Draft version March 8, 2020
Typeset using LaTeX twocolumn style in AASTeX62

# Radial Distributions of Dwarf Satellite Systems in the Local Volume

(Accepted March 8, 2020)

Submitted to ApJ

## ABSTRACT

The radial distribution of low mass satellites around a Milky Way (MW)-like host is an important benchmark for simulations of small-scale structure as the distribution is sensitive to disruption of subhalos by the central disk and can indicate whether disruption is artificial (ie. numeric) or physical in origin. Using a recent sample of 12 well-surveyed satellites systems around MW-like hosts in the Local Volume, we investigate the radial distribution of satellites and compare with the distributions predicted by modern ΛCDM cosmological simulations. The observed systems are generally complete to $M_V < -9$ and within 150 projected kpc. We consider multiple simulations, including big box cosmological simulations and high resolution zoom in simulations of a single MW sized halo. We focus on the concentration of the radial distributions and find that, overall, 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. We show this using several different metrics of the central 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 discuss possible causes for the discrepancy, including artificial disruption of subhalos due to resolution effects. However, we demonstrate that adding back in a population of artificially lost subhalos would have important ramifications for what stellar to halo mass relation is allowed by observations, requiring the relation to be steeper than generally predicted in hydrodynamic simulations.

*Keywords:* methods: observational – techniques: photometric – galaxies: distances and redshifts – galaxies: dwarf

## 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; 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; Garrison-Kimmel et al. 2017b; Kelley et al. 2019; 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 ΛCDM, and, hence, it is critical to understand them fully.

Previous comparisons between observations of the MW satellites and the predictions from ΛCDM simulations have produced somewhat baffling results. Comparisons with the classical satellites ($M_* \gtrsim 10^5$ M$_\odot$) have indicated that the MW satellites are significantly more radially concentrated than the most massive dark matter (DM) subhalos in dark matter only (DMO) simulations (Lux et al. 2010; Yniguez et al. 2014). Many studies have argued that reionization can help in this regard as the distribution of *luminous* subhalos will be more centrally concentrated than the overall population of subhalos (e.g. Moore 2001; Taylor et al. 2004; Kravtsov et al. 2004; Font et al. 2011; Starkenburg et al. 2013; Bar-

Corresponding author: Scott G. Carlsten
scottgc@princeton.edu

ber et al. 2014; Dooley et al. 2017). This is generally attributed to the fact that the earliest forming subhalos will be the ones most likely to be luminous and are more concentrated near the host. While the halo mass scale at which reionization starts to significantly suppress galaxy formation is still relatively unclear, 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. While the fully hydrodynamic simulations presented in Samuel et al. (2020) showed better agreement with the MW radial distribution than Yniguez et al. (2014), the MW was still substantially more concentrated in several metrics. Unfortunately, the results with the classical satellites are fundamentally limited by the low statistics offered by the $\sim 10$ MW classical satellites (depending on how 'classical' is defined).

One way to increase statistics is to consider the ultra-faint dwarf (UFD) satellites of the MW as well, of which there is now $\sim 50$ known. The MW UFDs seem to also be more concentrated than DMO simulations would indicate, particularly when disruption by the baryonic disk of the host is included. Kim et al. (2018) found that in the simulations of Garrison-Kimmel et al. (2017b) there simply weren't enough subhalos near the host that survive the enhanced disruption from the disk to host the known MW UFDs at small radii, assuming a reasonable stellar halo mass relation (SHMR). A similar conclusion was reached by Graus et al. (2019). Graus et al. (2019) noted that the close-in MW UFDs could be explained if *very* low mass subhalos were populated. These low mass subhalos are well 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's still unclear how galaxies could form in these halos. One observational prediction for this model is that if luminous galaxies populate these low mass subhalos near the MW, they presumably would further out in the MW virial volume. Thus there should be a very large number of UFDs in the outskirts of the MW virial volume awaiting discovery.

Using the UFDs to test simulation predictions with observations comes with its own problems, however. The observational consensus of UFDs is very radially incomplete due to their faintness (e.g. Koposov et al. 2008; Walsh et al. 2009) and this might be biasing the observed radial distribution. Also, as the UFDs are likely residing in quite low mass subhalos, resolution of the simulations becomes a major concern.

In this paper, we take a complementary approach to the work presented above by studying the observed radial distributions of classical-like satellites around many hosts in the Local Volume (LV). By comparing multiple MW-like hosts together, we are able to get 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 allow us to explore how much the simulation results depend on the properties of the simulation, 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? (Yniguez et al. 2014; Willman et al. 2004)).

This paper is structured as follows: in §2 we overview the observational data sets of LV satellite systems that we use, in §3 we overview the different simulation suites that we compare with, in §4 we present the results of comparing observations with models, in §5 we discuss possible caveats related to the observations and simulations, in §6 we show the implications of this work on the stellar halo mass relation (SHMR) allowed by observations, in §7 we present a detailed comparison of our work with previous work (as mentioned above), and finally in §8 we conclude and outline directions for future work.

## 2. DATA

In this work, we compare simulations with a sample of well-characterized observed satellite systems around massive hosts in the Local Volume ($D \lesssim 10$ Mpc). 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 handful of massive hosts have been surveyed to the point that 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. We use the compilation of LV satellite systems given in Carlsten et al. (submitted). That work uses the catalog of satellite candidates around ten LV massive hosts from Carlsten et al. (2019c) and confirms satellites using distances measured via surface brightness fluctuations (SBF) (Carlsten et al. 2019a). The SBF analysis was very successful for five of the ten hosts surveyed by Carlsten et al. (2019c) (NGC 1023, NGC 4258, NGC 4631, M51, M104). These hosts are mostly complete down to a satellite luminosity of $M_V \sim -9$ within the inner $\sim 150$ projected kpc (see Carlsten et al. 2019c, 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 (either including them as satellites or not). A sixth host, NGC 4565, had good area coverage and deep survey data, but due to its larger distance ($D = 11.9$ Mpc) its SBF results were ambiguous for all but the brightest candidates. All of the candidates brighter than $M_V = -12$ were confirmed by SBF or redshift and there were no ambiguous candidates at this luminosity. Thus, we include this host as well, but we note that its satellite system is only characterized down to $M_V = -12$.

In addition to the six hosts from Carlsten et al. (submitted), 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 Carlsten et al. (submitted). The MW satellite list comes from McConnachie (2012) and uses distances from Fritz et al. (2018). We include the disrupting Sgr dSph but we note that including this in the comparisons with models is a little dubious, 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. (2017), Danieli et al. (2017), Carlsten et al. (2019b), and Bennet et al. (2019).

A detailed discussion of the completeness of each of these satellite systems can be found in Carlsten et al. (submitted) 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. M81 is likely complete to better than $M_V \lesssim -9$ within the inner projected 250 kpc. CenA and M101 are 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. For the six hosts from Carlsten et al. (submitted), we assume the satellite systems are complete to $M_V \sim -9$ (with the exception of NGC 4565), and we generally use the actual survey footprints (see Fig 1 of Carlsten et al. (2019c)) to characterize the areal completeness. For some of the comparisons, we do make the assumption that these surveys are complete out to 150 projected kpc. This is certainly optimistic but these hosts all have coverage over at least 70% of the inner 150 projected kpc area (see Carlsten et al. 2019c, for detailed numbers of the coverage), and the covered volume will be even more.

