

## Symphony: Cosmological Zoom-in Simulation Suites over Four Decades of Host Halo Mass

ETHAN O. NADLER,<sup>1,2</sup> PHILIP MANSFIELD,<sup>3,4,5</sup> YUNCHONG WANG,<sup>3,4,5</sup> XIAOLONG DU,<sup>1</sup> SUSMITA ADHIKARI,<sup>6</sup> ARKA BANERJEE,<sup>6</sup> ANDREW BENSON,<sup>1</sup> ELISE DARRAGH-FORD,<sup>3,5,4</sup> YAO-YUAN MAO,<sup>7</sup> SEBASTIAN WAGNER-CARENA,<sup>3,5,4</sup> RISA H. WECHSLER,<sup>3,5,4</sup> AND HAO-YI WU<sup>8</sup>

<sup>1</sup>Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>2</sup>Department of Physics & Astronomy, University of Southern California, Los Angeles, CA, 90007, USA

<sup>3</sup>Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA

<sup>4</sup>SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

<sup>5</sup>Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA

<sup>6</sup>Department of Physics, Indian Institute of Science Education and Research, Homi Bhabha Road, Pashan, Pune 411008, India

<sup>7</sup>Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA

<sup>8</sup>Department of Physics, Boise State University, Boise, ID 83725, USA

### ABSTRACT

We present Symphony, a compilation of 262 cosmological, cold-dark-matter-only zoom-in simulations spanning four decades of host halo mass, from  $10^{11}$ – $10^{15} M_{\odot}$ . This compilation includes three existing simulation suites at the cluster and Milky Way–mass scales, and two new suites: 39 Large Magellanic Cloud-mass ( $10^{11} M_{\odot}$ ) and 49 strong-lens-analog ( $10^{13} M_{\odot}$ ) group-mass hosts. Across the entire host halo mass range, the highest-resolution regions in these simulations are resolved with a dark matter particle mass of  $\approx 3 \times 10^{-7}$  times the host virial mass and a Plummer-equivalent gravitational softening length of  $\approx 9 \times 10^{-4}$  times the host virial radius, on average. We measure correlations between subhalo abundance and host concentration, formation time, and maximum subhalo mass, all of which peak at the Milky Way host halo mass scale. Subhalo abundances are  $\approx 50\%$  higher in clusters than in lower-mass hosts at fixed sub-to-host halo mass ratios. Subhalo radial distributions are approximately self-similar as a function of host mass and are less concentrated than hosts’ underlying dark matter distributions. We compare our results to the semianalytic model GALACTICUS, which predicts subhalo mass functions with a higher normalization at the low-mass end and radial distributions that are slightly more concentrated than Symphony. We use UNIVERSE MACHINE to model halo and subhalo star formation histories in Symphony, and we demonstrate that these predictions resolve the formation histories of the halos that host nearly all currently observable satellite galaxies in the universe. To promote open use of Symphony, data products are publicly available at <http://web.stanford.edu/group/gfc/symphony>.

**Keywords:** Dark matter (353); Galaxy abundances (574); *N*-body simulations (1083); Galaxy dark matter halos (1880); Computational methods (1965)

### 1. INTRODUCTION

The hierarchical formation and nonlinear evolution of self-gravitating dark matter systems, or “halos,” underpins our modern understanding of cosmic structure. Numerical simulations, which remain the most widely used and accurate technique to model structure formation, are particularly important for resolving “subhalos” that reside within larger host halos (e.g., see Zavala & Frenk 2019 for a review). Recent studies at the forefront of structure formation and galaxy evolution depend on the properties of subhalos with low masses relative to their host halos—and the relatively faint satellite galaxies that reside within them—in a wide range of cosmic

environments (e.g., Geha et al. 2017; Kallivayalil et al. 2018; Gilman et al. 2020b; Meneghetti et al. 2020; Adhikari et al. 2021). Thus, a unified simulation suite that captures subhalo populations at high resolution across the entire observationally relevant range of host and subhalo masses is timely.

“Zoom-in” simulations, in which a small region of a larger, lower-resolution “parent” simulation is resimulated at higher resolution, are useful for resolving subhalo populations in a cosmological context (Katz & White 1993; Bertschinger 2001). Most “zoom-ins” focus on a specific host halo (or pair of host halos; e.g., Garrison-Kimmel et al. 2014) and resimulate a region that contains all particles that eventually reside within that halo (its “Lagrangian volume”) at higher resolution than in the parent box. This is done using initial conditions generated with nested regions of increasing refinement (Jenkins 2010; Hahn & Abel 2011). This approach ensures that the formation and evolution of the host halo’s dark

matter structure, including its subhalos, are captured at high resolution (e.g., Oñorbe et al. 2014). The zoom-in technique has facilitated an enormous range of studies within  $\Lambda$  cold dark matter (CDM) in both  $N$ -body and hydrodynamic contexts (e.g., see Vogelsberger et al. 2020 for a recent review), and recent zoom-in simulations have increasingly included nonstandard dark matter or cosmological physics (e.g., see Banerjee et al. 2022 and references therein). However, existing zoom-in simulations leave room for improvements in several crucial areas.

First, the majority of cosmological zoom-in simulations focus on host halos with masses similar to the Milky Way (e.g., Diemand et al. 2008; Springel et al. 2008; Garrison-Kimmel et al. 2014; Mao et al. 2015; Griffen et al. 2016; Sawala et al. 2016; Wetzel et al. 2016; Samuel et al. 2020; Poole-McKenzie et al. 2020) or galaxy clusters (e.g., Gao et al. 2012; Barnes et al. 2017; Cui et al. 2018). Although many exceptions exist (e.g., Wang et al. 2015; Dutton et al. 2016; Fiacconi et al. 2016, 2017; Despali et al. 2019; Richings et al. 2021), zoom-in suites at other mass scales typically include only a few distinct hosts, precluding analyses of their subhalo populations that capture host-to-host scatter (see Figure 1). Such analyses are crucial in order to characterize subhalo population statistics as a function of host halo mass, which are often simply extrapolated from Milky Way-mass hosts in semianalytic models (e.g., Dooley et al. 2017; Gilman et al. 2020a). Second, existing zoom-in simulations at different mass scales often vary in resolution (e.g., in terms of how many particles comprise the host halo) and are performed or analyzed with codes that differ in detail. This limits the feasibility of unified subhalo population analyses that simultaneously cover a wide range of host mass and robustly quantify host-to-host scatter at high resolution.

To bridge these gaps, we present Symphony, the first statistical compilation of cosmological zoom-in simulations of host halos with masses from  $10^{11}$ – $10^{15} M_{\odot}$ . Symphony includes 39 zoom-ins at the  $10^{11} M_{\odot}$  LMC-mass scale (the “LMC” suite), 45 zoom-ins at the  $10^{12} M_{\odot}$  Milky Way-mass scale (the “Milky Way” suite), 49 strong lens analog zoom-ins at the  $10^{13} M_{\odot}$  group-mass scale (the “Group” suite), 33 zoom-ins at the  $\approx 5 \times 10^{14}$  low-mass cluster scale (the “Low-mass Cluster,” or “L-Cluster” suite), and 96 zoom-ins at the  $10^{15} M_{\odot}$  cluster scale (the “Cluster” suite), for a total of 262 distinct simulations. Each suite is run with comparable, high resolution relative to its host halo mass, such that hosts consist of greater than  $10^6$  particles and subhalos are well resolved down to  $\approx 10^{-4}$  times the host mass, on average. Thus, Symphony enables precise, self-consistent measurements of subhalo population statistics and their host-to-host scatter over four decades of host halo mass.

From an observational perspective, Symphony’s new LMC and Group suites are relevant given recent advances in our ability to probe substructure in both regimes. Specifically, the combination of recent astrometric measurements (e.g. Gaia Collaboration et al. 2018) and photometric observations of Milky Way satellite galaxies (e.g., from the Dark Energy Survey; Abbott et al. 2018) indicate that several nearby ultra-

faint dwarf galaxies are satellites of the LMC (Kallivayalil et al. 2018; Patel et al. 2020); upcoming facilities are also expected to detect satellites of LMC-mass systems throughout the Local Volume (e.g., Mutlu-Pakdil et al. 2021). On the Group scale, high-resolution imaging (e.g., Nierenberg et al. 2020) has enabled measurements of dark matter substructure within early-type galaxy strong lenses (e.g., Végetti et al. 2012; Hezaveh et al. 2016; Hsueh et al. 2020; Gilman et al. 2020a,b). These developments underscore the need for high-resolution simulations of dark matter substructure at the corresponding host halo mass scales. Meanwhile, Symphony’s Milky Way-mass, Low-mass Cluster, and Cluster suites serve as benchmarks for analyses of subhalo evolution within the Milky Way, its analogs, and galaxy clusters throughout the universe. Zoom-ins from these suites have been resimulated with hydrodynamic or nonstandard dark matter physics (e.g., Martizzi et al. 2016; Nadler et al. 2020a, 2021a; Bhattacharyya et al. 2022; Mau et al. 2022), and the Symphony data release will facilitate further work along these lines.

From a theoretical perspective, Symphony’s large dynamic range is desirable for calibrating semianalytic structure formation models, which are usually constrained using large-volume cosmological simulations (e.g., Benson 2017) or Milky Way-mass zoom-ins (e.g., Pullen et al. 2014; Yang et al. 2020). Furthermore, empirical and semianalytic models of galaxy evolution have been applied to zoom-in simulations (e.g., Starkenburg et al. 2013; Lu et al. 2016; Newton et al. 2018; Wang et al. 2021; Chen et al. 2022; Kravtsov & Manwadkar 2022), and this technique can be extended to other host masses using Symphony. Here, we connect to both modeling approaches by comparing Symphony to predictions from the semianalytic structure formation model GALACTICUS (Benson 2012; Pullen et al. 2014) and by using the empirical galaxy-halo connection model UNIVERSEMACHINE (Behroozi et al. 2019; Wang et al. 2021) to predict star formation histories (SFHs) for Symphony halos and subhalos. By combining Symphony and UNIVERSEMACHINE, we capture the relation between dark matter accretion and SFHs for nearly all currently observable satellite galaxies in the universe, with the exception of the faintest known satellites in nearby clusters. In addition, Symphony’s new LMC suite provides robust predictions for the population statistics of halos below the threshold of galaxy formation (e.g., Nadler et al. 2020b; Munshi et al. 2021), which will be crucial to understand in order to search for deviations from  $\Lambda$ CDM on small scales using forthcoming datasets (e.g., see Bechtol et al. 2022 and references therein).

This paper is organized as follows. Section 2 describes the Symphony zoom-in simulations. Sections 3 and 4, respectively, study the properties of Symphony host halos and subhalo populations. Section 5 compares Symphony results to predictions from GALACTICUS. Section 6 applies UNIVERSEMACHINE to all Symphony simulations and presents the resulting central and satellite galaxy SFH predictions. Section 7 discusses science enabled by the Symphony suites and compilation. Section 8 concludes. Throughout, Sym-

phony results are presented using a color scheme of pink (LMC), blue (Milky Way), green (Group), gold (L-Cluster), and red (Cluster).

## 2. SYMPHONY SIMULATIONS

This section provides an overview of Symphony, including information about parent and zoom-in simulations, initial conditions, simulation parameters, zoom-in host halo selection criteria, and convergence properties.

### 2.1. Parent Simulations

Host halos for our LMC, Milky Way, and Group suites were selected from the same parent simulation, c125-1024, which has a side length of  $125 \text{ Mpc } h^{-1}$  with 1024 particles per side; we also use a higher-resolution version of this box, c125-2048, for certain tests (see Mao et al. 2015 for details). c125-1024 and its corresponding zoom-in suites were run with cosmological parameters of  $h = 0.7$ ,  $\Omega_m = 0.286$ ,  $\Omega_\Lambda = 0.714$ ,  $\sigma_8 = 0.82$ , and  $n_s = 0.96$  (Hinshaw et al. 2013). The Bryan & Norman (1998) virial overdensity in this cosmology corresponds to  $\Delta_{\text{vir}} \approx 99$  times the critical density of the universe at  $z = 0$ .

The L-Cluster hosts were selected from a parent simulation of side length  $1 \text{ Gpc } h^{-1}$  with 1024 particles per side, which we refer to as 1000-1024A (see Bhattacharyya et al. 2022 for details). This simulation uses cosmological parameters of  $h = 0.7$ ,  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ ,  $\sigma_8 = 0.85$ , and  $n_s = 0.96$ , corresponding to  $\Delta_{\text{vir}} \approx 101$  at  $z = 0$  (Bhattacharyya et al. 2022).

The Cluster hosts were selected from the Carmen simulation (McBride et al. 2009a), which has a side length of  $1 \text{ Gpc } h^{-1}$  with 1120 particles per side, hereafter referred to as 1000-1120B. This simulation uses cosmological parameters of  $h = 0.7$ ,  $\Omega_m = 0.25$ ,  $\Omega_\Lambda = 0.75$ ,  $\sigma_8 = 0.8$ , and  $n_s = 1$ , corresponding to  $\Delta_{\text{vir}} \approx 94$  at  $z = 0$  (Wu et al. 2013c,b). Thus, the L-Cluster and Cluster zoom-in simulations adopt different cosmological parameters compared to both each other and to the LMC, Milky Way, and Group suites. These differences have a minor impact on our results, but should be kept in mind when comparing suites in detail.

