

The Local Group’s mass: probably no more than the sum of its parts

Till Sawala¹,^{1,2} Meri Teeriaho¹ and Peter H. Johansson¹

¹*Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2, FI-00014 Helsinki, Finland*

²*Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*

Accepted 2023 March 17. Received 2023 February 13; in original form 2022 October 13

ABSTRACT

The total mass of the Local Group and the masses of its primary constituents, the Milky Way (MW) and M31, are important anchors for several cosmological questions. Recent independent measurements have consistently yielded halo masses close to $10^{12}M_{\odot}$ for the MW, and $1\text{--}2 \times 10^{12}M_{\odot}$ for M31, while estimates derived from the pair’s kinematics via the ‘timing argument’ have yielded a combined mass of around $5 \times 10^{12}M_{\odot}$. We analyse the extremely large UCHUU simulation to constrain the mass of the Local Group and its two most massive members. First, we demonstrate the importance of selecting pairs whose kinematics reflect their mutual interactions. Adopting the observed separation and radial velocity, we obtain a weighted posterior of $75_{-40}^{+65} \text{ km s}^{-1}$ for the uncertain transverse velocity. Via Gaussian process regression, we infer a total mass of $3.2_{-0.9}^{+1.2} \times 10^{12}M_{\odot}$, significantly below the timing argument value. Importantly, the remaining uncertainty is not rooted in the analysis or observational errors, but in the irreducible scatter in the kinematics–mass relation. We further find a mass for the less massive halo of $0.9_{-0.3}^{+0.6} \times 10^{12}M_{\odot}$ and for the more massive halo of $2.3_{-0.9}^{+1.0} \times 10^{12}M_{\odot}$, consistent with independent measurements of the masses of MW and M31, respectively. Incorporating the MW mass as an additional prior let us constrain all measurements further and determine that the MW is very likely less massive than M31.

Key words: methods: numerical – methods: statistical – Galaxy: kinematics and dynamics – Local Group – dark matter.

1 INTRODUCTION

The total mass of our own Milky Way (MW), the neighbouring Andromeda galaxy (M31), and the Local Group (LG) are key to the interpretations of many observations. For example, the abundance of massive MW satellites is consistent with Lambda cold dark matter (Λ CDM) predictions only if its virial mass is not significantly higher than $\sim 10^{12}M_{\odot}$ (Boylan-Kolchin, Bullock & Kaplinghat 2012; Sawala et al. 2015). More generally, through the predicted amount of substructure, the mass of the MW also places constraints on the dark matter particle mass (Kennedy et al. 2014; Lovell et al. 2014). If the mass of the MW or M31 is much lower than $\sim 10^{12}M_{\odot}$, abundance matching would imply that they are clear outliers in the galaxy–halo relation (Guo et al. 2011). The MW also contains the best-studied ‘plane of satellite’ (Pawlowski 2018), whose significance as a challenge to the dark matter paradigm and its nature as a possibly long-lived disc may depend both on the LG’s mass assembly history (Libeskind et al. 2010) and the depth of the potential in which the satellites evolve (Sawala et al. 2022). Beyond satellite galaxies, the dwarf galaxies in the LG and their comparison to galaxies in other environments have also received attention, and the total mass of the LG is important for understanding the environment in which they form (e.g. Font et al. 2022).

Much attention has therefore been devoted in recent years to measuring the mass of the MW, M31, and the LG using different techniques. However, while measurements of each individual galaxy, and of the MW in particular, have found masses of the order of $\sim 10^{12}M_{\odot}$, dynamical measurements of the total mass typically yield

masses of $\sim 5 \times 10^{12}M_{\odot}$, significantly higher than the sum of both individual masses.

In this work, we use kinematics and masses of LG analogues from the UCHUU simulation (Ishiyama et al. 2021), the largest N -body simulation of sufficient resolution, to determine the relationship between the observable kinematics and the total mass. We show how the estimates can be further constrained assuming independent measurements for either the MW or M31, and how individual constraints, combined with the kinematics and the mass ratio distribution, can be used to infer that the mass of the MW is below that of M31. We also revisit the ‘Timing Argument’, and explore whether there is, in fact, a discrepancy between the kinematics of the LG and the mass estimates for individual objects.

This paper is organized as follows. In Section 2, we review previous measurements of the individual masses and of the total mass of the LG, highlighting the apparent discrepancy. In Section 3, we describe the simulations and our selection of LG analogues. In Section 4, we introduce ‘tidal dominance’ and ‘force ratio’ as important criteria for linking the mass to the kinematics. We present our results for estimating the mass of the LG using Gaussian regression in Section 5. In Section 6, we compute probability densities for the masses of the LG, MW, M31 and the mass ratios, based on a probability-weighted analysis. Finally, we conclude with a summary in Section 7.

2 PREVIOUS MEASUREMENTS

The flurry of measurements of the MW mass over the past 3 yr reflect the significance of the result, as well as the wealth of newly available data, particularly from *Gaia* DR2 and, most recently, *Gaia* DR3.

* E-mail: till.sawala@helsinki.fi

Excluding estimates that derive the mass of the MW from LG kinematics, recent estimates of the virial mass of the MW include $1.17_{-0.15}^{+0.21} \times 10^{12} M_{\odot}$ using *Gaia* DR2 satellite dynamics (Callingham et al. 2019), $1.54_{-0.44}^{+0.75} \times 10^{12} M_{\odot}$ using combined *Gaia* DR2 and *HST* kinematics of globular clusters (Watkins et al. 2019), $1.51_{-0.40}^{+0.45} \times 10^{12} M_{\odot}$ using *Gaia* DR2 proper motions of satellite galaxies (Fritz et al. 2020), and $0.89_{-0.08}^{+0.1} \times 10^{12} M_{\odot}$ derived from the kinematics of disc and halo stars in the galkin catalogue (Karukes et al. 2020).

Further mass estimates for the MW include $1.23_{-0.18}^{+0.21} \times 10^{12} M_{\odot}$ derived from *Gaia* DR2 satellite kinematics and a simulation-based distribution function (Li et al. 2020), $1.08_{-0.14}^{+0.20} \times 10^{12} M_{\odot}$ from the *Gaia* DR2 rotation curve (Cautun et al. 2020), $0.822 \pm 0.052 \times 10^{12} M_{\odot}$ from observations of classical Cepheids (Ablimit et al. 2020), $1.16 \pm 0.24 \times 10^{12} M_{\odot}$ [including the mass of the Large Magellanic Cloud (LMC)] from a large sample of halo stars (Deason et al. 2021), $1.19_{-0.32}^{+0.49} \times 10^{12} M_{\odot}$ from a Bayesian estimate using dwarf galaxy kinematics from multiple sources (Slizewski et al. 2022), $1.26_{-0.22}^{+0.40} \times 10^{12} M_{\odot}$ from high-velocity RR-Lyrae stars (Prudil et al. 2022), $1.08_{-0.11}^{+0.12} \times 10^{12} M_{\odot}$ from the H3 survey and *Gaia* EDR3 (Shen et al. 2022), and finally $1.1_{-0.1}^{+0.1} \times 10^{12} M_{\odot}$ from *Gaia* EDR3 proper motions of satellites when compared to Λ CDM simulations (Rodriguez Wimberly et al. 2022). We refer the reader to Wang et al. (2020) for a comprehensive review of earlier results. Analyses from several different tracers, and particularly from the latest studies using the most precise *Gaia* EDR3 observations, consistently point towards an MW virial mass that is very close to $10^{12} M_{\odot}$.

While M31 does not enjoy the benefit of *Gaia* proper motions, there exist nevertheless a number of studies of its mass. Again, excluding measurements that use LG kinematics, previous results include $1.4 \pm 0.4 \times 10^{12} M_{\odot}$ derived from satellite kinematics (Watkins, Evans & An 2010), $1.2_{-0.7}^{+0.9} \times 10^{12} M_{\odot}$ derived from kinematics of M31 dwarf spheroidals (Tollerud et al. 2012), $\sim 1.05_{-0.15}^{+0.15} \times 10^{12} M_{\odot}$ (combining the DM and stellar mass) from SED fitting together with the rotation curve and the kinematics of outer globular clusters and satellite galaxies (Tamm et al. 2012), $2.0_{-0.3}^{+0.4} \times 10^{12} M_{\odot}$ from kinematics of the Giant southern stream (Fardal et al. 2013), $1.35_{-0.15}^{+0.15} \times 10^{12} M_{\odot}$ (Veljanoski et al. 2013) and $1.4_{-0.2}^{+0.2} \times 10^{12} M_{\odot}$ (Veljanoski et al. 2014), both derived from outer halo globular clusters, and $0.8 \pm 0.1 \times 10^{12} M_{\odot}$ estimated from high-velocity planetary nebulae (Kaffe et al. 2018). Compared to the MW, the scatter in mass estimates is larger, but most estimates are in the range of $1\text{--}2 \times 10^{12} M_{\odot}$.

Combining the best individual estimates for the masses of the MW and M31 thus leads to a total combined mass in the range of $2\text{--}3 \times 10^{12} M_{\odot}$. These estimates also suggest a mass ratio which is close to unity, and a scenario where the mass of M31 is lower than that of the MW is certainly not ruled out. Given the difference in stellar mass, it is still commonly assumed that the mass of M31 is larger than that of the MW. However, it is worth noting that abundance matching implies higher masses for both haloes (Guo et al. 2010), which also implies considerable scatter in the stellar–halo mass relation in order to be consistent with these mass estimates.

Meanwhile, estimates for the combined mass using the timing argument in Λ CDM yield significantly higher results. Examples of direct timing argument measurements using only the radial velocity are $5.27 \times 10^{12} M_{\odot}$ (Li & White 2008), $5.58_{-0.72}^{+0.85} \times 10^{12} M_{\odot}$ (van der Marel & Guhathakurta 2008), and $4.27_{-0.45}^{+0.45} \times 10^{12} M_{\odot}$ (van der Marel et al. 2012). Accounting also for cosmic bias, van der Marel et al. (2012) measure $4.93 \pm 1.6 \times 10^{12} M_{\odot}$. Meanwhile Partridge,

Lahav & Hoffman (2013) found $4.73 \pm 1.03 \times 10^{12} M_{\odot}$ incorporating Dark Energy into the timing argument. Benisty et al. (2022) measure $5.6_{-1.2}^{+1.6} \times 10^{12} M_{\odot}$ accounting for the presence of the LMC, and Hartl & Strigari (2022) measure $4.75_{-2.41}^{+2.22} \times 10^{12} M_{\odot}$. Accounting for the velocity of the MW within the LG system (Petersen & Peñarrubia 2020), Chamberlain et al. (2023) obtain $4.5_{-0.8}^{+0.8} \times 10^{12} M_{\odot}$ using Cepheid and *Gaia* data.

McLeod et al. (2017) used artificial Neural Networks to infer the mass of the LG based on LG analogues from a cosmological simulation with a volume of $\sim 0.2 \text{ Gpc}^3$ and found a total mass of $4.9_{-1.4}^{+1.3} \times 10^{12} M_{\odot}$ when assuming a low transverse velocity of $17 \pm 17 \text{ km s}^{-1}$ (van der Marel et al. 2012).

On the whole, these results are significantly higher, and also appear inconsistent with the masses of the MW and M31 measured from internal kinematic tracers, including neutral H I gas, the stellar disc, or the motions of halo stars, globular clusters, or satellite galaxies.

Fig. 1 gives an overview of these measurements. It is worth noting that some previous results that are not limited to pair kinematics appear to bridge the gap. Benisty et al. (2022) report that the Timing Argument mass reduces to $3.4 \times 10^{12} M_{\odot}$ when accounting for cosmic bias, i.e. taking into account the halo mass function. Peñarrubia et al. (2014) obtain an even lower value of $2.3 \pm 0.7 \times 10^{12} M_{\odot}$ from a Bayesian model of galaxies extending out to a distance of 3 Mpc.

3 SIMULATIONS

This work is based entirely on publicly available data from the UCHUU collisionless cosmological simulation (Ishiyama et al. 2021), which samples a volume of nearly 26 Gpc^3 with 12 800³ particles. In total, at $z = 0$, the UCHUU simulation contains ~ 250 million haloes in the mass range of $0.5\text{--}5 \times 10^{12} M_{\odot}$, and a particle mass of $4.83 \times 10^8 M_{\odot}$ implies that the lowest mass haloes we consider are resolved with more than 1000 particles. UCHUU uses cosmological parameters corresponding to the Planck 2018 results, namely, $\Omega_0 = 0.3089$, $\Omega_b = 0.0486$, $\Omega_{\Lambda} = 0.6911$, $h = 0.6774$, $n_s = 0.9667$, and $\sigma_8 = 0.8159$. The simulation outputs have been processed using the ROCKSTAR phase-space structure finder. All results reported in this paper are expressed in terms of physical coordinates, distances in kpc, masses in solar masses (M_{\odot}), and velocities in km s^{-1} . For the halo mass, we use the most common definition $M_{200, c}$, i.e. the mass contained within a sphere enclosing a mean density of 200 times the critical density.

