mass earlier than their M31s (by 0.25 Gyr) but attribute this difference to mass. We, on the other hand, find that our MWs also have a shorter average half-mass time than isolated MW-mass

galaxies. The difference between MWs and M31s we find is also larger, $\langle t_{50,M31} \rangle - \langle t_{50,MW} \rangle = 0.6$ Gyr. This could plausibly be explained by the separation between our LG pairs, which is 806^{+124}_{-133} kpc, whereas Carlesi et al. (2020) allow for a much larger range of 0.5 – 1.9 Mpc. This may imply a larger separation between MW and M31 on average in their sample (see e.g. Table 7 of Carlesi et al. 2016), which would reduce the effect of M31 on the MW.

These findings may have implications for the formation of the Galactic disc, as explored by Semenov et al. (2024) using the TNG50 simulations. It appears that relatively few MW-mass disc galaxies in typical environments assembled a significant fraction of their mass within the first few billion years and formed a disc that has survived until the present day (fewer than 10%). Nevertheless, this seems to have happened in the real MW. However, we find that once the LG structure and its local environment are

correctly reproduced, such behaviour is common. Increasing the chance of such an early assembly history requires (i) a quiet environment surrounding the LG and (ii) a relatively large mass ratio of M31 to the MW; both factors reduce the likelihood of significant late-time growth.

Possible improvements to the analysis presented here would include using a larger sample to increase the numerical precision of our estimates. This could also allow us to be stricter in selecting close analogues of the MW and M31, for example by identifying LMC-like objects that also reproduce the LMC's present-day distance and approximate orbit. For a high-quality match, however, a preferable approach would be to add the LMC (or M33) as an extra constraint to the W24 methodology. Another possibility would be to constrain the spin orientation of the inner MW and M31 haloes to match the observed orientations of the two galaxies. A further obvious improvement would be to carry out hydrodynamical galaxy formation simulations, as they would allow us to follow the mass growth of all the different constituents of the galaxies, i.e. cold gas, hot gas, and stars, as well as the dark matter. We are currently working on such a suite of simulations and hope to report on them in future work.

Data availability

Movies associated to Figs. 4 and 6 are available at <https://www.aanda.org>

Acknowledgements. This work has been financially supported by a Spinoza Prize from NWO to AH. EW thanks Dylan Nelson for providing access to the TNG simulations, Anna Genina and Victor Forouhar Moreno for helpful discussions, and Akshara for her moral support. We are grateful to the referee for a constructive report. The Gadget simulations were performed on the Freya cluster at the Max Planck Computing and Data Facility. The constrained simulations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at the PDC Center for High Performance Computing, KTH Royal Institute of Technology, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. JJ and GL acknowledge support from the Simons Foundation through the Simons Collaboration on “Learning the Universe”. This work was made possible by the research project grant “Understanding the Dynamic Universe”, funded by the Knut and Alice Wallenberg Foundation (Dnr KAW 2018.0067). Additionally, JJ acknowledges financial support from the Swedish Research Council (VR) through the project “Deciphering the Dynamics of Cosmic Structure” (2020-05143). GL acknowledges support from the CNRS-IEA “Manticore” project. This work is done within the Aquila Consortium (<https://www.aquila-consortium.org/>). *Software:* During this work, we have made use of various software packages. This includes BORG (Jasche & Wandelt 2013; Jasche & Lavaux 2019), GADGET-4 (Springel et al. 2021), JAX (Bradbury et al. 2018), NUMPYRO (Phan et al. 2019), XARRAY (Hoyer & Hamman 2017; Hoyer et al. 2023), NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), H5PY (<https://www.h5py.org/>), ASTROPY (Astropy Collaboration 2018, 2022), JUG (Coelho 2017).

References

Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, 156, 123
 Astropy Collaboration (Price-Whelan, A. M., et al.) 2022, *ApJ*, 935, 167
 Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, 478, 611