Properties for all 12 hosts considered in this work can be found in Carlsten et al. (submitted) and Carlsten et al. (2019c). These are all massive hosts with stellar mass ranging from roughly 1/2 that of the MW to $\sim 10\times$ that of the MW. As discussed in Carlsten et al. (submitted), the hosts naturally split in 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 group are all very similar to the MW, and we will refer to these as 'MW-like' or 'MW-analogs'. These include the MW, M31, M94, M101, NGC 4631, NGC 4258, NGC 4565, and M51. We estimate the halos masses of these hosts are in the range $\sim 0.8 - 3 \times 10^{12}$ 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 $\sim 3 - 8 \times 10^{12}$ M$_\odot$. It is important to compare the observed hosts with simulated hosts of similar mass, and so we often consider each group of observed hosts separately and compare each individually with the appropriate simulated hosts.

## 3. MODELS

We compare the observations with a wide variety of different simulations both to improve the statistics in the simulated systems and to see how the properties of the simulated systems depend on the specifics of the simulation. We include both big-box cosmological simulations that include many MW-like hosts in the simulated volume but at a low resolution and zoom-in cosmological simulations that include only a single MW-sized host in each simulated volume but at a much higher resolution.

For all of the simulation suites except for the zoom hydrodynamic simulations, we just use the dark matter

(DM) halo catalogs from the simulations. To populate these subhalos with luminous galaxies we could use an abundance matching (AM) relation. In particular, we could use a stellar halo mass relation (SHMR) to assign a stellar mass to each subhalo. Carlsten et al. (submitted) found decent agreement between the observed satellite LFs in the current sample and the simulated LFs assuming common SHMRs found in the literature. However, there is still significant uncertainty in what the true SHMR is (cf. §6). Therefore, in order to keep the results regarding the radial distribution as independent as possible from the assumed SHMR, we do not use a SHMR to populate subhalos. Instead, we generally select the $n$ most massive subhalos 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 subhalos masses, we always consider the peak virial mass over the subhalo's history, $M_{\rm peak}$.

For the cosmological simulation, we use the public IllustrisTNG-100[1] simulation (Nelson et al. 2019; Pillepich et al. 2018; Springel et al. 2018; Nelson et al. 2018; Naiman et al. 2018; Marinacci et al. 2018). The baryonic mass resolution of TNG is $\sim 10^6$ M$_\odot$which means that the satellites of the mass we are interested in will not be resolved. However, the DM particle mass of $7.5 \times 10^6$ M$_\odot$ means that subhalos hosting the satellites we are focusing on ($M_{\rm vir} \sim 5 \times 10^9$ M$_\odot$) will be fairly resolved. From the $100^3$ Mpc$^3$ box, we select host halos as described in Carlsten et al. (submitted). In this volume there are $\sim 1600$ MW-like hosts and $\sim 300$ small-group sized hosts. Note that 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.

For the zoom-in simulations, we use three separate simulation suites. The first is the ELVIS[2] (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 is significantly higher resolution than IllustrisTNG. The second zoom simulation we include is the PhatELVIS[3] suite (Kelley et al. 2019) of DMO simulations. These simulations are distinct from the ELVIS suite both by being higher resolution (DM mass resolution of $3 \times 10^4$ M$_\odot$) and also that they account 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 subalo catalogs, as described above. The final zoom simulation we include is the full hydrodynamic simulations from the NIHAO project[4] (Buck et al. 2019). There are only 6 hosts in this suite, but the simulations are very high resolution with DM particle mass of $\sim 10^5$ M$_\odot$and star particle mass of $\sim 10^4$ M$_\odot$. Additionally, the simulations are hydrodynamic so we take the luminous galaxy catalog directly from the simulations. The lowest mass satellite we consider in this paper is $M_* > 10^5$ M$_\odot$which will be at least marginally resolved in the NIHAO results. All of the zoom-in simulations are of hosts 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. The 'small-group' hosts are compared solely with the IllustrisTNG simulations.

## 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 simulations against the models in many different ways, using different 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 IllustrigTNG results but occasionally also consider the other simulation suites to show consistency. We start by considering the normalized 2D (projected) radial distributions of all of the observed hosts 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 this section, we explore the projected (2D) radial distribution of satellites of the observed LV hosts and compare with the simulated systems.

In Figure 1, we show the radial profiles for each of the 12 observed hosts that we consider in this work.

[1] https://www.tng-project.org/data/

[2] http://localgroup.ps.uci.edu/elvis/index.html

[3] http://localgroup.ps.uci.edu/phat-elvis/

[4] http://www2.mpia-hd.mpg.de/~buck/#sim_data

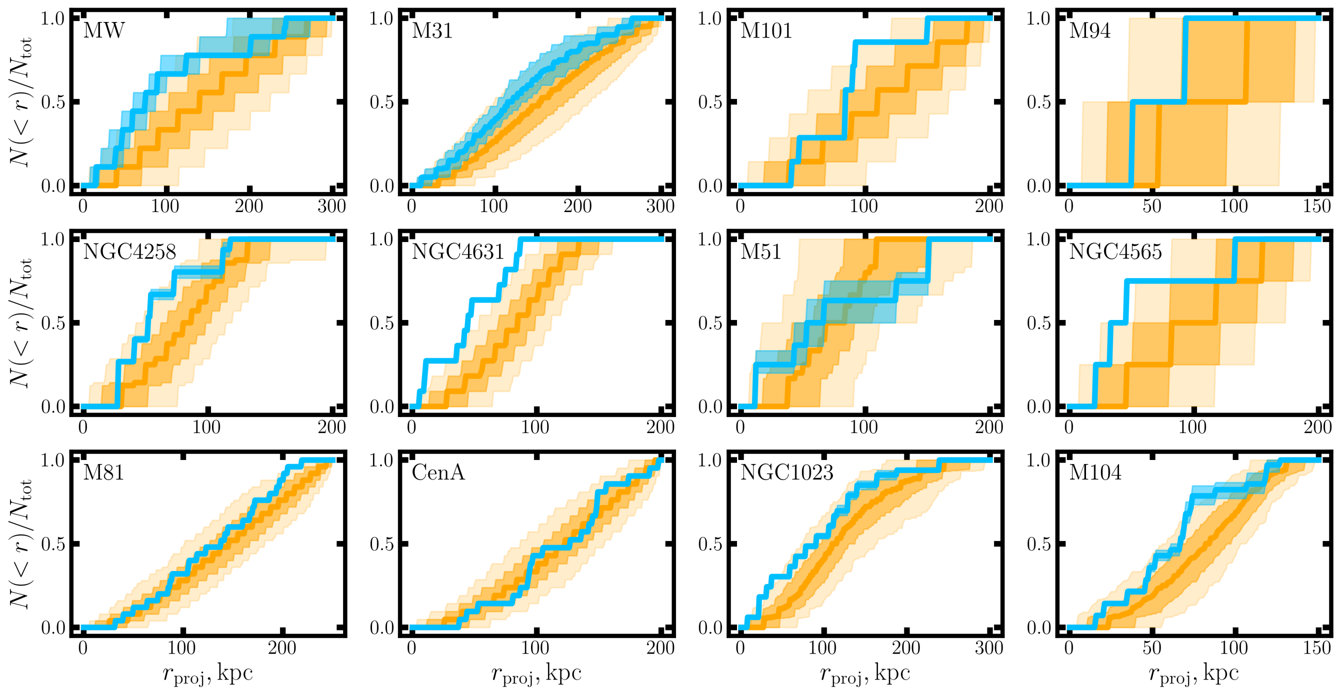


**Figure 1.** The 2D (projected) radial distributions of satellites ($M_V < -9$) around the 12 LV hosts considered in this work. All profiles are normalized by the total number of satellites in the surveyed area. Observed systems are shown in blue while the analogous simulated systems from IllustrisTNG are shown in orange. The 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 selected 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.