3.1 Definition of LG analogues

Observations accurately measure the distance between the centres of the MW and Andromeda as $r = 765 \pm 11 \text{ kpc}$ (Chamberlain et al. 2023), and the radial velocity as $v_r = 109.3 \pm 4.4 \text{ km s}^{-1}$ (van der Marel et al. 2012). The transverse velocity is less certainly known, and some measurements strongly differ. Using only *HST* proper motions, van der Marel et al. (2012) found $v_t = 17 \pm 17 \text{ km s}^{-1}$, which would imply an extremely radial orbit. Later combining *HST* and *Gaia* DR2 proper motions, van der Marel et al. (2019) find a somewhat higher value and measure $v_t = 57_{-31}^{+35} \text{ km s}^{-1}$. This measurement constitutes a weighted average of an *HST*-only value of $36_{-26}^{+39} \text{ km s}^{-1}$, and a *Gaia*-only value of $133_{-68}^{+70} \text{ km s}^{-1}$. Salomon et al. (2016) derived an even greater transverse velocity of $v_t = 164.4 \pm 61.8 \text{ km s}^{-1}$ by studying the projected line-of-sight velocities of M31 satellites, and comparing the results to those found in simulations. Salomon et al. (2021) measure $v_t = 82.4 \pm 31.2 \text{ km s}^{-1}$ using *Gaia* DR3 proper motions of blue main sequence

stars in M31. We adopt here $v_t = 57 \pm 35 \text{ km s}^{-1}$, very similar to the combined result of van der Marel et al. (2019). Whenever we Monte-Carlo sample the uncertainties, or construct probability-weighted estimates, we assume that the probability distributions for the true values of r , v_r , and v_t are Gaussian, subject to the additional constraints $r > 0$ and $v_t > 0$.

Our broadest category of LG analogues is defined as a pair of haloes that satisfy all of the following criteria:

- (i) each M_{200} mass within $[0.5 \dots 5] \times 10^{12} M_\odot$
- (ii) separation, r , within $[500 \dots 1000]$ kpc
- (iii) radial velocity, v_r , within $[-250 \dots 0]$ km s^{-1}
- (iv) transverse velocity, v_t , within $[0 \dots 250]$ km s^{-1}
- (v) no further halo more massive than $0.5 \times 10^{12} M_\odot$ within 2 Mpc

It is worth noting that the kinematics constrain primarily the total mass of the LG. The kinematics also allow, for example, a high mass halo paired with a halo below the defined mass range, so we must consider the lower limit of $0.5 \times 10^{12} M_\odot$ for each halo as an additional, fixed prior. As we will discuss, this has implications for the inferred mass ratio, but is, however, well-motivated from other observations.

In total, the UCHUU simulation contains ~ 1.4 million systems defined using the criteria above. In Fig. 2, we show the distribution of these LG analogues within the volume. The panel on the left shows the entire volume of nearly 26 Gpc^3 . The uniquely large volume of the UCHUU simulation at this resolution allows us to retain sufficiently large sample sizes, even when additional constraints are imposed.

It should also be noted that the LG, whether in the real Universe or in cosmological simulations, is not a true two-body system. It includes mass beyond the respective r_{200} of its two main haloes, to which different tracers may be sensitive to different degrees (Peñarrubia & Fattahi 2017). Throughout this paper, we adopt the convention of defining the mass of the LG as the sum of the two halo masses, or $M_{\text{LG}} = M_1 + M_2$. We distinguish between the more massive halo whose mass we label M_1 and the less massive halo of mass M_2 . We do not generally associate them with the MW or M31 specifically, except where we additionally constrain the MW halo mass in Section 6.3, where we label the mass of the MW as M_{MW} and that of M31 as M_{M31} .

4 TRUE LG ANALOGUES

The degree to which the mass of the LG is related to the kinematics of the MW-M31 orbit depends on the degree to which the motion of each object is affected by their common gravitational field, or by other haloes. For example, the ‘Timing argument’ assumes that the system evolves in complete isolation. Previous works (e.g. Li & White 2008; Carlesi et al. 2016; Fattahi et al. 2016) that selected LG analogues from cosmological simulations have typically excluded systems that contain a third object within some radius, either above the lower limit of the range of halo masses considered or, less restrictively, more massive than the less massive of the pair.

However, we find that this is not sufficient to ensure that the dynamics of the simulated LG analogues are free of undue outside influences, i.e. greater than assumed to be the case for the real LG. In this section, we introduce two new criteria, *tidal dominance*, and the external-to-internal *force ratio*. We use them together to identify *true* LG pairs, i.e. those pairs governed by internal kinematics to a similar extent as the real LG.

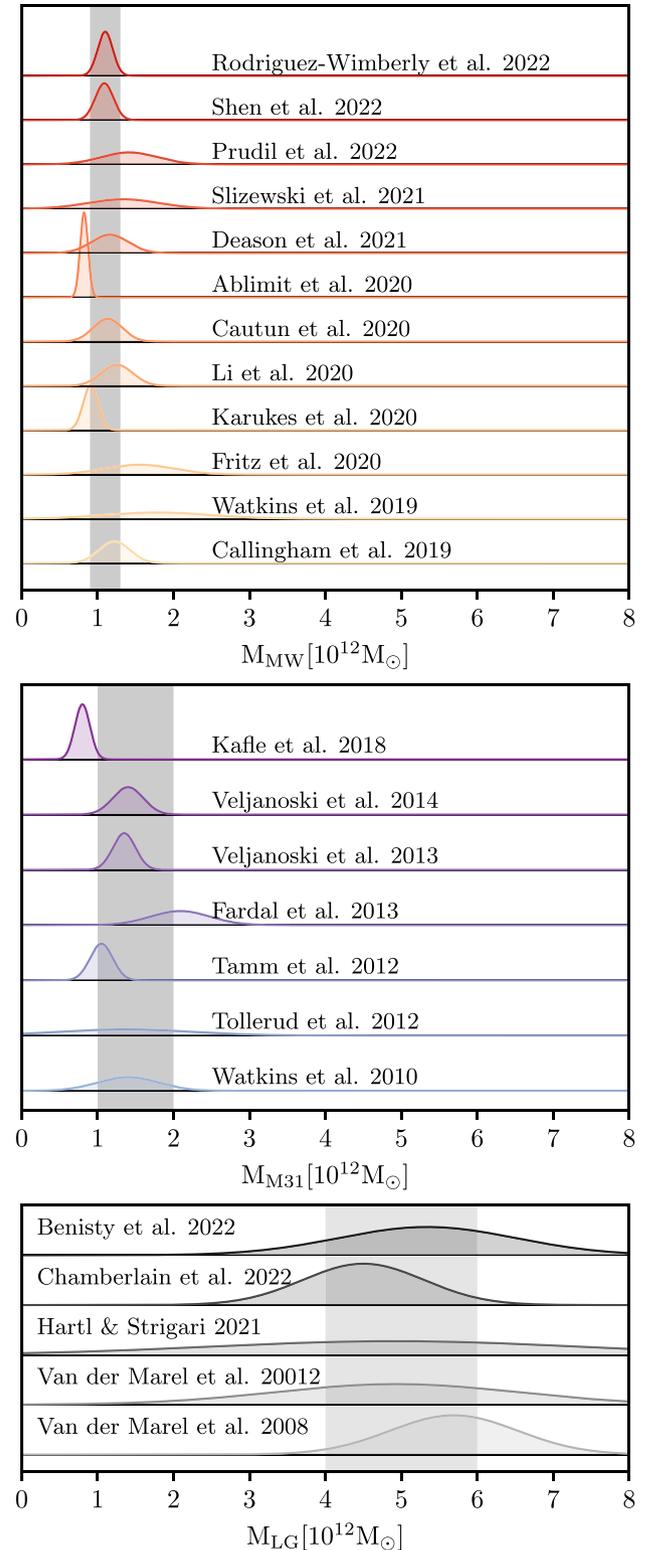


Figure 1. Recent direct estimates for the virial mass of the MW (top), M31 (centre), and estimates using LG kinematics for the combined mass (bottom). Each literature estimate is represented by a (skewed) normal distribution. The grey bands on the top two panels denote $M_{\text{MW}} = 1.1^{+0.2}_{-0.2} \times 10^{12} M_\odot$ and $M_{\text{M31}} = 1.5^{+0.5}_{-0.5} \times 10^{12} M_\odot$, respectively, which we use as additional constraints in some measurements. The estimates for the mass of the LG, shown in the bottom panel, are significantly higher than the sum of the masses of the two constituents.

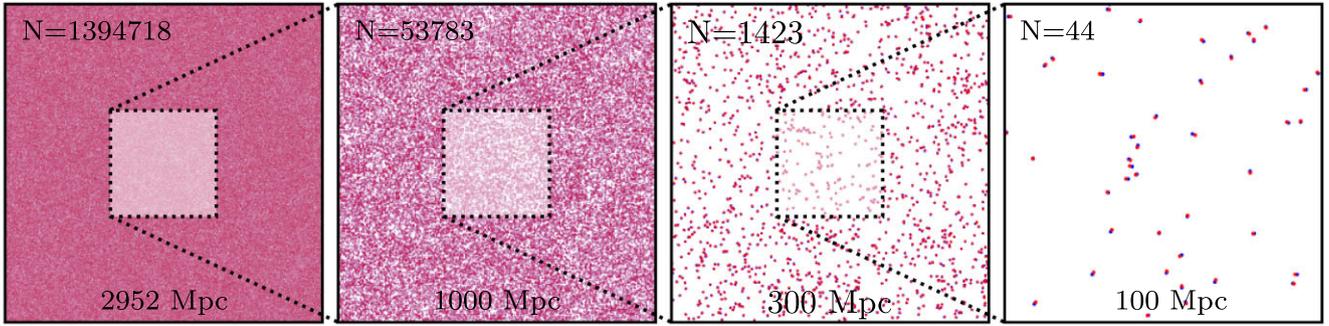


Figure 2. Position of LG analogues in the UCHUU simulation. The blue and red dots denote the positions of the MW and M31 analogues, respectively. Each panel shows a cubic volume; the scale bar corresponds to one-third of the side length. The total volume of UCHUU, shown in the far left panel, is 25.7 Gpc^3 and contains ~ 1.4 million LG analogues. For comparison, the remaining panels show zoom-ins of the full volume, with the second panel from the left comprising a 1 Gpc^3 volume, equal in volume to the P-Millennium simulation; the third panel from the left depicts a volume of 0.027 Gpc^3 equal in volume to Illustris-TNG-300, and the final panel shows a volume of 0.001 Gpc^3 , equal in volume to DOVE/Eagle-100.

4.1 Tidal dominance

The UCHUU halo catalogues contain information about the hierarchy of tidal forces. For each halo, we can identify a more massive halo that exerts the greatest tidal force. This allows us to identify LG analogues in which the tidal field of one LG member is dominated by the other, as opposed to those for which the tidal fields of both haloes are dominated by other objects. We refer to LG analogues in the first category as ‘tidally dominant’ and the remainder as ‘tidally subdominant’.

It is expected that the relations between kinematics and mass only hold for tidally dominant systems. Indeed, this is the case. In Fig. 3, we plot distributions of UCHUU LG analogue masses as functions of the three kinematic parameters. The left panel shows total mass, M_{LG} , as a function of r and v_r , while the right panel shows M_{LG} as a function of r and v_t . For *tidally dominant* systems, shown on top, there are clear correlations, as also shown in previous studies (e.g. Fattahi et al. 2016): at fixed separation, LG analogues with more negative radial velocity or higher transverse velocity tend to have higher masses, although some scatter is also apparent.

For *tidally subdominant* systems, shown on the bottom and comprising ~ 8 per cent of the total sample, the kinematics are primarily governed by forces outside the system itself. Consequently, the mass of the system is largely uncorrelated with the kinematics. In particular, we find a large number of low-mass tidally subdominant systems with kinematics similar to the LG not due to internal processes, but due to random, external perturbations.

We refer to the subsets of tidally dominant and subdominant LG analogues as \mathcal{T}_{dom} and \mathcal{T}_{sub} , respectively, and remove the subdominant systems from most of our subsequent analysis.

4.2 Force ratio

We also consider the gravitational forces exerted by the two LG members on each other, and compare them to the forces exerted on each member by every other halo above $2 \times 10^{11} M_{\odot}$ within 5 Mpc of the LG centre. We parametrize the ratio between the maximum force exerted on to either LG member by any of these haloes, and the force exerted by the other member, as the force ratio:

$$\mathcal{F} = \max \left(\frac{\max(M_i/r_{1i}^2), \max(M_k/r_{2j}^2)}{M_2/r^2}, \frac{\max(M_k/r_{2j}^2)}{M_1/r^2} \right), \quad (1)$$

where M_i and r_{1i} are the mass and distance of a halo from one member of the pair, and M_j and r_{2j} are the mass and distance of a halo from the other member. A low value of \mathcal{F} means that the forces which the two LG members exert on each other dominate the gravitational interactions of both haloes, while a high value of \mathcal{F} means that, for at least one of the two LG members, at least one other halo exerts a significant force.