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, *MNRAS*, 398, 1150
 Boylan-Kolchin, M., Springel, V., White, S. D. M., & Jenkins, A. 2010, *MNRAS*, 406, 896
 Boylan-Kolchin, M., Besla, G., & Hernquist, L. 2011, *MNRAS*, 414, 1560
 Boylan-Kolchin, M., Bullock, J. S., & Kaplinghat, M. 2012, *MNRAS*, 422, 1203
 Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, *JAX: Composable Transformations of Python+NumPy Programs*
 Buch, D., Nadler, E. O., Wechsler, R. H., & Mao, Y.-Y. 2024, *ApJ*, 971, 79
 Buck, T. 2020, *MNRAS*, 491, 5435
 Buck, T., Obreja, A., Ratcliffe, B., et al. 2023, *MNRAS*, 523, 1565
 Busha, M. T., Marshall, P. J., Wechsler, R. H., Klypin, A., & Primack, J. 2011, *ApJ*, 743, 40
 Carlesi, E., Sorce, J. G., Hoffman, Y., et al. 2016, *MNRAS*, 458, 900
 Carlesi, E., Hoffman, Y., Gottlöber, S., et al. 2020, *MNRAS*, 491, 1531
 Ciucă, I., Kawata, D., Ting, Y.-S., et al. 2024, *MNRAS*, 528, L122
 Coelho, L. P. 2017, *J. Open Res. Software*, 5, 1
 Deason, A. J., & Belokurov, V. 2024, *New Astron. Rev.*, 99, 101706
 D’Souza, R., & Bell, E. F. 2018, *Nat. Astron.*, 2, 737
 Evans, T. A., Fattahi, A., Deason, A. J., & Frenk, C. S. 2020, *MNRAS*, 497, 4311
 Fattahi, A., Navarro, J. F., Sawala, T., et al. 2016, *MNRAS*, 457, 844
 Gaia Collaboration (Brown, A. G. A., et al.) 2018, *A&A*, 616, A1
 Gaia Collaboration (Vallenari, A., et al.) 2023, *A&A*, 674, A1
 Garrison-Kimmel, S., Wetzel, A., Hopkins, P. F., et al. 2019, *MNRAS*, 489, 4574
 Hammer, F., Puech, M., Chemin, L., Flores, H., & Lehnert, M. D. 2007, *ApJ*, 662, 322
 Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357
 Helmi, A. 2020, *ARA&A*, 58, 205
 Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Nature*, 563, 85
 Hoyer, S., & Hamman, J. 2017, *J. Open Res. Software*, 5, 1
 Hoyer, S., Roos, M., Joseph, H., et al. 2023, <https://doi.org/10.5281/zenodo.10023467>
 Hunter, J. D. 2007, *Comput. Sci. Eng.*, 9, 90
 Jasche, J., & Lavaux, G. 2019, *A&A*, 625, A64
 Jasche, J., & Wandelt, B. D. 2013, *MNRAS*, 432, 894
 Libeskind, N. I., Carlesi, E., Grand, R. J. J., et al. 2020, *MNRAS*, 498, 2968
 Liu, L., Gerke, B. F., Wechsler, R. H., Behroozi, P. S., & Busha, M. T. 2011, *ApJ*, 733, 62
 Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, 154, 94
 Mao, Y.-Y., Geha, M., Wechsler, R. H., et al. 2024, *ApJ*, 976, 117
 McCall, M. L. 2014, *MNRAS*, 440, 405
 Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, *ApJ*, 901, 48
 Nelson, D., Pillepich, A., Springel, V., et al. 2019a, *MNRAS*, 490, 3234
 Nelson, D., Springel, V., Pillepich, A., et al. 2019b, *Comput. Astrophys. Cosmol.*, 6, 2
 Phan, D., Pradhan, N., & Jankowiak, M. 2019, ArXiv e-prints [arXiv:1912.11554]
 Pillepich, A., Nelson, D., Springel, V., et al. 2019, *MNRAS*, 490, 3196
 Pillepich, A., Sotillo-Ramos, D., Ramesh, R., et al. 2024, *MNRAS*, 535, 1721
 Renaud, F., Agertz, O., Andersson, E. P., et al. 2021, *MNRAS*, 503, 5868
 Santistevan, I. B., Wetzel, A., El-Badry, K., et al. 2020, *MNRAS*, 497, 747
 Santos-Santos, I. M. E., Fattahi, A., Sales, L. V., & Navarro, J. F. 2021, *MNRAS*, 504, 4551
 Sawala, T., Frenk, C. S., Fattahi, A., et al. 2016, *MNRAS*, 457, 1931
 Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, 446, 521
 Semenov, V. A., Conroy, C., Smith, A., Puchwein, E., & Hernquist, L. 2024, ArXiv e-prints [arXiv:2409.18173]
 Sotillo-Ramos, D., Pillepich, A., Donnari, M., et al. 2022, *MNRAS*, 516, 5404
 Springel, V., Wang, J., Vogelsberger, M., et al. 2008, *MNRAS*, 391, 1685
 Springel, V., Pakmor, R., Zier, O., & Reinecke, M. 2021, *MNRAS*, 506, 2871
 Stewart, K. R., Bullock, J. S., Wechsler, R. H., Maller, A. H., & Zentner, A. R. 2008, *ApJ*, 683, 597
 Toth, G., & Ostriker, J. P. 1992, *ApJ*, 389, 5
 Vasiliev, E. 2023, *Galaxies*, 11, 59
 Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nat. Methods*, 17, 261
 Wempe, E., Lavaux, G., White, S. D. M., et al. 2024, *A&A*, 691, A348