### 2.2. Zoom-in Simulations

For all Symphony suites, zoom-in initial conditions were generated using MUSIC (Hahn & Abel 2011), simulations were run with GADGET-2 (Springel 2005), and halo catalogs and merger trees were, respectively, generated using ROCKSTAR and CONSISTENT-TREES (Behroozi et al. 2013b,c). To promote open use of Symphony, halo catalogs and merger trees for our 262 fiducial-resolution zoom-ins are available at <http://web.stanford.edu/group/gfc/symphony>; furthermore, a subset of particle snapshots is publicly available.

Figure 1 compares the Symphony compilation to existing zoom-in suites, Figure 2 visualizes five host halos in each suite at  $z = 0$ , and Table 1 lists the numerical properties of our five suites. As shown in Figure 1, Symphony provides a unique combination of host statistics, resolution, and dynamic range compared to existing zoom-in suites (see Ap-

pendix A for references). We note that a handful of cosmological simulations (e.g., VSMDPL from the MultiDark suite and Uchuu; Klypin et al. 2016; Ishiyama et al. 2021) contain more high-resolution Group, L-Cluster, and/or Cluster-mass hosts than Symphony; we discuss the pros and cons of zoom-ins relative to such simulations in Section 7.

The environmental properties of Symphony’s LMC, Milky Way, and Group hosts differ from typical halos of these masses to varying degrees because of the isolation criteria used to select the host halos, which are described below for each suite. For example, relative to all halos in the corresponding mass ranges from the Erebos simulations (Diemer & Kravtsov 2014, 2015), we find that 3%, 87%, and 69% of objects satisfy our isolation criteria for the LMC, Milky Way, and Group suites, respectively. We plan to characterize the impact of Symphony hosts’ environments on their formation histories and subhalo populations in future work.

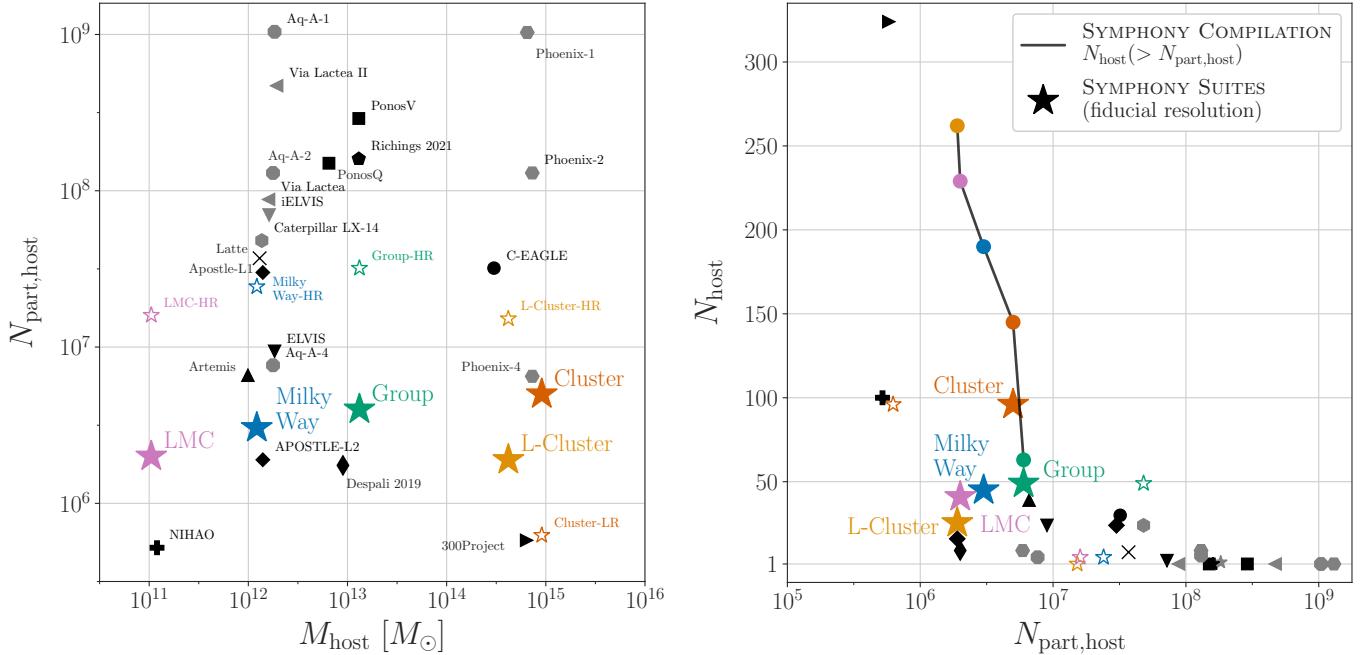
#### 2.2.1. LMC-mass Suite

Host halos for the LMC suite were chosen from a  $z = 0$  virial mass range of  $10^{11.02 \pm 0.05} M_\odot$  in c125-1024, which is comparable to LMC’s halo mass (Erkal et al. 2019; Shipp et al. 2021). Hosts in this mass range were selected subject to the constraint that no more massive halo is found within a radius of  $3 \text{ Mpc } h^{-1}$  in the parent box; thus, although we abbreviate these hosts as “LMC”s, they are similar to the actual LMC in terms of total mass but not environment. Just 3% of halos within this mass range meet this selection criterion. Zoom-in initial conditions for each system were generated with five refinement regions, yielding an equivalent of 16,384 particles per side for the most refined region. The highest-resolution region for each simulation corresponds to the Lagrangian volume containing particles within five times the virial radius of the host halo in the parent box at  $z = 0$ . The dark matter particle mass in the highest-resolution region is  $m_{\text{part}} = 5.0 \times 10^4 M_\odot$ , and the comoving Plummer-equivalent gravitational softening in this region is  $80 \text{ pc } h^{-1}$ , corresponding to 0.011 times the mean interparticle spacing.

Thirty-nine host halos were resimulated, resulting in a distribution of  $m_{\text{part}}/M_{\text{host}}$  with a median and standard deviation of  $(4.8 \pm 0.8) \times 10^{-7}$  and an  $\epsilon/R_{\text{vir},\text{host}}$  distribution of  $(9.3 \pm 0.5) \times 10^{-4}$ , where  $M_{\text{host}}$  and  $R_{\text{vir},\text{host}}$  denote the  $z = 0$  host halo virial mass and radius, respectively. Over the LMC suite, zoom-in host halo masses at  $z = 0$  differ from the target hosts in the parent box by  $1.1\% \pm 11\%$ , which represents the largest amount of scatter relative to the target host masses among our c125-1024 suites. We attribute this scatter to the fact that the target hosts in the parent box are less well resolved than the target hosts in any other suite.

#### 2.2.2. Milky Way-mass Suite

The halos in the Milky Way suite were first presented in Mao et al. (2015); we refer the reader to this work for a complete description of the original suite. As noted in Mao et al. (2022), the Mao et al. (2015) simulations used  $n_s = 1.0$  to generate initial conditions, while the parent box



**Figure 1.** Left panel: median number of particles per host halo versus the median host mass for Symphony (stars) and a representative sample of existing cosmological zoom-in simulation suites (markers). The five Symphony suites presented in this work are: LMC (magenta), Milky Way (blue), Group (green), L-Cluster (gold), and Cluster (red); Symphony resimulation suites are shown as small unfilled stars. Existing zoom-in suites with (without) hydrodynamic resimulations are shown in black (gray). Median host masses and particle counts are taken from each existing suite without converting to a unified virial mass definition. Host-to-host scatter for suites with multiple simulations is not shown. See Appendix A for references. Right panel: number of host halos per zoom-in suite versus the median number of particles per host. Only target host halos are included in the y-axis counts; however, many zoom-ins (including those in Symphony) often contain halos of similar masses in addition to the host in the highest-resolution region. The black line shows the cumulative number of zoom-in simulations combined over all five fiducial-resolution zoom-in suites in the Symphony compilation, i.e.,  $N_{\text{host}}(> N_{\text{part},\text{host}})$ . Note that the Group point along the cumulative line is offset vertically from the Group suite for visual clarity.

used  $n_s = 0.96$ . We therefore resimulate this suite using  $n_s = 0.96$ , and we exclusively present these new results here.<sup>1</sup>

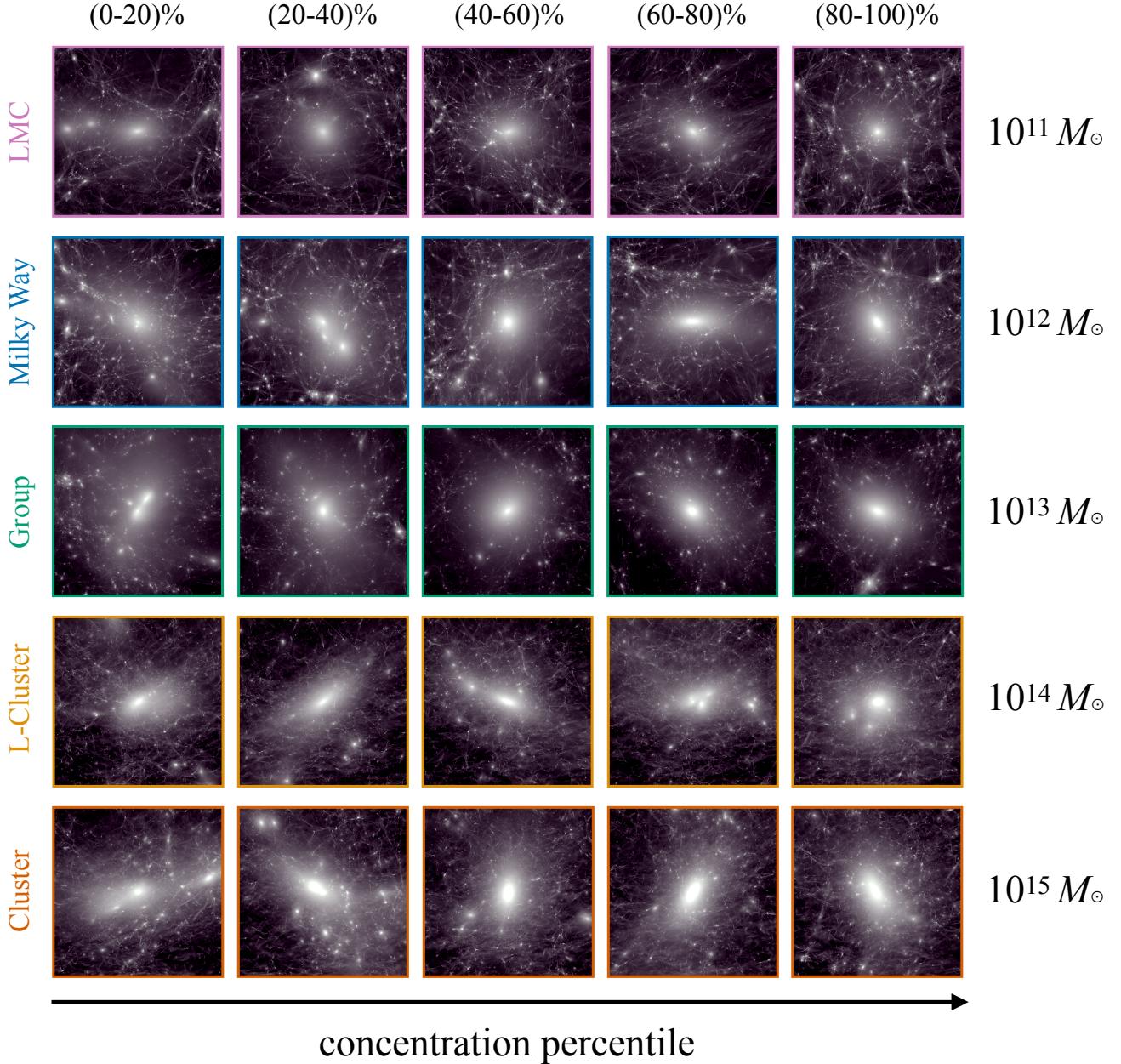
Host halos for the Milky Way suite were chosen from a  $z=0$  virial mass range of  $10^{12.09 \pm 0.02} M_{\odot}$  in c125-1024, and thus fall within observational estimates of the Milky Way's dark matter halo mass (e.g., see Bland-Hawthorn & Gerhard 2016; Callingham et al. 2019 and references therein). Hosts in this mass range were selected such that they were not within  $4R_{\text{vir,host}}$  of any more massive halo in the parent box. Eighty-seven percent of halos in this mass range in the parent box meet this isolation criterion, and as shown in Mao et al. (2015), this sample has a somewhat wider concentration distribution than the mass-selected sample. Thus, although we abbreviate these hosts “Milky Ways,” they are similar to the actual Milky Way in terms of total mass but not necessarily in terms of environment. In fact, the Milky Way's large-scale environment is unusual, and constrained simulations would be required to match it in detail (e.g., see McCall 2014; Carlesi et al. 2016; Neuzil et al. 2020; McAlpine et al.

2022 and references therein). In addition, these hosts do not necessarily satisfy constraints on the Milky Way's formation history, such as the recent infall of the LMC and an early *Gaia*-Enceladus-like merger. D. Buch et al. (2023, in preparation) presents a zoom-in suite of Milky Way-mass hosts selected from c125-1024 and using the same resolution that satisfy additional “Milky Way-like” constraints.

Zoom-in resimulation initial conditions for each Milky Way-mass host were generated with four refinement regions, yielding an equivalent of 8192 particles per side in the most refined region. The highest-resolution region for each simulation corresponds to the Lagrangian volume containing particles within  $10R_{\text{vir,host}}$  of the host halo in the parent box at  $z=0$ . The dark matter particle mass in the highest-resolution regions is  $m_{\text{part}} = 4.0 \times 10^5 M_{\odot}$  and the comoving Plummer-equivalent gravitational softening is  $170 \text{ pc } h^{-1}$ , corresponding to 0.011 times the mean interparticle spacing.