The kinematics of systems with low force ratios are determined by the LG itself, while the kinematics of systems with high force ratios are strongly affected by haloes outside the pair. We classify our LG analogues according to \mathcal{F} as either ‘internally dominated’, with $\mathcal{F} < 0.5$ comprising approximately two-thirds of all systems, and ‘significantly externally affected’ with $\mathcal{F} \geq 0.5$, comprising the remaining approximately one-third of the sample.

In Fig. 4, we show the correlations of mass and kinematics for LG analogues in the three categories. We find that there is a clear separation: halo pairs with low force ratios have the strongest correlation with the least amount of scatter, while halo pairs with high force ratios show a weaker correlation between mass and kinematics, and significantly more scatter. The real LG most likely contains no halo with a mass above $2 \times 10^{11} M_{\odot}$. With a distance to Centaurus A of 3.4 Mpc, approximately five times the distance between the MW and M31, and its mass approximately $10\times$ greater than that of the MW or M31, we estimate $\mathcal{F}_{\text{LG}} \sim 0.4$.

Throughout the rest of this paper, we will refer to the set of LG analogues that are both ‘tidally dominant’ and have force ratios $\mathcal{F} < 0.5$, as *true* pairs.

4.3 Distributions of kinematic variables

In Fig. 5, we show the distributions of the three kinematic variables, r , v_r , and v_t . On all panels, the grey bands indicate the assumed 1σ and 2σ observational errors, as discussed in Section 3.1. The pink solid lines show the distributions for the full set of LG analogues, and the thick blue solid lines show the distributions for the ‘true’ LG analogues. The grey lines show the distributions for the LG analogues not included in the true sample: dotted for tidally subdominant systems and dashed for those with $\mathcal{F} \geq 0.5$.

It can be seen that all three variables are distributed differently for the different subsets. True LG analogues have, on average, smaller separation than those in the $\mathcal{F} \geq 0.5$ or \mathcal{T}_{sub} subsets, and smaller than the full sample, where these two subsets are included. This is expected because, all else being equal, the internal forces are

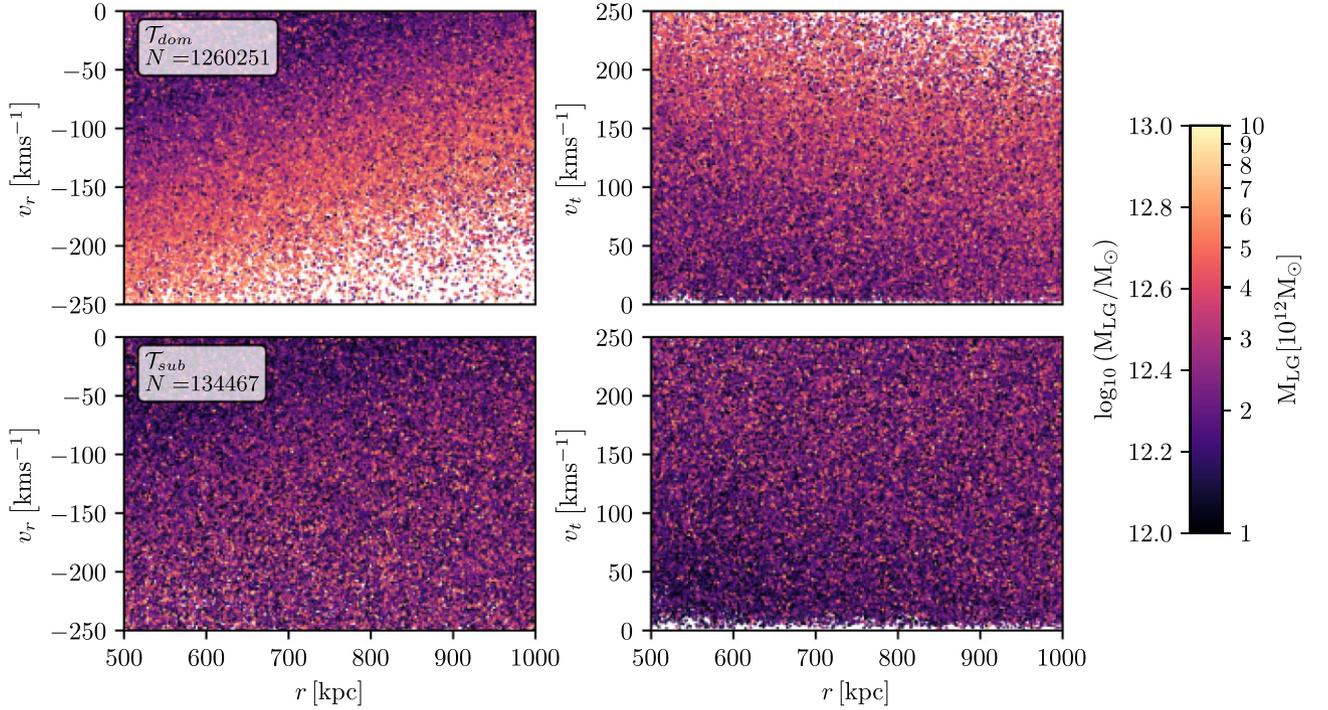


Figure 3. The LG mass as a function of r and v_r (left column), and as a function of r and v_t (right column). The top row shows results for *tidally dominant* LG analogues. There is a clear correlation between mass and the kinematic parameters: at fixed separation, a greater magnitude of v_r or v_t both correlate with greater mass. At fixed v_r , a greater separation, r , correlates with greater mass, but at fixed v_t , the mass is not strongly dependent on separation. The bottom row shows results for *tidally subdominant* LG analogues, that is, where a third object either inside or outside the LG exerts the greatest tidal force on the lower mass halo. The number of these systems increases with separation, the masses are overall lower, and largely uncorrelated with the kinematics of the pair.

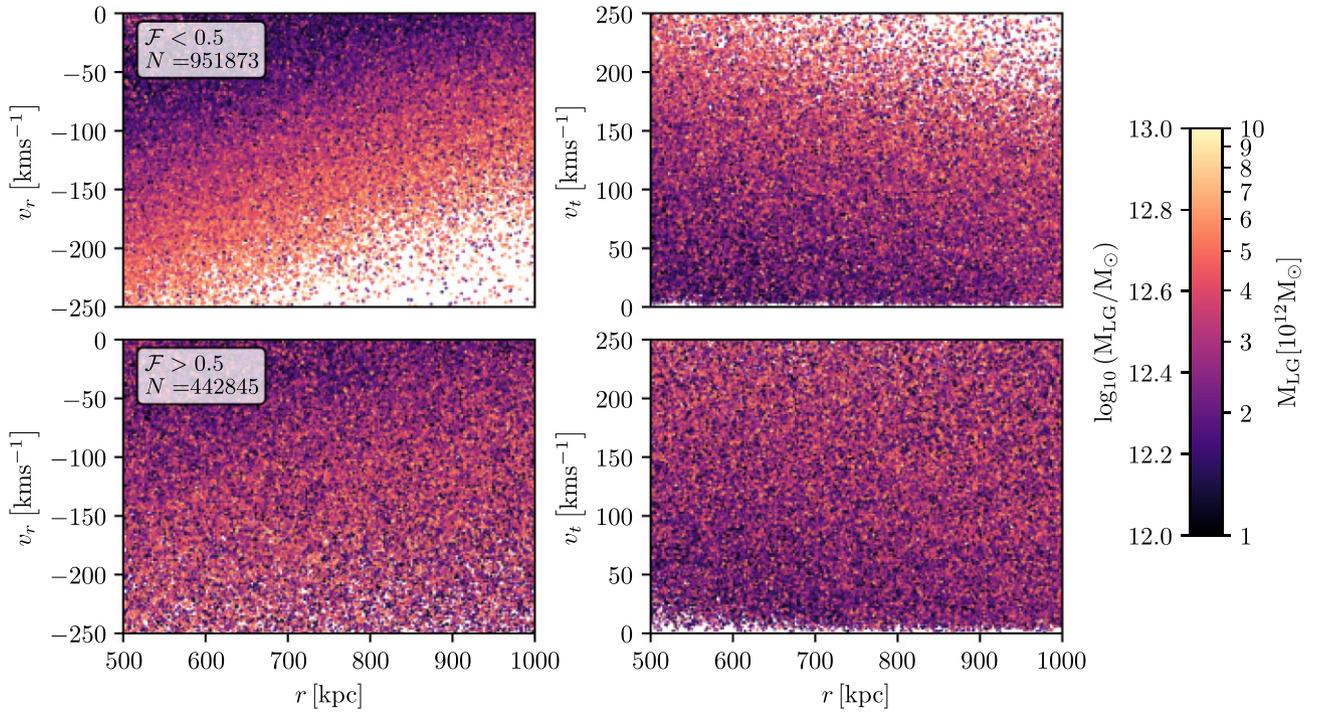


Figure 4. The LG mass as a function of r and v_r (left column), and as a function of r and v_t (right column), similar to Fig. 3. The top row shows results for low force ratio systems ($\mathcal{F} < 0.5$), while the bottom row shows high force ratio systems ($\mathcal{F} > 0.5$). In low force ratio systems, the gravitational fields of the two haloes dominate their overall gravitational interactions, resulting in a clear correlation between kinematics and mass. In high force ratio systems, other haloes cause significant perturbations to the gravitational field, blurring this correlation. The increased scatter with increased \mathcal{F} also means that regions only sparsely populated in the top panels, such as very low v_r or very high v_t , contain significantly more systems in the bottom panels.

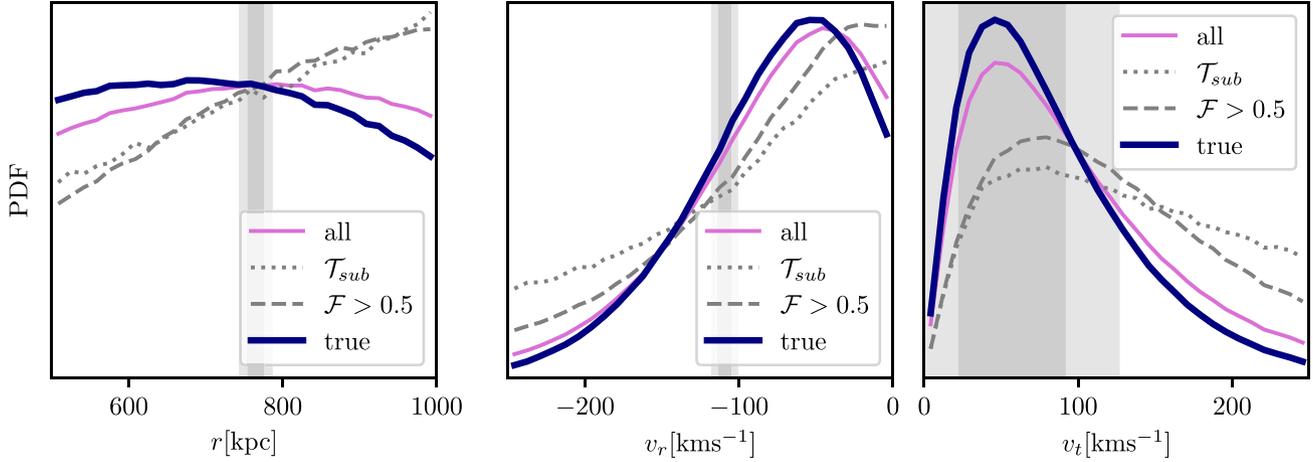


Figure 5. Distributions of observables, r (left), v_r (centre), and v_t (right) in our LG analogues. On all panels, the shaded vertical bands indicate $\pm 1\sigma$ and $\pm 2\sigma$ of the observations. The pink lines show the full sample of LG analogues that match the kinematic criteria. The grey lines show LG analogues we exclude from our analysis: for \mathcal{T}_{sub} (dotted lines), none of the LG members dominate the other’s tidal force, while for $\mathcal{F} > 0.5$ (dashed lines), a third halo exerts a gravitational force on one of the LG members which is larger than 0.5 of the force exerted by the other member. Both of these subsets contain haloes with very high or low radial, and high transverse velocities, which are not related to the internal kinematics. When haloes in either of these two subsets are removed, the remainder (‘true pairs’, thick blue lines) is more likely to match each of the observed values, and shows a stronger correlation between mass and kinematics.

relatively stronger for lower separation systems. The fraction of systems that match the observations for r is nearly identical in all subsets.

The differences in radial and transverse velocity are more pronounced: the set of true LG analogues contain significantly more objects with either very negative ($v_r < -150 \text{ km s}^{-1}$) or very small ($v_r > -50 \text{ km s}^{-1}$) radial velocities, or with high ($v_t > 100 \text{ km s}^{-1}$) transverse velocities, compared to the $\mathcal{F} \geq 0.5$ and \mathcal{T}_{sub} subsets. In fact, the majority of the full sample of LG analogues that meet our selection criteria, and that have either very negative radial velocities or very high transverse velocities, are not ‘true’ LG analogues governed by mutual interactions, but are systems whose kinematics are caused by other objects. It is worth noting that removing these objects causes the ‘true’ LG analogues to have a significantly higher fraction of objects within the observational uncertainties.

