



Radial Distributions of Dwarf Satellite Systems in the Local Volume

Scott G. Carlsten¹ , Jenny E. Greene¹ , Annika H. G. Peter^{2,3,4} , Johnny P. Greco^{4,6} , and Rachael L. Beaton^{1,5,7} ¹Department of Astrophysical Sciences, 4 Ivy Lane, Princeton University, Princeton, NJ 08544, USA; scottgc@princeton.edu²Department of Physics, The Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210, USA³Department of Astronomy, The Ohio State University, 140 W. 18th Ave., Columbus, OH 43210, USA⁴Center for Cosmology and AstroParticle Physics (CCAPP), The Ohio State University, Columbus, OH 43210, USA⁵The Observatories of the Carnegie Institution for Science, 813 Santa Barbara St., Pasadena, CA 91101, USA

Received 2020 June 3; revised 2020 September 4; accepted 2020 September 6; published 2020 October 20

Abstract

The radial spatial distribution of low-mass satellites around a Milky Way (MW)-like host is an important benchmark for simulations of small-scale structure. The distribution is sensitive to the disruption of subhalos by the central disk and can indicate whether the disruption observed in simulations of MW analogs is artificial (i.e., numerical) or physical in origin. We consider a sample of 12 well-surveyed satellite systems of MW-like hosts in the Local Volume ($D < 12$ Mpc) that are complete to $M_V < -9$ and within 150 projected kpc. We investigate the radial distribution of satellites and compare with Λ CDM cosmological simulations, including big-box cosmological simulations and high-resolution zoom-in simulations of MW-sized halos. We find that the observed satellites are significantly more centrally concentrated than the simulated systems. Several of the observed hosts, including the MW, are $\sim 2\sigma$ outliers relative to the simulated hosts in being too concentrated, while none of the observed hosts are less centrally concentrated than the simulations. This result is robust to different ways of measuring the radial concentration. We find that this discrepancy is more significant for bright, $M_V < -12$ satellites, suggestive that this is not the result of observational incompleteness. We argue that the discrepancy is possibly due to artificial disruption in the simulations, but, if so, this has important ramifications for what relation between stellar mass and halo mass is allowed in the low-mass regime by the observed abundance of satellites.

Unified Astronomy Thesaurus concepts: Dark matter (353); Local Group (929); Milky Way Galaxy (1054); Dwarf galaxies (416)

1. Introduction

One important observational benchmark with which to test models of small-scale structure formation is the radial distribution of dwarf satellites around the Milky Way (MW) and MW-like hosts. The radial distribution of luminous, low-mass satellites is sensitive to the physics of reionization (e.g., Kravtsov et al. 2004; Ocvirk & Aubert 2011; Dooley et al. 2017) and to the disruption of the subhalos that host the satellites by the central primary (e.g., D’Onghia et al. 2010; Brooks et al. 2013; Garrison-Kimmel et al. 2017b; Kelley et al. 2019; Richings et al. 2020; Samuel et al. 2020). Of particular importance is understanding whether the disruption of subhalos is physical or an artificial feature of the simulations (due to, e.g., low-resolution effects; van den Bosch et al. 2018; van den Bosch & Ogiya 2018). Tidal stripping and disruption of subhalos are both integral parts of the baryonic solutions to the well-known “small-scale challenges” to Λ CDM (e.g., Zolotov et al. 2012; Brooks & Zolotov 2014; Garrison-Kimmel et al. 2017b, 2019), and hence it is critical to understand them fully.

Previous comparisons between observations of the MW satellites and the predictions from Λ CDM simulations have produced somewhat baffling results. Comparisons with the classical satellites ($M_* \gtrsim 10^5 M_\odot$) indicate that the MW satellites are significantly more radially concentrated than the most massive dark matter (DM) subhalos in dark-matter-only (DMO) simulations (Kravtsov et al. 2004; Lux et al. 2010; Yniguez et al. 2014). Many studies have argued that reionization can help in this regard because the distribution

of *luminous* subhalos will be more centrally concentrated than the overall population of subhalos (e.g., Moore 2001; Kravtsov et al. 2004; Taylor et al. 2004; Font et al. 2011; Starkenburg et al. 2013; Barber et al. 2014; Dooley et al. 2017). The earliest forming subhalos will be the ones most likely to be luminous and they are more concentrated near the host. While the halo mass scale at which reionization starts to significantly suppress galaxy formation is still uncertain, recent simulations indicate that it is at (or well below) the low-mass end of halos expected to host classical-sized satellites (Sawala et al. 2016; Wheeler et al. 2019). Thus, reionization does not appear to be a viable explanation for the concentration of the MW classical satellites. The fully hydrodynamic simulations presented in Samuel et al. (2020) showed a similar result to Yniguez et al. (2014) with the MW still substantially more concentrated than simulations in several metrics. Unfortunately, the results in comparison to the classical satellites are fundamentally limited by the low statistics offered by the ~ 10 MW classical satellites.

One way to increase statistics is to consider the ultrafaint dwarf (UFD) satellites of the MW as well, of which there are now ~ 50 known. The MW UFDs seem also to be more concentrated than DMO simulations would predict, particularly when disruption by the baryonic disk of the host is included. Investigating the simulations of Garrison-Kimmel et al. (2017b), Kim et al. (2018) found that there were not enough subhalos that survived the enhanced disruption by the disk to host the known MW UFDs at small radii, assuming a reasonable stellar-to-halo mass relation (SHMR). A similar conclusion was reached by Graus et al. (2019), who noted that the close-in MW UFDs could be explained if *very* low-mass subhalos were populated. These low-mass subhalos are well

⁶ NSF Astronomy & Astrophysics Postdoctoral Fellow.⁷ Hubble Fellow.

below the usual cutoff for luminous satellites due to reionization suppression of galaxy formation (e.g., Bullock et al. 2000; Somerville 2002; Okamoto et al. 2008; Okamoto & Frenk 2009). It is still unclear how galaxies could form in these halos, but if they do, there is no reason to suspect they would be centrally concentrated. Thus, there should be a very large number of UFDs in the outskirts of the MW virial volume awaiting discovery (see also Kim et al. 2018, for estimates).

Using the UFDs to test simulation predictions with observations comes with its own problems, however. The observational census of UFDs is radially incomplete due to their intrinsic faintness and small size (e.g., Koposov et al. 2008; Walsh et al. 2009) and this is biasing the observed radial distribution. Also, as the UFDs likely reside in very low-mass subhalos, the resolution of the simulations becomes a major concern. In the end, the MW is still only one system, and it is difficult to draw broad conclusions on the radial distribution of satellites from the MW alone.

In this paper, we take a complementary approach to the work presented above by studying the observed radial distributions of classical-mass satellites around many hosts in the Local Volume (LV; $D < 12$ Mpc). By comparing multiple MW-like hosts together, we are able to build far better statistics than with the MW alone. The basic question that this paper tries to answer is “How well do the spatial distributions of dwarf satellite systems created in modern simulations match those observed for the MW and MW analogs in the LV?” We compare the observed systems to a wide range of recent simulations, including big-box cosmological hydrodynamic simulations with many simulated MW-like hosts, high-resolution DMO zoom simulations (both with and without an included disk potential), and high-resolution hydrodynamic zoom simulations that include only a handful of MW-like hosts. Using these very different simulations allows us to explore how much the results of the simulation depend on its properties, including resolution. In this paper, we thus perform the first comparison between a *population* of observed satellite systems and a *population* of simulated analogs. This has only recently been made possible with the creation of all of these simulation suites and the observations required to characterize the satellite systems of nearby MW analogs. By studying the radial distributions of a population of satellite systems, it might also be possible to learn about the scatter between satellite systems (i.e., Why is the radial distribution of M31 so different from that of the MW? Willman et al. 2004; Yniguez et al. 2014).

This paper is structured as follows: in Section 2 we present the observational data sets of LV satellite systems, in Section 3 we list the different simulation suites that we compare against, and in Section 4 we show the comparison of observations with models. In Section 5, we discuss possible causes for the discrepancy that we find, including possible caveats related to both the observations and simulations, and we conclude and outline directions for future work in Section 6.

2. Sample of Satellite Systems

In this work, we compare simulations with a sample of 12 well-characterized observed satellite systems around massive hosts in the Local Volume. Characterizing these satellite systems has been the result of the combined effort of multiple groups over the last several years. Still, due to the inherent faintness of dwarf satellites and the large areas that need to be

surveyed, only a small number of massive hosts have been surveyed to the point where the inventory of “classical”-sized satellites is likely complete for a large fraction of the host’s virial volume. Most of the difficulty is in measuring the distances to these low-mass galaxies to confirm that they are actually physically associated to a specific host. Several previous studies (e.g., Bennet et al. 2019; Carlsten et al. 2019a, 2020b) have shown that the contamination from background interlopers along the line of sight can be significant ($>80\%$).

For this study, we use the compilation of LV satellite systems given in Carlsten et al. (2020b, hereafter Paper II). That work uses the catalog of satellite candidates around 10 LV massive hosts from Carlsten et al. (2020a, hereafter Paper I) and confirms satellites using distances measured via surface brightness fluctuations (SBF; Carlsten et al. 2019b). Six of the hosts we consider here were characterized in this work (NGC 1023, NGC 4258, NGC 4631, M51, M104, and NGC 4565).⁸ The first five of these hosts are mostly complete down to a satellite luminosity of $M_V \sim -9$ and surface brightness of $\mu_{0,V} \sim 26.5$ mag arcsec⁻² within the inner ~ 150 projected kpc (see Paper I for more detailed quantification of completeness). For these hosts, there were still a few candidates where the SBF results were ambiguous. We treat these candidates as “possible” satellites and always give a spread of possible results (both including them as satellites and not). The sixth host, NGC 4565, had good area coverage and deep survey data, but due to its larger distance ($D = 11.9$ Mpc) and worse seeing, the SBF distance results were ambiguous for all but the brightest candidates. All of the candidates brighter than $M_V = -12$ were confirmed by SBF or redshift such that there were no ambiguous candidates in NGC 4565 at this luminosity. Thus, we include NGC 4565 as well, but we note that its satellite system is only characterized down to $M_V = -12$.

For NGC 4631, we have used DECaLS (Dey et al. 2019) with Gemini+GMOS follow-up to identify more satellites outside the original footprint of Paper I. Its satellite system is now likely complete to ~ 200 kpc. We describe the extra data in Appendix A.

In addition to the six hosts from Paper II, we include the six other nearby hosts that have had their satellites well surveyed in previous work. These six are the MW, M31, Centaurus A (CenA; NGC 5128), M81, M94, and M101. The specific lists of satellites that we consider in each of these systems can be found in the appendix of Paper II. The MW satellite list comes from McConnachie (2012) and uses distances from Fritz et al. (2018). We do include the dSph Sgr, but note that because it is currently disrupting, objects like it may no longer be identified in simulations by halo finders. It is not clear what the correspondence is between simulations and observations for various stages of tidal disruption, and whether systems like this can be meaningfully compared with simulations. Overall, our results do not change qualitatively with the inclusion of Sgr, and we discuss this more below. The M31 satellites come from McConnachie et al. (2018) with distances primarily from Weisz et al. (2019). The CenA satellites come from the compilations of Müller et al. (2019) and Crnojević et al. (2019). The satellites of M81 are taken from Chiboucas et al. (2013). M94 satellites come from Smercina et al. (2018). Finally, the M101 satellite system comes from the work of Bennet et al.

⁸ The other four hosts considered in that work had significantly less survey area coverage and are not considered here.

(2017, 2019), Danieli et al. (2017), and Carlsten et al. (2019a). These works surveyed the satellites out to ~ 200 kpc. We have used DECaLS (Dey et al. 2019) imaging to identify two more satellites of M101, which we have confirmed with SBF, completing its satellite system out to 300 kpc. We describe this in Appendix A.

A detailed discussion of the completeness of each of these satellite systems can be found in Paper II and references therein. In brief, we assume that the MW and M31 systems are complete to classical satellites ($M_V \lesssim -8$) within the inner 300 kpc. M101 is complete to $M_V \sim -8.5$ within the inner 300 kpc. M81 is likely complete to better than $M_V \lesssim -9$ within the inner projected 250 kpc. CenA is likely complete within the inner projected 200 kpc to about $M_V \lesssim -9$. M94 is complete within only the inner 150 kpc at this luminosity limit. For the six hosts from Paper II, we assume that the satellite systems are complete to $M_V \sim -9$ (with the exception of NGC 4565 noted above), and we use the actual survey footprints (see Figure 1 of Paper I) to characterize the areal completeness.