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 which uses a luminosity limit of $M_V < -12$). All profiles are normalized by the total number of satellites in the surveyed area. Each host is compared with the simulated systems from IllustrisTNG. The MW-like hosts are compared to the analogous hosts in Illustris, and the same for the small-group hosts. 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 (these are described above). 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 high enough precision to probe the 3D structure of the group. SBF distances are accurate to $\sim 15\%$ while HST TRGB distances are accurate to $\sim 5\%$. At a host distance of 7 Mpc, these correspond to uncertainties of $\sim 1000$ 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 of the virial volume of the host. 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 that 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 Carlsten et al. (submitted) 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, the observed hosts appear to be generally more concentrated than the simulated

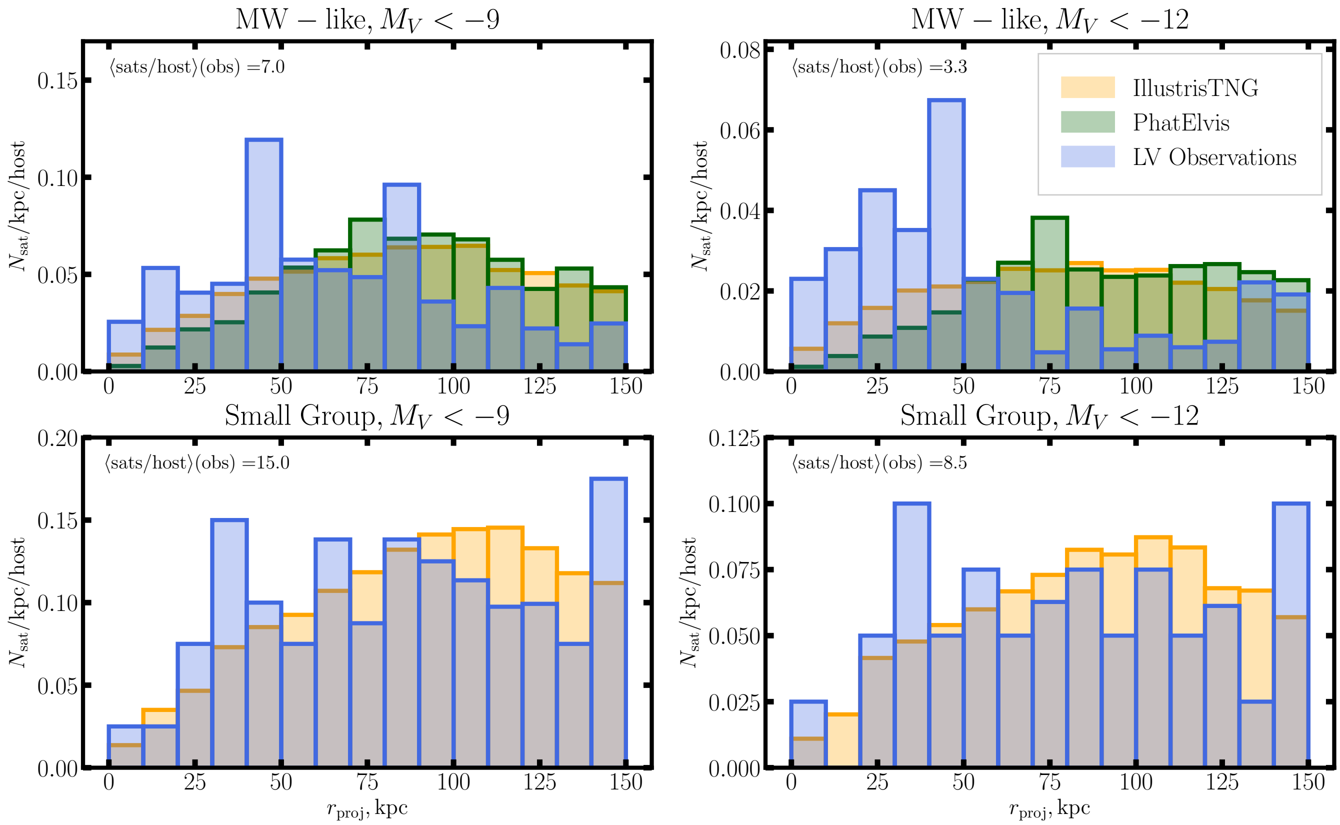


**Figure 2.** Histogram of the projected satellite separation ($r_{proj}$) for all observed hosts together compared against the simulated hosts. The observation histograms are normalized such that the total area is the average number of satellites per host. These numbers are given in the top left corner of each panel. The top panels are for the MW-like hosts while the bottom are for the more massive 'small-group' hosts. The left panels are for all satellites $M_V < -9$ while for the right panels, only $M_V < -12$ satellites are included.

hosts. Several of the observed hosts (e.g. the MW, NGC 4258, NGC 4631, NGC 4565, and NGC 1023) have their profiles at or just within the $-2\sigma$ (i.e. more centrally concentrated) scatter in the simulations while no host is correspondingly outside the $+2\sigma$ (i.e. less concentrated) scatter in the simulations. It appears that the 'small-group' hosts are less discrepant with the simulations. Both M81 and CenA closely follow the median simulated profile.

Another comparison we do with the simulations is to simply consider the histogram of the satellites' projected separations from their hosts. In Figure 2, we show the distribution of all satellite projected separations across all hosts combined. The histograms are normalized such that the total area under the curve is the average number of satellites per host. Only satellites within $r_{proj} < 150$ kpc are included. We 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. 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 simulations are forward modelled to include the effect of the survey footprints of the observed hosts in a similar way as 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 all the way out to 150 kpc. For the 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.

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 the simulated hosts.

To assess the significance of the discrepancy with the MW-like hosts, we use a two-sample 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 $5.6 \times 10^{-5}$ for all ($M_V < -9$) satellites and $3.6 \times 10^{-7}$ for the bright ($M_V < -12$) satellites. When compared to the TNG simulations, we find median $p$-values of 0.0026 and $7.5 \times 10^{-5}$, respectively.

### 4.2. *3D Radial Distributions*

Figure 3 shows the 3D radial distributions of the classical ($M_V < -8$) 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 $r < 100$ kpc, to compare with the results of Yniguez et al. (2014). 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. For the NIHAO hosts, the hydrodynamic results are used, and satellites with $M_V < -8$ are selected. There are two interesting things to note about from this plot. 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 are not affecting the results, and that even the low resolution IllustrisTNG results are converged at this mass scale. We consider this point in more detail in §5. The biggest difference between the simulations appears to be 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) put into the simulations. ELVIS has no added disk. Both the Illustris and NIHAO simulations should have some of this effect because the host will form a disk in these hydrodynamic simulations, but it is possible that many of the Illustris hosts do not have as massive of a disk as the MW forms. On the other hand, the NIHAO suite might just not have enough hosts to see this effect.