Forty-five host halos were resimulated, resulting in a distribution of  $m_{\text{part}}/M_{\text{host}}$  with a median and standard deviation of  $(3.2 \pm 0.2) \times 10^{-7}$  and an  $\epsilon/R_{\text{vir,host}}$  distribution of  $(8.7 \pm 0.1) \times 10^{-4}$ . Over the Milky Way suite, zoom-in host halo masses at  $z=0$  differ from the target hosts in the parent box by  $1.0 \pm 3\%$ .

<sup>1</sup> This change to  $n_s$  alleviates the discrepancies in subhalo merger timing relative to the parent box identified in Wang et al. (2021).



**Figure 2.** Visualizations of the projected dark matter density at  $z = 0$ , centered on five host halos from each of our zoom-in suites. The host halo mass scale of each zoom-in suite increases from top to bottom. Within each suite, columns show hosts randomly selected from concentration quintiles, such that concentration increases from left to right. Each visualization spans its host's virial radius in projection. Visualizations were created using the phase-space tessellation method described in Kaehler (2017) and Kaehler (2018).

### 2.2.3. Group Suite

Host halos for the Group suite were chosen from a virial mass range of  $10^{12.86 \pm 0.10} M_{\odot}$  at  $z = 0.5$  in c125-1024. These hosts are selected in a narrow mass range at  $z = 0.5$  to more directly relate them to observed strong gravitational lenses studied in recent substructure analyses (e.g., Gavazzi et al. 2007; Auger et al. 2010; Gilman et al. 2020a). Hosts in this mass range were selected subject to the constraint that no halo more massive than  $10^{13} M_{\odot} h^{-1}$  is found within a ra-

dius of  $3 \text{ Mpc } h^{-1}$  at  $z = 0$  in the parent box. Sixty-nine percent of halos in this mass range meet this selection criterion. Zoom-in resimulation initial conditions for each system were generated with three refinement regions, yielding an equivalent of 4096 particles per side for the most refined region. The highest-resolution region for each simulation corresponds to the Lagrangian volume containing particles within  $10R_{\text{vir,host}}$  of the host halo in the parent box at  $z = 0$ . The dark matter particle mass in the highest-resolution re-

**Table 1.** Properties of the Five Symphony Simulation Suites.

Zoom-in Suite	Parent Simulation	$N_{\text{sims}}$	$M_{\text{host}} (M_{\odot})$	$M_{\text{sub,min}} (M_{\odot})$	$M_{\text{sub/host,min}}$	$m_{\text{part}} (M_{\odot})$	$\epsilon (\text{pc } h^{-1})$	Reference
LMC-mass	c125-1024	39	$10^{11.02 \pm 0.05}$	$1.5 \times 10^7$	$2.7 \times 10^{-4}$	$5.0 \times 10^4$	80	This work
Milky Way-mass	c125-1024	45	$10^{12.09 \pm 0.02}$	$1.2 \times 10^8$	$1.1 \times 10^{-4}$	$4.0 \times 10^5$	170	Mao et al. (2015)
	Group	49	$10^{13.12 \pm 0.11}$	$9.9 \times 10^8$	$1.1 \times 10^{-4}$	$3.3 \times 10^6$	360	
Low-mass Cluster	1000-1024a	33	$10^{14.62 \pm 0.11}$	$6.5 \times 10^{10}$	$2.4 \times 10^{-4}$	$2.2 \times 10^8$	1200	Bhattacharyya et al. (2022)
	Cluster	96	$10^{14.96 \pm 0.03}$	$5.3 \times 10^{10}$	$6.7 \times 10^{-5}$	$1.8 \times 10^8$	3295	

NOTE—The first column lists the name of each suite, the second column lists the name of each parent simulation, the third column lists the number of simulations per suite used in our analysis, the fourth column lists the median and standard deviation of the target host halo virial mass distribution, the fifth and sixth columns list the minimum well-resolved subhalo virial mass and sub-to-host halo mass ratio at  $z = 0$ , the seventh column lists the dark matter particle mass in the highest-resolution zoom-in region, the eighth column lists the comoving Plummer-equivalent force softening scale in the same region, and the ninth column lists the original reference for each suite. Host halo properties are listed at  $z = 0$  for the Group suite; at  $z = 0.5$ , when these hosts were selected from their parent simulation, the host mass range is  $10^{12.86 \pm 0.10} M_{\odot}$ .

gion is  $m_{\text{part}} = 3.3 \times 10^6 M_{\odot}$  and the comoving Plummer-equivalent gravitational softening is  $360 \text{ pc } h^{-1}$ , or 0.012 times the mean interparticle spacing.

Forty-nine host halos were resimulated, resulting in a distribution of  $m_{\text{part}}/M_{\text{host}}$  with a median and standard deviation of  $(2.4 \pm 0.5) \times 10^{-7}$  and an  $\epsilon/R_{\text{vir,host}}$  distribution of  $(8.3 \pm 0.6) \times 10^{-4}$ ; these hosts are resolved with slightly higher particle counts at  $z = 0.5$ , when they were selected from the parent box. Over the Group suite, zoom-in host halo masses at  $z = 0$  differ from the target hosts in the parent box by  $0.0 \pm 3\%$ .

#### 2.2.4. Low-mass Cluster Suite

The L-Cluster suite was first presented in Bhattacharyya et al. (2022), along with a corresponding suite of self-interacting dark matter zoom-in simulations that are not included in this work. We refer the reader to Bhattacharyya et al. (2022) for a complete description of this suite, and we summarize its properties here.

Host halos for the L-Cluster suite were chosen from a  $z = 0$  virial mass range of  $10^{14.62 \pm 0.11} M_{\odot}$  in 1000-1024a. Fifty hosts were selected randomly from all halos in the parent box in this range with no additional isolation criteria. Zoom-in resimulation initial conditions for each system were generated with four refinement regions, yielding an equivalent of 8192 particles per side for the most refined region. The highest-resolution region for each simulation corresponds to the Lagrangian volume containing particles within  $10 \text{ Mpc } h^{-1}$  of the host halo in the parent box at  $z = 0$ . The dark matter particle mass in the highest-resolution region is  $m_{\text{part}} = 2.2 \times 10^8 M_{\odot}$  and the comoving Plummer-equivalent gravitational softening is  $1200 \text{ pc } h^{-1}$ , corresponding to 0.010 times the mean interparticle spacing.

Thirty-three host halos were resimulated, resulting in a distribution of  $m_{\text{part}}/M_{\text{host}}$  with a median and standard deviation of  $(5.9 \pm 1.1) \times 10^{-7}$  and an  $\epsilon/R_{\text{vir,host}}$  distribution of  $(8.8 \pm 0.7) \times 10^{-4}$ . The L-Cluster zoom-in host mass distribution at  $z = 0$  is consistent with that in the parent box.

#### 2.2.5. Cluster Suite

The Cluster suite (or “Rhapsody”) was first presented in Wu et al. (2013c,b); we refer the reader to this work for a complete description of this suite, and we summarize its properties here.

Host halos for the Cluster suite were chosen from a  $z = 0$  virial mass range of  $10^{14.96 \pm 0.03}$  in 1000-1120B. Ninety-six hosts were selected randomly from all halos in the parent box in this range with no additional isolation criteria applied. Zoom-in resimulation initial conditions for each system were generated with four refinement regions, yielding an equivalent of 8192 particles per side in the most refined region. The highest-resolution region for each simulation is chosen to be 40% larger than the Lagrangian volume containing all friends-of-friends particles of the host halo in the parent box at  $z = 0$ . The dark matter particle mass in the highest-resolution region is  $m_{\text{part}} = 1.8 \times 10^8 M_{\odot}$  and the comoving Plummer-equivalent gravitational softening is  $3250 \text{ pc } h^{-1}$ , corresponding to 0.027 times the mean interparticle spacing.

Ninety-six host halos were resimulated and presented in Wu et al. (2013c,b); we analyze the same systems here. These hosts have a distribution of  $m_{\text{part}}/M_{\text{host}}$  with a median and standard deviation of  $(2.0 \pm 0.1) \times 10^{-7}$  and an  $\epsilon/R_{\text{vir,host}}$  distribution of  $(18 \pm 0.3) \times 10^{-4}$ . The Cluster zoom-in host mass distribution at  $z = 0$  is consistent with that in the parent box.

#### 2.3. Convergence Limits

In Appendix B, we use resimulations of Symphony zoom-ins to test the impact of numerical parameters on subhalo population statistics. These tests include five high-resolution resimulations for each of the LMC, Milky Way, and Group suites, one high-resolution resimulation from the L-Cluster suite (Bhattacharyya et al. 2022), 96 low-resolution resimulations from the Cluster suite (Wu et al. 2013c,b), and three Milky Way-mass hosts that were resimulated using a wide range of force softening scales and time-stepping criteria. From these tests, we conclude that:

- Subhalo mass functions (SHMFs) evaluated using virial mass at  $z = 0$ ,  $M_{\text{sub}}$ , are converged (i.e., differ by less than 10% at varying resolution) for subhalos with greater than 300 particles at  $z = 0$  in all Symphony suites.<sup>2</sup>
- SHMFs evaluated using peak virial mass,  $M_{\text{peak,sub}}$ , are converged for subhalos that are far from the centers of their hosts at  $z = 0$  and have greater than  $\approx 1000$  particles when  $M_{\text{peak,sub}}$  is achieved, and become less well converged with decreasing distance. This is consistent with previous studies (van den Bosch et al. 2018; van den Bosch & Ogiya 2018; Mansfield & Avestruz 2021), which demonstrate the difficulty of measuring converged subhalo radial distributions at a given  $M_{\text{peak,sub}}$ .
- Subhalo maximum circular velocity functions are highly dependent on force softening scale, as shown in previous convergence studies (e.g., Ludlow et al. 2019; Mansfield & Avestruz 2021), but our choice of  $\epsilon$  for each suite avoids both suppression of inner densities due to excessive softening and runaway time integration errors due to insufficient softening.
- The time-stepping criterion used in our simulations is well converged, consistent with the results of previous studies (e.g., Ludlow et al. 2019).

Thus, SHMFs for each Symphony suite are converged for subhalos with virial mass at  $z = 0$  of

$$M_{\text{sub}} > 300m_{\text{part}}, \quad (1)$$

corresponding to  $1.5 \times 10^7 M_{\odot}$ ,  $1.2 \times 10^8 M_{\odot}$ ,  $9.9 \times 10^8 M_{\odot}$ , and  $6.5 \times 10^{10} M_{\odot}$ , and  $5.3 \times 10^{10} M_{\odot}$  for the LMC, Milky Way, Group, L-Cluster, and Cluster suites, respectively.

In turn, sub-to-host halo mass ratios within each suite are converged above a conservative limit of  $300m_{\text{part}}/\text{min}(M_{\text{host}})$ , corresponding to  $2.7 \times 10^{-4}$ ,  $1.1 \times 10^{-4}$ ,  $1.1 \times 10^{-4}$ ,  $2.4 \times 10^{-4}$ , and  $6.7 \times 10^{-5}$  for the LMC, Milky Way, Group, L-Cluster, and Cluster suites, respectively. Analyses that combine all Symphony simulations across all suites should therefore focus on subhalos with

$$\tilde{M}_{\text{sub}} \equiv M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}. \quad (2)$$

This limit is driven by the lowest-mass host in our LMC suite and can be decreased by a factor of a few for analyses using subsets of Symphony suites and/or simulations. In particular, analyses that combine a subset of Symphony zoom-ins from one or multiple suites can be performed using the maximum  $M_{\text{sub}}/M_{\text{host}}$  limit among the subset determined by Equation 1. For analyses of individual zoom-ins, this reduces to an  $M_{\text{sub}} > 300m_{\text{part}}$  cut.

<sup>2</sup> To calculate subhalo properties, ROCKSTAR only uses particles that survive one unbinding pass after phase-space groups are determined.

**Table 2.** Properties of Symphony Host Halos.

Zoom-in Suite	$c_{\text{host}}$	$a_{1/2,\text{host}}$	$N_{\text{sub,resolved}}$	$\tilde{N}_{\text{sub,resolved}}$
LMC-mass	$12.2 \pm 2.7$	$0.46 \pm 0.11$	$57 \pm 17$	$39 \pm 9$
Milky Way–mass	$10.8 \pm 3.5$	$0.46 \pm 0.12$	$83 \pm 18$	$32 \pm 9$
	$9.0 \pm 2.1$	$0.50 \pm 0.11$	$136 \pm 55$	$37 \pm 8$
Low-mass Cluster	$5.0 \pm 1.7$	$0.70 \pm 0.11$	$104 \pm 42$	$57 \pm 11$
	$5.3 \pm 1.3$	$0.64 \pm 0.10$	$210 \pm 31$	$56 \pm 11$

NOTE—The first column lists the name of each suite, and the second, third, fourth, and fifth columns list the median and  $1\sigma$  host-to-host scatter of the zoom-in host halo’s concentration, half-mass scale factor, and the abundance of surviving subhalos at  $z = 0$  above resolution limits of  $M_{\text{sub}} > 300m_{\text{part}}$  and  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$ , respectively.