As in Paper II, we only consider satellites that have surface brightness above $\mu_{0,V} \sim 26.5$ mag arcsec $^{-2}$. Satellites with significantly lower surface brightness (even with total luminosity of $M_V \sim -9$) are detectable from resolved stars around the MW and M31, but these would not be detectable in the majority of the LV hosts we consider here. Additionally, like Sgr, several of the extremely low surface brightness (LSB) satellites (e.g., AndXIX, Collins et al. 2020) are clearly undergoing tidal stripping, and it is unclear whether these satellites are appropriate to include in the comparison with simulations below. The subhalos hosting such stripped systems might not be recognized by the halo finders used in the simulations, as discussed more below. Since Sgr is much more massive (2–3 orders of magnitude), it is more likely than these objects to be recognized as a subhalo.

Properties for all 12 hosts considered in this work can be found in Paper I and Paper II. These are all massive hosts with stellar mass ranging from roughly half that of the MW to $\sim 5 \times$ that of the MW. As discussed in Paper II, the hosts naturally split into two groups based on halo mass. The categorization of each host into these two groups was based on stellar mass, circular rotation speed, and any available estimate of the halo mass from satellite dynamics. The low-mass groups are all very similar to the MW, and we will refer to them as “MW-like” or “MW analogs.” These include the MW, M31, M94, M101, NGC 4631, NGC 4258, NGC 4565, and M51. We estimate that the halo masses of these hosts are in the range $\sim 0.8\text{--}3 \times 10^{12} M_\odot$. We refer to the more massive hosts as “small group” hosts and include M81, CenA, NGC 1023, and M104. These hosts correspond to halo masses in the range $\sim 3\text{--}8 \times 10^{12} M_\odot$. It is important to compare the observed hosts with simulated hosts of similar mass; we often consider each group of observed hosts separately and compare each individually with the appropriate simulated hosts.

This sample of 12 hosts is representative of the overall sample of massive hosts in the Local Volume. It includes nine star-forming spiral galaxies, two ellipticals, and one lenticular. We note that this sample is almost *volume-limited* for hosts within $D \sim 8.5$ Mpc. Our study contains eight out of 11 massive ($M_\star \gtrsim 0.5 M_\star^{\text{MW}}$) hosts with $|b| > 15^\circ$ within this volume. The missing hosts are NGC 5236 (M83), NGC 253, and NGC 4826 (M64).

3. Models

We compare the observations with a wide variety of simulations both to improve the statistics and to see how the inferences we draw regarding these systems depend on the specifics of the simulation. We consider both big-box cosmological simulations that contain many MW-like hosts but at a low resolution, and zoom-in cosmological simulations that focus on only a single MW-sized host but at a much higher resolution. We include DMO simulations, DMO simulations with an added disk potential, and fully hydrodynamic simulations. This allows us to explore the effect of subhalo disruption by the central disk on the spatial distribution of satellites.

For all of the simulation suites except for the zoom-in hydrodynamic simulations, we only use the DM halo catalogs from the simulations. To populate these subhalos with luminous galaxies we could use an SHMR to assign a stellar mass to each subhalo. Paper II found decent agreement between the observed satellite luminosity functions in the current sample and the simulated luminosity functions assuming either SHMR of Garrison-Kimmel et al. (2017a), two commonly adopted SHMRs from the literature. However, there is still significant uncertainty in what the true SHMR is for this mass range. Therefore, in order to keep the results regarding the radial distribution as general as possible, we do not use an SHMR to populate subhalos. Instead, we select the n most massive subhalos that fall in the survey footprints, where n is the number of observed satellites (either in a particular observed host or averaged over several observed hosts). More specific details are given with each comparison. When considering subhalo masses, we always consider the peak virial mass over the subhalo’s history, M_{peak} .

For the cosmological simulation, we use the public IllustrisTNG-100⁹ simulation (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018, 2019; Pillepich et al. 2018; Springel et al. 2018). We use the DM halo catalogs from the full hydrodynamic simulation run and not the DMO simulation run. The full hydro simulation should include the effect of subhalo disruption by the baryonic disk of the host. The baryonic mass resolution of TNG is $\sim 10^6 M_\odot$, which means that the satellites of the luminosity we are interested in ($M_V < -9$ mag) are not resolved. However, the DM particle mass of $7.5 \times 10^6 M_\odot$ means that subhalos hosting the satellites we are focusing on ($\sim 5 \times 10^9 M_\odot$) will contain ~ 1000 particles at infall. From the 100^3 Mpc 3 box, we select host halos as described in Paper II. In brief, when we compare the simulations individually with each observed host, we select the simulations to have stellar mass consistent with the observed host, using the central host’s stellar mass predicted in the hydrodynamic results. Each TNG host is given a probability to be included given by a Gaussian distribution in log stellar mass that is centered on the stellar mass of the observed host with spread 0.2 dex. We assume 0.2 dex is an appropriate estimate of the error in determining the stellar mass of nearby massive galaxies. The resulting distribution of stellar masses of the selected TNG hosts is peaked at the observed stellar mass of the LV host but allows some spread due to measurement uncertainty. When comparing with all observed hosts together, we select TNG hosts based on halo mass in the range $0.8 \times 10^{12} M_\odot < M_{200} < 8 \times 10^{12} M_\odot$.

⁹ <https://www.tng-project.org/data/>

For the zoom-in simulations, we use three separate simulation suites. These offer significantly higher resolution than IllustrisTNG, which allows us to explore the effect of resolution on the model predictions. The first is the ELVIS¹⁰ (Garrison-Kimmel et al. 2014) suite of DMO simulations. This suite includes 24 isolated MW-sized hosts and a further 24 that are in a paired Local Group (LG) configuration. We consider all 48 of these hosts in the same way. With a DM mass resolution of $1.6 \times 10^5 M_\odot$, this simulation has significantly higher resolution than IllustrisTNG. The second zoom simulation we include is the PhatELVIS¹¹ suite (Kelley et al. 2019) of DMO simulations. These simulations are distinct from the ELVIS suite both by having even higher resolution (DM mass resolution of $3 \times 10^4 M_\odot$) and by accounting for the enhanced disruption of subhalos due to the baryonic disk of the host. A gravitational potential grown to match that of the MW’s disk is artificially put into the simulations. PhatELVIS includes 12 simulated hosts. Both the ELVIS and PhatELVIS suites are DMO, and so we just use the DM subhalo catalogs, as described above. The final zoom simulation we include is the full hydrodynamic simulations from the NIHAO project¹² (Buck et al. 2019). There are only six hosts in this suite, but the simulations have very high resolution with DM particle mass of $\sim 10^5 M_\odot$ and star particle mass of $\sim 10^4 M_\odot$. Since the simulations are hydrodynamic we take the luminous galaxy catalog directly from the simulations. The lowest-mass satellite we consider in this paper has $M_* > 10^5 M_\odot$, which will be at least marginally resolved in the NIHAO results. This is a similar resolution to that used in the FIRE-2 simulations recently presented by Samuel et al. (2020). All of the zoom-in simulations are of hosts with roughly the halo mass of the MW, so we only compare these simulations to the MW-like observed hosts. Specific details (including masses) of each of the simulated zoom-in hosts can be found in the respective publications.

4. Results

In this section, we compare the observed satellite systems to the simulated systems. Throughout this paper, we focus primarily on the shape of the radial distribution of satellites and not the absolute radial distribution. We compare the observations with the models in many different ways, using distinct metrics, to try to get a complete understanding of how well they agree. Due to the large number of simulated hosts, we primarily compare the observations to the IllustrisTNG results, and secondarily consider the other simulation suites to show consistency. We start by considering the normalized 2D (projected) radial distributions of all of the observed hosts compared to the simulations. We then compare the normalized 3D distribution of satellites around the MW and M31 (the only observed hosts where we have 3D information on the satellite positions). Finally, we explore different ways of parameterizing the shape of the radial distribution.

4.1. 2D Projected Radial Distributions

In Figure 1, we show the radial profiles for each of the 12 observed hosts that we consider in this work. We only consider satellites brighter than $M_V < -9$ and assume that all hosts are

complete in luminosity down to this level (except NGC 4565, for which we use a luminosity limit of $M_V < -12$). Each host is compared with the simulated systems from IllustrisTNG. The IllustrisTNG hosts are selected based on stellar mass to be consistent with each observed host, as described in Section 3. We also show the radial profiles predicted by the DMO ELVIS simulation suite. The simulated systems are mock-observed at the distance of each host, and the observational area selection function for each host is used to select which simulated subhalos would be observed. We select subhalos whose line-of-sight (LOS) distance from the observer is within 500 kpc of the host. This accounts for the fact that for most of the hosts, the distances available for the satellites are not of high enough precision to probe the 3D structure of the group. SBF distances are accurate to $\sim 15\%$ (e.g., Carlsten et al. 2019b), while Hubble Space Telescope (HST) distances using the tip of the red giant branch (TRGB) are accurate to $\sim 5\%$ (e.g., Danieli et al. 2017; Bennet et al. 2019). At a host distance of 7 Mpc, these correspond to uncertainties of ~ 1000 and 300 kpc, respectively. It is entirely possible that some of the “confirmed” satellites of these hosts are actually near-field galaxies that project onto the host but are outside its virial volume. Thus, we account for the LOS uncertainty by including subhalos within 500 kpc LOS of the host. For the MW and M31, we have detailed 3D positions for the satellites, and we use those to mock “re-observe” these systems at a distance of 7 Mpc. This allows us to explore the effect of the observing angle on the radial profile. For the hosts from Paper II that have some unconfirmed candidate satellites, the uncertainty in membership is accounted for as a spread in possible radial profiles. Specifically, each possible combination of the unconfirmed members is considered, and that many radial profiles are generated. We plot the median and $\pm 1\sigma$ spread in these profiles.

Since the scatter in the profile between hosts will depend on how many satellites each host has, for a fair comparison we select the same number of subhalos from each simulated host as is observed for a particular observed host. The n most massive subhalos (considering peak mass) that fall in the survey footprint are selected as the luminous satellites, where n is the number of observed satellites for a specific host.

While there clearly is significant scatter between the observed hosts in Figure 1, they appear to be generally more concentrated than the simulated hosts. Several of the observed hosts (e.g., the MW, M101, NGC 4258, NGC 4631, NGC 4565, and NGC 1023) have their profiles at or just within the -2σ (i.e., more centrally concentrated) scatter in the simulations whereas no host is correspondingly outside the $+2\sigma$ (i.e., less concentrated) scatter in the simulations. The “small-group” hosts are less discrepant with the simulations. Indeed, both M81 and CenA closely follow the median simulated profile.

Another way to assess the concentration of the satellite population is the histogram of the satellites’ projected separations, r_{proj} , from their hosts. In Figure 2, we show the distribution of all projected separations of satellites across all hosts combined. The histograms are normalized such that the total area under the curve is the average number of satellites per host.¹³ Only satellites within $r_{\text{proj}} < 150$ kpc are included. We

¹⁰ <http://localgroup.ps.uci.edu/elvis/index.html>

¹¹ <http://localgroup.ps.uci.edu/phat-elvis/>

¹² http://www2.mpi-hd.mpg.de/~buck/#sim_data

¹³ The bins are simply chosen to be 10 kpc wide. While different binning schemes could significantly change the appearance of the distributions, we discuss the significance of the disagreement of the distributions using metrics that do not require binning below.