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 outside of the $2\sigma$ regions of all of the simulations. This confirms the result of Yniguez et al. (2014). Most hosts have significant populations of satellites outside of $r = 150$ kpc whereas the MW has only 3. M31, on the other hand, appears to have a fairly characteristic radial distribution compared to the simulations. Yniguez et al. (2014) argue that this unusual concentration indicates that some undiscovered MW satellites exist far out in the virial volume, awaiting discovery. We address the point of completeness of the MW satellite census in §5.1 and argue that this is unlikely.

### 4.3. *Satellite Concentration*

To explore a different metric of 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 $R_{\rm half}$ is calculated essentially as the median satellite projected separation 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 plot. 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 (contours). For the Illustris hosts, we use the stellar mass of the host reported by the hydrodynamic simulation results. The average observed number of satellites within 150 kpc and with $M_V < -9$ (including the projection effects of the MW and M31 and the effect of uncertain membership) is 7 per MW-like host and 15 per small-group host. Thus, we select the 7 most massive subhalos for each TNG MW-like host and the 15 most massive for each small-group host to compare with the observed satellites. The Illustris hosts are all mock observed at $D = 7$ Mpc and include both the MW-like hosts and the small-group 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 7 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 $\sim 90$ kpc for the simulations but is closer to $\sim 60 - 70$ kpc for the observed systems. The observed hosts are systematically more centrally concentrated than the simulated hosts. 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 satel-

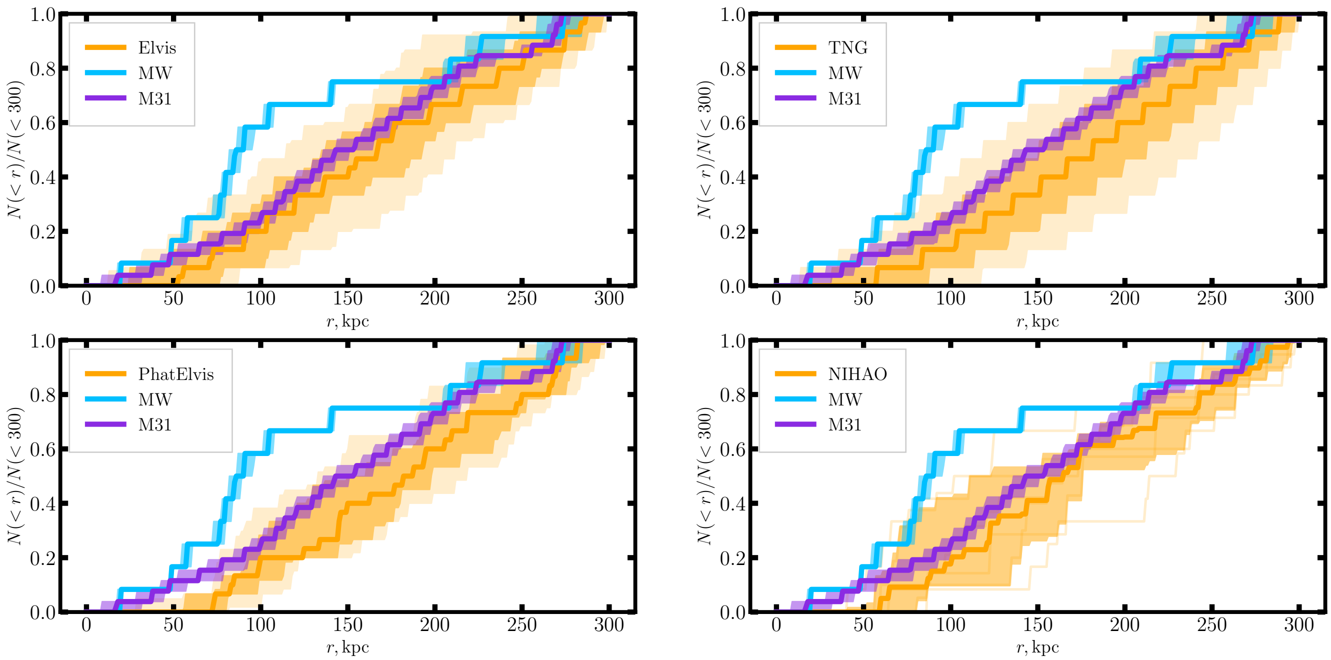


**Figure 3.** The 3D radial distribution of MW and M31 satellites ($M_V < -8$) compared with four different simulation suites. For the DMO simulations, the 15 most massive subhalos are selected as satellites. The profiles are cumulative and normalized by the number of satellites within 300 kpc. The shaded 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 6 hosts in the NIHAO suite, they are plotted individually, and the thick line shows the median.

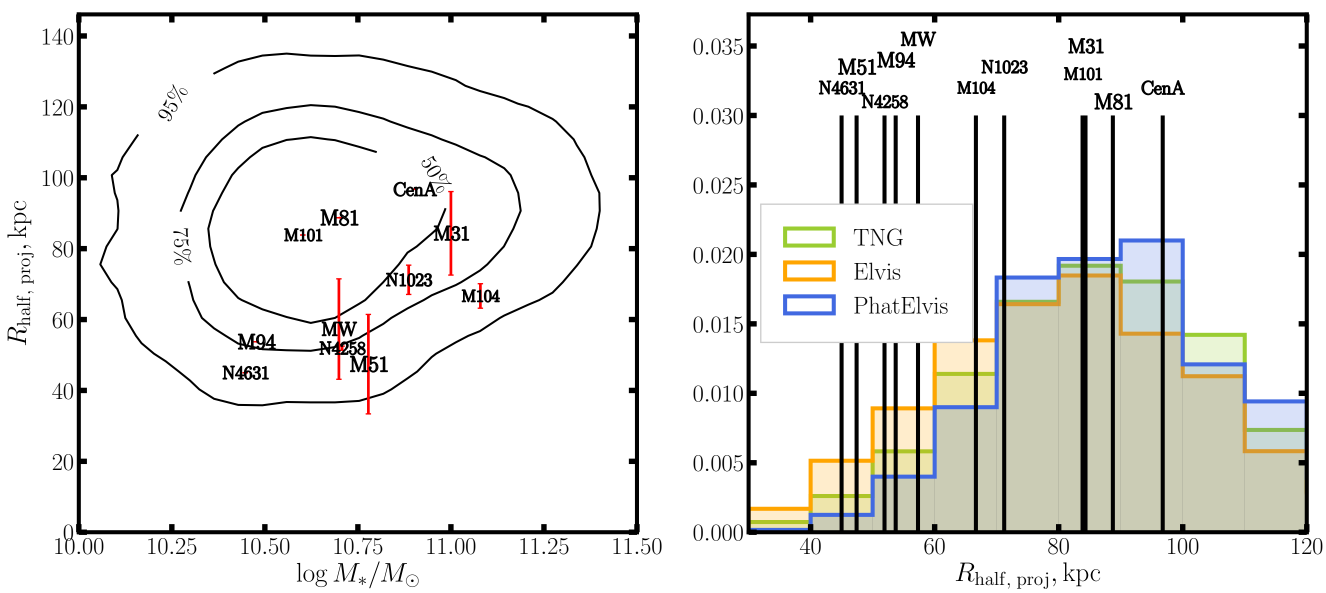


**Figure 4.** The radius encompassing half of the satellites for 11 observed hosts versus the simulated hosts. Only satellites brighter than $M_V < -9$ and within a projected 150 kpc of their host are considered. On the left, the median satellite radius is plotted against the stellar mass of the host. The contours in the background show the results for the IllustrisTNG hosts. The errorbars for M31 and the MW show the effect of different observing angles while the errorbars 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 separation for the hosts are compared against the simulations in histogram form. The results from the ELVIS and PhatELVIS suites are also shown.