### 3. SYMPHONY HOST HALOS

We now present the properties of Symphony host halos (Section 3.1), including concentrations and formation times (Section 3.2), correlations between secondary host halo properties and the abundance of surviving subhalos at  $z = 0$  (Section 3.3), mass accretion histories (Section 3.4), and density profiles (Section 3.5).

We defer from providing updated fits to standard halo relations (e.g., the mass–concentration relation) due to the diverse environmental properties of our zoom-in suites. A more detailed comparison of our results to measurements from existing cosmological and zoom-in simulations, and an exploration of additional host halo properties (e.g., host halo shapes) are beyond the scope of this work.

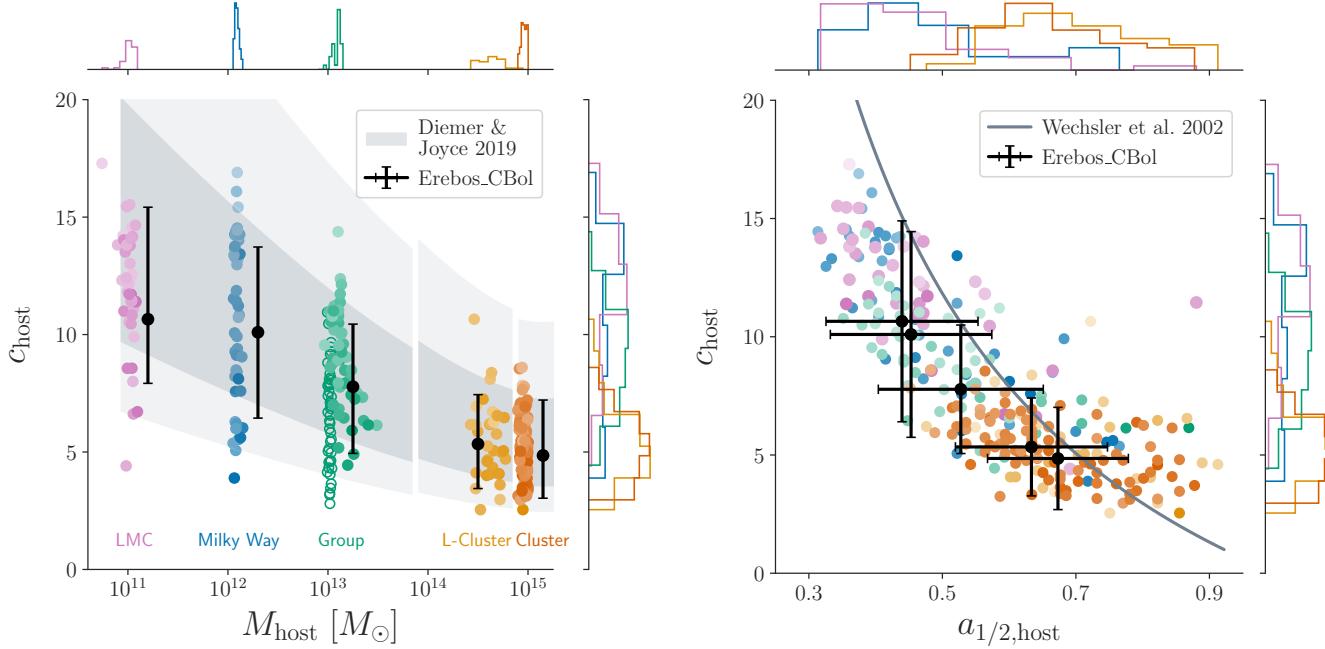
#### 3.1. Overview

Table 2 lists the median and standard deviation of the concentration and half-mass formation time distribution for hosts in each Symphony suite, along with subhalo abundance above our absolute (Equation 1) and normalized (Equation 2) convergence limits.

Symphony host halos are typically resolved with several million particles, the large majority of which are the highest-resolution particle type in each suite. We verify this by measuring the distance from the center of the host within which the highest-resolution particles contribute  $> 90\%$  of the total mass. This “contamination radius” is always several times (and often about an order of magnitude) larger than the virial radius of the host halo, consistent with previous findings for the original Milky Way suite (Wang et al. 2021).

The LMC, Milky Way, and Group hosts are resolved with  $\gtrsim 10^5$  high-resolution particles at  $z \lesssim 3$ , while the L-Cluster and Cluster hosts are resolved with comparable particle counts at  $z \lesssim 1.5$ .<sup>3</sup> Furthermore, hosts in all suites are typically resolved with  $\gtrsim 300$  particles at  $z \lesssim 7$ . These properties

<sup>3</sup> Note that the LMC, Milky Way, and Group halo catalogs extend to  $z \approx 20$ , while the L-Cluster and Cluster halo catalogs extend to  $z \approx 12$ .



**Figure 3.** Left panel: joint and marginal distributions of  $z = 0$  halo virial mass and NFW concentration for hosts in our zoom-in simulation suites. Each point represents a unique zoom-in simulation with particle mass of  $m_{\text{part}} \approx 3 \times 10^{-7} M_{\text{host}}$  and is shaded according to the abundance of subhalos with  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$ ; darker colors indicate hosts with higher subhalo abundance relative to the rest of the hosts in each suite. Unfilled Group points show host mass and concentration values at  $z = 0.5$ , when these halos were mass-selected in their parent simulation. The three dark (light) gray bands show the  $1\sigma$  ( $2\sigma$ ) uncertainty on the median mass–concentration relation from Diemer & Joyce (2019) for our LMC/Milky Way/Group, L-Cluster, and Cluster cosmological parameters, respectively. Black points and error bars show medians and  $1\sigma$  scatter from the Erebos simulations (Diemer & Kravtsov 2014, 2015) for halos in each suite’s mass range that satisfy the corresponding isolation criteria; these measurements differ from the Diemer & Joyce (2019) relation for the LMC, Milky Way, and Group suites due to these cuts. Right panel: joint and marginal distributions of host half-mass scale factor and NFW concentration. The gray line shows the prediction for the  $c_{\text{host}}(a_{1/2,\text{host}})$  relation from Wechsler et al. (2002), for reference; this model is calibrated using  $\sigma_8 = 1.0$ , which contributes to the normalization offset, and is not re-fit here. Manual inspection of individual mass accretion histories (MAHs) indicates that points with large  $a_{1/2,\text{host}}$  for a given  $c_{\text{host}}$  recently experienced a major merger.

enable the robust measurements of hosts’ formation histories and dark matter structure presented below.

### 3.2. Host Halo Concentrations and Formation Times

Figure 3 shows the host halo mass, concentration, and formation time distributions for all Symphony suites. In particular, the left panel of Figure 3 shows our hosts’ mass–concentration ( $M_{\text{host}} - c_{\text{host}}$ ) relation, and the right panel shows the relationship between  $c_{\text{host}}$  and  $a_{1/2,\text{host}}$  for our hosts. Here,  $c_{\text{host}} \equiv R_{\text{vir},\text{host}}/R_{\text{s},\text{host}}$  is the Navarro–Frenk–White (NFW) concentration (Navarro et al. 1997), and  $a_{1/2,\text{host}}$  is the scale factor at the time when a host first reached half its  $z = 0$  mass. For the purposes of this subsection, we assume that Symphony hosts are adequately described by NFW profiles; we explore this assumption in Section 3.5. Both the  $M_{\text{host}} - c_{\text{host}}$  and  $a_{1/2,\text{host}} - c_{\text{host}}$  relations are well studied (e.g. Navarro et al. 1997; Wechsler et al. 2002; van den Bosch et al. 2014; Ludlow et al. 2016) and our results are qualitatively consistent with earlier findings that at  $z = 0$ ,  $c_{\text{host}}$  tends to decrease with increasing halo mass and increasing  $a_{1/2,\text{host}}$ .

We quantitatively compare Symphony results against the Diemer & Joyce (2019)  $M_{\text{host}} - c_{\text{host}}$  relation in the left panel

of Figure 3. To account for Symphony hosts’ environmental properties, we also compare against results from the Erebos\_CBoL cosmological simulation suite (Diemer & Kravtsov 2014, 2015) using identical mass and isolation cuts to our zoom-in suites. Specifically, we compare each suite against the largest Erebos box that resolves halos in the corresponding host mass range.<sup>4</sup> Note that the Diemer & Joyce (2019)  $M_{\text{host}} - c_{\text{host}}$  relation is calibrated using the Erebos suite, meaning that the difference between the Diemer & Joyce (2019) and Erebos results in Figure 3 is due to environmental selection effects.

Symphony hosts display lower median concentrations and smaller scatter than the sample of halos from cosmological simulations used to measure the mass–concentration relation

<sup>4</sup> Thus, we compare the LMC, Milky Way, Group, L-Cluster, and Cluster suites to Erebos\_CBoL63, L125, L250, L1000, and L2000, respectively, based on the “0%” convergence criteria from Mansfield & Avestruz (2021). Note that Erebos\_CBoL adopts a different cosmology compared to any of our suites, with  $\Omega_m = 0.27$ ,  $h = 0.7$ , and  $\sigma_8 = 0.82$ ; however, its cosmology is the closest to our simulations out of the suites in Mansfield & Avestruz (2021) that resolve LMC-mass halos.

in Diemer & Joyce (2019), particularly for LMC-mass halos (see Diemer & Joyce 2019 for a comparison to other mass–concentration relation models and simulation measurements). This follows because our isolation criteria require that Symphony’s LMC, Milky Way, and Group hosts are not close to more massive halos. Thus, our host samples should not include “splashback” subhalos, which have previously passed through the virial radius of a larger host and form the entire high-concentration tail of the cosmological mass–concentration relation (Mansfield & Kravtsov 2020). Furthermore, our isolation cuts also remove many nonsplashback halos whose growth has slowed due to tidal truncation or the high velocity dispersion of dark matter particles in dense environments. The trend toward lower concentration that we report is expected for zoom-in simulations in general: accurately simulating splashback halos at high resolution requires simulating their more massive hosts at identical mass resolution and thus much higher particle counts, which is often impractical.

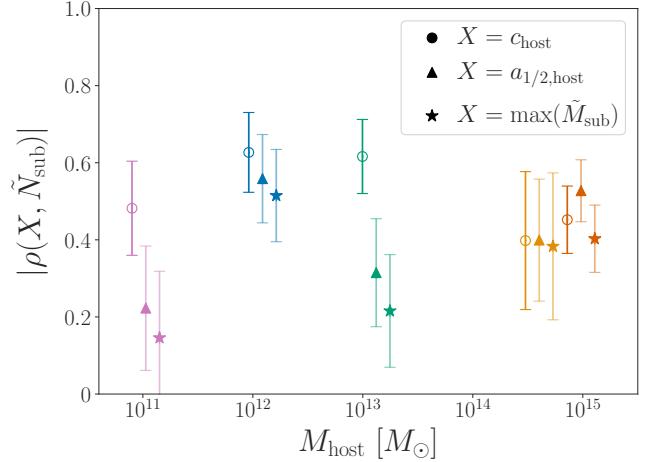
With appropriate isolation criteria imposed, Symphony’s LMC, Milky Way, Group, and L-Cluster concentration distributions are consistent with Erebus halos: two-sample Kolmogorov–Smirnov (K-S) tests yield  $p > 0.01$  in each case. Meanwhile, the Cluster distributions are statistically distinguishable ( $p < 10^{-5}$ ) because Symphony lacks the high-concentration tail measured in Erebus. No isolation criteria were imposed on our Cluster hosts; thus, selection effects cannot explain this difference. However, we note that our Cluster hosts were simulated with a cosmology that is different from the Erebus\_CBo suite, and that this is a plausible explanation for the difference.

For all suites, Symphony displays  $\approx 30\%$  smaller scatter in concentration relative to Erebus, with typical values of  $\sigma_{\log c} \approx 0.12$  dex rather than  $\sigma_{\log c} \approx 0.16$  dex. This discrepancy can be understood as a result of Erebus’ lower resolution: the Erebus halos we match to each Symphony suite typically have  $\approx 1 \times 10^3$ – $4 \times 10^3$  particles. Assuming that any additional scatter relative to Symphony results from finite-resolution effects, Benson et al. (2019) predicted that Symphony halos resolved with typical Erebus particle counts will display  $\sigma_{\log c} \approx 0.18$  dex, which is similar to the scatter Diemer & Joyce (2019) reported. Explicit resolution tests on Erebus halos would be needed to confirm this explanation; if it holds, our results imply that  $\sigma_{\log c} \approx 0.12$  dex is closer to a converged estimate of the concentration scatter in a sample without environmental constraints.

Finally, we note that recent major mergers may also play a role in setting the concentration scatter (e.g., Wang et al. 2020b), which is hinted at by the increased concentration in  $c_{\text{host}}$  at large  $a_{1/2,\text{host}}$  in Figure 3. We discuss the mass–concentration relation scatter further when comparing to GALACTICUS predictions in Section 5.2.