Appendix A: Up-sampling of initial condition fields

In W24 we utilised a zoom scheme that involved splitting the initial condition field into two parts, a field for the low-resolution large box, s_{LR} , and a field for the high-resolution zoom region, s_{HR} . The zoom region has a box size of 10 Mpc, with 64 elements on a side, giving s_{HR} a mesh cell size of $\frac{10 \text{ Mpc}}{64} = 156.25 \text{ kpc}$. For our re-simulations, we used a grid twice as fine for the zoom region while leaving the outer region unchanged. Therefore, we up-sampled the s_{HR} field, from its original 64^3 size to a 128^3 grid, preserving all the Fourier modes of the original s_{HR} field, and randomly generating the modes at frequencies larger than the Nyquist frequency (per axis) of s_{HR} .

To accomplish this, we combined two fields: our original field s_{HR} , which sets the larger-scale modes, and a randomly generated field ($s_{\text{ss}} \sim \mathcal{N}(0, 1)$) that sets the smallest-scale Fourier modes. To ensure that the final field is real-valued, we used the Hartley transform, which guarantees a real-valued output for a real-valued input. The Hartley transform (\hat{s}) of a real field (s) was computed from the Fourier transform ($\hat{s} = \mathcal{H}s = \text{Re } \mathcal{F}s - \text{Im } \mathcal{F}s$). We set our new field's modes by taking \hat{s}_{HR} modes at frequencies lower than the Nyquist frequency of s_{HR} (the large scales) and from \hat{s}_{ss} at higher frequencies (the small scales). At the Nyquist frequencies, the modes are split equally across the positive and negative Fourier modes in the new field, to ensure that, if $s_{\text{ss}} = 0$, the process corresponds to pure trigonometric interpolation. Hence, when one, two, or three indices are $\pm i_{\text{Ny}}$, the contribution of \hat{s}_{HR} is weighted by $a = \frac{1}{2}, \frac{1}{4},$ and $\frac{1}{8}$, respectively. Since the variance of a mode $X = a\mathcal{N}(0, 1) + b\mathcal{N}(0, 1)$ is given by $\sigma^2 = a^2 + b^2$, to conserve unit variance at these modes, the coefficient b for the small-scale field \hat{s}_{ss} was set to $b = \sqrt{1 - a^2}$, resulting in $\sqrt{\frac{3}{4}}, \sqrt{\frac{15}{16}},$ and $\sqrt{\frac{63}{64}}$ for the three cases, respectively. Summarising, the modes of the new $s_{\text{HR}}^{\text{new}}$ field are:

$$\hat{s}_{\text{HR}}^{\text{new}} = \begin{cases} \hat{s}_{\text{HR}} & \text{if the absolute values of all indices are} \\ & \text{less than } i_{\text{Ny}}, \\ \hat{s}_{\text{ss}} & \text{if the absolute value of any index} \\ & \text{exceeds } i_{\text{Ny}}, \\ \frac{1}{2}\hat{s}_{\text{HR}} + \sqrt{\frac{3}{4}}\hat{s}_{\text{ss}} & \text{if exactly one index is } \pm i_{\text{Ny}}, \\ \frac{1}{4}\hat{s}_{\text{HR}} + \sqrt{\frac{15}{16}}\hat{s}_{\text{ss}} & \text{if exactly two indices are } \pm i_{\text{Ny}}, \\ \frac{1}{8}\hat{s}_{\text{HR}} + \sqrt{\frac{63}{64}}\hat{s}_{\text{ss}} & \text{if all indices are } \pm i_{\text{Ny}}, \end{cases}$$

where the indices range from $-N_{\text{ss}}/2$ to $N_{\text{ss}}/2$, $i_{\text{Ny}} = N_{\text{HR}}/2$ is the Nyquist index of the \hat{s}_{HR} field, and the last three cases apply only if no index exceeds i_{Ny} in absolute value.

Appendix B: Auto-correlation analysis

In this work we used initial conditions inferred from Hamiltonian Monte Carlo chains with relatively long auto-correlation lengths. As a result, some quantities extracted from the simulations may exhibit auto-correlation, which would increase the numerical error in the posterior means reported in this paper. For independent samples of a quantity Q , the standard deviation of the mean is $\sigma_{\langle Q \rangle} = \frac{\sigma_Q}{\sqrt{N}}$. For correlated samples, this becomes

$\sigma_{\langle Q \rangle} = \frac{\sigma_Q}{\sqrt{N_{\text{eff}}}}$, where N_{eff} is the effective sample size. The effective sample size is calculated as $N_{\text{eff}} = N/(1 + 2 \sum_{d=1}^{\infty} \rho(d))$, where $\rho(d)$ is the auto-correlation function of Q at lag d .