lite spatial distributions. The ELVIS hosts are slightly more centrally concentrated than the IllustrisTNG or PhatELVIS hosts, but that is easily explained by the lack of any central disk. A few of the most massive hosts (M81, CenA, and M31) appear to have spatial distributions characteristic of the simulated hosts, but the majority of the observed hosts are more concentrated and *none* are less concentrated.

We have checked if the comparisons will change if we scale all the satellite separations by the virial radii of the hosts. In Figure 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, especially when we split the hosts into the MW-like and small-group categories. The simulated hosts will have the same spread in virial radii. Still, it is useful to investigate if the increased concentration in satellites that we see compared to the simulated hosts is due to the 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. *Summary*

In this section, we performed many different comparisons between the radial distributions of observed satellites in the LV and analogous simulated systems in modern cosmological simulations. While some of the observed hosts seem to have very similar radial profiles as the simulated hosts, on average the observed hosts are more concentrated. In Figures 1 and 3, we showed this by directly comparing the normalized radial distributions. Normalizing the profiles is essential in seeing the the difference between observations and the simulations. In Figure 2, we compared the distribution of satellite projected separation, $r_{\rm proj}$, between the observed hosts and the simulations. For the MW-like hosts, the distribution of observed satellites is significantly more centrally concentrated than the simulated hosts. This is particularly noticeable compared to the PhatELVIS simulations and when only the bright ($M_V < -12$) satellites are considered. Next, we used the half-satellite radius (the radius enclosing half of the satellites), $R_{\rm half}$, to quantify the level of concentration of the radial profiles. Figure 4 shows that the population of observed hosts is shifted to smaller values of $R_{\rm half}$. While a few of the observed hosts show the average value for the simulations, most of the observed hosts are more concentrated, and none are less so.

In the rest of this paper, we discuss these results in more detail. In §5 we discuss all the possible problems with either the observations or the simulations that might be contributing to such a strong discrepancy. In §7 we compare our results with previous work on this subject, highlighting where our approach is different.

## 5. POSSIBLE EXPLANATIONS

In this section, we discuss possible caveats to the observations and simulations that might explain the discrepancy that we observe. We start by discussing the observations then move to the simulations.

### 5.1. *Observations*

The most obvious possible problem with the observations is that of incompleteness. 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 focusing mostly on the central host. Indeed several of the hosts of Carlsten et al. (2019c) had dithering patterns that led to somewhat deeper exposures around the host than further out. However, there are two lines of reasoning indicating that this is not significantly affecting the observed radial profiles. First is that the completeness checks of Carlsten et al. (2019c) (see their Figure 3) indicate that the catalogs of satellites drop from $\sim 100\%$ completeness to $\sim 0\%$ over roughly half a magnitude in either surface brightness or total magnitude of the satellite. These checks are the average over the entire survey footprint. If there was a significant difference in depth between different areas of the survey, then we'd expect the completeness to drop off more shallowly. Second, Figure 2 shows that the discrepancy does not get better (in fact it gets quite a bit worse) when we only consider the brighter satellites.

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-sized satellites would have to be in the zone of avoidance around the MW disk to have evaded discovery. 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; Walsh et al. 2009; Drlica-Wagner et al. 2019). The issue is that obscuration by the disk does not preferentially hide distant satellites so it is unclear if this would really decrease the concentration of the MW satellites. It would, however, increase the scatter of the simulated profiles, perhaps making the MW less of an outlier. We find that Figure 3 hardly changes if we select only subhalos that are $> 15\,$deg above or below a randomly oriented disk in the simulated hosts (as viewed from the center of the host). Even in this case, the MW remains a significant outlier from the simulations.

Another possible issue with the observations is the confirmation of satellites with SBF distances. While the SBF approach has been shown to work well in the past (e.g. Carlsten et al. 2019b; Bennet et al. 2019), 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 Carlsten et al. (submitted). However, including background galaxies will make the radial profiles less centrally concentrated not more.

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 somewhat of a lower bound.

### 5.2. *Simulations*

The most likely possible problems with the simulations stem from resolution effects. On the surface, the fact that we get very similar answers across the different simulations suites (once we keep in mind 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’. Orphan galaxies represent the possibility that the DM halo associated with a luminous satellite 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 as it has no corresponding DM subhalo. Previous works (e.g. Gao et al. 2004; Newton et al. 2018; Bose et al. 2019) 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 dynamical friction timescale to represent the time they take to merge 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, however. Orphan galaxies are clearly only 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 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 ($\sim$ 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) and the satellite would likely either be 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. However, it is unclear whether this is the appropriate thing to do either. 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. 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 as these galaxies do not correspond to the population of observed satellites.

The second 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) provides new criteria to judge whether the tidal evolution of a subhalo is trustworthy or whether it is affected by discreteness noise. For a subhalo on a circular orbit with a radius of 0.2 times the host virial radius, the subhalo 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 van den Bosch & Ogiya (2018) criteria are for circular orbits in a NFW tidal field. Presumably even higher resolution would be required for the non-circular orbits characteristic of cosmological simulations in the presence of the tidal field of the central disk. Artificial disruption of subhalos could easily explain the discrepancy we find if many of the simulated subhalos near the host galaxy were artificially destroyed.

## 6. IMPLICATIONS FOR THE SHMR

We have shown that observations show a significant population of close in satellites that are not present in DMO simulations. While it is unclear how to resolve this discrepancy, as discussed in the previous section, an appealing option is that the simulations are missing subhalos, either from artificial over-merging or incompleteness in the halofinder 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 effect what stellar halo mass relation (SHMR) is allowed by the observed abundances of satellites. In Carlsten et al. (submitted), we showed that the richness of the LV satellite systems could be well matched by the SHMR of Garrison-Kimmel et al. (2017a). This SHMR agrees well with the predictions from recent hydrodynamic simulations of dwarf galaxy formation. However, 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 PhatELVIS suite. Due to its resolution and inclusion of a disk potential, this is the simulation suite that should best describe MW-sized hosts. Thus, we only consider the 8 MW-sized observed hosts in this comparison. 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 in the halo catalogs. We insert the orphans with a projected separation from the host drawn from a Gaussian with mean 30 kpc and standard deviation of 10 kpc. These numbers are chosen to match the peak of close-in observed bright satellites visible in Figure 2. We acknowledge this is clearly unphysical. This is just a toy model which 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 between $10^{8.5} < M < 10^{11.8}$ 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 as lower mass subhalos are most likely lost preferentially if this is due to resolution. We by-hand 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 $\sim$ 100 subhalos in this mass range within the inner 150 projected kpc leads to fair agreement with the observations. Figure 5 shows this result on the left. The observed distribution of bright ($M_V < -12$) satellites are shown along with the distribution of the 3 most massive subhalos before and after subhalo injection. Recall that there are, on average, 3 satellites in this luminosity range per observed host, and so we select the 3 most massive subhalos as the simulated satellites. Note that we do not do any survey area corrections for the simulations in this test; we assume the observed systems are all complete to a projected 150 kpc. This is why the PhatELVIS results in Figure 5 look a little different from Figure 2.