### 3.3. Correlations between Secondary Host Halo Properties and Subhalo Abundance

Next, we study the relationship between hosts’ secondary properties (i.e., properties beyond host halo mass) and the



**Figure 4.** Spearman correlation coefficients between subhalo abundance and concentration (circles), half-mass scale factor (triangles), and the sub-to-host halo mass ratio of the largest subhalo (stars) for hosts in each Symphony suite. The unfilled circles indicate that we show the absolute value of the anticorrelation with host concentration; the other correlation coefficients are positive. Error bars indicate  $1\sigma$  jackknife uncertainties. Note that the horizontal error bars are smaller than the markers, and that the three error bars for each suite are offset horizontally for visual clarity. Subhalo abundance is measured above our conservative sub-to-host halo convergence limit of  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$  for all suites.

abundance of surviving subhalos at  $z = 0$ . Figure 4 shows that the abundance of subhalos above our minimum sub-to-host halo mass ratio (Equation 2),

$$\tilde{N}_{\text{sub}} \equiv N_{\text{sub}}(\tilde{M}_{\text{sub}} > 2.7 \times 10^{-4}), \quad (3)$$

anticorrelates with host concentration at (roughly) fixed host mass within each suite. These anticorrelations between  $\tilde{N}_{\text{sub}}$  and  $c_{\text{host}}$  are significant (Spearman- $\rho$  tests yield  $p < 0.01$  within each suite) and, with the exception of the L-Cluster suite, are much stronger than the residual correlations with host mass. The strength of these anticorrelations peaks at the Milky Way host halo mass scale, with  $\rho = -0.48, -0.63, -0.62, -0.41$ , and  $-0.45$  for the LMC, Milky Way, Group, L-Cluster, and Cluster suites, respectively. We caution that this trend with host mass is only significant at the  $1\sigma$  level given the jackknife uncertainties on the measurement for each suite (see Figure 4). These findings are consistent with previous results (e.g., Zentner et al. 2005)—including several based on the original Milky Way suite (Mao et al. 2015; Fielder et al. 2019; Nadler et al. 2021b)—and extend them to a wider range of host halo masses at a fixed, high-resolution sub-to-host halo mass ratio.

For the LMC and Group suites, subhalo abundance anticorrelates more strongly and significantly with host concentration than it correlates with host formation time, which we parameterize by  $a_{1/2,\text{host}}$ . Meanwhile, the correlations between subhalo abundance and these two host properties are consistent with one another for the remaining suites. We note that

Cluster subhalo abundances are slightly more sensitive to formation time than host concentration. This may be due to the prevalence of late major mergers in these systems (see the right panel of Figure 3) and associated increases in subhalo abundance (e.g., D’Souza & Bell 2021).

Several factors likely contribute to the dependence of these correlations on host halo mass. First, subhalos of early-forming hosts accrete earlier and are stripped for longer than subhalos of late-forming hosts, on average. This leads to a positive correlation between  $N_{\text{sub}}$  and  $a_{1/2,\text{host}}$ ; in turn,  $a_{1/2,\text{host}}$  anticorrelates with  $c_{\text{host}}$ . Second, the tidal radius of a subhalo orbiting a host becomes smaller as the enclosed mass within the subhalo’s orbit,  $M_{\text{host}}(< r)$ , becomes larger and as the slope of the host’s mass profile,  $d \ln M_{\text{host}} / d \ln r$  becomes steeper (e.g., van den Bosch et al. 2018). For a fixed orbital radius and host mass, higher-concentration hosts have larger enclosed masses and steeper mass profiles than lower-concentration hosts. Thus, subhalos of higher-concentration hosts have smaller tidal radii and higher mass-loss rates on average, causing an anticorrelation between  $N_{\text{sub}}$  and  $c_{\text{host}}$ ; in turn,  $c_{\text{host}}$  anticorrelates with  $a_{1/2,\text{host}}$ . Because  $N_{\text{sub}}$  anticorrelates more strongly with  $c_{\text{host}}$  than it correlates with  $a_{1/2,\text{host}}$  for our LMC, Milky Way, and Group suites, the mechanism related to the host’s mass profile may be more important than the infall time effect for these host halo masses. On the other hand,  $c_{\text{host}}$  might simply trace accretion histories better than single-epoch measurements like  $a_{1/2,\text{host}}$  (e.g., Chue et al. 2018); a detailed follow-up study that considers hosts’ entire accretion histories would be required to disentangle these effects further.

Figure 4 also shows that the mass of each host’s largest subhalo (measured in terms of the maximum sub-to-host halo mass ratio in each zoom-in) correlates with subhalo abundance in a similar manner to formation time. Correlations between subhalo abundance and the properties of the largest surviving subhalo are particularly relevant because observations of Milky Way analogs indicate that the luminosity of the brightest surviving satellite (which is related to the “magnitude gap”) correlates more strongly with satellite abundance than the central’s stellar mass (Geha et al. 2017; Mao et al. 2021). We intend to study this effect by applying galaxy-halo connection models to Symphony in future work.

### 3.4. Mass Accretion Histories

Figure 5 shows the normalized and un-normalized mean MAHs for host halos in each Symphony suite. As suggested by the half-mass formation times in Figure 3 and demonstrated by previous studies (e.g., Wechsler et al. 2002), the normalized MAHs demonstrate that lower-mass hosts build up their mass earlier, on average. Our fiducial-resolution cut conservatively requires that halos have more than 300 particles at any redshift: more than 95% of hosts meet this cut at  $z < 11.8$  (LMC),  $z < 8.7$  (Milky Way),  $z < 8.5$  (Group),  $z < 7.4$  (L-Cluster), and  $z < 6.7$  (Cluster).

The dependence of accretion history on host mass is strongest for the transition from the Group to L-Cluster suites, and is not statistically significant between the L-

Cluster and Cluster suites. To quantify this, we jackknife resample hosts within each suite and compare the ratios of the jackknifed MAHs among suites. The LMC vs. Milky Way and Milky Way vs. Group MAHs are distinguishable at  $1\sigma$  for  $z \gtrsim 2$  given these uncertainties, and are consistent at later times. Meanwhile, the L-Cluster and Cluster MAHs are statistically consistent at all redshifts we resolve. Note that our Group, Milky Way, and particularly LMC hosts occupy underdense environments relative to typical halos of these masses due to our isolation criteria (see Section 2.2), which biases these hosts toward later formation times than average (e.g., see Mansfield & Kravtsov 2020 for a review). Thus, the dependence of formation time on host halo mass without selection effects would be stronger than we report, although comparisons to MAHs in c125-1024 indicate that this systematic bias is much smaller than the host-to-host scatter in formation times within Symphony suites.

Because of the narrow mass ranges from which our hosts were selected in their parent simulations, the un-normalized MAHs display  $\lesssim 0.1$  dex host-to-host scatter near  $z = 0$  for all suites (see Table 1 for the host halo mass distributions). Note that the Group suite displays its smallest host-to-host scatter near  $z = 0.5$ , corresponding to the redshift at which these hosts were mass-selected in their parent box.

MAHs have been studied extensively using both empirical models calibrated to cosmological simulations (e.g., Wechsler et al. 2002; McBride et al. 2009b; Wu et al. 2013c; Hearn et al. 2021) and (semi)analytic prescriptions (e.g., van den Bosch 2002; van den Bosch et al. 2014; Correa et al. 2015). We find that the one-parameter exponential model introduced in Wechsler et al. (2002),

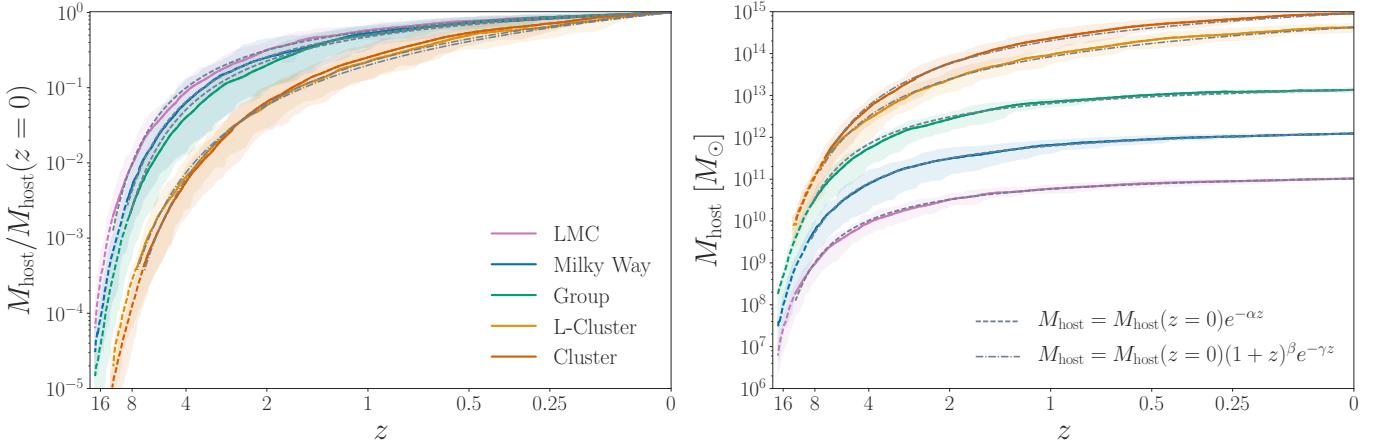
$$M_{\text{host}}(z) = M_{\text{host}}(z=0)e^{-\alpha z}, \quad (4)$$

fits the mean LMC, Milky Way, and Group MAHs well, with exponents of  $\alpha = 0.58$ ,  $0.68$ , and  $0.78$ , respectively. This model cannot accurately fit the L-Cluster and Cluster MAHs over the entire redshift range, consistent with the results of Wu et al. (2013c) using the Cluster suite. Instead, the two-parameter exponential plus power-law model from Tasitsiomi et al. (2004) and adopted by McBride et al. (2009b),

$$M_{\text{host}}(z) = M_{\text{host}}(z=0)(1+z)^{\beta} e^{-\gamma z}, \quad (5)$$

fits the L-Cluster and Cluster MAHs well, with best-fit parameters of  $(\beta, \gamma) = (-1.34, 0.69)$ , and  $(-0.81, 0.94)$ , respectively.<sup>5</sup> The GALACTICUS model we compare to in Section 5 predicts MAHs that agree well with the full Symphony distributions, indicating that semianalytic models calibrated to simulations can capture hosts’ MAH distributions over several decades of  $z = 0$  halo mass.

<sup>5</sup> We find that the one-parameter exponential plus power-law model from Wu et al. (2013c) does not fit the mean L-Cluster and Cluster MAHs as well as the McBride et al. (2009b) model.



**Figure 5.** Left panel: MAHs for the host halos in our five zoom-in simulation suites normalized to the  $z = 0$  virial mass for each host. Solid lines show mean MAHs for all hosts in each suite, and shaded bands show the corresponding 68% host-to-host scatter about the mean. Lines transition to dashed at the redshift above which  $> 5\%$  of the hosts in each suite are resolved with fewer than 300 particles. Jackknife resampling of the LMC, Milky Way, and Group suites indicates that the difference between normalized MAHs is marginally significant at  $z \gtrsim 2$  and not significant at later times; the difference between L-Cluster and Cluster MAHs is not statistically significant (see Section 3.4 for details). Dashed (dotted-dashed) gray lines show best-fit MAHs using the Wechsler et al. (2002) exponential model and the Tasitsiomi et al. (2004) exponential plus power-law model adopted by McBride et al. (2009b); best-fit parameters are provided in the text. Right panel: same as the left panel, but not normalized to  $z = 0$  host halo virial mass.

### 3.5. Host Halo Density Profiles

Next, we study the density profiles of Symphony host halos. We begin by reviewing standard halo profiles to contextualize our results. The NFW profile is given by

$$\rho(r) \propto r^{-1}(1+r)^{-2} \quad (6)$$

for  $x \equiv r/r_s$ , where  $r_s$  is the radius where the logarithmic slope is -2 (Navarro et al. 1997). The amplitude of the NFW profile is fixed for a given  $M_{\text{host}}$  and  $r_s$ , yielding a single-parameter fit in terms of  $r_s$ . The Einasto profile is given by

$$\frac{d \ln \rho}{d \ln r} \propto r^\alpha, \quad (7)$$

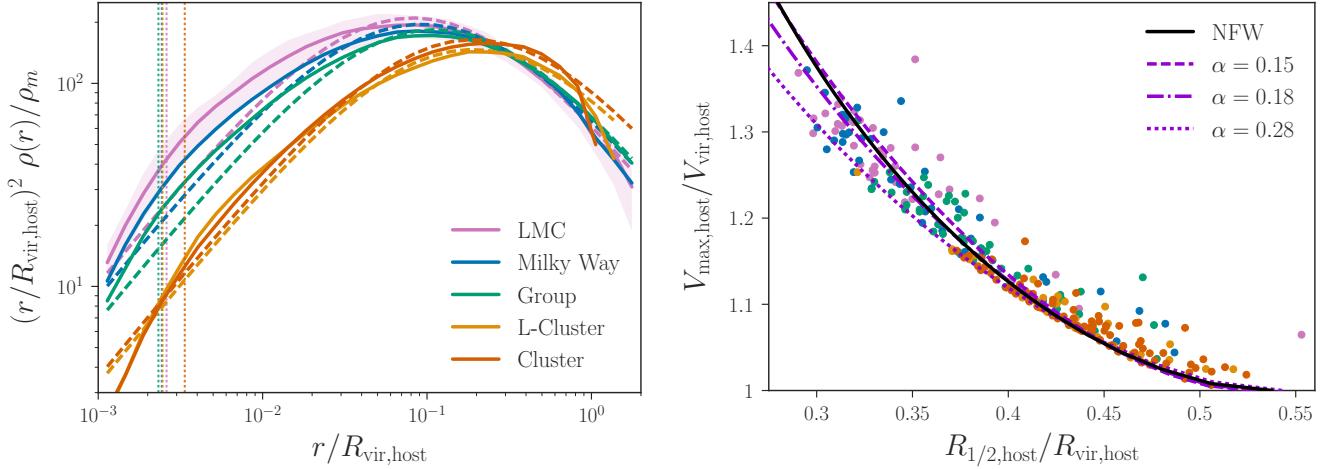
which becomes a two-parameter fit in terms of  $r_s$  and  $\alpha$  when  $M_{\text{host}}$  is fixed. Einasto profiles give more accurate descriptions of inner halo profiles than NFW profiles, even if  $\alpha$  is held fixed to values in the range  $\alpha \approx 0.17\text{--}0.18$  (e.g., Navarro et al. 2004; Wang et al. 2020a). We denote the maximum value of a halo’s rotation curve,  $V(< r) = \sqrt{GM(< r)/r}$ , by  $V_{\text{max}}$ , and we denote the radius at which  $M(< r) = M/2$  by  $R_{1/2}$ ; for hosts (subhalos), these variables are labeled by “host” (“sub”) subscripts. When scaled by the virial circular velocity and virial radius, respectively, both become robust, model-independent measures of halo concentration. All one-parameter halo models, including the NFW and Einasto profiles, predict a specific relationship between these halo properties.