4.4 The transverse velocity

Interestingly, Forero-Romero & Sierra-Porta (2022) recently reported that the average transverse velocity of LG analogues is higher in simulations with large box sizes, pointing to a limitation of studying LG analogues in relatively small volume simulations. The largest simulation they consider, ABACUS SUMMIT, has a box size almost identical to UCHUU, albeit at eight times lower mass resolution. Interestingly, they find a transverse velocity with a median and 1σ -equivalent uncertainty of $105^{+94}_{-59} \text{ km s}^{-1}$, significantly higher than our results: $v_t = 81^{+74}_{-46} \text{ km s}^{-1}$ for all LG analogues, and $v_t = 72^{+65}_{-40} \text{ km s}^{-1}$ for true pairs. If we further incorporate the observed values and uncertainties on r and v_r as priors, we obtain a nearly identical weighted posterior of $v_t = 75^{+65}_{-40} \text{ km s}^{-1}$ for the LG transverse velocity, consistent with the value of $v_t = 57^{+35}_{-31} \text{ km s}^{-1}$ of van der Marel et al. (2019).

The difference between our result and those of Forero-Romero & Sierra-Porta (2022) could be partly due to the fact that their selection of haloes with $v_{\text{max}} = 200 \dots 260 \text{ km s}^{-1}$ led them to consider more massive LG analogues which, as we will discuss in Section 6.4, tend to have higher transverse velocities. Perhaps more significantly, due

to the fact that they include pairs with separations up to 1.5 Mpc, their sample almost certainly contains a large fraction of unbound pairs, or pairs whose kinematics are not determined internally, but through interactions with other haloes. For example, we find $v_t = 118^{+82}_{-68} \text{ km s}^{-1}$ for tidally subdominant pairs, and $v_t = 107^{+78}_{-59} \text{ km s}^{-1}$ for pairs with high force ratios. Both of these are close to the value obtained by Forero-Romero & Sierra-Porta (2022). However, they may not be a good prior for the transverse velocity of the LG.

5 THE LG MASS VIA GAUSSIAN PROCESS REGRESSION

We estimate the mass of the LG, $M_{\text{LG}} = M_{\text{MW}} + M_{\text{M31}}$, using Gaussian process regression (GPR), a machine learning algorithm that provides an estimate of both the most likely value, and the uncertainty (standard deviation) of the estimate. Specifically, we infer the relationship between the inputs (kinematics) and the target (LG mass), i.e. the conditional mass distribution given the observed kinematics.

GPR is a type of supervised machine learning, meaning that it first ‘learns’ the connection between the input and the target variables from a given dataset where both are known. In our case, the regressor learns the correlation between the kinematics (input) and the LG mass (target) from sets of ‘true’ LG analogues from the UCHUU simulation. Afterwards, the trained regressor can predict the LG mass from a given new set of kinematics. During the training, functions are fitted to the training data, whose coefficients are optimized iteratively. Without further constraints, there is an infinite number of possible functions that can fit a given set of data. This sets GPR, as a form of non-parametric regression, apart from parametric regression models, where the functional form is fixed, and only function parameters are optimized.

The goal of the regression is to find a probability distribution of all functions that fit the data, and from that, predict both the mean and variance of the mass for a given set of kinematics, over the entire range of possible input values. In our case, we construct the possible functions as the dot products (inner product) of a radial basis

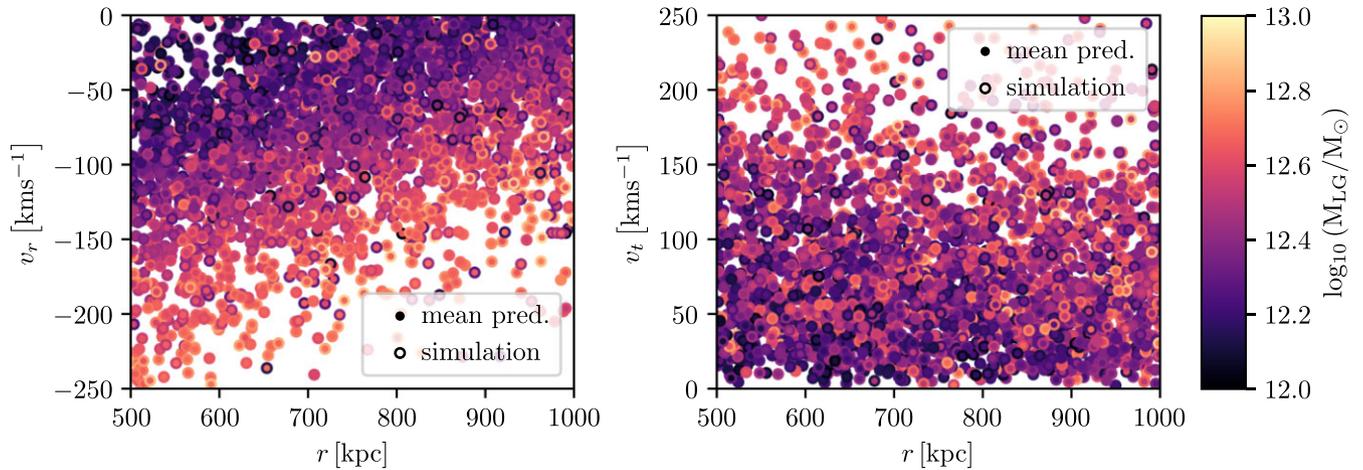


Figure 6. LG mass measured in the simulation (symbol edges) and mean predicted mass using Gaussian regression (symbol centres) as a function of the observed kinematics, for a random subset of \mathcal{T}_{dom} , $\mathcal{F} < 0.5$ LG analogues. Similar colours for the centre and edge mean that the regression predicted a mass similar to that in the simulation. The regression correctly captures both the overall mass gradients detectable by eye, as well as, in many cases, the masses of apparent outliers.

function, a constant, and a white noise term. The regression begins with a uniform prior on the probabilities of all possible functions. Subsequently, the probability distribution is successively updated based on new information in the form of training data. The posterior distribution of the possible function is then used to predict both the mean and the variance (or standard deviation) (Wang 2020). In this work, we use the SCIKIT-LEARN (Pedregosa et al. 2011) implementation of GPR, GAUSSIANPROCESSREGRESSOR.

Fig. 6 shows the mean prediction of the LG mass obtained via GPR (symbol centres) in the $r - v_r$ plane (left panel), and in the $r - v_t$ plane (right panel), compared to the values measured for the same systems in the simulation (symbol edges). It can be seen that the GPR predicts the gradients in mass identifiable by eye, but also, in many cases, correctly predicts the masses of points that appear as outliers in each individual 2D plane, due to the fact that it uses the full, 3D kinematic information in its estimate. However, as demonstrated by the data points with differing edge- and centre colours, the regression is not perfect.

5.1 Predicted uncertainty

The algorithmic complexity of GPR of $\mathcal{O}(N^3)$ prohibits us from using the full data for training. However, GPR is quite efficient even for small sample sizes. In Fig. 7, we show the convergence behaviour of the regression. We train the regressor multiple times on N_s data points, and for each instance, we obtain an independent estimate of the LG mass via Monte-Carlo sampling of the observations (see Section 5.2). The precision of the regression increases with the number of data points: for a sample size of $N_s = 100$, we find $M_{LG} = 3.14 \pm 0.27 \times 10^{12} M_\odot$, while for a sample size of $N_s = 3200$, we find a median value of $M_{LG} = 3.18 \pm 0.06 \times 10^{12} M_\odot$. The uncertainty of $\pm 0.06 \times 10^{12} M_\odot$ represents the uncertainty of our prediction of the *median* value of the LG mass distribution. As we explain below, however, it is quite different from the uncertainty of our prediction of the mass of the LG.

In Fig. 8, we plot the predicted mass as a function of mass in the simulation, with symbols coloured, as in Fig. 9, according to the mass in the simulation in the symbol centres, and according to the mean prediction at the symbol edges. It can be seen that the range of mean predictions is narrower than the range in the simulations:

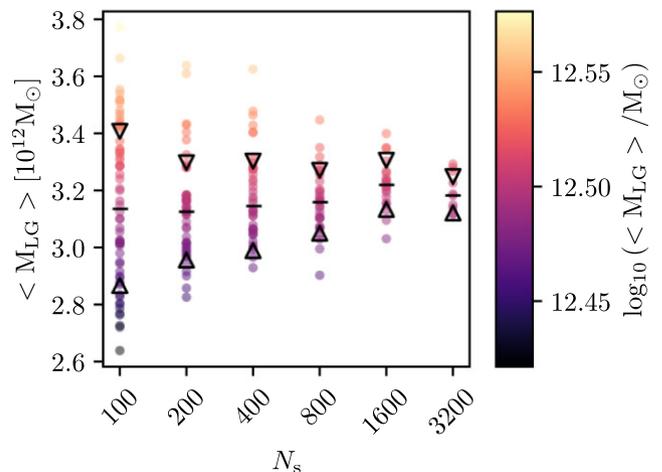


Figure 7. Individual predicted median values of M_{LG} from independent Gaussian process regressors trained on N_s data points, after Monte-Carlo sampling the observational uncertainty. The black horizontal lines indicate the median of the values plotted for each sample size, and upwards and downwards triangles indicate $\pm 1\sigma$ scatter. As the number of points in the sample increases, the scatter in the predictions is reduced. At $N_s = 3200$, we obtain an expectation value for the median of $M_{LG} = 3.18 \pm 0.06 \times 10^{12} M_\odot$.

the ensemble of mean predictions lacks both the high- and low-mass outliers found in the simulation. The mean prediction underpredicts the mass of the most massive objects, and overpredicts the masses for the least massive ones. This behaviour is not unexpected: indeed, if the connection between kinematics and mass contains some amount of random scatter, the extremely low- and high-mass points are likely to fall, respectively, on the low- and high-mass side of the scatter. While the mean regression cannot, by definition, reproduce the random behaviour for individual objects, we can account for the presence of scatter by including a white noise term in the kernel function.

One key advantage of the Gaussian regression over some other regression models (such as decision trees or neural networks) is that it automatically provides an estimate of the uncertainty, or error, of each estimate along with the mean estimate. We show the predicted

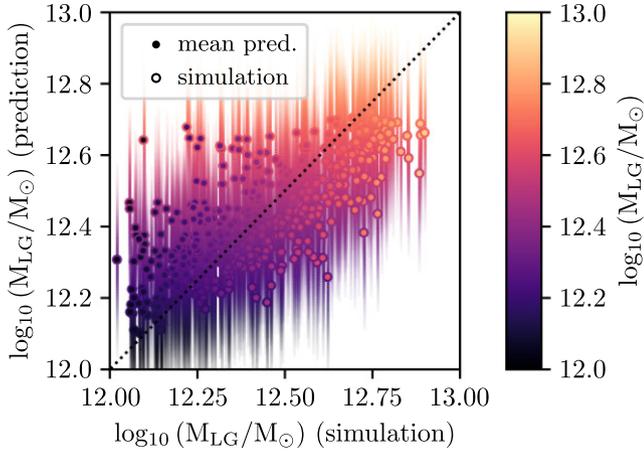


Figure 8. The LG mass predicted using Gaussian regression (y-axis) compared to the mass measured in the simulation (x-axis). As in Fig. 6, symbol centres are coloured according to the mean prediction; symbol edges are coloured according to the mass in the simulation. Error bars extend to $\pm 1\sigma$ of the uncertainty of each prediction. There is a clear correlation between predicted and measured mass, with no apparent bias. Some amount of regression-to-the-mean is also apparent: for low masses, the Gaussian regression tends to overpredict the mass, while for high masses, it underpredicts it. Consequently, the range of mean predictions is narrower than the full range of masses in the simulations. However, this apparent shortfall can be rectified when the uncertainty of the prediction is taken into account.

1σ -uncertainties for each data points as vertical lines in Fig. 8. The range of values within the uncertainties around the mean predictions shown in Fig. 8 extends beyond that of the mean values alone. For a full prediction that encapsulates both the predicted trend and the irreducible uncertainty, we therefore convolve each predicted mean with a normal distribution of the predicted standard deviations.