To investigate the effect on the SHMR, we fit for the low-mass slope of the SHMR before and after subhalo injection. We use the SHMR parameterization of Garrison-Kimmel et al. (2017a) which fixes the SHMR to that of Behroozi et al. (2013) at the high mass end and allows the power law slope at low mass ($M_{\rm halo} < 10^{11.5}$) to vary. Specifically the slope is defined as $\alpha$, where $M_\star \propto M_{\rm halo}^\alpha$. To do this fit, we use MCMC sampling. For each observed host, the observed satellite system is binned in 1 magnitude wide bins from $M_V = -21$ to $M_V = -9$. For each iteration of the MCMC, the halo catalogs of the simulated hosts are fed through the SHMR and each satellite is assigned a $M_V$ (assuming $M_\star/L_V = 1.2$). Using the simulations, we calculate the average expected number of satellites in each $M_V$ bin. The likelihood of the model is then given by the Poisson likelihood of the observed number of satellites in each bin. In particular, the likelihood we use is:

$$\mathcal{L} = \prod_{\rm hosts} \prod_{\rm bins} \frac{e^{-\lambda_i} \lambda_i^{x_{h,i}}}{x_{h,i}!} \tag{1}$$

where $\lambda_i$ is the expected (model) number of satellites in bin $i$ and $x_{h,i}$ is the observed number of satellites in bin $i$ around host $h$. To deal with the uncertainties regarding the membership of some of the satellite candidates, for each MCMC iteration, we include an uncertain satellite or not in the observational sample with a 50-50 chance. To deal with the uncertainty of the viewing angle on the MW's and M31's satellite system, a different viewing angle is used for these two observed hosts for each MCMC iteration.

We fix the scatter in the SHMR to be 0.5 dex below $10^{11.5}$ M$_\odot$ and 0.2 dex above. This is a fairly arbitrary choice, but we are most interested in the SHMR slope.

We assume a flat prior for the low mass slope in the range [1,6].

The results are shown in the right side of Figure 5. We find the best fitting low mass slope to be $1.77 \pm 0.05$ for the case without orphans and $2.43 \pm 0.10$ for the case with added subhalos. There is a clear effect in the low mass slope from adding in these subhalos, as expected. Figure 5 shows these two best fitting SHMRs along with that of Behroozi et al. (2013). Also shown are various recent simulation results from the literature. The Buck et al. (2019) results are for dwarfs produced in several zoom simulations of the formation of a MW-sized galaxy. We plot the 'field' subsample of dwarfs (beyond $2.5R_{\rm vir}$ from the host) as they will show the least effect from tides from the host. This is important for these simulations because we only have access to the $z = 0$ virial mass of the subhalos and not to $M_{\rm peak}$. The other simulation results are all for isolated dwarfs. The four points from Agertz et al. (2020) show the results from four different runs with different physics. In order of decreasing stellar mass, the runs are 'No feedback', 'weak feedback', 'fiducial', 'fiducial with radiative transfer'. We refer to Agertz et al. (2020) for more detail on the different physical prescriptions. We also show the dynamical constraints on the stellar to halo mass ratio from the observed kinematics of the MW dSphs from Read & Erkal (2019) and the constraints on the SHMR from modelling the population of the MW UFDs (in a similar way to how we are fitting the SHMR with the classical satellites) from Nadler et al. (2019).

There are three main points to emphasize from Figure 5. First, the inclusion of extra 'lost' subhalos does significantly affect the allowed SHMR. Without the extra subhalos, the SHMR low-mass slope is **1.8**, consistent with the result of Garrison-Kimmel et al. (2017a). We note that this consistency is somewhat of a coincidence, however. Garrison-Kimmel et al. (2017a) used the ELVIS suite which has more subhalos per host because it does not include a disk and the hosts are, on average, more massive than the PhatELVIS host halos. If we repeat our fit with ELVIS, we get a slope of **2.15**, which is steeper because of the extra subhalos. Garrison-Kimmel et al. (2017a) found 1.8 to be the best fitting slope possibly because they were considering only the MW and M31, and M31 is by far the richest of the observed 'MW-like' hosts. The average satellite richness is lower in our expanded sample of hosts, making the best-fitting SHMR slope steeper. With the extra subhalos we get a slope of **2.4** with the PhatELVIS suite. Second, the simulation results show significant spread. Even different simulations within the same project (e.g. FIRE) seem to show large spreads in the resulting stellar to halo mass ratio for simulated galaxies. Finally, most of the simulation results seem to agree better with a shallower SHMR slope. The exception are the simulated galaxies by Wheeler et al. (2015, 2019) which appear to to have somewhat lower stellar mass at the same halo mass as many of the other simulations.

While this experiment is clearly unphysical in the way potentially 'lost' subhalos are added back into the simulations, it does show that the radial distribution of satellites and the SHMR allowable by observations are interrelated. Adding in 'orphan' galaxies to bring the simulated radial distributions into agreement with observations will significantly affect what SHMR is able to fit the observations. In particular, the SHMR will need to be steeper, bringing it somewhat out of general agreement with current hydrodynamic simulations.

## 7. COMPARISON WITH PREVIOUS WORK

In §4, we found that the observed satellite systems around 12 massive hosts in the LV are significantly more centrally concentrated than comparable simulated hosts in the four suites of ΛCDM simulations that we consider. Something as simple as the radial distribution of satellites around the MW has been the subject of much prior work. In this section, we go through some of the recent previous work and compare their approach with ours. As we will see, some of the previous work found that the satellites of the MW are more concentrated than models while other work found good agreement between observations and their models. In these cases, we explain why we find such a different result. Finally, we compare our results with that of the SAGA Survey (Geha et al. 2017). This work is complementary to ours in the sense that they are surveying only the brightest satellites but around many more hosts than are in the LV and makes an interesting point of comparison.

Similar to our results, many previous works have found that the classical satellites of the MW are significantly more centrally concentrated than the massive DM subhalos in DMO simulations. Yniguez et al. (2014) showed that the MW classical satellites were more concentrated than any of the ELVIS simulated hosts. We confirm this result in Figure 3 and show that it holds for all the other simulation suites we consider.