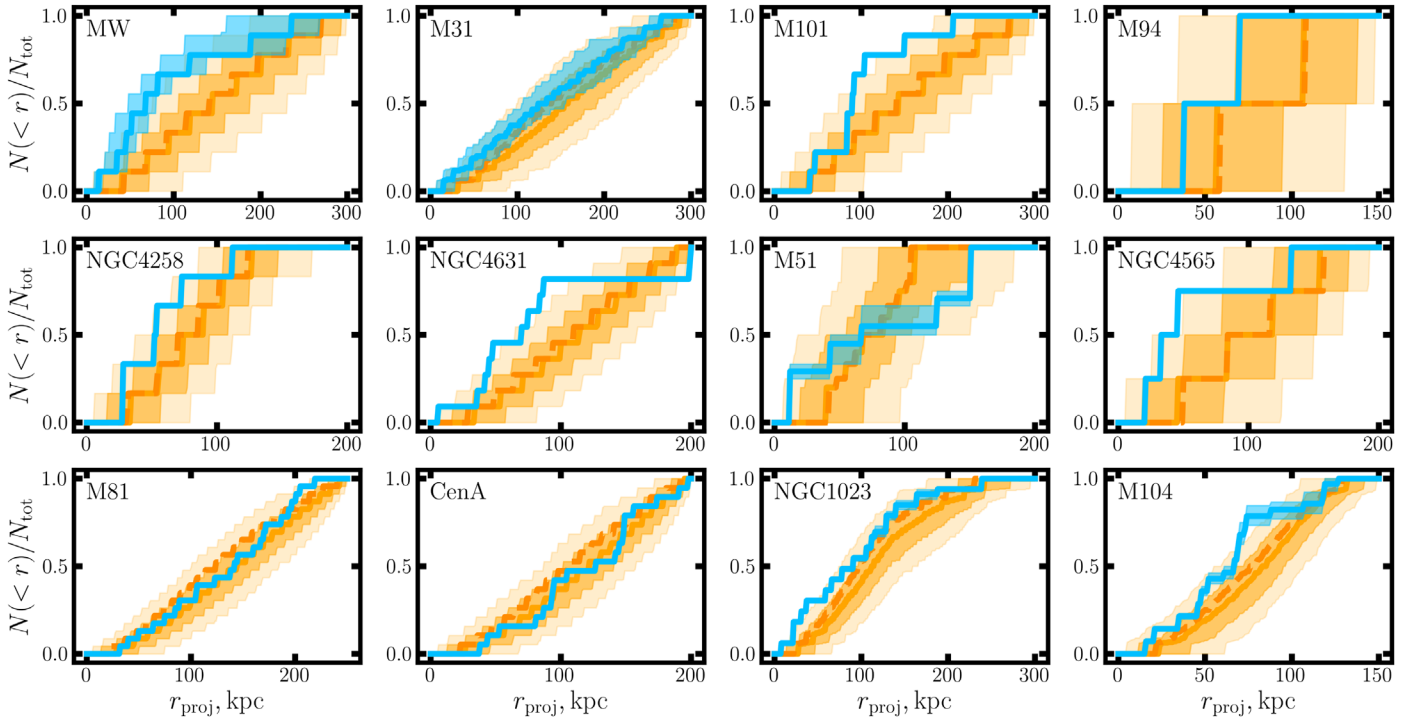


Figure 1. The 2D (projected) cumulative radial distributions of satellites ($M_V < -9$) around the 12 LV hosts considered in this work. All profiles are normalized by the total number of satellites in the surveyed area. Observed systems are shown in blue while the analogous simulated systems from IllustrisTNG are shown in orange. The dark orange dashed line shows the median profile for the ELVIS DMO simulations. The simulations are forward-modelled using the area completeness of the surveys for each host. For the simulations, the n most massive subhalos falling in the survey region are identified as the luminous satellites, where n is the observed number of satellites in the survey region for that specific host. For the MW and M31, the shaded region shows the effect of different projection angles on the radial profile. For the other LV hosts, the shaded region encompasses any uncertainty in membership of candidate satellites without distance information. The bottom row of hosts are the more massive “small-group” hosts. Note that the range of the x -axis is different for different hosts, depending on the radial coverage of its satellite census.

consider the MW-like and small-group hosts separately and look at all ($M_V < -9$) satellites and just the brighter ($M_V < -12$) satellites. This luminosity threshold is motivated by the completeness limit of the SAGA Survey (Geha et al. 2017) and enables comparison with this survey, because it is the closest in spirit to this one. The histograms of the observations are averaged over the viewing angle for the MW and M31 and also averaged over the uncertainty in membership for some candidates.

The IllustrisTNG hosts are selected based on halo mass. Halos in the mass range $0.8 \times 10^{12} M_\odot < M_{200} < 3 \times 10^{12} M_\odot$ are compared with the “MW-like” observed hosts, and halos in the range $3 \times 10^{12} M_\odot < M_{200} < 8 \times 10^{12} M_\odot$ are compared to the “small-group” hosts. The simulations are forward-modelled to include the effect of the survey footprints of the observed hosts in a similar way to that in Figure 1. For each simulated host, one of the observed hosts is selected at random, and that simulated host is forward-modelled through the area selection function of that observed host. The decrease in satellites at large r_{proj} in the simulated hosts is largely due to some of the observed hosts not being surveyed fully to 150 kpc. For each NIHAO and PhatELVIS host, 50 different viewing angles are taken. For the IllustrisTNG and PhatELVIS simulations, the n most massive subhalos in the survey footprint are selected as the satellites, where n is the average number of observed satellites above the luminosity limit. The average number of observed satellites is given in the upper left corner of each histogram in Figure 2, along with the average total number of satellites among the hosts shown in each panel. The total number of satellites is not necessarily an integer due

to the effects of different projecting angles for the MW and M31 and uncertain membership for some of the other hosts. We use the NIHAO hydro results to take satellites above the same luminosity threshold as used for the observations.

The observed MW-like hosts have noticeably shifted distributions of r_{proj} compared to the simulations. Restricting to only the bright ($M_V < -12$) satellites makes this discrepancy *significantly* more noticeable. On the other hand, the more massive “small-group” hosts have observed r_{proj} distributions that are only slightly flatter than those of the simulated hosts.

To assess the significance of the discrepancy with the MW-like hosts, we use a two-sample Kolmogorov–Smirnov (KS) test between the observations and simulations. To account for the different viewing angles of the MW and M31 and uncertain membership, we consider many different realizations of the observed hosts (with different viewing angles and different memberships) and calculate the KS statistic with the ensemble of PhatELVIS simulations. We find median p -values that the two samples are drawn from the same distribution of 5×10^{-3} for all ($M_V < -9$) satellites and 5×10^{-4} for the bright ($M_V < -12$) satellites, when we just consider one viewing angle per PhatELVIS host. When compared to the TNG simulations, we find similar median p -values of 4×10^{-3} and 5×10^{-4} , respectively. The PhatELVIS and TNG simulations do show fair agreement with each other with p -values that their radial distributions are drawn from the same underlying distribution of ~ 0.2 (using one viewing angle per PhatELVIS host).

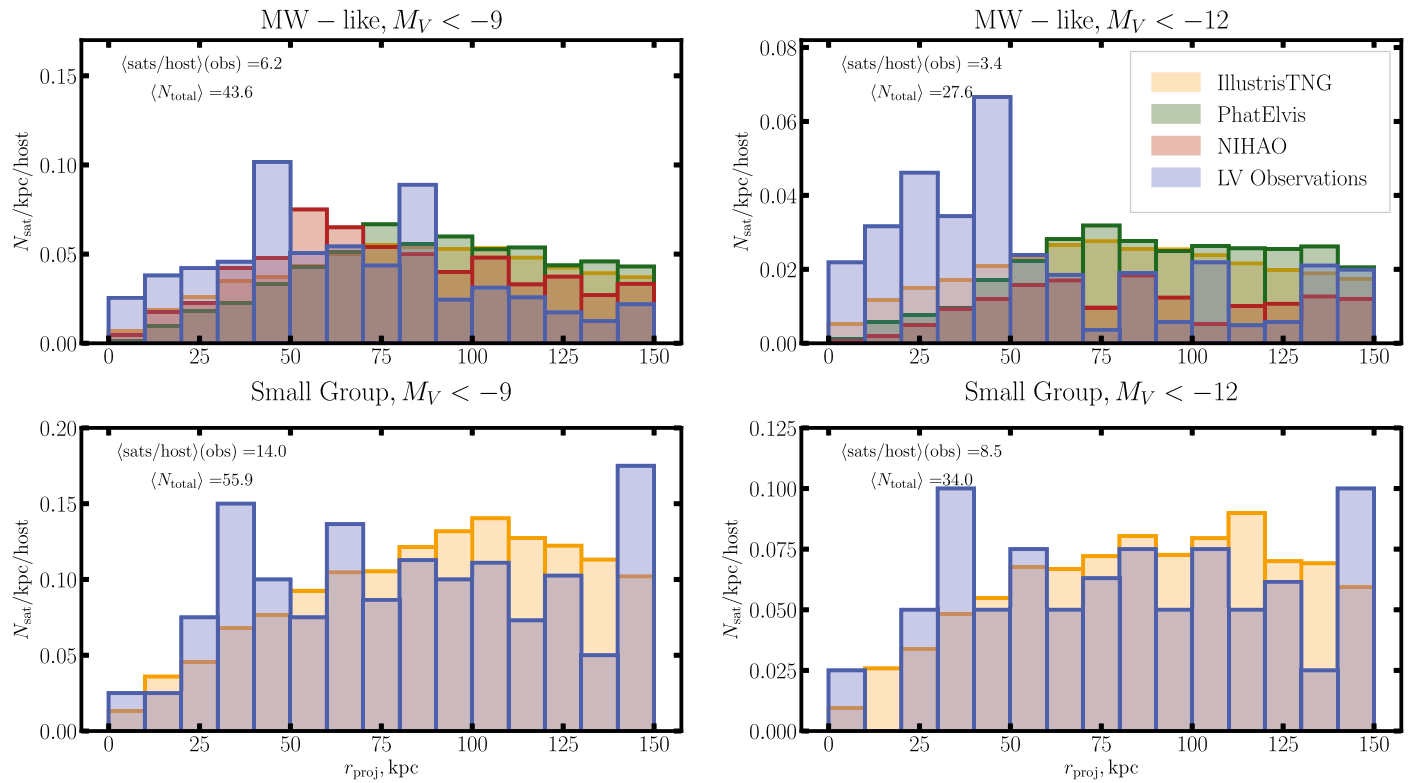


Figure 2. Histogram of the projected satellite separation (r_{proj}) for all observed hosts together compared against the simulated hosts. The observation histograms are normalized such that the total area is the average number of satellites per host. These numbers are given in the top left corner of each panel, along with the average total number of observed satellites contributing to each histogram (which is not necessarily an integer due to the effects of different projecting angles for the MW and M31). The top panels are for the MW-like hosts while the bottom ones are for the more massive “small-group” hosts. The left panels are for all satellites $M_V < -9$, while only $M_V < -12$ satellites are included in the right panels. We restrict to satellites within the inner 150 projected kpc, which is the extent of the radial coverage for most of the hosts.

4.2. 3D Radial Distributions

Figure 3 shows the 3D radial distributions of the classical ($M_V < -8$)¹⁴ satellites around M31 and the MW compared to the four simulation suites that we consider in this paper. The distributions are cumulative and normalized to the number of satellites within $r < 300$ kpc. For the DMO simulations, the 15 most massive (peak mass) subhalos are selected for each host, roughly in between the number of classical satellites around the MW and that around M31. The TNG hosts are selected based on halo mass in the range $0.8 \times 10^{12} M_\odot < M_{200} < 3 \times 10^{12} M_\odot$. For the NIHAO hosts, the hydrodynamic results are used, and satellites with $M_V < -8$ are selected.

There are two interesting things to note from this plot. The first is that the different simulation results look remarkably similar to each other. Ostensibly this means that at the satellite mass we focus on, the resolution of the simulations is not affecting the results (or is affecting them all in the same way), and that even the low-resolution IllustrisTNG results are converged at this mass scale. We consider this point in more detail in Section 5. The biggest difference between the simulations is that the PhatELVIS simulated hosts have significantly fewer satellites within 100 kpc than the other simulation suites. This is due to subhalo disruption by the disk potential that Kelley et al. (2019) injected into the simulations; ELVIS, however, has no added disk. Both the Illustris and

NIHAO simulations should have this effect because the host will form a disk in these hydrodynamic simulations, so it is unclear why the PhatELVIS simulations show a much more pronounced deficit of subhalos in the inner regions.

The second thing to note from Figure 3 is that the MW’s satellite distribution is significantly more centrally concentrated than any of the simulation results. It is well outside the 2σ regions for all of the simulations. This confirms the result of Yniguez et al. (2014), who used the ELVIS simulations. Most hosts have significant populations of satellites outside $r = 150$ kpc whereas the MW has only two in this range of luminosity and surface brightness. M31, on the other hand, appears to have a radial distribution fully consistent with the simulations. Both Yniguez et al. (2014) and Samuel et al. (2020) argue that this unusual concentration indicates that undiscovered MW satellites exist far out in the virial volume, awaiting discovery. We address the point of completeness of the MW satellite census in Section 5.1 and argue that this is unlikely and, even if true, unlikely to actually help ease the discrepancy.

The fact that the MW is unusually centrally concentrated in both 2D (projected) and 3D is suggestive that the other overly concentrated LV hosts would be similarly concentrated in 3D. Interestingly, we note that the MW appears to be a *significantly* more extreme outlier in 3D than in 2D, which is also suggestive of what the other LV hosts might look like in 3D.

¹⁴ Note that we still implement the $\mu_{0,V} < 26.5 \text{ mag arcsec}^{-2}$ cut for consistency. The result, however, is unchanged if we include satellites of all surface brightnesses.