In Fig. 9, we compare the distribution of LG masses measured directly in the simulation with those obtained by applying the Gaussian regression to the kinematics measured in the simulations for the same objects. On the top panel, the open histogram shows the full data in the tidally dominant, $\mathcal{F} < 0.5$ sample, i.e. the ‘true’ LG analogues. The red histogram shows the distribution of mean predictions of the regression, while the blue histogram shows the distribution after convolution of the predicted means with the predicted uncertainties. As discussed above, the mean predictions do not fully account for the scatter, but after the uncertainty is taken into account, the resulting distribution reproduces the shape of the data. On the bottom panel, the data are reduced to the subset for which each observable, r , v_r , and v_t , is limited to $\pm 3\sigma$ around the observational value, where σ is the quoted observational error. Here, the difference between the distribution of masses in the data and in the distribution of mean predictions is even more pronounced: while the distribution of the mean predictions has $M_{\text{LG}} = 3.4_{-0.2}^{+0.3} \times 10^{12} M_{\odot}$, the data have $M_{\text{LG}} = 3.5_{-1.1}^{+1.2} \times 10^{12}$, a similar median, but much greater scatter. After convolution with the predicted uncertainty, the prediction has $M_{\text{LG}} = 3.4_{-0.9}^{+1.3} \times 10^{12} M_{\odot}$, reproducing well both the median and the scatter in the data.

5.2 A prediction of the LG mass

To predict the mass of the actual LG based on our regression, we use the observed values of r , v_r , and v_t as inputs to the regressor trained on the simulation data. In Fig. 10, we show the resulting probability density for the LG mass. To account for the observational

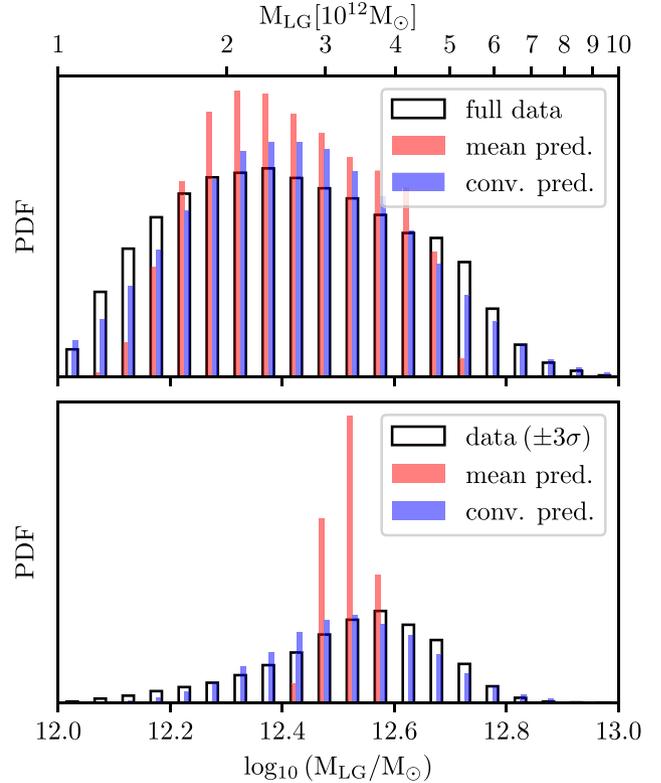


Figure 9. Distribution of LG masses in the simulation (black open histograms), predicted means of the GPR (red histograms), and after the predicted means are convolved with the respective predicted uncertainties (blue histograms). The top panel shows the results for all LG analogues in the tidally dominant, $\mathcal{F} < 0.5$ sample. The bottom panel shows the results for only those LG analogues that are within $\pm 3\sigma$ for each of the kinematic variables on which the prediction is based, r , v_r , and v_t . It can be seen that the distributions of mean predictions are more concentrated than those of the simulation data, lacking both high- and low-mass outliers. By contrast, after convolving the data with the predicted means with the predicted uncertainties, the distribution of the data is reproduced well. This effect is particularly pronounced when, as shown in the bottom panel, the range of input variables is narrowed, in this case to $\pm 3\sigma$, where σ is the default observational uncertainty.

uncertainty, we create Monte-Carlo samples for 10^6 combinations of the observables, each following a normal distribution with mean and standard deviations inferred from the observations, as described in Section 3.1. We truncate the Monte Carlo samples to exclude nonsensical values, i.e. $r < 0$ or $v_t < 0$, but this has only a minor effect on the results.

The thick red line in Fig. 10 shows the resulting mass distribution using the mean values of the regressor with the quoted observational uncertainties. The predicted mass is $3.43_{-0.2}^{+0.1} \times 10^{12} M_{\odot}$ – an apparently impressively precise result. As indicated by the dashed and dotted red lines, the precision is limited only by the observational errors: the uncertainty increases to $\pm 0.3 \times 10^{12} M_{\odot}$ when the observational errors are assumed to be twice as high (dashed red line) and becomes infinitesimally small when the assumed errors go to zero (dotted vertical line). Notably, as we will show in Section 6, this behaviour is similar to an analysis based on the classical ‘timing argument’. However, it is not a true reflection of the uncertainty of the estimate of the mass based on the kinematics.

The blue lines in Fig. 10 show, for the same Monte Carlo samples with the same assumed observational uncertainties, the

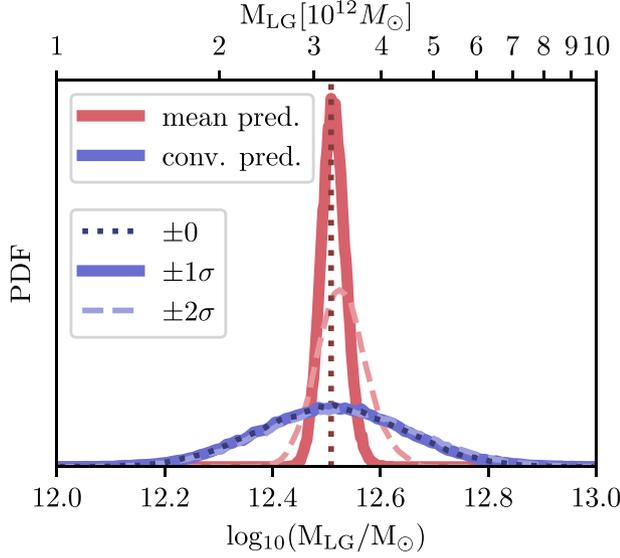


Figure 10. LG mass predicted by the Gaussian regression, when using Monte Carlo samples of the observed kinematics within the stated uncertainties as inputs. The red lines show the distribution of the mean predictions, and the blue lines show the distribution after convolving each mean prediction with its associated uncertainty. It can be seen that the mean prediction appears to be very precise, but this belies the true uncertainty in the LG mass–kinematics relation. Only after convolving with the uncertainties, we get a realistic estimate, and find a mass of $3.2_{-0.9}^{+1.2} \times 10^{12} M_{\odot}$ when assuming standard $\pm 1\sigma$ observational errors. While the uncertainty of the mean prediction depends strongly on the assumed observational errors (dotted lines: zero error, solid lines: the default errors, dashed lines: twice the default errors), the uncertainty of the convolved prediction is dominated by the uncertainty in the kinematics–mass relation.

LG mass distribution after the predicted means are convolved with the predicted uncertainties. These estimates account both for the observational errors and for the uncertainty in the relation between kinematics and mass. With the standard, $\pm 1\sigma$ observational errors, we obtain a mass of $M_{\text{LG}} = 3.2_{-0.9}^{+1.2} \times 10^{12} M_{\odot}$ – a similar mass with a larger, but more realistic estimated uncertainty. Comparing this

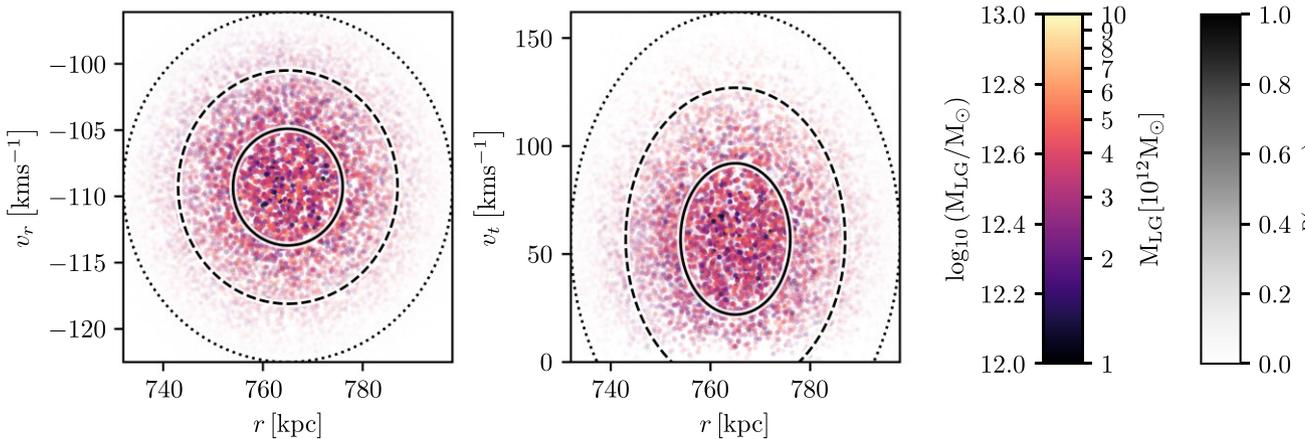


Figure 11. LG analogues, coloured by mass, as a function of r and v_r (left column), and as a function of r and v_t (right column). The opacity of each data point (alpha value) is computed proportional to the probability that it represents the observed values of r , v_r , and v_t , computed using a 3D Gaussian kernel function (equation 3) in the (r, v_r, v_t) -space based on the observed values and their uncertainties. Although we include the full sample of tidally dominant systems with $\mathcal{F} < 0.5$ in our probability weighted estimate, for clarity, the axis ranges only extend out to $\pm 3\sigma$ for each variable. For illustration, ellipses corresponding to 1σ , 2σ , and 3σ in the two 2D-spaces (r, v_r) and (r, v_t) are overplotted.

result to the one shown by the red lines, and discussed above, it is clear that the precision of estimating the LG mass is not limited by the current observational error, but by the fundamental scatter in the relation between the kinematics and mass. Indeed, when we bring the assumed observational error to zero, as shown by the dotted blue line, the mass estimate hardly changes.

6 DIRECT PROBABILITY WEIGHTED ESTIMATES

The uniquely large sample of LG analogues in UCHUU, including many whose kinematics closely match the observations, also allows us to directly estimate the most likely mass of the LG and its members along with the associated uncertainties. For this purpose, we consider each LG analogue as a measurement of the observed LG, weighted according to a Gaussian kernel which describes the likelihood that it represents the observations.

For each LG analogue, i , we compute the distance to the observed kinematics as

$$\sigma_i = \sqrt{\sigma_{r,i}^2 + \sigma_{v_r,i}^2 + \sigma_{v_t,i}^2}, \quad (2)$$

where $\sigma_{x,i} = (x_i - x_{\text{obs}})/\sigma_{x,\text{obs}}$, with x_{obs} the mean observed value, and $\sigma_{x,\text{obs}}$ the associated error, respectively, for each x in r, v_r, v_t .

We then compute the relative probability that an LG analogue, i , represents the observations as

$$P_i = \frac{1}{(2\pi)^{n/2}} e^{-\sigma_i^2/2}, \quad (3)$$

where n is the number of independent observables and σ_i is as defined in equation (2). We use these probabilities as weights in computing the probability distributions of the posteriors given by the observables.

When we include additional constraints, such as the observed mass of the MW or M31, we include these as additional observables, x , and compute the associated distances, σ_i , and probabilities, P_i , accordingly.

In Fig. 11, we show the distributions of the kinematic properties (v_r and r on the left, v_t and r on the right) of LG analogues, with the colour hue according to the LG mass (as in Figs 3 and 4, and the

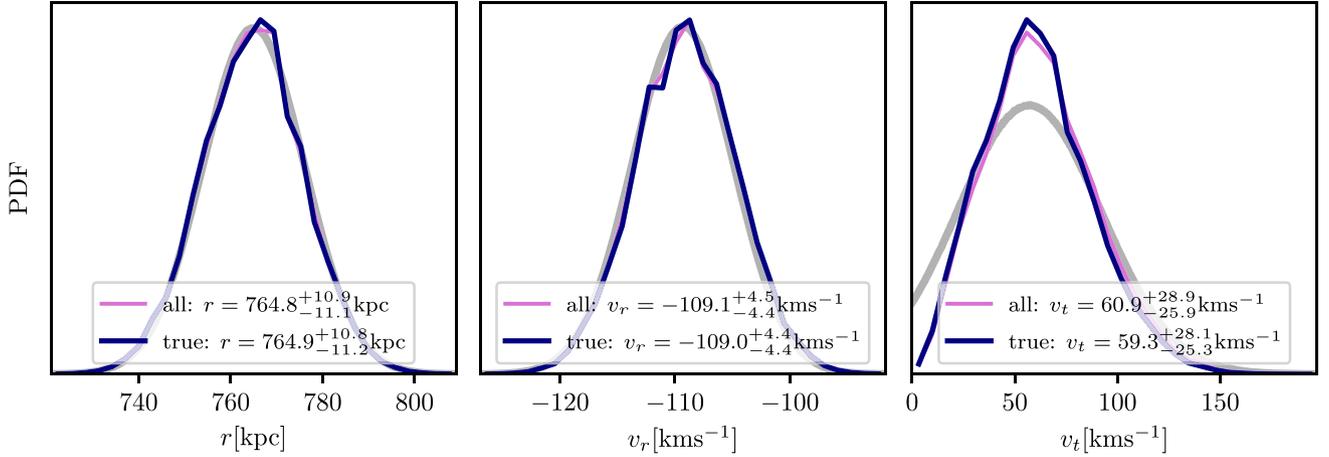


Figure 12. Distributions of the kinematic variables r , v_r , and v_t measured in the simulation with each LG analogue weighed according to its probability to represent the observed LG. On all panels, the purple lines show the weighted estimate for the full set of LG analogues, blue lines show the weighted estimate for the ‘true’ LG analogues (tidally dominant and low force ratio). The grey lines indicate the assumed observational uncertainty of each variable. The distributions of r and v_r , weighted by the overall probability, closely match the distributions assumed for the observations for each variable individually. The distributions of v_t , weighted by the overall probability, are less likely to contain values close to zero than the pure observational uncertainty would predict.

alpha-value, or opacity, according to the Gaussian kernel estimate of the relative probability that they represent the observed LG. In total, there are 2921 LG analogues with $\sigma(i) < 1$, 7129 with $\sigma(i) < 2$ and 11 537 with $\sigma(i) < 3$, allowing us to make robust estimates of the distributions, not limited by sample size.