The left panel of Figure 6 shows Symphony host halo density profiles as a function of distance from the host center, in units of the virial radius, measured directly from the particle snapshots and stacked over all hosts in each suite. Dashed lines show the mean of the best-fit NFW profiles reported by

ROCKSTAR for each suite, which are consistent with our direct measurements at the 10% level for  $r/R_{\text{vir,host}} \gtrsim 10^{-2}$ . For distances between the “convergence radius” of  $\approx 2.8\epsilon$  (e.g., Ludlow et al. 2019) and  $r/R_{\text{vir,host}} \approx 10^{-2}$ , Symphony profiles are systematically denser than predicted by ROCKSTAR’s NFW fit, consistent with previous measurements from cosmological and zoom-in simulations (e.g., Navarro et al. 2010; Ludlow et al. 2013). Furthermore, host concentrations estimated from direct profile measurements for each suite are consistent with the median values derived from ROCKSTAR, in line with previous studies using the original version of our Milky Way and our Cluster suites (Wu et al. 2013b; Fielder et al. 2020).

The right panel of Figure 6 compares the integrated profile properties  $V_{\text{max,host}}$  and  $R_{1/2,\text{host}}$  against the predictions of various halo models. The solid and dashed curves in this panel show the  $V_{\text{max,host}}/V_{\text{vir,host}} - R_{1/2,\text{host}}/R_{\text{vir,host}}$  relation predicted for NFW profiles and for Einasto profiles across a range of  $\alpha$  values. The points show Symphony host halos, color-coded by suite. At high host halo masses (i.e., mainly for L-Cluster and Cluster halos), Symphony hosts consistently lie above the NFW and Einasto curves, with  $V_{\text{max,host}}/V_{\text{vir,host}}$  values up to  $\approx 10\%$  higher than those models at fixed  $R_{1/2,\text{host}}/R_{\text{vir,host}}$ . Meanwhile, at lower host masses and typical values of  $R_{1/2,\text{host}}/R_{\text{vir,host}}$ , the simulation measurements scatter both above and below the NFW and Einasto predictions.

We hypothesize that the scatter toward larger values of  $V_{\text{max,host}}/V_{\text{vir,host}}$  for high-mass hosts may be caused by recent major mergers, which heat the inner regions of host halos. At all masses, halo triaxiality (e.g., Jing & Suto 2002) and substructure (e.g., Fielder et al. 2020) also contribute to the scatter about the predictions for smooth, spherically symmet-



**Figure 6.** Left panel: mean density profiles of Symphony host halos normalized to the mean cosmological matter density at  $z = 0$ , as a function of distance from the host center in units of the virial radius, for each suite (solid lines). Dashed lines show the mean of the best-fit NFW profiles reported by ROCKSTAR for each suite. Shaded bands indicate 68% host-to-host scatter, which is only shown for the LMC suite for visual clarity. Dotted vertical lines show approximate convergence radii of  $2.8\epsilon$  (e.g., Ludlow et al. 2019) for each suite, below which density profile measurements may not be converged. Right panel: maximum circular velocity normalized to the virial velocity versus half-mass radius normalized to the virial radius for all host halos in our five zoom-in suites. Points are colored by suite following the color scheme in the left panel. The black line shows the relation for an NFW profile, and purple lines show relations for Einasto profiles with  $\alpha = 0.15$  (solid),  $\alpha = 0.18$  (dotted-dashed), and  $\alpha = 0.28$  (dotted).

ric NFW and Einasto profiles. Studying these effects across our entire range of host halo mass is an interesting avenue for future work. In addition, Symphony density profile measurements will enable studies of the detailed correlation between host halos’ dark matter structure and subhalo population statistics, both at (roughly) fixed host halo mass within each suite and over several decades of host mass using the entire compilation.

#### 4. SYMPHONY SUBHALO POPULATIONS

We now present results for subhalo populations (Section 4.1), including SHMFs (Section 4.2) and radial distributions (Section 4.3); several auxiliary characteristics of the subhalo populations, including infall time and tidal stripping distributions, are presented in Appendix C.

Again, we defer from providing updated fits to standard subhalo population statistics (e.g., the SHMF), and we leave an investigation of additional subhalo population properties (e.g., subhalo orbital parameters) to future work.

##### 4.1. Overview

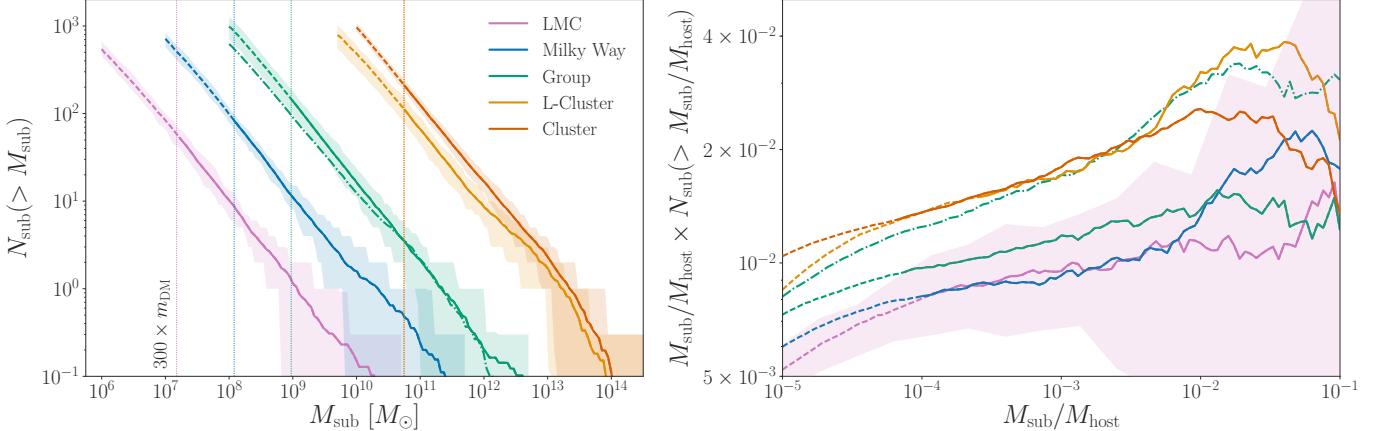
Subhalos above the Symphony convergence limit contain at least 300 particles at  $z = 0$  and typically have fallen into their hosts several gigayears ago with  $\approx 50\%$  more particles than at  $z = 0$  (see Appendices C.1–C.2). Appendix B demonstrates that the abundance of Symphony subhalos above this limit is converged at the  $\approx 10\%$  level; because even higher resolution is required to track the inner structure of stripped subhalos accurately (e.g., Errani & Navarro 2021), we focus on SHMFs and radial distributions here, and we study the

convergence properties of subhalo maximum circular velocity functions in Appendix B.3.

For the LMC, Milky Way, Group, L-Cluster, and Cluster suites, subhalos above our sub-to-host halo convergence limit contribute 10%, 12%, 18%, 21%, and 16% of the total mass within their host halo’s virial radius, respectively. On average, 84%, 81%, 86%, 69%, and 73% of these objects are “first-order” subhalos, meaning that they do not lie within the virial radius of any large halo except for the main host (see Springel et al. 2008 for a discussion of subsubstructure measurement subtleties).

##### 4.2. Subhalo Mass Functions

The left panel of Figure 7 shows the mean cumulative subhalo virial mass function (SHMF) and host-to-host scatter for each of our five suites. Applying a conservative limit of  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$  to all suites, which ensures that all subhalos in all hosts contain more than 300 particles, we find that the mean cumulative SHMFs are well described by power laws with slopes of  $-0.93 \pm 0.01$  (LMC),  $-0.92 \pm 0.01$  (Milky Way),  $-0.91 \pm 0.01$  (Group),  $-0.91 \pm 0.01$  (L-Cluster), and  $-0.88 \pm 0.01$  (Cluster), where errors represent  $1\sigma$  uncertainties on the best-fit slope of the mean SHMF, and slopes are fit over the range  $2.7 \times 10^{-4} < M_{\text{sub}}/M_{\text{host}} < 10^{-3}$  assuming Poisson uncertainties. Thus, we do not detect a systematic mass trend among the LMC/Milky Way/Group SHMF slopes. Note that the L-Cluster and Cluster slopes are difficult to interpret in detail relative to the LMC, Milky Way, and Group results due to differences in these simulations’ cosmological parameters (see Section 2). Our SHMF slopes are broadly consistent with previous measurements from cos-



**Figure 7.** Left panel: subhalo mass functions evaluated using virial mass at  $z = 0$  for our five zoom-in simulation suites. Solid lines show mean SHMFs stacked over each suite, and shaded bands show the corresponding 16th–84th percentile of the host-to-host scatter. The dotted-dashed Group result shows the SHMF at  $z = 0.5$ , when the corresponding host halos were mass-selected in their parent box. Vertical dotted lines show the convergence limit for each simulation, calculated as 300 times the dark matter particle mass in the highest-resolution zoom-in region, and lines transition from solid to dashed at the median  $M_{\text{sub}}/M_{\text{host}}$  threshold corresponding to the convergence limit for each suite. Right panel: same as the left panel, but scaled to highlight the dependence on host mass. Subhalo masses are normalized by  $M_{\text{host}}$  to remove most of the linear dependence of subhalo abundance on host mass, and abundances are scaled by  $M_{\text{sub}}/M_{\text{host}}$  to reduce the dynamic range of the y-axis. In this panel, host-to-host scatter is only shown for the LMC suite for visual clarity.

ological simulations and semianalytic models (e.g., van den Bosch & Jiang 2016; Mao et al. 2015; Benson 2020), accounting for both the host-to-host scatter and error of our slope measurements within each suite.

At fixed subhalo mass, the SHMF amplitude is roughly self-similar, scaling linearly with host mass. It is difficult to quantify this scaling precisely because subhalo abundance within each suite exhibits significant host-to-host scatter; meanwhile, comparisons between suites at fixed subhalo mass necessarily involve subhalos with different numbers of particles, which may suffer from numerical issues to different extents. Nonetheless, our results are broadly consistent with previous studies of the subhalo abundance–host mass scaling (e.g., Giocoli et al. 2008; Ishiyama et al. 2009, 2013; Jiang & van den Bosch 2016). These studies found that hosts at different masses have nearly self-similar subhalo mass functions, but that less-massive hosts tend to have lower SHMF amplitudes, likely because they disrupt their subhalos more efficiently. To our knowledge, this is the first explicit confirmation of this scaling using LMC and Group-mass zoom-in samples (however, see, e.g., Moliné et al. 2023 for a recent measurement from cosmological simulations).

To more directly compare subhalo abundances among Symphony suites, the right panel of Figure 7 shows normalized SHMFs, where subhalo masses are divided by the mass of each host halo to remove most of the linear dependence of subhalo abundance on host mass, and the resulting SHMFs are scaled by  $M_{\text{sub}}/M_{\text{host}}$  to reduce the dynamic range. The turnover at high sub-to-host halo mass ratios likely reflects the exponential cutoff in the SHMF (e.g., Gao et al. 2011); however, given the rarity of high-mass subhalos in Sym-

phony, we are not able to make decisive statistical statements about this regime.

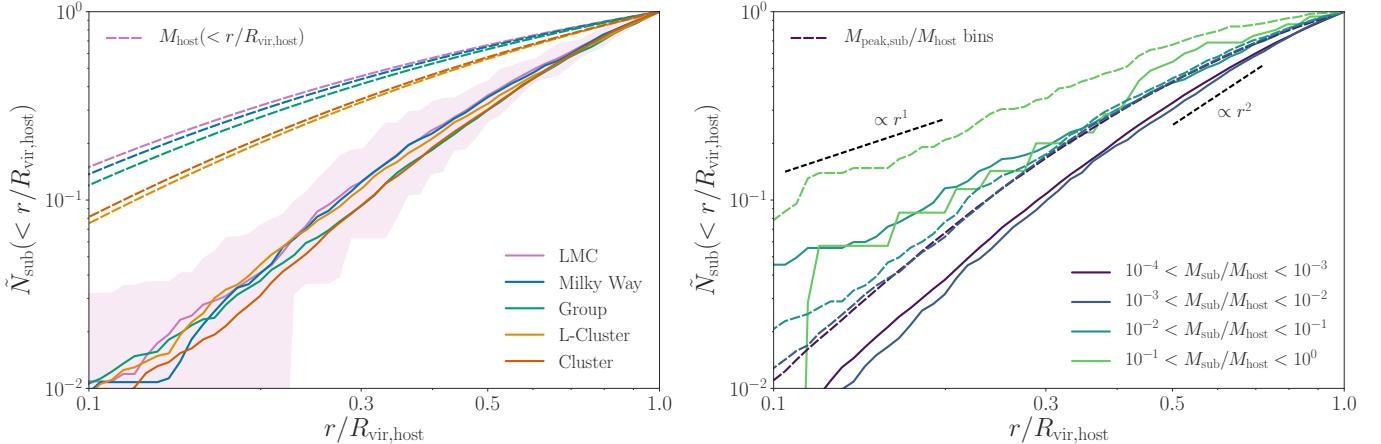
At fixed sub-to-host halo mass ratio, normalized subhalo abundances are consistent within the host-to-host scatter among the LMC, Milky Way, and Group suites. Meanwhile, the L-Cluster and Cluster hosts exhibit  $\approx 2$  times higher normalized subhalo abundances than the lower-mass suites at fixed mass ratio, which is broadly consistent with previous studies of the “evolved” subhalo mass function (e.g., Giocoli et al. 2008). This may result from the systematically later infall times for subhalos of the L-Cluster and Cluster hosts (see Appendix C.1), which leaves less time for these objects to be stripped below a fixed sub-to-host halo mass ratio. Leveraging the dynamic range of Symphony’s host and subhalo populations to explore the evolution of the SHMF as a function of host halo mass is an interesting avenue for future work.