Many previous works argued that the *luminous* subhalos are more centrally concentrated than the DM subhalo population as a whole. The physical basis of this 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

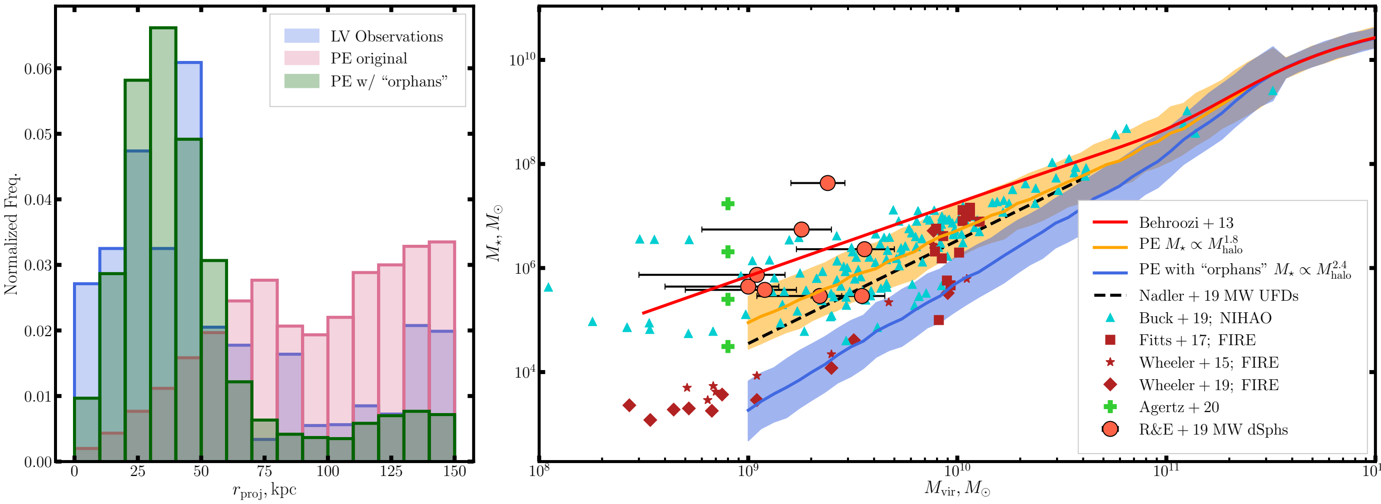


**Figure 5.** One possibility to resolve the discrepancy in radial distributions between simulated and observed hosts is that the simulations are missing many subhalos, either due to artificial disruption or halofinder incompleteness. This figure shows the effect of adding these subhalos back into the simulation. *Left:* The radial distribution of the PhatELVIS (PE) simulations after these 'orphan' galaxies are added back in. See text for details about how they are added. *Right:* We fit for the low mass slope of the SHMR both with and without these extra subhalos. Including them leads to a much steeper best fit SHMR. Also shown are the results from various hydrodynamic simulations including Buck et al. (2019); Fitts et al. (2017); Wheeler et al. (2015, 2019); Agertz et al. (2020). We show as well the observational constraints on the MW dSphs from Read & Erkal (2019) and the constraint from modelling the MW UFD population from Nadler et al. (2019).

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 (SAM) of galaxy formation coupled with DMO simulations and found that the radial distribution of model satellites was similarly concentrated as 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 SAM of Starkenburg et al. (2013) implemented the effect of reionization using the filtering mass approach of Gnedin (2000) which is now known to over-predict 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 possibly is a result of the stronger effect of reionization in their implementation. 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 AM relation we use indicates that satellites of the luminosity we consider ($M_V \sim -10$) are hosted by halos of mass $\gtrsim 5 \times 10^9 \mathrm{M}_\odot$which are above the effects of reionization.

Very recently, Samuel et al. (2020) extensively compared the satellite radial profiles of the MW and M31 to simulated hosts in the FIRE project. They argued that the simulated hosts had radial profiles that were fairly consistent with that of the MW and M31. Our results are not inconsistent with theirs for two reasons. First, in this paper, we compared a population of observed systems to a population of simulated analogs which has significantly improved statistics over Samuel et al. (2020). If we only compared the MW and M31 to a handful of simulated hosts, the disagreement would be less noticeable. Second, we find the disagreement only when we normalize the radial profiles, focusing on the shape. Samuel et al. (2020) primarily consider the absolute radial profile, the spread of which in the model is able to encompass the MW. If, instead, the normalized profiles of the FIRE hosts were compared with the MW, the result might be different. Samuel et al. (2020) did explore the shape of the radial profile using 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. In this paper, we opt to use the simpler $R_{\rm half}$ as a measure of the concentration of a profile. For systems with few satellites, it becomes difficult to define the radius that contains 90% of satellites. Furthermore, Samuel et al. (2020) found that when normalizing the

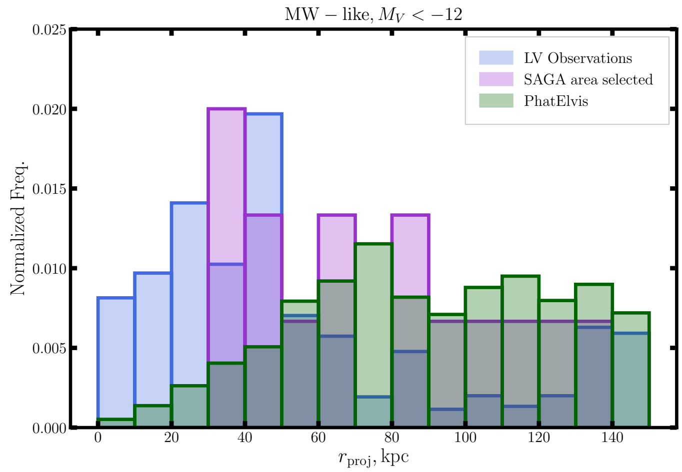


**Figure 6.** The distribution of projected separations for the LV observed satellite systems compared with both the simulated hosts of the PhatELVIS project and the SAGA Survey. The normalization of the histograms is arbitrary. Both the SAGA and PhatELVIS hosts are passed through the observational selection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

simulated profiles to that of the MW at $r = 150$ kpc, all of the simulated hosts had more satellites at $r > 150$ kpc than the MW has. They used this to argue that there could be some classical-sized satellites yet undiscovered in the periphery of the MW virial volume. We have discussed the possibility of this above in §5.1.

In the final part of this section, we compare our results to the radial profiles inferred from the SAGA Survey (Geha et al. 2017). This survey is observing MW-analogs in the distance range $20 - 40$ Mpc. As such, it is only sensitive to the bright satellites ($M_V \lesssim -12$) but will have many more observed hosts than what is possible in the LV. The first paper presented 8 complete hosts. While this is less statistics than currently possible for the LV systems (particularly because the LV hosts are complete to much fainter satellites), it is interesting to compare their radial profiles with those inferred in this work. Figure 6 shows the distribution of projected separations for the MW-like LV hosts and for bright ($M_V < -12$) satellites compared with the confirmed SAGA satellites and the profiles predicted from the PhatELVIS hosts. In order to focus on the shape of the distributions, all histograms are normalized the same. Both the SAGA satellites and PhatELVIS hosts are passed through the observational selection functions of the LV hosts. A random LV host is selected for each SAGA and PhatELVIS host and that host's selection function (based on survey area) is used.