In Fig. 12, we show the histograms of the distributions of kinematic variables of LG analogues, weighed using the Gaussian distance in terms of r , v_r , and v_t . It can be seen that, for r and v_r , both the median values and the standard deviations of the distributions are very close to the assumed observational values. This reflects the fact that the relative observational errors are so small that the distributions of v_r and r among the simulated LG analogues over these intervals are close to uniform. As discussed above, the observational uncertainty of v_t is considerably larger, and given that v_t is bounded from below by zero, the resulting distribution is asymmetric and skewed slightly towards higher values.

6.1 Kinematically derived probability distributions

In Fig. 13, we show the distributions of all four covariants, obtained using a probability weighted average: the total mass (top left), the mass ratio (top right), the lower of the two masses (bottom left), and the greater of the two masses (bottom right). On each panel, we show both the distribution for the full set of LG analogues, and for the subset of analogues that are both tidally dominant and with a low force ratio (true pairs).

For these ‘true pairs’, we find a total mass of $M_{\text{LG}} = 3.45^{+1.09}_{-1.12} \times 10^{12} M_{\odot}$. Both the centre of the distribution and the uncertainty are very similar to the result obtained in Section 5 using GPR. We find a mass of $M_2 = 0.9^{+0.58}_{-0.30} \times 10^{12} M_{\odot}$ for the lower mass halo, and $M_1 = 2.33^{+1.00}_{-0.85} \times 10^{12} M_{\odot}$ for the higher mass halo.

The median total mass is considerably lower than most of the estimates using the timing argument. Notably, both individual masses and their uncertainties are consistent with the independent mass estimates for the MW and M31 listed discussed in Section 2, at least if the MW is the less massive of the pair. We explore the mass-order in more detail in the next section.

It is also worth noting that the most likely (modal) mass of the lower mass halo, M_2 , is near the lower limit of the mass range, $0.5 \times 10^{12} M_{\odot}$. This is a consequence of the steep mass function in Λ CDM, and the fact that the kinematics constrain only the total mass, on which the lower mass halo has a comparatively small effect. As mentioned in Section 3.1, the lower limit on M_1 that we selected must be considered as an additional prior.

6.2 Comparison to the timing argument

In Fig. 14, we show the probability distribution function for the probability-weighted estimate of M_{LG} , the mass of the LG. The dark blue solid line shows the result we obtain using the $\pm 1\sigma$ uncertainties described in Section 3.1, $r = 765 \pm 11$ kpc, $v_r = 109.3 \pm 4.4$ km s $^{-1}$, and $v_t = 57 \pm 35$ km s $^{-1}$. The light blue dashed line shows the result if we instead increase all observational errors by a factor of 2: $r = 765 \pm 22$ kpc, $v_r = 109.3 \pm 8.8$ km s $^{-1}$, and $v_t = 57 \pm 70$ km s $^{-1}$. Confirming the results obtained through GPR discussed in Section 5, the distribution function for the LG mass is almost unchanged. The true uncertainty in the mass of the LG is not limited by the observational errors in r , v_r , or v_t , but by the scatter in the kinematics–mass relation, i.e. the scatter in mass of objects with identical kinematics.

The other four lines in Fig. 14 show the results of applying the timing argument, either to the kinematics of the probability weighted LG analogues in the simulation (red lines), or to Monte Carlo samples of the observations (grey lines). We compute timing argument masses by numerically solving the following four equations (Marel & Guhathakurta 2008), where a is the semimajor axis of the MW-M31 orbit, e is the eccentricity, and η is the eccentric anomaly (which parametrizes the orbital phase). The parameters r , v_r , and v_t are the separation, radial, and transverse velocities, respectively, and t_0 is the age of the Universe.

$$r = a(1 - e \cos \eta), \quad (4)$$

$$t_0 = \left(\frac{a^3}{GM} \right)^{1/2} (\eta - e \sin \eta), \quad (5)$$

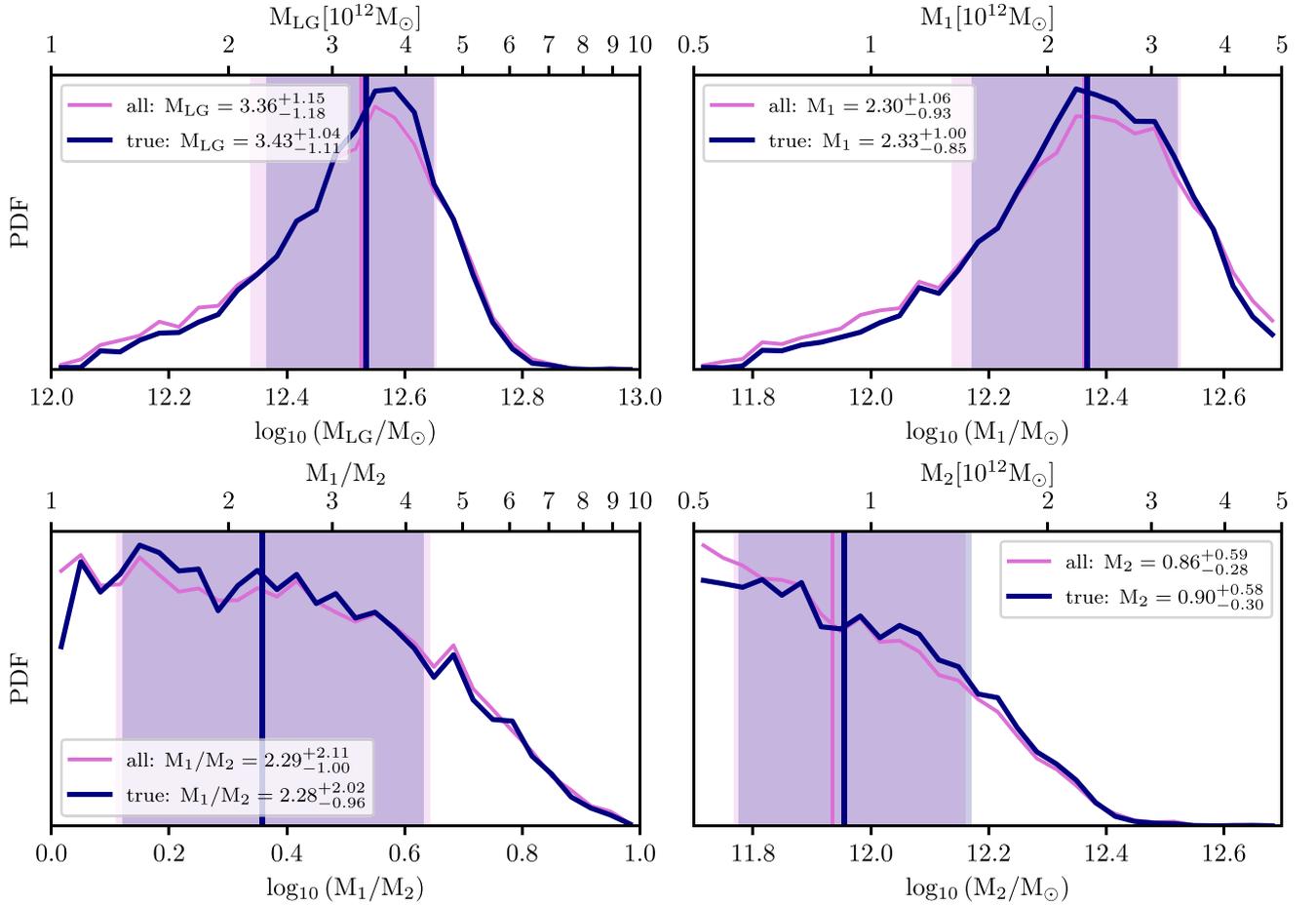


Figure 13. Clockwise from the top left: posterior distributions of the total mass (M_{LG} , top left), mass of the more massive halo (M_1 , top right), mass of the less massive halo (M_2 , bottom right), and mass ratio ($M_{\text{M31}}/M_{\text{MW}}$, bottom left), when each LG analogue either from the full sample (purple lines) or the sample fulfilling the tidal dominance and force ratio criteria (dark blue lines) is weighted according to its likelihood to represent the observed LG, and assuming that M31 is the more massive of the pair. The shaded areas indicate the ranges between the 16th and 84th percentiles. Using this weighted sample, we obtain a median total mass of $3.43 \times 10^{12} M_{\odot}$, and a median MW mass of $0.90 \times 10^{12} M_{\odot}$, in line with the observations. We also obtain a median M31 mass of $2.33 \times 10^{12} M_{\odot}$ and a median mass ratio of 2.28.

$$v_r = \left(\frac{GM}{a} \right)^{1/2} \frac{e \sin \eta}{1 - e \cos \eta}, \quad (6)$$

$$v_t = \left(\frac{GM}{a} \right)^{1/2} \frac{(1 - e^2)^{1/2}}{1 - e \cos \eta}. \quad (7)$$

Note that the timing argument only applies to bound LG analogues and assumes that the LG is on its first approach. Both for the Monte Carlo samples and for the probability weighted kinematics, we also compute the distributions using either the quoted $\pm 1\sigma$ uncertainty (solid lines), or increasing all assumed observational errors by a factor of 2 (dashed lines). By comparing the grey and red lines, we see that the probability weighted kinematics of the LG analogues in the simulations closely match the Monte Carlo samples, resulting in very similar distributions for the timing argument mass. By comparing the solid and dashed lines, we see that the precision of the timing argument estimate increases for smaller assumed errors. Indeed, if the observed errors tend to zero, the timing argument yields a unique solution. However, this does not reflect the true uncertainty in the relation between mass and kinematics measured in the simulations, as shown by the blue lines. We also note that the timing argument

masses are offset from those measured directly in the simulations, suggesting that the timing argument is biased.

6.3 The MW mass as an additional prior

Given the spate of recent measurements of the MW mass, we also investigate the dependence of LG properties when the MW mass is imposed as an additional constraint. The results are not purely LG kinematic estimates, and cannot be used directly to reveal (or alleviate) tensions between individual estimates and those obtained using pure kinematics. However, they may provide more accurate mass estimates, and a comparison to the purely kinematic results can also illuminate whether the different mass estimates are compatible.

As discussed in the introduction, the mass of the MW is very well constrained to be close to $10^{12} M_{\odot}$ using *Gaia* DR3 observations, among others. Assuming a value of $1.1_{-0.2}^{+0.2} \times 10^{12} M_{\odot}$ as a fourth observable in equations (2) and 3, we obtain new distributions for the masses of the individual objects, and for the mass of the LG, as shown in Fig. 15.