In W24, the auto-correlation function of the initial condition fields, $\rho_s(d)$, was found to follow an exponential decay, $\rho(d) = e^{-d/L}$, with auto-correlation length L . Fitting the model to the 12

Hamiltonian Monte Carlo chains resulted in an effective sample size of $N_{\text{eff}} = 20$ for the initial condition field.⁶ For this study we ran the same 12 chains for longer; this produced an effective sample size of the initial condition field of $N_{\text{eff}} = 32$. From these chains, we extracted 224 realisations.

It is important to note that for quantities sensitive to the behaviour on small scales (e.g. below 1 Mpc), the correlation lengths between samples will be smaller; hence, the 224 realisations could be seen in this context as (quasi-)independent. We explain below how to quantify the numerical uncertainties on the various quantities discussed in this paper taking into account the quasi-independence of our samples.

As an example, we determined the effective sample size of a_{50} . We assumed an exponential auto-correlation function (Ae^{-d/L_s}) and that auto-correlation lengths (L_s) on small scales are a factor f smaller than those for the initial conditions, i.e. $L = f \times L_s$. To estimate L_s , we fitted the auto-correlation function using NumPyro. We assumed a_{50} follows a log-normal distribution, $\ln a_{50} \sim \mathcal{N}(\mu, \sigma^2)$, with an exponentially declining auto-correlation function:

$$\rho_{a_{50}}(d) = \langle (a_{50}[i] - \mu)(a_{50}[i + d] - \mu) \rangle = Ae^{-d/L_s}. \quad (\text{B.1})$$

We find that $f = 0.70 \pm 0.17$ and $A = 0.41 \pm 0.7$, giving an effective sample size of $N_{\text{eff}} = 73 \pm 13$. The Monte Carlo standard error on a_{50} is 0.01 for both the MW and M31, and the earlier average assembly reported in Section 3.1 is significant at the 1.7% level.

For the a_{50} of the mass bound to the main sub-halo, the difference between MW and M31 is even more significant, at 0.4%. In this case, we removed the 65 cases in which the MW and M31 ended up in the same FoF group, to avoid potential biases due to the differing behaviour in the Subfind unbinding procedure in the same-FoF-group and different-FoF-group scenarios. Specifically, the Subfind unbinding procedure only considers particles within the same FoF group, meaning the tidal effect of M31 lowers the MW mass in the same-FoF-group case, but not in the different-FoF-group scenario.

Appendix C: Un-normalised mass growth histories

Instead of the normalised growth histories shown in Fig. 1, one can also consider the un-normalised growth histories, which are shown in Fig. C.1. Note that the MW and M31 have very similar mass assembly histories at early times, diverging only around redshift 4.

⁶ This differs from the $N_{\text{eff}} = 27$ reported in W24 due to a programming error in the effective sample size calculation in that paper. The results of W24 are unaffected, as all quoted uncertainties reflect the posterior widths, which are independent of the effective sample size. The corrected effective sample size of $N_{\text{eff}} = 20$ inflates the predictive uncertainties on inferred quantities by 2.5%, compared to 1.7% with $N_{\text{eff}} = 27$. This difference remains negligible.

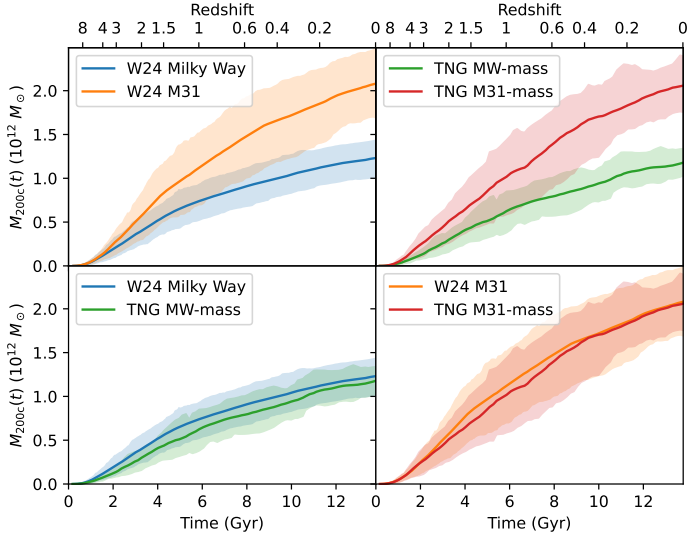


Fig. C.1. Top-left panel: Like Fig. 1 but un-normalised. Note that at early times, the MW and M31 had a very similar mass growth history, diverging from $z \sim 4$. The larger scatter in the present-day M31 mass reflects the larger (observational) uncertainties on its mass from W24. Top-right panel: Same but for MW-mass and M31-mass galaxies from an unconstrained simulation. Bottom panels: Comparisons of the MWs and M31s from the constrained simulations with isolated MW- or M31-mass haloes.