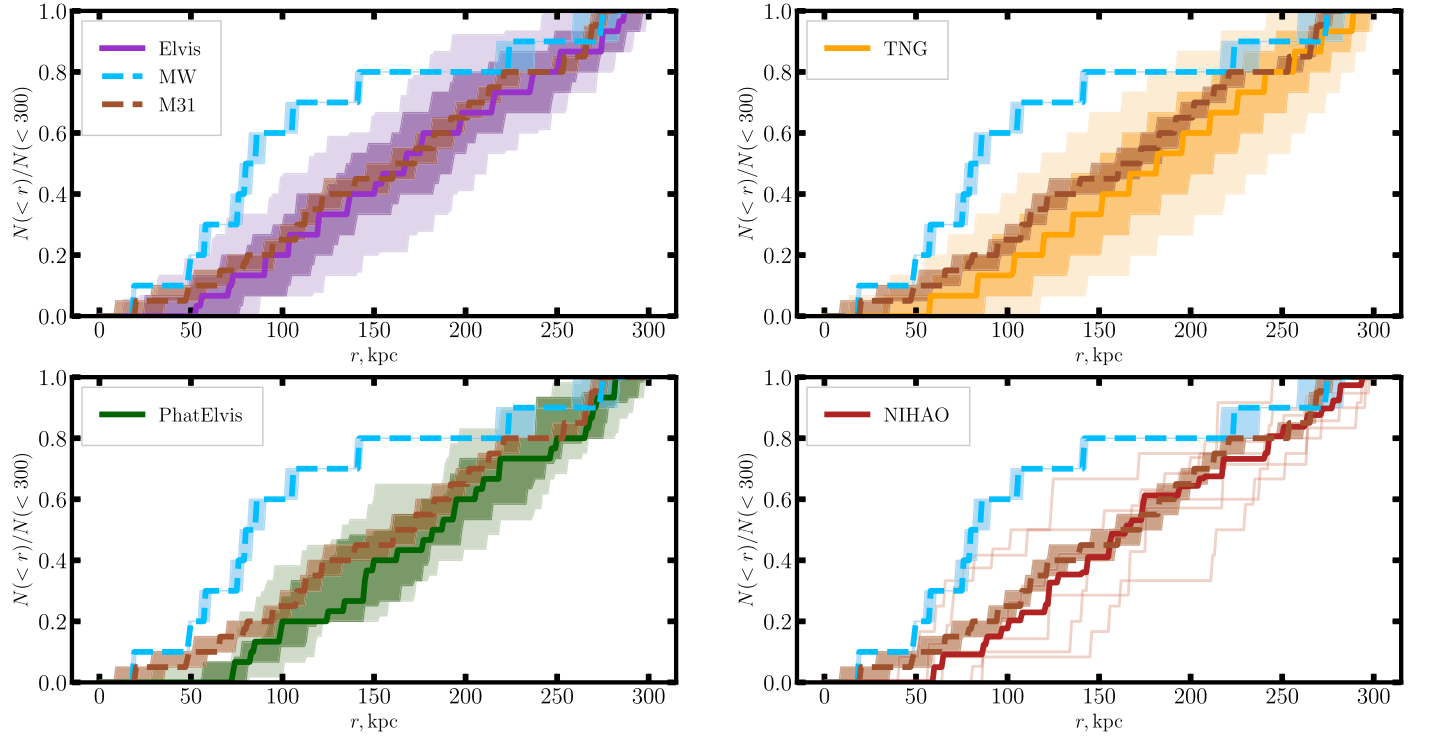


Figure 3. The 3D radial distribution of MW and M31 satellites ($M_V < -8$) compared with four different simulation suites. For the DMO simulations, the 15 most massive subhalos are selected as satellites. The profiles are cumulative and normalized by the number of satellites within 300 kpc. The shaded bands for M31 and the MW account for distance uncertainties in the satellites. The shaded bands for the simulations denote the $\pm 1\sigma$ and $\pm 2\sigma$ spread in the simulation results. As there are only six hosts in the NIHAO suite, they are plotted individually, and the thick line shows the median.

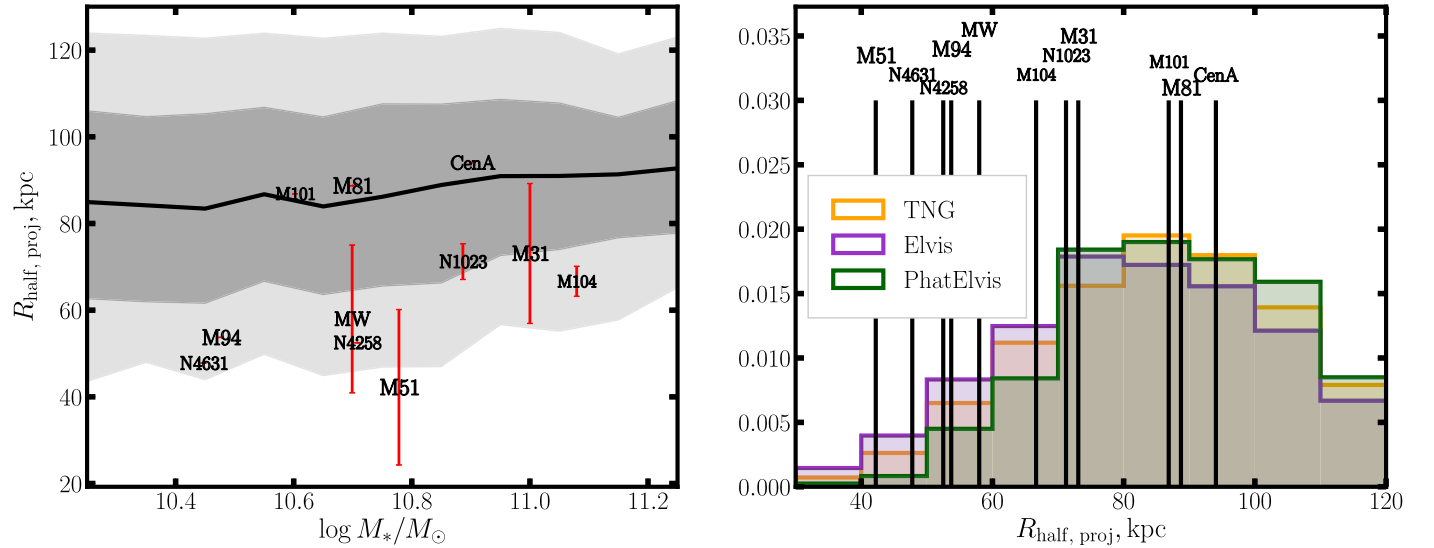


Figure 4. The radius encompassing half of the satellites for 11 observed hosts vs. the simulated hosts. Only satellites brighter than $M_V < -9$ and within a projected 150 kpc of their host are considered. On the left, the median satellite radius is plotted against the stellar mass of the host. The contours in the background show the $\pm 1\sigma$ and $\pm 2\sigma$ results for the IllustrisTNG hosts. The error bars for M31 and the MW show the effect of different observing angles while the error bars on the other hosts indicate the effect of uncertain membership for the subset of candidate satellites without distance measurements. On the right, the median satellite separations for the hosts are compared against the simulations in histogram form. The results from the ELVIS and PhatELVIS suites are also shown.

4.3. Satellite Concentration and Host Stellar Mass

To explore a different metric describing the shape of the radial profiles, in Figure 4 we show the radius that contains half of the satellites for the observed and simulated hosts. This R_{half} is calculated, essentially, as the median projected separation of satellites from the host. To compare all observed hosts together, we only consider satellites within a projected 150 kpc from their host. Only satellites brighter than $M_V < -9$ are

considered, and thus we do not include NGC 4565 in this assessment. On the left, the half-satellite radius is plotted against the stellar mass of the host. The observed hosts (points) are compared with the simulated IllustrisTNG hosts in the background. The IllustrisTNG hosts are selected by halo mass in the range $0.8 \times 10^{12} M_\odot < M_{200} < 8 \times 10^{12} M_\odot$. The average observed number of satellites within 150 kpc and with $M_V < -9$ (including the projection effects of the MW and M31

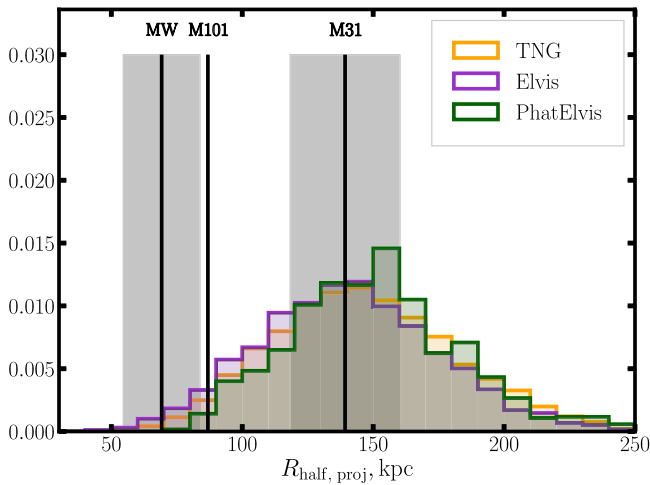


Figure 5. The same visualization as in the right panel of Figure 4 but for satellites within 300 projected kpc. Only the MW, M31, and M101 are shown because these are the only observed systems complete at these radii. The shaded bands show the $\pm 1\sigma$ spread in $R_{\text{half, proj}}$ due to assuming different viewing angles for the MW and M31.

and the effect of uncertain membership) is 6 per MW-like host and 14 per small-group host. Thus, we select the six most massive subhalos for each TNG MW-like host ($0.8 \times 10^{12} M_{\odot} < M_{200} < 3 \times 10^{12} M_{\odot}$) and the 14 most massive for each small-group host ($3 \times 10^{12} M_{\odot} < M_{200} < 8 \times 10^{12} M_{\odot}$) to compare with the observed satellites. For the IllustrisTNG hosts, we use the stellar mass of the host reported by the hydrodynamic simulation results. The IllustrisTNG hosts are all mock-observed at $D = 7$ Mpc (roughly the average distance of the observed hosts). For each simulated host, one of the observed hosts is chosen at random and that host’s survey area selection function is applied to the simulated host. On the right, the half-satellite radii of the observed hosts are compared in histogram form against the IllustrisTNG, ELVIS, and PhatELVIS simulated hosts. For the ELVIS and PhatELVIS hosts, the six most massive subhalos in the survey footprints are selected, and 100 viewing angles are taken for each host.

From Figure 4, we see that for satellites within a projected 150 kpc of their host, the median satellite separation is ~ 90 kpc for the simulations but is closer to ~ 50 – 70 kpc for the observed satellite systems. The observed hosts are systematically more centrally concentrated than the simulated hosts at the 1σ – 2σ level. The higher-mass observed hosts have their satellites at somewhat larger radii, in agreement with Figure 1. The three different simulation suites that we consider all show similar spatial distributions of satellites. The ELVIS hosts are slightly more centrally concentrated than the IllustrisTNG or PhatELVIS hosts, but that is easily understood due to the lack of any central disk. A few of the most massive hosts (M81, CenA, and M31) have spatial distributions characteristic of the simulated hosts, but the majority of the observed hosts are more concentrated, and, most importantly, *none* are less concentrated.

In Figure 5, we show the distribution of $R_{\text{half, proj}}$ for the simulations and observed systems when including satellites to 300 projected kpc (\sim the virial radius). Only the MW, M31, and M101 are included because no other observed system is complete to these radii. We include satellites with $M_V < -8.5$ mag, which is the estimated completeness for M101. The average number of satellites among these three hosts is 12, and

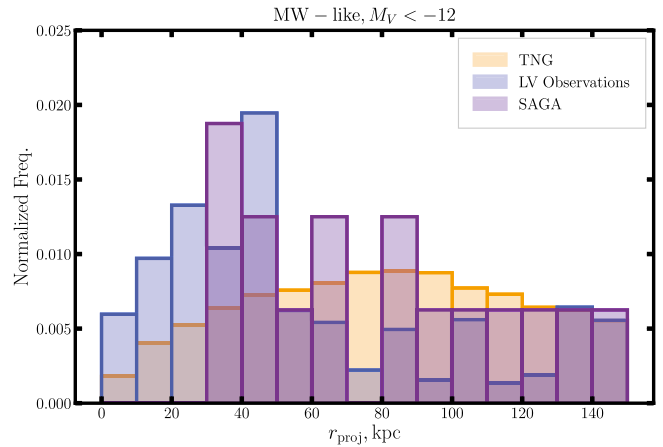


Figure 6. The distribution of projected separations for the LV observed satellite systems compared with both the simulated hosts of the IllustrisTNG project and the SAGA Survey. The normalization of the histograms is arbitrary. The IllustrisTNG hosts are processed through the observational selection functions of the LV hosts using the survey area footprints of those hosts.

so, for this comparison, we draw the 12 most massive subhalos from the simulations for each host. When including the full system of satellites, the discrepancy between the MW’s and M101’s systems and the simulated systems becomes much more significant than in Figure 4. This clearly demonstrates the importance of satellite searches that cover the *full* virial extent of the hosts.

As a check, we evaluate whether our conclusions change if, instead of comparing at a fixed radius, we normalize the satellite separations by the individual virial radii of the hosts. In Figures 1–4, we have considered the satellite separations in physical distances. This is justified by the fact that the hosts we consider all have similar expected virial radii, particularly so when we split the hosts into the MW-like and small-group categories. Still, it is useful to investigate whether the increased satellite concentration relative to the simulated hosts is due to some observed hosts having small virial radii (and hence naturally more compact satellite configurations). We found that the results are qualitatively the same when scaling by the host virial radius.