### 4.3. Subhalo Radial Distributions

The left panel of Figure 8 shows the normalized subhalo radial distributions for each Symphony suite,

$$\tilde{N}_{\text{sub}}(< r/R_{\text{vir,host}}) \equiv \frac{N_{\text{sub}}(< r/R_{\text{vir,host}})}{N_{\text{sub}}(< R_{\text{vir,host}})} \quad (8)$$

using a resolution cut of  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$ . We also measure normalized radial distributions stacked over all suites in the right panel of Figure 8. To investigate how the radial distribution depends on subhalo mass, we bin these stacked measurements according to  $M_{\text{sub}}/M_{\text{host}}$  (solid lines) and  $M_{\text{peak,sub}}/M_{\text{host}}$  (dashed lines). Note that the  $M_{\text{sub}}/M_{\text{host}}$  bins extend slightly below our fiducial sub-to-host halo resolution cut; however, our qualitative results are not highly sensitive to this lower limit. Meanwhile, note that the results



**Figure 8.** Left panel: cumulative subhalo radial distributions as a function of distance from the host halo center, in units of the host halo virial radius, for our five zoom-in simulation suites. Solid lines show mean radial distributions stacked over each suite, and the shaded band shows the 16th–84th percentile for the LMC suite. Only subhalos above a resolution cut of  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^4$  are used to compute the radial distribution for each zoom-in, and each radial distribution is normalized to the total number of subhalos above this cut within the host halo’s virial radius. Dashed lines show mean enclosed mass profiles, normalized by host halo virial mass, stacked over hosts in each suite. Right panel: normalized subhalo radial distributions, stacked over all suites, in bins of sub-to-host halo mass ratio binned by  $z = 0$  (solid) and peak (dashed) subhalo virial mass. Dashed black lines indicate normalized radial distributions proportional to  $r$  and  $r^2$  for reference.

binned by  $M_{\text{peak},\text{sub}}/M_{\text{host}}$  include subhalos at small distances whose statistics are not formally converged at any mass ratio (see Appendix B.2), and should therefore be interpreted with caution.

The left panel of Figure 8 shows that Symphony’s total radial subhalo distributions (above a  $z = 0$  sub-to-host halo mass ratio cut) are systematically less concentrated than their hosts’ enclosed mass profiles. This is consistent with previous studies using cosmological simulations (e.g., Zentner et al. 2005; Springel et al. 2008), and may be related to the withering and artificial disruption of subhalos, even with large peak particle counts (see Appendix B.2, and Green et al. 2021). Interestingly, the shapes of Symphony radial subhalo distributions do not noticeably depend on host mass. A host mass dependence may be expected if subhalos trace their hosts’ underlying dark matter density profile; this would predict that lower-mass hosts have more centrally concentrated subhalo distributions given their more concentrated dark matter density profiles, as indicated by the dashed lines in the left panel of Figure 8 (also see Figure 6). However, lower-mass hosts also tidally strip their subhalos more efficiently and over longer timescales (see Appendices C.1–C.2), which potentially counteracts the trend due to underlying density profiles. Disentangling these effects will require a dedicated follow-up study.

The right panel of Figure 8 demonstrates that subhalos with low masses relative to their hosts exhibit the least centrally concentrated radial distributions; at higher mass ratios, dynamical friction causes subhalos to sink toward the host center more efficiently. Near hosts’ outer regions, the normalized radial profiles for low-mass subhalos scale as  $r^2$ . Note that radial distributions are generally shallower (and therefore the radial density profiles are steeper) when binned in terms of peak rather than  $z = 0$  subhalo virial mass, con-

sistent with previous studies (e.g., Nagai & Kravtsov 2005; Kravtsov 2010). However, because our simulations’ convergence properties are poorer when using peak rather than  $z = 0$  subhalo masses (see Appendix B.2), the radial distributions binned by  $M_{\text{peak},\text{sub}}$  should be interpreted with caution, particularly at low peak particle counts.

## 5. COMPARISON WITH GALACTICUS

To place our work in the context of semianalytic models, we compare our host halo properties and subhalo populations to those predicted by the GALACTICUS structure and galaxy formation model (Benson 2012; Pullen et al. 2014).<sup>6</sup> GALACTICUS’ structure formation modules combine prescriptions for building merger trees using extended Press-Schechter theory with analytic, physically motivated models for halo and subhalo evolution.

We begin by briefly describing the GALACTICUS model and implementation used to generate these predictions (Section 5.1); we then compare host halo concentration distributions (Section 5.2), subhalo mass functions (Section 5.3), and radial distributions (Section 5.4) for all suites. We comment on avenues for future work that combine Symphony with GALACTICUS and other semianalytic models (Section 5.5).

### 5.1. GALACTICUS Realizations of Symphony Systems

We generate 10 realizations of GALACTICUS merger trees and subhalo populations for each individual Symphony zoom-in simulation, using appropriate cosmological parameters and a resolution matched to the particle mass of each

<sup>6</sup> In particular, we use the version of GALACTICUS corresponding to commit bc4ecd2 in <https://github.com/galacticusorg/galacticus>.

suite. These merger trees have  $z = 0$  host halo virial masses matched to each Symphony host; thus, the GALACTICUS and Symphony host halo mass distributions match by construction. To model halo concentration, the Ludlow et al. (2016) concentration model is applied to the formation histories of halos with progenitors in each merger tree using the best-fit parameters from Benson et al. (2019), and the Diemer & Joyce (2019) mass-concentration relation with a scatter of 0.16 dex is used for halos without progenitors. Thus, our comparisons of Symphony and GALACTICUS host halo concentration distributions in Section 5.2 only rely on the Ludlow et al. (2016) implementation in GALACTICUS because all hosts have progenitors above the resolution threshold.

We use an updated version of the GALACTICUS subhalo evolution model presented in Yang et al. (2020), which was calibrated to match  $z = 0$  SHMFs, maximum circular velocity ( $V_{\max,\text{sub}}$ ) functions, and the  $M_{\text{sub}}-V_{\max,\text{sub}}$  relation from the ELVIS (Garrison-Kimmel et al. 2014) and Caterpillar (Griffen et al. 2016) simulations; here, we use the version of the model calibrated to Caterpillar. The subhalo evolution model we use includes a treatment of tidal stripping and is not subject to artificial subhalo disruption present in simulations. In addition, we use a new prescription for the orbital evolution of sub-subhalos (X. Du & A. Benson 2023, in preparation), which was not included in Yang et al. (2020); this model leads to a slightly more concentrated radial subhalo distribution near the host center ( $r/R_{\text{vir},\text{host}} \lesssim 0.1$ ) but does not significantly affect the SHMF or radial distribution predictions we present.

Note that, in GALACTICUS,  $M_{\text{sub}}$  is calculated from each subhalo's remaining bound mass at  $z = 0$ , and  $M_{\text{peak},\text{sub}}$  is identical to the virial mass at first infall onto the host halo,  $M_{\text{acc},\text{sub}}$ , because the evolution of subhalos before accretion (including pre-infall tidal stripping) is not modeled. However, in Symphony,  $M_{\text{peak},\text{sub}}$  is typically  $\approx 30\%$  higher than  $M_{\text{acc},\text{sub}}$  and occurs  $\approx 2$  Gyr earlier for subhalos in all Symphony suites (consistent with, e.g., Behroozi et al. 2014). Modeling pre-infall subhalo evolution in GALACTICUS is beyond the scope of our comparison, but constitutes an interesting area for future work.

## 5.2. Host Halo Concentrations

The left panel of Figure 9 summarizes our comparisons between Symphony, GALACTICUS, and Erebus mass-concentration relations. In particular, we show cumulative distributions of host concentration for all Symphony zoom-ins, GALACTICUS realizations, and Erebus halos in the LMC, Group, and Cluster mass ranges. Because GALACTICUS assumes virial equilibrium when computing host concentrations along each merger tree, we also test the effects of applying a virial relaxation cut to the Symphony hosts. In particular, we use the relaxation criterion in Ludlow et al. (2016), which excludes hosts that more than double their mass within the last 3.7 Gyr (i.e.,  $a_{1/2,\text{host}} \gtrsim 0.75$ ). This cut has a negligible effect on Symphony concentration distributions, which is expected because the LMC, Milky Way, and Group suites have very few unrelaxed hosts, while the unre-

laxed L-Cluster and Cluster hosts are not significant outliers in concentration (see the right panel of Figure 3).

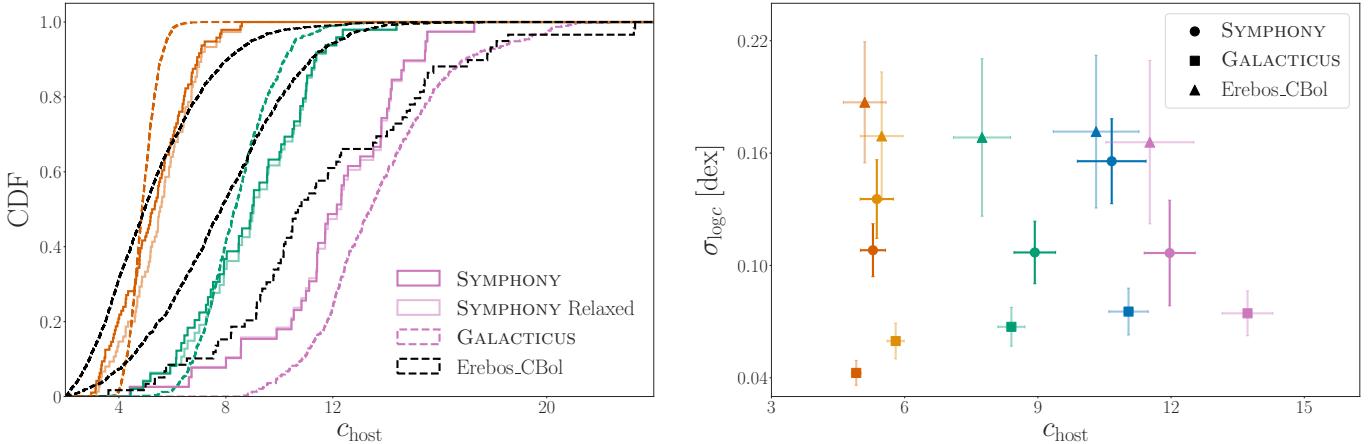
Symphony and GALACTICUS concentration distributions are statistically distinguishable: two-sample K-S tests yield  $p < 0.01$  for all suites, with the largest discrepancy for the Cluster suite. Interestingly, GALACTICUS predicts an even lower median concentration for Cluster hosts than Symphony, which in turn predicts a concentration distribution that is shifted low compared to Erebus (see Section 3.2). Furthermore, as shown in the right panel of Figure 9, GALACTICUS predicts smaller concentration scatter within each suite compared to Symphony, with typical values of  $\sigma_{\log c} \approx 0.06$  dex. The direction of this discrepancy in scatter is expected because the Ludlow et al. (2016) model underpredicts concentration scatter when applied to semianalytic merger trees, although it predicts the expected scatter when applied to merger trees extracted from cosmological simulations (Benson et al. 2019).

Several effects may contribute to the discrepancies between Symphony and GALACTICUS concentration distributions, and particularly to Symphony's larger concentration scatter, including: (1) GALACTICUS merger trees do not account for the dependence of halo concentration on environment, which increases concentration scatter (Benson et al. 2019); (2) Symphony hosts' density profiles are not spherical or smooth (i.e., they contain substructure), and they are not perfectly fit by NFW profiles, both of which increase concentration scatter (Benson et al. 2019; Fielder et al. 2020); (3) Symphony halos are not in perfect virial equilibrium, as assumed by the Ludlow et al. (2016) model, even after applying the relaxation cut. Exploring the difference between the concentration distributions of Symphony and GALACTICUS hosts—for example, by applying the Ludlow et al. 2016 model to the Symphony merger trees to predict concentration distributions, or by removing substructure from Symphony hosts in our density profile and concentration measurements—is left for future work.