We consider both the case where the MW, at this mass, is the less massive or more massive halo of the pair. The first scenario, ‘regular mass order’, is shown by the dark blue lines in Fig. 15. As-

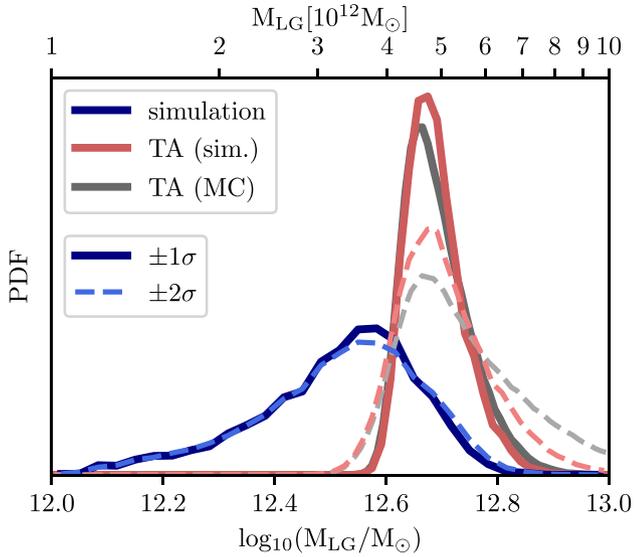


Figure 14. Probability distribution of the LG mass. Blue, as measured in the simulation, weighted by the probability of representing the observations according to the assumed uncertainty. Red: as inferred from the timing argument for the same sample in the simulation with the same weights. Grey: as inferred from the timing argument, drawing Monte-Carlo samples with the same assumed means and uncertainties. The solid lines show the quoted $\pm 1\sigma$ observational errors, and the dashed lines show the effect of increasing the assumed observational errors to 2σ . The precision of the timing argument estimate is limited only by the observational errors. By contrast, in the simulations, the uncertainty is dominated by the underlying scatter in the relation between mass and kinematics, imposing a fundamental limit on the precision.

suming $M_{MW} = 1.1^{+0.2}_{-0.2} \times 10^{12} M_{\odot}$ leads to an MW mass of $M_{MW} = 1.05^{+0.19}_{-0.21} \times 10^{12} M_{\odot}$, an M31 mass of $M_{M31} = 2.28^{+0.78}_{-0.97} \times 10^{12} M_{\odot}$, a total mass of $3.36^{+0.93}_{-0.81} \times 10^{12} M_{\odot}$, and a mass ratio of $2.18^{+1.17}_{-0.78}$. Reassuringly, all values are consistent with the values obtained purely from kinematics, showing that the kinematic analysis yields values compatible with the individual measurements of the MW mass. Using the measured MW mass as an additional observational constraint has increased the accuracy of the predictions for all three components. Under the assumption that the MW is the less massive of the pair, an MW mass of $1.1^{+0.2}_{-0.2} \times 10^{12}$ is not in conflict with the LG kinematics. To the contrary, the MW mass is slightly above the median value inferred purely from kinematics.

When the MW is assumed to be the more massive of the two haloes (orange lines in Fig. 15), the three other results change significantly. The total mass of the LG is reduced to $1.90^{+0.39}_{-0.31} \times 10^{12} M_{\odot}$, while the mass of M31 is now $0.73^{+0.28}_{-0.16} \times 10^{12} M_{\odot}$, and the mass ratio is $M_{MW}/M_{M31} = 1.48^{+0.48}_{-0.34}$ ($M_{M31}/M_{MW} = 0.68^{+0.2}_{-0.17}$). All these are outside the range predicted from the kinematics alone. The strong preference for an M31 mass below $10^{12} M_{\odot}$ and a total mass below $2 \times 10^{12} M_{\odot}$ despite the kinematics in this ‘inverted mass ordering’ stem from the fact that mass ratios very close to 1:1 are rare.

The probability density functions shown in Fig. 15 show the estimated values of the masses and mass ratio given the assumed kinematics, and under the assumptions that an MW of $1.1^{+0.2}_{-0.2} \times 10^{12}$ is either the less or more massive halo of the pair. They do not, by themselves, predict which assumption for the mass ordering is more likely to be correct. To estimate the relative probability for a ‘regular’ or ‘inverted’ mass order given the observed kinematics and

their associated uncertainties, we compare the sum of the weighted probability densities for both scenarios, which also accounts for the mass function and mass ratio distribution in the simulations. We find a ratio of 3.6, implying a 78 per cent chance that an MW of $1.1^{+0.2}_{-0.2} \times 10^{12}$ is less massive than M31.

6.4 The inverse problem: possible kinematics for a given mass

The kinematics strongly constrain the mass of the LG, but among the 1.4 million LG analogues, we find examples of LGs with any mass in the range $1\text{--}10 \times 10^{12} M_{\odot}$ for nearly every combination of the kinematics. Part of the preference for a low-mass system is simply due to the fact that in Λ CDM, low-mass haloes and by extension halo pairs, are far more prevalent than high-mass haloes. It may therefore be illuminating to also consider the inverse question: given an assumed LG mass, what is the range of plausible kinematics?

In Fig. 16, we show the distributions of the kinematics for LG analogues with masses of $\sim 2, 3, 4, 5,$ and $6 \times 10^{12} M_{\odot}$. Each mass range is limited to $\pm 0.2 \times 10^{12} M_{\odot}$, small enough to clearly separate different mass ranges, and large enough to include at least 10^4 systems in each mass range.

Also shown are ellipses centred on the observed values and extending to the $\pm 1\sigma$ observational uncertainty. We find that the combination of separation, r , and transverse velocity, v_t , does not sufficiently discriminate between different masses over this interval. Interestingly, this is not only due to the large observational error, but also due to the fact that the 1σ scatter in v_t ranges from $\sim 20\text{--}90 \text{ km s}^{-1}$ at $M = 2 \times 10^{12} M_{\odot}$ to $\sim 40\text{--}150 \text{ km s}^{-1}$ at $M = 6 \times 10^{12} M_{\odot}$. Each mass range includes a significant overlap with the observed value and uncertainty, and v_t values as high as 125 km s^{-1} are not uncommon for LG masses of up to $4 \times 10^{12} M_{\odot}$. Higher mass pairs tend to have greater v_t values, which in addition to the effects mentioned in Section 4.3 may partly explain the higher value reported by Forero-Romero & Sierra-Porta (2022), and values as high as 110 km s^{-1} can certainly not be ruled out. Only a value of v_t as high as 164 km s^{-1} , as suggested by Salomon et al. (2016), is disfavoured (under 5 per cent probability in combination with the observations of r and v_r) if the LG mass is not in excess of $4 \times 10^{12} M_{\odot}$.

By contrast, the combination of separation, r , and radial velocity v_r , strongly depends on the mass. For LG analogues with masses of $2 \times 10^{12} M_{\odot}$ and a separation of $\sim 770 \text{ kpc}$, the magnitude of the observed radial velocity is 2σ above the median, while for LG analogues of $6 \times 10^{12} M_{\odot}$, it is 1.2σ below the median. The observed kinematics are most likely to occur for LG analogues of just under $4 \times 10^{12} M_{\odot}$. However, the fact that low-mass LGs are more common (and, to a smaller extent, the relatively low transverse velocity) pushes the most likely LG mass down to $\sim 3.3 \times 10^{12} M_{\odot}$.

7 CONCLUSION

We have inferred the halo masses of the MW and M31, their combined mass, and mass ratio, by analysing kinematically selected LG analogues from the UCHUU cosmological simulation. To establish the relation between mass and kinematics, we distinguish between ‘true pairs’ in which the interactions between the MW and M31 dominate the gravitational force, and those whose kinematics are strongly affected by external influences. We find that this distinction significantly changes the distributions of kinematic variables.

Assuming Λ CDM, limiting the range of halo masses to $0.5\text{--}5 \times 10^{12} M_{\odot}$, and considering only true pairs where the gravitational field and tidal field are primarily due to the MW and M31

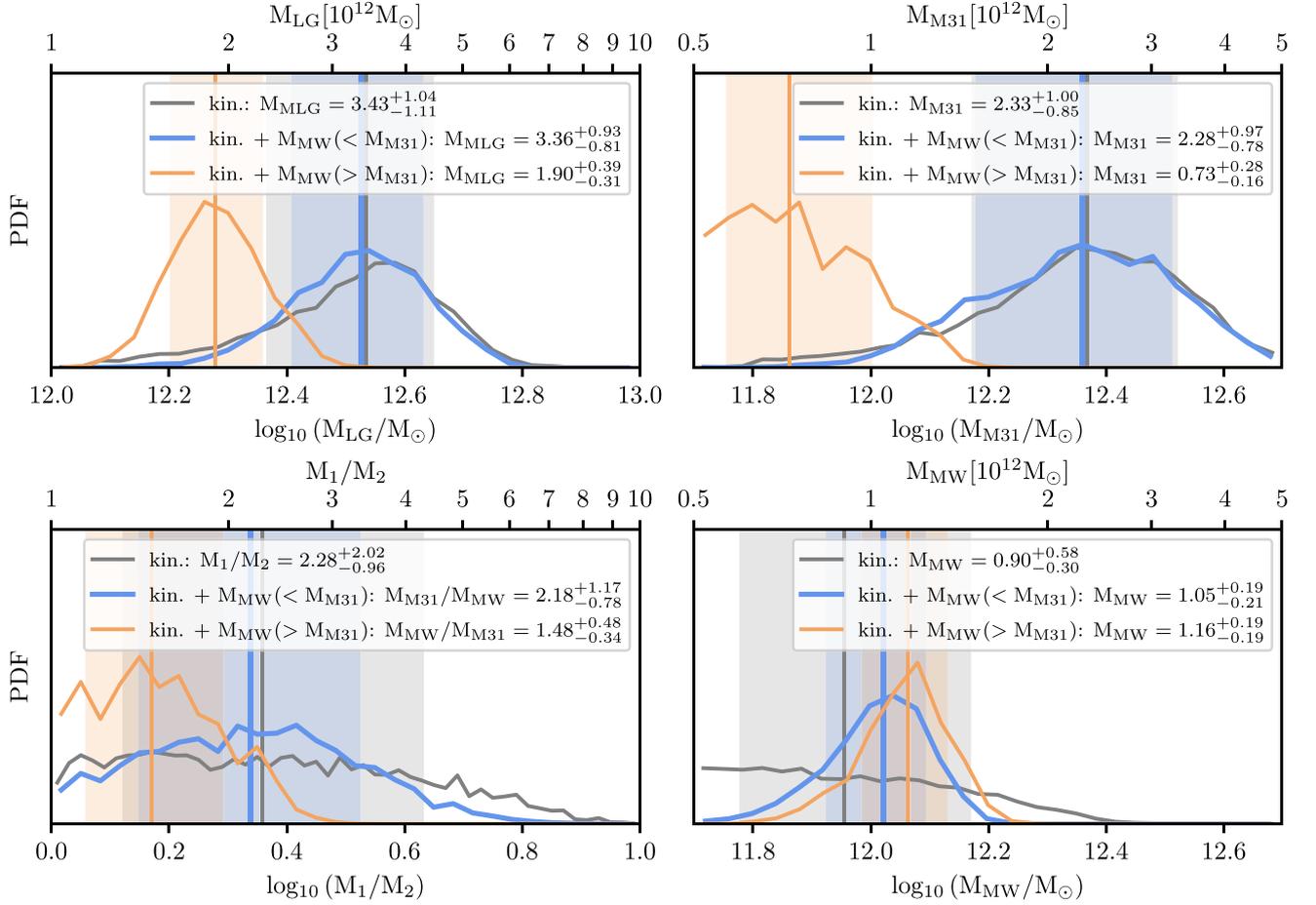


Figure 15. Distributions of the total mass, mass of M31, mass of the MW, and the mass ratio, measured in the simulation with each LG analogue weighed according to its likelihood to represent the observed LG, either assuming only the observed kinematics (grey, as the dark blue lines in Fig. 13), or also assuming $1.1^{+0.2}_{-0.2} \times 10^{12} M_{\odot}$ for the mass of the MW, as either the less massive of the pair ('regular mass order', in blue), or the more massive of the pair ('inverted mass order', in orange). The shaded areas show the regions between the 16th and 84th percentiles. Assuming that an MW of $\sim 1.1 \times 10^{12} M_{\odot}$ is the less massive galaxy yields very similar results for the total mass, mass ratio, and mass of M31, with slightly reduced errors. By contrast, assuming an MW of $\sim 1.1 \times 10^{12} M_{\odot}$ to be the more massive galaxy results in a much lower total mass, a much lower M31 mass, and a mass ratio much closer to unity.

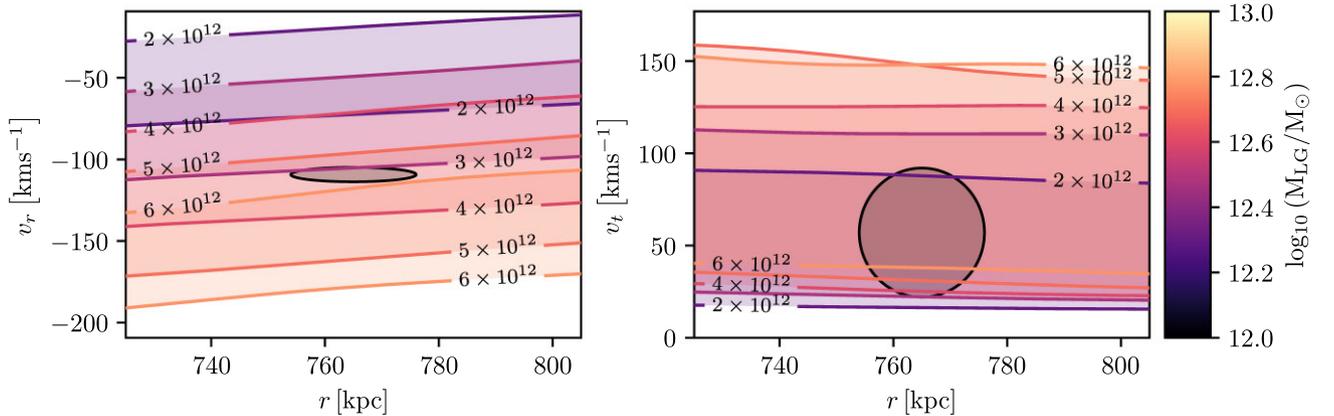


Figure 16. Distributions of kinematic parameters for LG analogues with masses of $2 \pm 0.2 \times 10^{12} M_{\odot}$, $3 \pm 0.2 \times 10^{12} M_{\odot}$, $4 \pm 0.2 \times 10^{12} M_{\odot}$, $5 \pm 0.2 \times 10^{12} M_{\odot}$, and $6 \pm 0.2 \times 10^{12} M_{\odot}$. On both panels, contour lines correspond to $\pm 1\sigma$ around the peak density, with the upper limits in v_r and v_t denoted on the left side, and the lower limits denoted on the right side of each panel. The ellipses on both panels show the observational uncertainty. In terms of v_r and r , as shown on the left panel, the observed kinematics are typical for LG analogues with a mass of $\sim 4 \times 10^{12} M_{\odot}$, and lie just outside the 1σ range for LG analogues with a mass of $3 \times 10^{12} M_{\odot}$ or $6 \times 10^{12} M_{\odot}$. The scatter in v_t is larger than that in v_r in every mass range, and all masses are compatible with the measured v_t value.

themselves, our analysis using GPR results in a total mass of $M_{\text{LG}} = 3.2_{-1.2}^{+0.9} \times 10^{12} M_{\odot}$. We obtain a nearly identical result using a probability weighted estimate.