4.4. Comparison to the SAGA Survey

In this section, we compare our results to the radial profiles inferred from the SAGA Survey (Geha et al. 2017). SAGA focuses on MW analogs over the distance range 20–40 Mpc. As such, SAGA is only sensitive to the brightest satellites ($M_V \lesssim -12$); in the end, however, it will have many more observed hosts than what is possible in the LV. Geha et al. (2017) presented results for the first eight completed hosts. While the statistical sample of satellites is larger using our LV systems (particularly because our LV hosts are complete to much fainter satellites), it is interesting to compare the SAGA radial profiles with those inferred in this work. Figure 6 shows the distribution of projected separations for the MW-like LV hosts for bright ($M_V < -12$) satellites compared both to the confirmed SAGA satellites and to the profiles predicted from the IllustrisTNG hosts; the latter are selected by halo mass in the range $0.8 \times 10^{12} M_{\odot} < M_{200} < 3 \times 10^{12} M_{\odot}$. To maintain our focus on the shape of the distributions, all histograms are normalized. As before, the IllustrisTNG hosts are passed through the observational selection functions of the LV hosts.

Interestingly, the radial distribution of the SAGA satellites is intermediate to the LV host satellites and the IllustrisTNG results. The SAGA surveys does show the same large excess of satellites at $r_{\text{proj}} \sim 40$ kpc; however, there are not many satellites within $R_{\text{proj}} < 30$ kpc in the current SAGA sample, possibly due to the targeting strategy based on the Sloan Digital Sky Survey and the distance of these hosts. It is possible that these very close-in satellites are confused with the outskirts of the hosts, and at the large distance of these hosts, it is hard to distinguish them. We note that the statistical sample of the SAGA satellites is still small and a firm conclusion cannot be drawn; in particular, the highest peak at $r_{\text{proj}} \sim 40$ kpc corresponds to only three SAGA satellites. Future SAGA results will provide confirmation of this trend.

4.5. Comparison to Recent Work

Recently, Samuel et al. (2020) extensively compared the satellite radial profiles of the MW and M31 to simulated hosts in the FIRE project. Samuel et al. (2020) argued that the simulated hosts had radial profiles that were consistent with those of the MW and M31. Although we draw different conclusions here, we would argue that the results are fairly similar. First, in our work, we have a statistical sample of hosts. If we only compared the MW and M31 to a handful of simulated hosts, the disagreement would be less egregious because M31’s radial distribution is consistent with the simulations. Second, we identify the disagreement *only* when we normalize the radial profiles specifically to focus on their shape. In contrast, Samuel et al. (2020) primarily consider the absolute radial profile, and the spread of the absolute radial profiles from the models does, in fact, encompass the MW observations. Particularly, this occurs because absolute radial profiles confuse the richness of a satellite system with its spatial distribution. Precisely because we do not use an SHMR to populate subhalos (instead, we take the same number of massive subhalos as the observed number of satellites), the richness of the simulated systems in our comparison exactly matches that of the observed hosts, by definition. But in the FIRE-2 simulations, the number of satellites per simulated MW-like host differs by a factor of a few between hosts. This spread is able to encompass the centrally concentrated profile of the MW; we demonstrate this explicitly in Appendix B.

Samuel et al. (2020) explored the shape of the radial profile of MW and M31 satellites with the ratio of the radius containing 90% of satellites to that containing 50%. With this metric, they found that the MW was more concentrated than all of their baryonic simulations, consistent with what we find. In this paper, we opt to use the simpler R_{half} as a measure of the concentration of a profile. Because we have systems with only a few satellites, it becomes difficult to define the radius that contains 90% of satellites.

Our census of MW satellites is different from that of Samuel et al. (2020). We include Sgr while they do not. Samuel et al. (2020) do include Crater II and Antlia II while we do not include these satellites due to their extremely low surface brightness levels that fall below our completeness estimate for the LV host sample that we adopt (Paper I; Paper II). These small differences have the effect of somewhat boosting the central concentration of MW satellites in our results. Were we to adopt the sample of Samuel et al. (2020), however, our conclusions about the relative concentration of the ensemble of

hosts would not change significantly because this change only applies to the MW.

Finally, Samuel et al. (2020) found that if the simulated profiles were normalized to that of the MW at $r = 150$ kpc, all of the simulated hosts had more satellites at $r > 150$ kpc than the MW. They use this to argue that there could be classical-sized satellites yet undiscovered in the periphery of the MW virial volume. We discuss the possibility of this below in Section 5.1.

4.6. Summary

In this section, we perform comparisons between the radial distributions of observed satellites in the LV and analogous simulated systems in a set of modern cosmological simulations. While some of the observed hosts (particularly the more massive “small-group” hosts) seem to have similar radial profiles to the simulated hosts, on average the observed hosts are more concentrated. In Figures 1 and 3, we demonstrated this discrepancy by directly comparing the normalized radial distributions of satellites in the observed and simulated hosts. Normalizing the profiles is essential to see the difference between observations and simulations. In Figure 2, we compared the distribution of satellite projected separation, r_{proj} , between the observed hosts and the simulations. For the MW-like hosts, the distribution of observed satellites is significantly more centrally concentrated than for the simulated hosts. This is particularly striking (i) compared to the PhatELVIS simulations and (ii) when only the bright ($M_V < -12$) satellites are considered, although we note that the discrepancy is clear for all simulation suites and considering all satellite luminosities. Next, we used the half-satellite radius (the radius enclosing half of the satellites), R_{half} , to quantify the concentration in the radial profiles. Figure 4 shows that the population of observed hosts is shifted to smaller values of R_{half} than the simulated hosts. While three of the observed hosts show the average value for the simulations, the rest of them are more concentrated, and none are less so.

5. Discussion

In this section, we explore possible causes of this discrepancy in the Λ CDM paradigm. We leave an exploration of non-vanilla cosmological explanations to future work. We start by discussing the observations then move to the simulations.

5.1. Observational Incompleteness

The most obvious possible problem with the observations is that of incompleteness. Here, we describe some of the possible concerns and argue that observational completeness is not the cause of the discrepancy.

The most significant worry is that the LV satellite surveys are less complete at large radii than they are closer to the host. This could be due, for instance, to the dithering pattern of the observations of the central host. Indeed several of the hosts of Paper I had dithering patterns in the CFHT data used there that led to somewhat deeper exposures near the host than at larger R_{proj} . However, there are three lines of reasoning indicating that this is not significantly affecting the observed radial profiles.

First, Figure 2 shows that the discrepancy does not get better (in fact it gets worse) when we consider only the brightest satellites, which are more likely to be complete.

Second, the completeness estimates from Paper I (see their Figure 3) indicate that the satellite catalogs drop from $\sim 100\%$ completeness to $\sim 0\%$ over roughly half a magnitude for both surface brightness and total magnitude of the satellite. These estimates average over the entire survey footprint, and if there was a significant difference in depth between different areas of the survey, then we would expect the completeness to drop off more gradually. From Paper I, the completeness limit only changed by ~ 1 mag (but always $M_V \gtrsim -9$) between hosts that had exposure times differing by up to a factor of two, also suggestive that dithering would not lead to a significantly different completeness in the inner and outer regions.

Finally, we have used DECaLS (Dey et al. 2019) to search for satellites in several of the systems included in the present work. The details of these searches will be presented elsewhere (see Appendix A for a description of our search of NGC 4631). DECaLS is an extremely wide-field imaging survey with uniform depth, and we can easily detect satellites down to $\sim \mu_{0,V} = 26$ mag arcsec $^{-2}$, roughly the limit of the current observational sample. With DECaLS we recover the satellites discovered in the CFHT data of Paper I, but do not find any extra satellites further out from the hosts, indicating that nonuniform depth in the CFHT data is not causing a radial bias in the satellite census.

Due to our position within the MW, the MW’s satellite census is particularly vulnerable to incompleteness at large radii. Previous authors (e.g., Yniguez et al. 2014; Samuel et al. 2020) have suggested that the concentration of MW satellites is indicative of incompleteness in the satellite census at large radii. Such missing classical-mass satellites, however, would all have to be in the zone of avoidance around the MW disk to have evaded discovery thus far. The rest of the sky has been searched at a depth that would easily discover classical satellites throughout the virial volume (e.g., Whiting et al. 2007; Koposov et al. 2008; Tollerud et al. 2008; Walsh et al. 2009; Drlica-Wagner et al. 2019). Obscuration by the disk should not preferentially hide distant satellites, and so it is unclear if this would actually decrease the concentration of the MW satellites. Not finding disk-obscured satellites, however, would increase the scatter of the simulated profiles, perhaps making the MW less of an outlier. Testing this, we find that Figure 3 hardly changes if we select only subhalos that are $>15^\circ$ above or below a randomly oriented disk in the simulated hosts (as viewed from the center of the host). Even excluding potentially obscured satellites, the MW remains a significant outlier from the simulations.

Another possible issue with the observations is the confirmation of satellites with SBF distances. While our SBF methodology and calibration have been verified with independent HST distances (Bennet et al. 2019; Carlsten et al. 2019a), the distances are much less secure than HST TRGB distances. It is possible that a few unrelated background interlopers are included in the confirmed satellites of Paper II. However, including background galaxies will make the radial profiles *less* centrally concentrated, not more. Additionally, we account for the possibility of dwarfs that are nearby ($\lesssim 1$ Mpc) to the host but actually outside the virial radius by using a generous cut along the line of sight when selecting subhalos in the simulations. SBF distances are not precise enough to distinguish these objects from bona-fide satellites within the virial radius of the host. From our experience with the simulations, we do not expect these to be a significant source

of contamination ($\lesssim 10\%$ of the total satellite sample) (see also Paper II).

When surveying nearby disk systems, especially face-on disk galaxies, it is possible that the surveys miss very close-in satellites because they project onto the face of the host galaxy. However, including these satellites would make the observed profiles *even more* centrally concentrated, not less. In other words, the concentration of the observed systems is a lower bound taking this into account.

5.2. Reionization

Many previous works have argued that the *luminous* subhalos are more centrally concentrated than the DM subhalo population as a whole. The physical basis of this argument is that reionization suppresses star formation in many halos, and the halos that form earliest have the highest chance of being luminous (we note that there are other criteria one could use to decide which subhalos are luminous, e.g., Hargis et al. 2014). These subhalos are more concentrated around the host. As mentioned in the Introduction, classical-sized satellites, which are the focus of this paper, should be above the scale at which reionization can keep a subhalo dark, and this is unlikely to affect the radial distribution of classical satellites.

Font et al. (2011) and Starkenburg et al. (2013) presented semi-analytic models of galaxy formation coupled with DMO simulations and found that the radial distribution of model satellites was concentrated similarly to those of the MW. However, Font et al. (2011) considered the radial distribution of UFDs along with the classical satellites. The spatial distribution of UFDs might certainly be biased by reionization, and that is causing the model satellite distribution to agree well with observations. The semi-analytic models of Starkenburg et al. (2013) implemented the effect of reionization using the filtering mass approach of Gnedin (2000), which is now known to overpredict the suppressing effect of reionization (e.g., Okamoto et al. 2008). The good agreement that Starkenburg et al. (2013) find between their model classical satellites and the observed classical satellites is possibly a result of the stronger suppressing effect of reionization in their implementation, causing a radial bias in the locations of luminous satellites even at the classical mass scale. Modern hydrodynamic simulations seem to support the idea that the suppressing effect of reionization should not be important at the mass scales of the classical satellites (Ocvirk et al. 2016; Sawala et al. 2016; Wheeler et al. 2019). The SHMR of Garrison-Kimmel et al. (2017a) indicates that satellites of the luminosity we consider ($M_V \sim -10$) are hosted by halos of mass $\gtrsim 5 \times 10^9 M_\odot$, which are above the effects of reionization.

Furthermore, the hydrodynamic simulations, which will include the effects of reionization, of the NIHAO project show similar radial distributions to the other DMO simulation suites we compare with, indicating that reionization is not particularly important at the satellite mass scale we deal with here.