Interestingly, the SAGA satellites are somewhat in between the LV host satellites and the PhatELVIS results. The SAGA surveys show the same large excess of satellites at $r_{\rm proj} \sim 40$ kpc. There are no very close-in satellites in the SAGA results, possibly due to the SDSS targeting and larger distance of these hosts. It is possible that these very close-in satellites are lost in the outer envelope of the hosts, and at the large distance of these hosts, it will hard to separate them. We note that the statistics for the SAGA satellites are rather low and a firm conclusion can not be drawn. The highest peak at $r_{\rm proj} \sim 40$ kpc corresponds to only three SAGA satellites. Future SAGA results will provide confirmation or not of this trend.

## 8. CONCLUSIONS

In this paper, we compared the observed satellite spatial distributions in 12 LV massive hosts to that predicted in various state-of-the-art cosmological simulations in the ΛCDM 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 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 6 MW-like hosts (NIHAO). Overall we find fairly good agreement between the simulations. Our main findings are as follows:

(i) 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.

(ii) 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 hosts such as M81 and CenA, have similar radial profiles to the simulated hosts, but *none* of the observed are less concentrated than the average simulated profile.

(iii) 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). This does not change if we express the satellite separations as fractions of the host virial radius (Figure **??**).

(iv) Figure 2 shows that the spatial distribution of the satellites is significantly different between the observed satellites and simulated. 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.

(v) The spatial distribution of satellites found in the SAGA Survey (Geha et al. 2017) seems to agree fairly well 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 §5, we discussed many possible causes for the discrepancy both on the observational side and on the simulation side. There does not seem to be any likely causes for this on the observational side, due to incompleteness. Observational completeness has been well quantified by the various works that have characterized the LV satellite systems. Unless these completeness levels are very erroneous, the observational results seem quite solid. We do note, however, that most of the discrepant observed systems (e.g. NGC 4631 and NGC 4258) come from the work of Carlsten et al. (2019c). Not including these systems makes the discrepancy much less severe. With that said, the MW which is arguably the observed host with the best completeness is quite discrepant with the simulations, and this is suggestive that there is more to this than simple observational incompleteness.

We also discussed possible issues with our application of 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 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 it is not detected by a subhalo finder, the luminous galaxy would be mostly destroyed and, hence, 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. This will have important ramifications for the allowable SHMR in this mass range. Carlsten et al. (submitted) 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 5) reproduces fairly well the overall number of observed satellites in these systems. In §6, we show that 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.

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 be the simulations that 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. 2019). We note that the model of Nadler et al. (2019) 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 on 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 ($r_{\rm proj} < 150$ kpc) satellites of the MW are more concentrated than simulated analogs but only at a $\sim 1\sigma$ level. In Appendix A, we show that when considering the entire satellite system out to 300 kpc, the discrepancy between $R_{\rm half}$ of the MW and that of the simulated systems becomes much more significant at $> 2\sigma$.

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 2019). 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 and NGC 4631) having similarly concentrated radial profiles makes this possible connection all the more intriguing.

## ACKNOWLEDGEMENTS

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.

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)

## REFERENCES

Agertz, O., Pontzen, A., Read, J. I., et al. 2020, MNRAS, 491, 1656

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

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57

Bennet, P., Sand, D. J., Crnojević, D., et al. 2019, arXiv e-prints, arXiv:1906.03230

—. 2017, ApJ, 850, 109

Bose, S., Deason, A. J., Belokurov, V., & Frenk, C. S. 2019, arXiv e-prints, arXiv:1909.04039

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., Green, J., Greco, J., & Beaton, R. submitted, ApJ

Carlsten, S. G., Beaton, R. L., Greco, J. P., & Greene, J. E. 2019a, ApJ, 879, 13

—. 2019b, ApJL, 878, L16

Carlsten, S. G., Greco, J. P., Beaton, R. L., & Greene, J. E. 2019c, arXiv e-prints, arXiv:1909.07389

Chiboucas, K., Jacobs, B. A., Tully, R. B., & Karachentsev, I. D. 2013, AJ, 146, 126

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

D'Onghia, E., Springel, V., Hernquist, L., & Keres, D. 2010, ApJ, 709, 1138

Dooley, G. A., Peter, A. H. G., Carlin, J. L., et al. 2017, Monthly Notices of the Royal Astronomical Society, 472, 1060

Drlica-Wagner, A., Bechtol, K., Mau, S., et al. 2019, arXiv e-prints, arXiv:1912.03302

Fitts, A., Boylan-Kolchin, M., Elbert, O. D., et al. 2017, MNRAS, 471, 3547

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

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
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 American Institute of Physics Conference Series, Vol. 586, 20th Texas Symposium on relativistic astrophysics, ed. J. C. Wheeler & H. Martel, 73–82
Müller, O., Rejkuba, M., Pawlowski, M. S., et al. 2019, arXiv e-prints, arXiv:1907.02012
Nadler, E. O., Wechsler, R. H., Bechtol, K., et al. 2019, arXiv e-prints, arXiv:1912.03303
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, Computational Astrophysics and Cosmology, 6, 2
Newton, O., Cautun, M., Jenkins, A., Frenk, C. S., & Helly, J. C. 2018, MNRAS, 479, 2853
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. 2019, MNRAS, 2774
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
Read, J. I., & Erkal, D. 2019, MNRAS, 487, 5799
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 Symposium, Vol. 220, Dark Matter in Galaxies, ed. S. Ryder, D. Pisano, M. Walker, & K. Freeman, 91
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
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., Oñorbe, J., Bullock, J. S., et al. 2015, MNRAS, 453, 1305
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

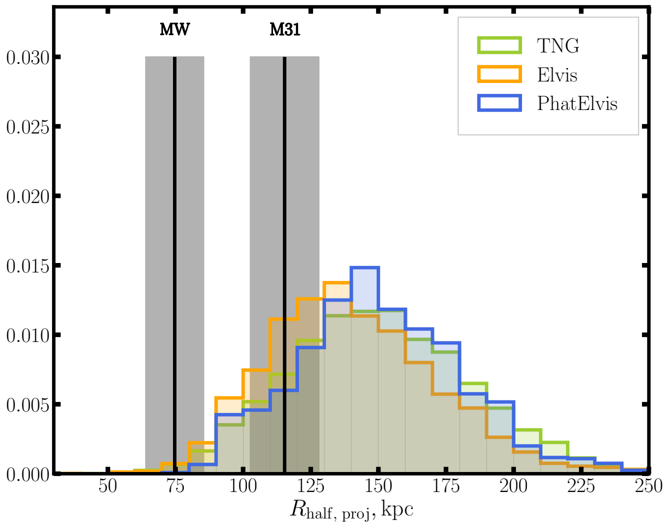


**Figure 7.** Same as Figure 4 but for satellites within 300 projected kpc. Only the MW and M31 are shown as these are the only observed systems complete at these radii. The shaded bands show the $\pm 1\sigma$ spread in $R_{\rm half,\,proj}$ due to taking different viewing angles.

# APPENDIX

## A. $R_{\rm HALF}$ FOR THE MW AND M31

In Figure 7, we show the distribution of $R_{\rm half,\,proj}$ for the simulations and observed systems when including satellites all the way out to 300 projected kpc. Only the MW and M31 are included as no other observed system is complete to these radii. In this comparison, the 15 most massive subhalos are selected for each host in the simulation as the luminous satellites. When including the full system of satellites, the discrepancy between the MW's system and the simulated systems becomes much more significant.