Our result for the mass of the LG is significantly lower than most previous results which apply the timing argument to the kinematics, but in agreement with Benisty et al. (2022), who obtain a total mass of $3.4 \times 10^{12} M_{\odot}$ when correcting the timing argument for cosmic bias. As discussed in Section 6.4, accounting for cosmic bias (the fact that there are more low-mass than high-mass haloes), which is folded into our analysis, is certainly important for determining the most likely mass. However, it is not clear that our agreement with Benisty et al. (2022) is more than incidental, and we leave a more detailed analysis of the timing argument to future work.

We find that the kinematics most strongly constrain the total mass of the pair, leaving the mass of the two haloes and their mass ratio less well constrained. But importantly, the kinematics are not in conflict with the individual measurements. In particular, contrary to previous claims, they do not point towards a mass of the MW above $10^{12} M_{\odot}$. In fact, a purely kinematic analysis gives a mass for the less massive halo of $0.9_{-0.31}^{+0.59} \times 10^{12} M_{\odot}$. When we assume an MW mass of $1.1 \pm 0.2 \times 10^{12} M_{\odot}$ as an additional prior, we also obtain essentially the same posteriors for the total mass, M31 mass, and mass ratio, assuming that the MW is the less massive of the pair. Lending strong support for the ‘regular’ mass order, if we assume that the MW is more massive, all the distributions shift significantly, showing strong tensions with the observations.

We also find that LG analogues with close to equal masses are quite unlikely, and find a mass ratio of $2.28_{-0.96}^{+2.02}$ using only kinematics. It is worth noting that value and its distribution are quite sensitive to the lower bound of $5 \times 10^{11} M_{\odot}$ for each halo: allowing lower mass LG members is bound to shift the mass ratio distribution towards even higher values, while assuming a mass of $\sim 1.1 \pm 0.2 \times 10^{12} M_{\odot}$ for the less massive halo gives a mass ratio of $2.18_{-0.78}^{+1.17} M_{\odot}$. For the mass of the more massive halo, we find $2.33_{-0.85}^{+1.00} \times 10^{12} M_{\odot}$ based only on pair kinematics, and we find $2.28_{-0.78}^{+0.97} \times 10^{12} M_{\odot}$ for the mass of M31, assuming it is the more massive halo in an LG with an MW of $1.1 \pm 0.2 \times 10^{12} M_{\odot}$.

It is worth noting that the pair kinematics only give a probability distribution for the mass of the LG. Both the GPR and our probability weighted estimate show significant tails towards higher and lower masses. To further constrain the mass (albeit forgoing the potential to reveal possible tensions), and reduce the residual scatter, additional independent tracers will need to be considered, which further set the LG apart from the systems we consider here, and which may contain additional information about its mass. This may include the observed stellar masses of the galaxies obtained through abundance matching (e.g. Guo et al. 2011), the presence and orbits of (massive) satellites (Tepper-García, Bland-Hawthorn & Li 2020; Wang et al. 2020; Benisty et al. 2022; Chamberlain et al. 2023), the assembly histories of the MW and M31, the properties of the local Hubble flow (e.g. Fattahi et al. 2016), and finally possibly the peculiar local topology surrounding the LG (e.g. Neuzil, Mansfield & Kravtsov 2020; Aragon-Calvo, Silk & Neyrinck 2023).

Importantly, the precision of our mass estimate is neither limited by statistics, nor by observational uncertainty. Instead, we find a fundamental and irreducible scatter in the relation between mass and kinematics of $\sim \pm 10^{12} M_{\odot}$. This leads us to conclude that previous works which reported a smaller uncertainty purely based on the parameters of the MW–Andromeda orbit, either through the timing argument or machine learning, did not fully account for this scatter. On the other hand, in order to obtain a more accurate mass

estimate, future work will need to include additional information in the analysis.

ACKNOWLEDGEMENTS

We thank Marius Cautun, Shihong Liao, Jorge Peñarubbia, and Kai Puolamäki for insightful discussions and advice, and the referee for their careful reading of our manuscript, and in particular, for providing directions for future work. We thank the creators of the UCHUU simulation, Ishiyama et al., for making their data publicly available.

TS acknowledges support from Academy of Finland grants 311049 and 339127, and from the European Research Council (ERC) Advanced Investigator grant DMIDAS (GA 786910). TS and MT were also supported by Academy of Finland grant 335607. PHJ acknowledges the support by the European Research Council via the ERC Consolidator Grant KETJU (no. 818930) and the Academy of Finland grant 339127.

This work used facilities hosted by the CSC—IT Centre for Science, Finland, and the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

We thank Instituto de Astrofísica de Andalucía (IAA-CSIC), Centro de Supercomputación de Galicia (CESGA) and the Spanish academic and research network (RedIRIS) in Spain for hosting UCHUU DR1 and DR2 in the Skies & Universes site for cosmological simulations. The UCHUU simulations were carried out on Aterui II supercomputer at Center for Computational Astrophysics, CfCA, of National Astronomical Observatory of Japan, and the K computer at the RIKEN Advanced Institute for Computational Science. The UCHUU DR1 and DR2 effort has used the skun@IAA RedIRIS and skun6@IAA computer facilities managed by the IAA-CSIC in Spain (MICINN EU-Feder grant EQC2018-004366-P).

We gratefully acknowledge the use of open-source software, including MATPLOTLIB (Hunter 2007), SCIKIT-LEARN (Pedregosa et al. 2011), SCIPY (Virtanen et al. 2020), and NUMPY (Harris et al. 2020).

DATA AVAILABILITY

This work is based entirely on public data. The UCHUU simulation data are available for download at <http://skiesanduniverses.org/Simulations/Uchuu/> (see Acknowledgements).

Our analysis uses the open source software listed in the acknowledgements. Our documented analysis script, which can be used to reproduce all figures and in this paper from the above public data, is available in the form of documented python scripts and a jupyter notebook at https://github.com/TillSawala/LG_Mass.

REFERENCES

- Ablimit I., Zhao G., Flynn C., Bird S. A., 2020, *ApJ*, 895, L12
 Aragon-Calvo M. A., Silk J., Neyrinck M., 2023, *MNRAS*, 520, L28
 Benisty D., Vasiliev E., Evans N. W., Davis A.-C., Hartl O. V., Strigari L. E., 2022, *ApJ*, 928, L5
 Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2012, *MNRAS*, 422, 1203
 Callingham T. M. et al., 2019, *MNRAS*, 484, 5453
 Carlesi E. et al., 2016, *MNRAS*, 458, 900
 Cautun M. et al., 2020, *MNRAS*, 494, 4291

- Chamberlain K., Price-Whelan A. M., Besla G., Cunningham E. C., Garavito-Camargo N., Peñarrubia J., Petersen M. S., 2023, *ApJ*, 942, 18
- Deason A. J. et al., 2021, *MNRAS*, 501, 5964
- Fardal M. A. et al., 2013, *MNRAS*, 434, 2779
- Fattahi A. et al., 2016, *MNRAS*, 457, 844
- Font A. S., McCarthy I. G., Belokurov V., Brown S. T., Stafford S. G., 2022, *MNRAS*, 511, 1544
- Forero-Romero J. E., Sierra-Porta D., 2022, *ApJ*, 939, 16
- Fritz T. K., Di Cintio A., Battaglia G., Brook C., Taibi S., 2020, *MNRAS*, 494, 5178
- Guo Q., White S., Li C., Boylan-Kolchin M., 2010, *MNRAS*, 404, 1111
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Harris C. R. et al., 2020, *Nature*, 585, 357
- Hartl O. V., Strigari L. E., 2022, *MNRAS*, 511, 6193
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Ishiyama T. et al., 2021, *MNRAS*, 506, 4210
- Kaffe P. R., Sharma S., Lewis G. F., Robotham A. S. G., Driver S. P., 2018, *MNRAS*, 475, 4043
- Karukes E. V., Benito M., Iocco F., Trotta R., Geringer-Sameth A., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 033
- Kennedy R., Frenk C., Cole S., Benson A., 2014, *MNRAS*, 442, 2487
- Li Y.-S., White S. D. M., 2008, *MNRAS*, 384, 1459
- Li Z.-Z., Qian Y.-Z., Han J., Li T. S., Wang W., Jing Y. P., 2020, *ApJ*, 894, 10
- Libeskind N. I., Yepes G., Knebe A., Gottlöber S., Hoffman Y., Knollmann S. R., 2010, *MNRAS*, 401, 1889
- Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, *MNRAS*, 439, 300
- McLeod M., Libeskind N., Lahav O., Hoffman Y., 2017, *J. Cosmol. Astropart. Phys.*, 2017, 034
- Neuzil M. K., Mansfield P., Kravtsov A. V., 2020, *MNRAS*, 494, 2600
- Partridge C., Lahav O., Hoffman Y., 2013, *MNRAS*, 436, L45
- Pawlowski M. S., 2018, *Mod. Phys. Lett. A*, 33, 1830004
- Peñarrubia J., Ma Y.-Z., Walker M. G., McConnachie A., 2014, *MNRAS*, 443, 2204
- Pedregosa F. et al., 2011, *J. Mach. Learn. Res.*, 12, 2825
- Petersen M. S., Peñarrubia J., 2020, *MNRAS*, 494, L11
- Peñarrubia J., Fattahi A., 2017, *MNRAS*, 468, 1300
- Prudil Z. et al., 2022, *A&A*, 664, A148
- Rodríguez Wimberly M. K. et al., 2022, *MNRAS*, 513, 4968
- Salomon J. B., Ibata R. A., Famaey B., Martin N. F., Lewis G. F., 2016, *MNRAS*, 456, 4432
- Salomon J. B., Ibata R., Reylé C., Famaey B., Libeskind N. I., McConnachie A. W., Hoffman Y., 2021, *MNRAS*, 507, 2592
- Sawala T. et al., 2015, *MNRAS*, 448, 2941
- Sawala T. et al., 2022, *Nat. Astron.*, published online at <https://doi.org/10.1038/s41550-022-01856-z>
- Shen J. et al., 2022, *ApJ*, 925, 1
- Slizewski A., Dufresne X., Murdock K., Eadie G., Sanderson R., Wetzel A., Jurić M., 2022, *ApJ*, 924, 131
- Tamm A., Tempel E., Tenjes P., Tihhonova O., Tuvikene T., 2012, *A&A*, 546, A4
- Tepper-García T., Bland-Hawthorn J., Li D., 2020, *MNRAS*, 493, 5636
- Tollerud E. J. et al., 2012, *ApJ*, 752, 45
- van der Marel R. P., Guhathakurta P., 2008, *ApJ*, 678, 187
- van der Marel R. P., Fardal M., Besla G., Beaton R. L., Sohn S. T., Anderson J., Brown T., Guhathakurta P., 2012, *ApJ*, 753, 8
- van der Marel R. P., Fardal M. A., Sohn S. T., Patel E., Besla G., del Pino A., Sahlmann J., Watkins L. L., 2019, *ApJ*, 872, 24
- Veljanoski J. et al., 2013, *ApJ*, 768, L33
- Veljanoski J. et al., 2014, *MNRAS*, 442, 2929
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Wang J., 2020, preprint ([arXiv:2009.10862](https://arxiv.org/abs/2009.10862))
- Wang W., Han J., Cautun M., Li Z., Ishigaki M. N., 2020, *Sci. China Phys. Mech. Astron.*, 63, 109801
- Watkins L. L., Evans N. W., An J. H., 2010, *MNRAS*, 406, 264
- Watkins L. L., van der Marel R. P., Sohn S. T., Evans N. W., 2019, *ApJ*, 873, 118

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