5.3. Effect of Host Assembly History

In this section, we explore whether different properties of the simulated hosts have an effect on their radial distribution of subhalos. It is possible that the relatively small sample of observed hosts are biased in some way (e.g., are located in early forming halos). There is significant literature discussing the impact of having a subhalo as massive as that hosting the

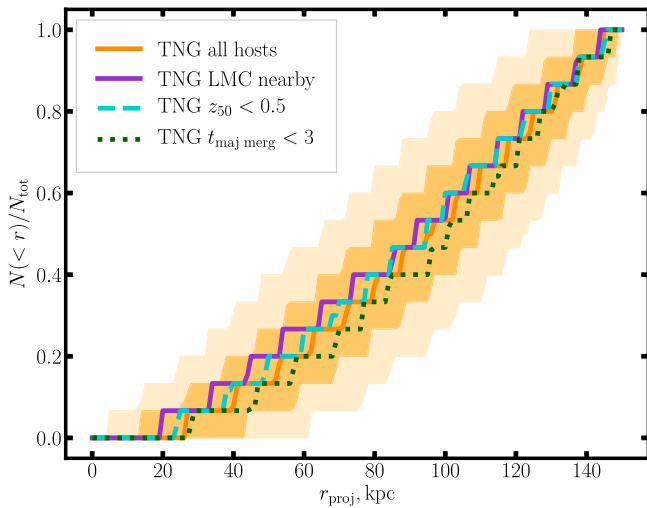


Figure 7. The radial distribution of the 15 most massive subhalos in IllustrisTNG hosts within 150 kpc. Orange shows the overall sample of “MW-like” hosts shown in the previous figures. Purple shows the radial distribution for hosts with a Magellanic Cloud-like satellite. Dashed turquoise shows the distribution for hosts that reached half of their current mass after $z = 0.5$. Dotted green shows the median profile for hosts that experienced a major merger in the last 3 Gyr.

LMC (e.g., Lu et al. 2016; Nadler et al. 2020) on the overall population of subhalos. So for one check, we select hosts from IllustrisTNG that have a Magellanic Cloud (MC)-like satellite. To do this, we select the subsample of “MW-like” halos that have a massive subhalo with $v_{\max} > 55 \text{ km s}^{-1}$ (following Lu et al. 2016) within the distance range $30 \text{ kpc} < D < 60 \text{ kpc}$ from the host. This corresponds to roughly 10% of all halos in the mass range $0.8 \times 10^{12} M_{\odot} < M_{200} < 3 \times 10^{12} M_{\odot}$. Figure 7 shows the radial profiles of these hosts compared to the overall sample of MW-like halos. The profiles are more centrally concentrated for hosts with an MC-like satellite but not nearly at the level of concentration required to explain the MW and other LV hosts. In Paper II, we found that the LV hosts did show a preponderance of bright (MC-like) satellites compared to the simulated systems using an SHMR to populate subhalos. It is unclear whether this is related to the overly concentrated satellite systems we find in this work. We also show the radial profiles for hosts that reached half of their current mass after $z = 0.5$ and hosts that experienced a major merger (mass ratio 3:1 or less) in the last 3 Gyr. These criteria were chosen to be about as selective as the presence of the LMC (roughly 10% of all halos). Neither of these selection criteria leads to a large difference in radial distribution of subhalos.

We emphasize again here that the current set of observed hosts is $\sim 80\%$ volume complete for massive ($M_{\star} \gtrsim 0.5 M_{\star}^{\text{MW}}$) hosts with $|b| > 15^{\circ}$ within $\sim 8.5 \text{ Mpc}$ so we believe it is unlikely that a bias due to a specific host formation history (or any other kind of bias) could be causing the discrepancy. The entire $D \lesssim 8.5 \text{ Mpc}$ volume would seemingly have to be an outlier, which seems unlikely. However, see Neuzil et al. (2020) for possible evidence that this volume is indeed a $\sim 2\sigma$ – 3σ outlier compared to simulations with its relatively high density of bright ($M_B < -20.5$) galaxies. With this said, 12 hosts is still a relatively small sample that does not fully span the possible range of environments and host properties, and an increased observational sample is urgently needed.

5.4. Incompleteness in the Simulations

The most likely possible problems with the simulations stem from resolution effects. On the surface, the fact that we get similar answers across the different simulation suites (once we recall that only some of the suites include a central disk) would seem to indicate that resolution is not playing a major role. The suites differ by roughly a factor of 200 in resolution between IllustrisTNG and PhatELVIS. However, it is possible that even if the simulation results are converged with respect to resolution, they might not be converged to the correct answer. We focus on two specific possible resolution-related problems in the application of the simulations.

The first possible problem in the application of the simulations is that we do not allow for the possibility of “orphan galaxies.” An orphan galaxy represents a luminous satellite when its DM halo becomes stripped to the point that it falls below the threshold of the subhalo finder. The luminous galaxy needs to be manually put back into the simulation and tracked because it has no corresponding DM subhalo. Previous works (e.g., Gao et al. 2004; Newton et al. 2018; Bose et al. 2020) have found that reproducing the radial distribution of luminous satellites in clusters and around the MW required inserting and tracking orphan galaxies. Different prescriptions are used to track the evolution of the orphan galaxies, but generally they are removed after a certain time to represent their merging into the central primary.

It is unclear whether inserting orphan galaxies is an appropriate thing to do for the comparisons we do in this work. Orphan galaxies are mainly important for subhalos near the resolution limit of the simulation, which is not the case for this work, where we are focused on the fairly massive classical-sized satellites. A typical classical satellite hosting a subhalo with mass $\sim 5 \times 10^9 M_{\odot}$ will be resolved with $> 10^5$ particles in the PhatELVIS simulations. For this subhalo to drop below the threshold of a subhalo finder (~ 20 particles), it has to be $> 99.9\%$ stripped. By this point a *significant* fraction of the stars would also be stripped (e.g., Peñarrubia et al. 2008; Errani et al. 2018) and the satellite would likely be either undetectable (due to very low surface brightness) or clearly tidally disturbed. Some of the observed satellites that we include are clearly undergoing tidal disruption (e.g., Sgr, NGC 5195, NGC 4627, M32, NGC 205, dw1240p3237) but the majority are not. Not including these satellites will clearly reduce the observed central concentration but not enough to make the discrepancy go away.

On the other hand, it is not clear that these disrupting systems would not be captured by the halo finders in the simulations. Sgr is estimated to still have a significant DM halo from the velocity dispersion of stars (Law & Majewski 2010) that would be easily resolved in the simulation suites we use. However, perhaps there is not enough contrast between its DM subhalo and the parent halo for the subhalo to be identified as a distinct structure by a halo finder. NGC 205 similarly is estimated to have a significant component of dark mass (Geha et al. 2006). Therefore, it does not seem legitimate to include a large population of orphan galaxies because these heavily stripped galaxies do not correspond to the population of observed satellites. More work is required to fully understand the correspondence between observations and simulations for disrupting satellites at various stages of disruption to determine when such systems will no longer be observable or detected by

halo finders in the simulations. With this said, we explore a model with a population of orphan galaxies in the next section.

The second, related possible problem with the simulations is the possibility of a significant fraction of the tidal disruption of subhalos in the simulations being artificial. van den Bosch et al. (2018) and van den Bosch & Ogiya (2018) argue that even for cosmological simulations that are “converged” and resolve subhalos with >100 particles, most of the tidal disruption of these subhalos is artificial. They cite discreteness noise and inadequate force softening as the main culprits. van den Bosch & Ogiya (2018) provide new criteria to judge whether the tidal evolution of a subhalo is trustworthy or whether it is affected by discreteness noise. A subhalo on a circular orbit with a radius of 0.2 times the host virial radius needs to be resolved with $>10^5$ particles. For the high-resolution PhatELVIS simulations, the subhalos that host bright classical-sized satellites would be resolved at roughly this level. However, the criteria of van den Bosch & Ogiya (2018) are for circular orbits in a Navarro–Frenk–White tidal field, and it is unclear how to extend this to noncircular orbits in the presence of a disk’s tidal field. Artificial disruption of subhalos could easily explain the discrepancy we find if many of the simulated subhalos near the host galaxy were artificially destroyed.

If the halo catalogs are missing objects, either orphans or artificially disrupted halos, then the total number of predicted halos would necessarily rise, leading to the need for a new SHMR. In Paper II, we found that the observed LV satellite systems have total abundances well matched by the simulations combined with standard abundance matching relations found in the literature—in particular, that of Garrison-Kimmel et al. (2017a). The SHMR of Garrison-Kimmel et al. (2017a) shows fair agreement with modern hydrodynamic simulations. However, if a significant number of artificially lost subhalos need to be added into the simulations, a much steeper slope or lower normalization of the SHMR will be required to not overproduce satellites in the simulations. We show this qualitatively with a toy model for orphan galaxies in the next section.

5.5. Toy Model for Orphan Galaxies

While it is unclear how to resolve the discrepancy found in this paper, as discussed in the last section, a possible option is that the simulations are missing subhalos, from either artificial overmerging or incompleteness in the halo finder perhaps due to resolution effects. If the simulation halo catalogs are missing a large population of close-in subhalos, putting these back in could help resolve the discrepancy. However, adding more subhalos would affect what SHMR is allowed by the observed abundances of satellites. In Paper II, we showed that the overall richness of the LV satellite systems could be well matched by the SHMR of Garrison-Kimmel et al. (2017a). However, the model predicted too few bright satellites and too many faint ones. This SHMR agrees well with the predictions from recent hydrodynamic simulations of dwarf galaxy formation. If a large population of subhalos is added into the simulations, this SHMR likely will no longer match the observations. In this section, we develop a toy model that inserts subhalos into the halo catalogs to match the observed radial distribution, and we investigate the effects on the SHMR allowed by observations.

For this test, we use the TNG simulations. We only consider the eight MW-sized observed hosts in this comparison and compare with TNG hosts selected by halo mass in the range $0.8 \times 10^{12} M_{\odot} < M_{200} < 3 \times 10^{12} M_{\odot}$. While other

treatments of orphan galaxies insert them and trace their movement in the host halo in a physically motivated way, we insert the orphan galaxies directly into the halo catalogs. We insert the orphans with a projected separation from the host drawn from a Gaussian with mean 35 kpc and standard deviation 15 kpc. These numbers are chosen to match the peak of close-in observed bright satellites visible in Figure 2. This is clearly unphysical and is simply a toy model that we are using to answer the question: “Would including the number of orphan galaxies required to bring the observed and simulated radial distributions into agreement affect the SHMR?” We leave an investigation of where the orphan galaxies would actually end up in the simulations to future work. Subhalo masses are drawn from a distribution in the range $10^{8.5} < M/M_{\odot} < 10^{11.8}$ using the well-known subhalo mass function $dN/dM \propto M^{-1.9}$ (e.g., Springel et al. 2008). Using this distribution assumes that subhalos are artificially lost equally at all masses. This is likely a lower bound because lower-mass subhalos are most likely lost preferentially if this is due to resolution. On the other hand, the more massive satellites will experience more dynamical friction, and may naturally be more concentrated than the whole satellite population. We manually adjust the number of subhalos that we add in to match the distribution of bright satellites (top right corner of Figure 2). We find that adding ~ 100 subhalos in this mass range within the inner 150 projected kpc leads to fair agreement with the observations. Figure 8 shows this result on the left. The observed radial distribution of bright ($M_V < -12$) satellites is shown along with the distribution of the three most massive subhalos before and after subhalo injection. Recall that there are, on average, three satellites in this luminosity range per observed host, and so we select the three most massive subhalos as the simulated satellites. Note that we do not make any survey area corrections for the simulations in this test; we assume the observed systems are all complete to a projected 150 kpc.

On the right of Figure 8, we show the resulting predicted abundance of $M_V < -9$ satellites within $r_{\text{proj}} < 150$ kpc using the halo catalogs with and without injected subhalos and the SHMR of Garrison-Kimmel et al. (2017a). Without any added subhalos, we see the result found by Paper II, i.e., that the SHMR of Garrison-Kimmel et al. (2017a) produces roughly the right number of satellites (including host-to-host scatter). However, with the added subhalos, the SHMR significantly overproduces satellites. A significantly steeper slope or lower normalization for the SHMR would be required. It is possible that the required SHMR would no longer be in agreement with the results of modern hydrodynamic simulations of dwarf galaxy formation, but we leave a detailed investigation of this to future work.

6. Conclusions

In this paper, we compare the observed satellite spatial distributions in 12 LV massive hosts to that predicted in various state-of-the-art cosmological simulations in the Λ CDM paradigm. While many previous works compared the radial distribution of satellites around the MW and/or M31 to simulated analogs, we are able to achieve much better statistics by considering the satellite systems of many MW-like galaxies in the LV. This has only recently been made possible by observational work characterizing the satellite systems of nearby MW analogs. We compare the observations to multiple

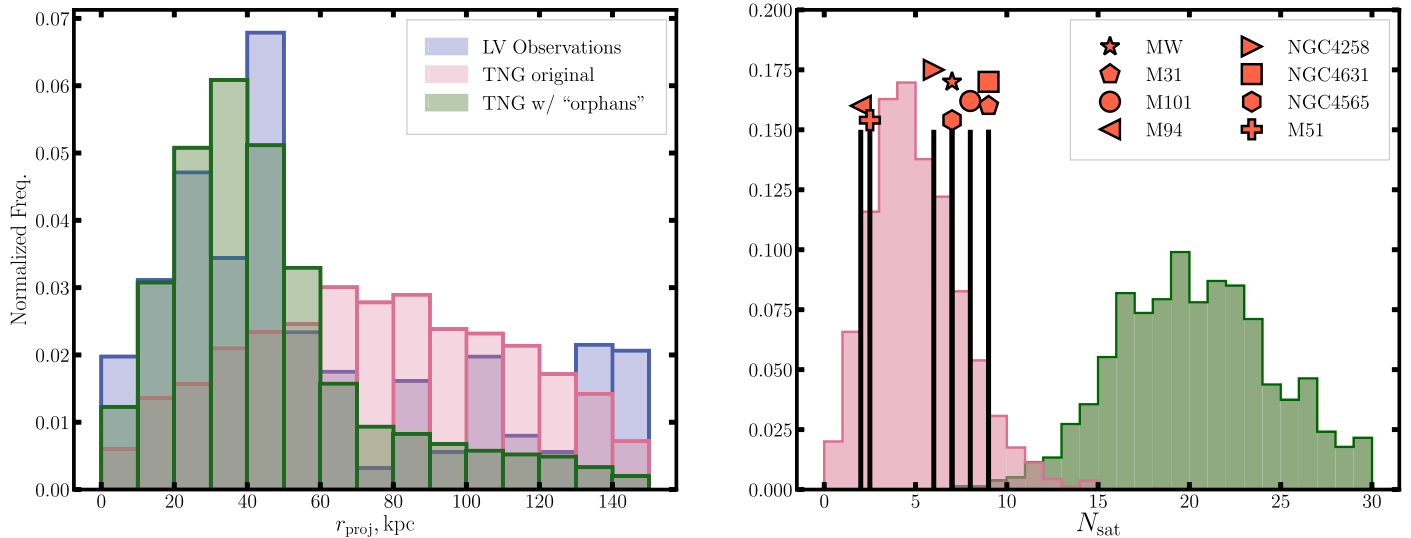


Figure 8. One possibility to resolve the discrepancy in radial distributions between simulated and observed hosts is that the simulations are missing many subhalos, due to either artificial disruption or incompleteness in the halo finder. This figure shows the effect of adding these subhalos back into the simulation. Left: the radial distribution of the IllustrisTNG simulations after these “orphan” galaxies are added back in. See text for details about how they are added. Right: the predicted abundance of $M_V < -9$ satellites within $r_{\text{proj}} < 150$ kpc using the SHMR of Garrison-Kimmel et al. (2017a) with and without the added subhalos. As expected, injecting subhalos would require a significantly steeper slope or lower normalization for the SHMR to avoid overproducing satellites.

different simulation suites. These include a big-box cosmological simulation (IllustrisTNG-100) that gives great statistics with >1000 MW-like hosts but at relatively low resolution, high-resolution DMO zoom-in simulations of several tens of MW-sized hosts both including the potential of a disk (PhatELVIS) and not (ELVIS), and a high-resolution fully hydrodynamic zoom-in simulation of six MW-like hosts (NIHAO). Overall we find fairly good agreement among the simulations. Our main findings are as follows.

1. We confirm previous findings that the classical satellites of the MW are significantly more concentrated at a $>2\sigma$ level than the massive subhalos of simulated analogs (Figure 3). We argue that this discrepancy is likely not resolved by either reionization or incompleteness in the census of MW classical-sized satellites.
2. We find that several of the other observed hosts in the LV have more concentrated radial profiles than the analogous simulated hosts at the $\sim 2\sigma$ level (Figure 1). A few of the observed hosts, particularly the more massive (“small-group”) hosts such as M81 and CenA, have similar radial profiles to the simulated hosts, but *none* of the observed hosts are less concentrated than the average simulated profile.
3. We use the median satellite separation for satellites within a projected 150 kpc as a metric for the concentration of the satellite spatial distribution. We find that the population of observed systems is systemically shifted to smaller radii compared to the simulated systems (Figure 4).
4. Figure 2 shows that the spatial distribution of the satellites is significantly different between the observed and simulated satellites. There is a significant population of close-in observed satellites that is missing in the simulations. This is particularly noticeable for the bright satellites ($M_V < -12$), where there are a large number of observed satellites at projected separations of 30–60 kpc but very few simulated satellites at these separations.

5. The spatial distribution of satellites found in the SAGA Survey (Geha et al. 2017) agrees with the distribution of observed LV satellites, although the statistics are too low to be conclusive (Figure 6).

Throughout this work we have been careful to consider the incompleteness and limitations of the observations. We have used the specific survey footprints for each observed host, where possible, to forward-model the simulated hosts.

In Section 5, we discussed many possible causes for the discrepancy both on the observational side and on the simulation side. Observational incompleteness cannot easily explain the discrepancy. The completeness has been well quantified by the various works that have characterized the LV satellite systems. We do note, however, that many of the discrepant observed systems (e.g., NGC 4631 and NGC 4258) come from the work of Paper I. Excluding these systems makes the discrepancy less severe. With that said, the MW is quite discrepant with the simulations. Since the MW is arguably the observed host with the best completeness, this is suggestive that there is more to this than simple observational incompleteness. On the other hand, the observational sample only contains 12 hosts, which makes it still a relatively small sample, and it might not fully span the relevant range of environments and formation histories that the cosmological simulations span.

We also discuss possible challenges in comparing with simulations, but no explanation is satisfactory. We discussed whether reionization or biased formation histories of the observed host halos could explain the discrepancy. Reionization has the effect of increasing the central concentration of luminous satellites but is unlikely to be effective at the satellite luminosities considered here. Different halo formation histories (e.g., late major merger, presence of an LMC-like companion at $z = 0$, etc.) do have a minor effect on the radial distribution of satellites but not enough to explain the observations. This is an important check because our hosts do have an unusually high prevalence of massive satellites (Paper II).

Finally we discussed possible issues in the simulations. In particular, we do not include the possibility of orphan galaxies

in the simulations. These are galaxies whose DM subhalo has dropped below the detection threshold of the subhalo finder and they need to be tracked “manually.” We argue that, at the subhalo masses of the satellites we consider, when the subhalo is stripped to the point where it is not detected by a subhalo finder, the luminous galaxy would be mostly destroyed, and hence it does not correspond to the observed satellites. We also discuss the possibility of significant artificial disruption in the simulations. This appears to be the most feasible cause of the discrepancy. If this is the cause, it will have important ramifications for the allowable SHMR in this mass range. Paper II found that the SHMR consistent with state-of-the-art hydrodynamic simulations from various projects (see, e.g., Garrison-Kimmel et al. 2017a, and Figure 8) reproduces fairly well the overall number of observed satellites in these systems. If there is a large population of subhalos that are getting artificially disrupted in the simulations (but in reality should still exist), then this SHMR will overproduce satellites and the relation will have to be steepened significantly or given a lower normalization.

The observed systems appear to disagree most with the simulated hosts in the PhatELVIS suite. Due to its high resolution and inclusion of a central disk, we expect the PhatELVIS suite to most realistically represent MW-like systems. This highlights the fact that disruption of DM substructure by a central disk is still not entirely understood. Our findings of too-concentrated satellite systems are similar to the results considering the UFDs of the MW, where it is difficult to reconcile the abundance of observed close-in UFDs with the dearth of close-in subhalos once disk disruption is accounted for (e.g., Kim et al. 2018; Graus et al. 2019; Nadler et al. 2020). We note that the model of Nadler et al. (2020) was unable to reproduce the radial distribution of UFDs around the MW even when including a population of orphan galaxies. Given the importance of tidal stripping and disruption in the baryonic resolution of both the “Missing Satellites” and “Too Big to Fail” problems of small-scale structure formation, it is crucial to fully understand the how the central disk affects the population of DM subhalos.

On the observational front, the way forward is still clearly to survey and characterize more satellite systems, and emphasis should be given to surveying a few systems out fully to the virial radius of the host. Much of the significance of the discrepancy of the MW’s radial distribution of classical satellites comes from the fact that it is surveyed out to the virial radius (300 kpc). Figure 4 shows that the inner ($r_{\text{proj}} < 150$ kpc) satellites of the MW are more concentrated than simulated analogs but only at a $\sim 1\sigma$ level. In Figure 5, we show that when considering the entire satellite system out to 300 kpc, the discrepancy between R_{half} of the MW and M101 and that of the simulated systems becomes much more significant at $\gtrsim 2\sigma$ – 3σ . A volume-limited sample of hosts in the LV would be a particularly powerful data set to compare with simulations.

Finally, we note that this radial distribution of satellites is not the only observed peculiarity of the spatial distribution of the MW satellites. It has long been known that the classical satellites are arranged in a thin plane (Kroupa et al. 2005), the likes of which are quite rare in cosmological simulations (e.g., Pawlowski et al. 2012; Pawlowski & Kroupa 2020). It is certainly possible that the unusual radial distribution of the MW satellites is related to their unusual planar configuration.

Other systems (e.g., NGC 4258, NGC 4631, and M101) having similarly concentrated radial profiles makes this possible connection all the more intriguing.

We thank J. Samuel and A. Wetzel for their willingness to share the FIRE-2 simulation results in an earlier version of this work. We also thank the anonymous referee for a supportive and helpful report.

Support for this work was provided by NASA through Hubble Fellowship grant #51386.01 awarded to R.L.B. by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. J.P.G. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-1801921. J.E.G. is partially supported by the National Science Foundation grant AST-1713828. S.G.C. acknowledges support by the National Science Foundation Graduate Research Fellowship Program under grant No. #DGE-1656466. A.H.G.P. is supported by National Science Foundation grant Nos. AST-1615838 and AST-1813628.

Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/IRFU, at the Canada–France–Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l’Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii.

Software: astropy (Astropy Collaboration et al. 2018).

Appendix A New Satellites of NGC 4631 and M101

We have extended the satellite census for NGC 4631 and M101 out to 200 and 300 kpc, respectively, using the extremely wide-field DECaLS imaging (Dey et al. 2019). Details of this satellite search will be given in a forthcoming paper (S. G. Carlsten et al. 2020, in preparation), but we give an overview here. We applied the LSB galaxy detection algorithm of Paper I to the g and r imaging of DECaLS to identify candidate satellites. We estimated the completeness of these satellite searches by injecting mock galaxies, following Paper I. We estimate the completeness as $M_V \sim -9$ mag and $\mu_{0,r} \sim 26$ mag arcsec $^{-2}$ throughout the search footprint. We then used deeper ground-based imaging to measure the surface brightness fluctuations of these galaxies and constrain their distances. The new satellites are listed in Table 1.

We identified two new candidate satellites around M101 in the 300 kpc radius footprint of our DECaLS search. One of these (dw1403p5338) is actually in the footprint of the search of Bennet et al. (2017) but was missed in that search. It was noted by van Dokkum et al. (2020) as a candidate satellite. We used the deep CFHT Legacy Survey imaging to measure its SBF distance, which is consistent with M101. We follow the criteria outlined in Paper II for “confirming” a candidate satellite. This satellite is below the $M_V \sim -9$ fiducial completeness limit that we use for most of this paper, and thus does not play a large role in the results. The other candidate satellite, dw1350p5441, was outside the CFHT Legacy Survey footprint but had archival HSC imaging that was deep enough to measure its SBF signal with high signal-to-noise ratio. It also is at the distance of M101.

We identified six new satellite candidates around NGC 4631 within 200 kpc but outside the search footprint of Paper I. We

Table 1
Newly Confirmed Satellites of NGC 4631 and M101

Host Name	Dwarf Name	R.A.	Decl.	M_V	SBF Distance (Mpc)	cz (km s ⁻¹)
M101	dw1403p5338	14:03:27.3	+53:37:52	-8.8	6.0 ± 0.8	...
M101	dw1350p5441	13:50:58.4	+54:41:21	-12.0	6.4 ± 0.8	...
NGC 4631	dw1248p3158	12:48:52.6	+31:58:13	-12.5	7.1 ± 0.8	...
NGC 4631	dw1236p3336	12:36:49.5	+33:36:47	-13.0	...	530

Note. M101 is at $D = 6.5$ Mpc (Beaton et al. 2019) and NGC 4631 is at $D = 7.4$ Mpc, with redshift $cz = 606$ km s⁻¹.

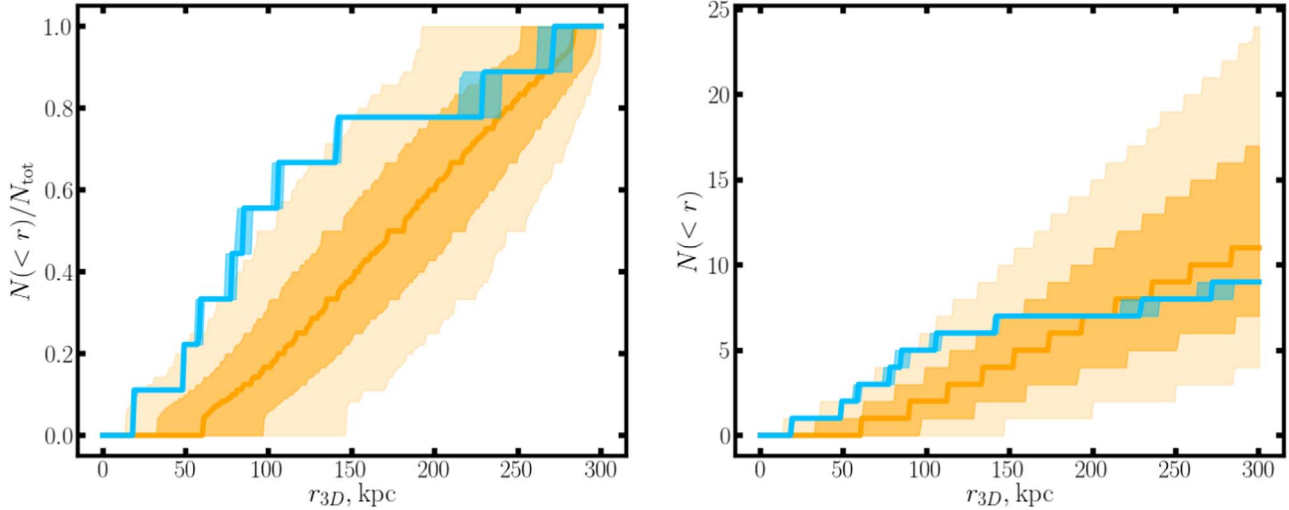


Figure 9. The radial distribution of MW satellites ($M_V < -9$) compared to the TNG simulations combined with the SHMR of Garrison-Kimmel et al. (2017a).

acquired deep ground-based images of four of these through the Gemini Fast-Turnaround Queue (Prop. ID: GN-2020A-FT-104; PI: S. Carlsten). The imaging was deep enough to constrain three of these to be background contaminants and confirm one (dw1248p3158) to be a real satellite. One of the remaining two candidates has a redshift compatible with being bound to NGC 4631, and we consider it a confirmed satellite. The remaining candidate was the faintest ($M_V \gtrsim -9.5$) and still is unconstrained. Thus, our completeness with full distance confirmation is roughly $M_V \sim -9.5$ over the inner 200 kpc for NGC 4631.

Appendix B

Normalized versus Absolute Radial Profiles

Throughout the paper, we consider the normalized radial profiles of satellites instead of the absolute profiles. Since we select the same number of satellites in the simulations as in the observed hosts, the satellite abundance is matched, by definition. However, if we were to use an SHMR from abundance matching to populate subhalos with luminous galaxies or to use hydrodynamic simulation results, the abundance of satellites would not necessarily be matched, and it would be important to normalize the radial profiles to make a detailed investigation of their shapes. This is part of the reason why our conclusions differ from that of Samuel et al. (2020), who considered largely the absolute (unnormalized) radial profiles. Absolute radial profiles conflate the problem of satellite abundance and spatial distribution, and the scatter in the simulated profiles is much larger and able to encompass (even very concentrated) observed profiles.

We show this in Figure 9 for the MW satellites ($M_V < -9$). We compare with the TNG simulations combined with the SHMR of Garrison-Kimmel et al. (2017a). We assume M/L_V of 1.2 and only select model satellites brighter than $M_V < -9$. We show both the normalized and absolute radial profiles. On the right, the simulated satellite abundance differs by a factor of a few between hosts, and this spread is able to encompass the MW’s centrally concentrated radial profile.

ORCID iDs

Scott G. Carlsten <https://orcid.org/0000-0002-5382-2898>
 Jenny E. Greene <https://orcid.org/0000-0002-5612-3427>
 Annika H. G. Peter <https://orcid.org/0000-0002-8040-6785>
 Johnny P. Greco <https://orcid.org/0000-0003-4970-2874>
 Rachael L. Beaton <https://orcid.org/0000-0002-1691-8217>

References

- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123
- Barber, C., Starkenburg, E., Navarro, J. F., McConnachie, A. W., & Fattahi, A. 2014, *MNRAS*, **437**, 959
- Beaton, R. L., Seibert, M., Hatt, D., et al. 2019, *ApJ*, **885**, 141
- Bennet, P., Sand, D. J., Crnojević, D., et al. 2019, *ApJ*, **885**, 153
- Bennet, P., Sand, D. J., Crnojević, D., et al. 2017, *ApJ*, **850**, 109
- Bose, S., Deason, A. J., Belokurov, V., & Frenk, C. S. 2020, *MNRAS*, **495**, 743
- Brooks, A. M., Kuhlen, M., Zolotov, A., & Hooper, D. 2013, *ApJ*, **765**, 22
- Brooks, A. M., & Zolotov, A. 2014, *ApJ*, **786**, 87
- Buck, T., Macciò, A. V., Dutton, A. A., Obreja, A., & Frings, J. 2019, *MNRAS*, **483**, 1314
- Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2000, *ApJ*, **539**, 517
- Carlsten, S. G., Beaton, R. L., Greco, J. P., & Greene, J. E. 2019a, *ApJL*, **878**, L16

- Carlsten, S. G., Beaton, R. L., Greco, J. P., & Greene, J. E. . 2019b, *ApJ*, **879**, 13
- Carlsten, S. G., Greco, J. P., Beaton, R. L., & Greene, J. E. 2020a, *ApJ*, **891**, 144
- Carlsten, S. G., Greene, J. E., Peter, A. H. G., Beaton, R. L., & Greco, J. P. 2020b, arXiv:2006.02443
- Chiboucas, K., Jacobs, B. A., Tully, R. B., & Karachentsev, I. D. 2013, *AJ*, **146**, 126
- Collins, M. L. M., Tollerud, E. J., Rich, R. M., et al. 2020, *MNRAS*, **491**, 3496
- Crnojević, D., Sand, D. J., Bennet, P., et al. 2019, *ApJ*, **872**, 80
- Danieli, S., van Dokkum, P., Merritt, A., et al. 2017, *ApJ*, **837**, 136
- Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *AJ*, **157**, 168
- D’Onghia, E., Springel, V., Hernquist, L., & Keres, D. 2010, *ApJ*, **709**, 1138
- Dooley, G. A., Peter, A. H. G., Yang, T., et al. 2017, *MNRAS*, **471**, 4894
- Drlica-Wagner, A., Bechtol, K., Mau, S., et al. 2019, arXiv:1912.03302
- Errani, R., Peñarrubia, J., & Walker, M. G. 2018, *MNRAS*, **481**, 5073
- Font, A. S., Benson, A. J., Bower, R. G., et al. 2011, *MNRAS*, **417**, 1260
- Fritz, T. K., Battaglia, G., Pawlowski, M. S., et al. 2018, *A&A*, **619**, A103
- Gao, L., De Lucia, G., White, S. D. M., & Jenkins, A. 2004, *MNRAS*, **352**, L1
- Garrison-Kimmel, S., Boylan-Kolchin, M., Bullock, J. S., & Lee, K. 2014, *MNRAS*, **438**, 2578
- Garrison-Kimmel, S., Bullock, J. S., Boylan-Kolchin, M., & Bardwell, E. 2017a, *MNRAS*, **464**, 3108
- Garrison-Kimmel, S., Wetzel, A., Bullock, J. S., et al. 2017b, *MNRAS*, **471**, 1709
- Garrison-Kimmel, S., Hopkins, P. F., Wetzel, A., et al. 2019, *MNRAS*, **487**, 1380
- Geha, M., Guhathakurta, P., Rich, R. M., & Cooper, M. C. 2006, *AJ*, **131**, 332
- Geha, M., Wechsler, R. H., Mao, Y.-Y., et al. 2017, *ApJ*, **847**, 4
- Gnedin, N. Y. 2000, *ApJ*, **542**, 535
- Graus, A. S., Bullock, J. S., Kelley, T., et al. 2019, *MNRAS*, **488**, 4585
- Hargis, J. R., Willman, B., & Peter, A. H. G. 2014, *ApJL*, **795**, L13
- Kelley, T., Bullock, J. S., Garrison-Kimmel, S., et al. 2019, *MNRAS*, **487**, 4409
- Kim, S. Y., Peter, A. H. G., & Hargis, J. R. 2018, *PhRvL*, **121**, 211302
- Koposov, S., Belokurov, V., Evans, N. W., et al. 2008, *ApJ*, **686**, 279
- Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004, *ApJ*, **609**, 482
- Kroupa, P., Theis, C., & Boily, C. M. 2005, *A&A*, **431**, 517
- Law, D. R., & Majewski, S. R. 2010, *ApJ*, **714**, 229
- Lu, Y., Benson, A., Mao, Y.-Y., et al. 2016, *ApJ*, **830**, 59
- Lux, H., Read, J. I., & Lake, G. 2010, *MNRAS*, **406**, 2312
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, *MNRAS*, **480**, 5113
- McConnachie, A. W. 2012, *AJ*, **144**, 4
- McConnachie, A. W., Ibata, R., Martin, N., et al. 2018, *ApJ*, **868**, 55
- Moore, B. 2001, in AIP Conf. Ser. 586, 20th Texas Symp. on Relativistic Astrophysics, ed. J. C. Wheeler (Melville, NY: AIP), **73**
- Müller, O., Rejkuba, M., Pawlowski, M. S., et al. 2019, *A&A*, **629**, 18
- Nadler, E. O., Wechsler, R. H., Bechtol, K., et al. 2020, *ApJ*, **893**, 48
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **477**, 1206
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **475**, 624
- Nelson, D., Springel, V., Pillepich, A., et al. 2019, *ComAC*, **6**, 2
- Neuzil, M. K., Mansfield, P., & Kravtsov, A. V. 2020, *MNRAS*, **494**, 2600
- Newton, O., Cautun, M., Jenkins, A., Frenk, C. S., & Helly, J. C. 2018, *MNRAS*, **479**, 2853
- Ocvirk, P., & Aubert, D. 2011, *MNRAS*, **417**, L93
- Ocvirk, P., Gillet, N., Shapiro, P. R., et al. 2016, *MNRAS*, **463**, 1462
- Okamoto, T., & Frenk, C. S. 2009, *MNRAS*, **399**, L174
- Okamoto, T., Gao, L., & Theuns, T. 2008, *MNRAS*, **390**, 920
- Pawlowski, M. S., & Kroupa, P. 2020, *MNRAS*, **491**, 3042
- Pawlowski, M. S., Pflamm-Altenburg, J., & Kroupa, P. 2012, *MNRAS*, **423**, 1109
- Peñarrubia, J., Navarro, J. F., & McConnachie, A. W. 2008, *ApJ*, **673**, 226
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, *MNRAS*, **475**, 648
- Richings, J., Frenk, C., Jenkins, A., et al. 2020, *MNRAS*, **492**, 5780
- Samuel, J., Wetzel, A., Tollerud, E., et al. 2020, *MNRAS*, **491**, 1471
- Sawala, T., Frenk, C. S., Fattahi, A., et al. 2016, *MNRAS*, **456**, 85
- Smercina, A., Bell, E. F., Price, P. A., et al. 2018, *ApJ*, **863**, 152
- Somerville, R. S. 2002, *ApJL*, **572**, L23
- Springel, V., Wang, J., Vogelsberger, M., et al. 2008, *MNRAS*, **391**, 1685
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, **475**, 676
- Starkenburg, E., Helmi, A., De Lucia, G., et al. 2013, *MNRAS*, **429**, 725
- Taylor, J. E., Silk, J., & Babul, A. 2004, in IAU Symp. 220, Dark Matter in Galaxies, ed. S. Ryder et al. (San Francisco, CA: ASP), **91**
- Tollerud, E. J., Bullock, J. S., Strigari, L. E., & Willman, B. 2008, *ApJ*, **688**, 277
- van den Bosch, F. C., & Ogiya, G. 2018, *MNRAS*, **475**, 4066
- van den Bosch, F. C., Ogiya, G., Hahn, O., & Burkert, A. 2018, *MNRAS*, **474**, 3043
- van Dokkum, P., Lokhorst, D., Danieli, S., et al. 2020, *PASP*, **132**, 074503
- Walsh, S. M., Willman, B., & Jerjen, H. 2009, *AJ*, **137**, 450
- Weisz, D. R., Dolphin, A. E., Martin, N. F., et al. 2019, *MNRAS*, **489**, 763
- Wheeler, C., Hopkins, P. F., Pace, A. B., et al. 2019, *MNRAS*, **490**, 4447
- Whiting, A. B., Hau, G. K. T., Irwin, M., & Verdugo, M. 2007, *AJ*, **133**, 715
- Willman, B., Governato, F., Dalcanton, J. J., Reed, D., & Quinn, T. 2004, *MNRAS*, **353**, 639
- Yniguez, B., Garrison-Kimmel, S., Boylan-Kolchin, M., & Bullock, J. S. 2014, *MNRAS*, **439**, 73
- Zolotov, A., Brooks, A. M., Willman, B., et al. 2012, *ApJ*, **761**, 71