## 5.3. Subhalo Mass Functions

The left panel of Figure 10 compares the ratio of the mean cumulative SHMF for each Symphony suite to the corresponding GALACTICUS predictions; Appendix D provides suite-by-suite SHMF comparisons. The left panel of Figure 10 demonstrates that Symphony and GALACTICUS SHMFs are consistent within the  $2\sigma$  Poisson error on the mean for  $M_{\text{sub}}/M_{\text{host}} \gtrsim 10^{-3}$  across all Symphony suites, but display a statistically significant  $\approx 25\%$  discrepancy at lower masses. In particular, GALACTICUS predicts  $\approx 25\%$  higher subhalo abundances than Symphony for  $M_{\text{sub}}/M_{\text{host}} \lesssim 10^{-3}$  in all suites, and lower subhalo abundances than Symphony at higher sub-to-host halo mass ratios in all except the Cluster suite. Both of these discrepancies are most severe for the Milky Way suite, and we discuss them in detail below.

The cumulative SHMF slopes predicted by GALACTICUS are  $-0.94 \pm 0.004$  (LMC),  $-0.94 \pm 0.001$  (Milky Way),  $-0.92 \pm 0.002$  (Group),  $-0.94 \pm 0.002$  (L-Cluster), and  $-0.92 \pm 0.002$  (Cluster) over the range  $2.7 \times 10^{-4} <$



**Figure 9.** Left panel: cumulative distributions of host halo concentration for Symphony hosts (solid), Symphony hosts that are virially relaxed according to the Ludlow et al. (2016) criterion (faint solid), GALACTICUS realizations (dashed), and Erebus halos (black dashed) in the LMC (pink), Group (green), and Cluster (red) mass ranges. For the LMC and Group suite comparisons, we only use Erebus halos that satisfy our hosts’ isolation criterion. Right panel: median host-to-host concentration scatter, measured as the standard deviation of the log  $c$  distribution in dex, versus median concentration for jackknifed samples of Symphony hosts (circles), GALACTICUS realizations (squares), and Erebus halos (triangles) in the LMC (pink), Milky Way (blue), Group (green), L-Cluster (gold), and Cluster (red) mass ranges. Error bars indicate  $1\sigma$  host-to-host jackknife uncertainties.

$M_{\text{sub}}/M_{\text{host}} < 10^{-3}$  assuming Poisson uncertainties. These slopes are a few percent steeper than the corresponding slopes measured in Symphony (see Section 4.2).

To interpret these results, we focus on the comparison between Symphony and GALACTICUS Milky Way predictions because, as described in Section 5.1, the GALACTICUS model we compare to accurately reproduces the normalization and slope of SHMFs from the Milky Way–mass Caterpillar zoom-ins (Yang et al. 2020). Thus, the Symphony–GALACTICUS Milky Way comparison can be understood by comparing Symphony and Caterpillar SHMFs, which we do systematically in Appendix E. In particular, the tests in Appendix E demonstrate that:

1. At high sub-to-host halo mass ratios,  $M_{\text{sub}}/M_{\text{host}} \gtrsim 10^{-2}$ , an upward fluctuation in the mean SHMF of Symphony’s target halos relative to the full sample of objects that pass the relevant mass and environmental cuts combines with a comparable downward fluctuation in the mean Caterpillar SHMF to yield a  $\approx 60\%$  (or  $\approx 2\sigma$ ) discrepancy. Both of these fluctuations are caused by the selection of the specific samples of 45 Symphony and 35 Caterpillar hosts, and manifest in both the zoom-in and parent cosmological simulations. Due to the low statistical strength, these fluctuations could be purely random, but they could also be related to the methods used to select target hosts after initial mass and environmental cuts are applied.
2. At intermediate sub-to-host halo mass ratios,  $10^{-3} \lesssim M_{\text{sub}}/M_{\text{host}} \lesssim 10^{-2}$ , an overabundance in the mean Caterpillar SHMF relative to cosmological simulations—which may, in part, be a fluctuation due to the selection of the specific Caterpillar sample—yields a  $\approx 25\%$  (or  $\approx 2\sigma$ ) discrepancy.

3. At low sub-to-host halo mass ratios,  $M_{\text{sub}}/M_{\text{host}} \lesssim 10^{-3}$ , the Symphony and Caterpillar SHMF slopes do not significantly differ. Thus, the overabundance relative to Symphony at intermediate subhalo masses propagates to very low sub-to-host halo mass ratios, resulting in the  $\approx 25\%$  discrepancy at the lowest resolved masses shown in Figure 10.

To accurately calibrate semianalytic models like GALACTICUS, SHMF biases among zoom-in host samples relative to all halos in cosmological volumes that pass the relevant mass and environmental cuts must be carefully modeled. We note that the effects of zoom-in host halo selection on the SHMF at high sub-to-host halo mass ratios can be identified relatively easily through comparisons between zoom-ins and the corresponding systems in their parent boxes, or successively higher-resolution resimulations thereof. On the other hand, detailed study would be required to understand the origin of discrepancies between zoom-ins at lower sub-to-host halo mass ratios, where zoom-in procedures or analyses themselves might influence the results.

Further investigation is needed to confirm whether discrepancies between Symphony and GALACTICUS predictions for the remaining Symphony suites are mainly inherited from the Milky Way comparison, or if they are affected by (1) selection effects or random fluctuations in other Symphony suites, and/or (2) the dependence of GALACTICUS predictions on host halo mass. We expect the effects of environmental cuts within each host halo mass range to be most severe for the LMC suite, relatively weak for the Group suite, and unimportant for the L-Cluster and Cluster suites. However, the effects of zoom-in host halo selection—i.e., differences between specific zoom-in host halo samples relative to all systems that pass the relevant mass and environmental cuts—

likely bias all suites’ SHMFs to some extent, with the possible exception of the Cluster suite, because all hosts in its parent box above a mass threshold were resimulated. In this context, it is reassuring that the Symphony–GALACTICUS SHMF discrepancies are less severe at high sub-to-host halo mass ratios for the non–Milky Way suites, and are least severe for the Cluster suite.

#### 5.4. Subhalo Radial Distributions

The right panel of Figure 10 compares the ratio of normalized distributions for each Symphony suite to the corresponding GALACTICUS predictions; Appendix D provides suite-by-suite comparisons and explores the dependence of GALACTICUS radial distributions on subhalo  $z = 0$  and peak mass. Symphony and GALACTICUS normalized radial distributions are consistent within the  $2\sigma$  Poisson error on the mean for subhalos down to our conservative sub-to-host halo resolution limit, at all at distances  $r/R_{\text{vir,host}} > 0.1$ ; we limit the comparison to these distances because we cannot explicitly verify that Symphony subhalo population statistics are converged at smaller radii (see Appendix B.2). Above our sub-to-host halo cut, Symphony radial profiles are less centrally concentrated compared to GALACTICUS. Furthermore, GALACTICUS radial distributions are also more concentrated when binned by  $M_{\text{peak,sub}}/M_{\text{host}}$ , consistent with the comparisons in Nadler et al. (2021b) for two hosts from the original Milky Way suite (see Appendix D). A dedicated study that models the effects of withering below the mass resolution limit and artificial disruption in Symphony will be needed to assess whether this discrepancy is physical or numerical.

Symphony radial distributions display a weak subhalo mass dependence for all  $M_{\text{sub}}/M_{\text{host}} \lesssim 10^{-2}$  and  $M_{\text{peak,sub}}/M_{\text{host}} \lesssim 10^{-1}$  (see the right panel of Figure 8), while GALACTICUS radial distributions are similar for all  $M_{\text{sub}}/M_{\text{host}}$  but become systematically more concentrated with increasing  $M_{\text{peak,sub}}/M_{\text{host}}$  (see Appendix D). Such discrepancies may indicate the need for more thorough calibration of the GALACTICUS subhalo evolution model used to generate these predictions, which was not calibrated to Caterpillar subhalo radial distributions. On the other hand, the differences in radial distributions as a function of sub-to-host halo mass ratio that we identify might require implementing additional subhalo evolution physics in GALACTICUS (for example, related to the strength and mass dependence of dynamical friction, or the efficiency of tidal stripping). We note that zoom-in host halo selection effects do not seem to significantly affect the radial distribution comparison, since the discrepancy between Symphony and GALACTICUS predictions is similar for all suites.

#### 5.5. Areas for Future Work

We anticipate several additional avenues for future work that combine Symphony with GALACTICUS and other semi-analytic structure formation models. First, the GALACTICUS model used here does not account for our hosts’ specific environments; however, Symphony’s LMC, Milky Way, and Group hosts generally occupy underdense regions due to the

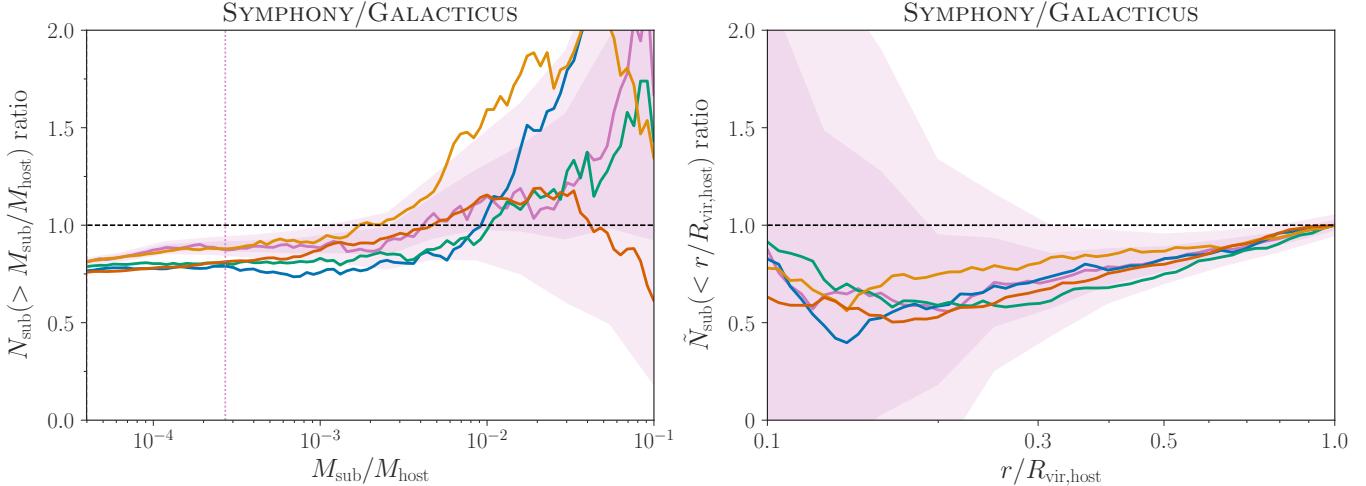
selection criteria described in Section 2. Building on Benson et al. (2019), it is interesting to consider how significantly the subhalo populations of hosts in particular environments, which have biased formation histories relative to the cosmic average, differ from typical subhalo populations at fixed host halo mass.

Second, it remains challenging to robustly resolve subhalos in cosmological simulations due to numerical effects, even at Symphony’s resolution. In particular, subhalos can be stripped below the mass resolution limit (or “wither”), which affects the low-mass end of simulated subhalo populations, and they can undergo artificial disruption, which may affect subhalos even with relatively high peak particle counts (e.g., see van den Bosch et al. 2018; van den Bosch & Ogiya 2018; Green et al. 2021). Furthermore, the internal structure of stripped subhalos can only be accurately resolved for subhalos with thousands of particles or more (e.g., see Errani & Navarro 2021). Semianalytic models including GALACTICUS are well suited to quantify the impact of withering and artificial disruption on subhalo population statistics because they are relatively inexpensive, which allows subhalo populations to be resolved at much higher resolution than in cosmological simulations (e.g., see Benson & Du 2022). Exploiting these advantages to quantify the impact of resolution effects and the convergence of subhalo population statistics in Symphony and its convergence resimulations (see Appendix B.1) represents an interesting area for future work.

Third, in addition to  $z = 0$  subhalo populations, we emphasize that Symphony resolves: (1) the redshift evolution of subhalo populations with high temporal resolution out to  $z \gtrsim 10$ , and (2) high-resolution halos out to many times the virial radius of each target host, including isolated halos, subhalos of lower-mass hosts, and splashback subhalos of the main host. Although we have not studied the redshift evolution of subhalo populations or the statistics of halos beyond the virial radius of the host in this work, Symphony data products provide opportunities to calibrate semianalytic models in these relatively under-explored regimes. The evolution of Symphony subhalo and halo populations is particularly informative for (re)calibrating semianalytic models of the conditional mass function and halo merger rates (e.g., Benson 2017). As mentioned above, performing such (re)calibration robustly will require modeling the impact of environment on halo and subhalo populations (e.g., using the peak background split methodology of Sheth & Tormen 1999), as well as secondary biases imprinted by zoom-in host halo selection from a mass and environmentally selected host sample.

## 6. APPLYING UNIVERSEMACHINE TO SYMPHONY

In this section, we use UNIVERSEMACHINE to model the SFHs of galaxies occupying halos and subhalos in the high-resolution regions of all Symphony simulations. We describe our UNIVERSEMACHINE model and zoom-in technique in Section 6.1, predictions for Symphony hosts’ SFHs in Section 6.2, predictions for the stellar mass–halo mass (SMHM) relation in Section 6.3, the observational relevance of our pre-



**Figure 10.** Left panel: the ratio of cumulative subhalo mass functions in each Symphony suite to corresponding GALACTICUS predictions. Solid lines show the ratio of median cumulative SHMFs stacked over all Symphony zoom-ins and GALACTICUS realizations in our LMC (pink), Milky Way (blue), Group (green), Low-Cluster (gold), and Cluster (red) mass ranges. Dark (light) shaded bands indicate the corresponding  $1\sigma$  ( $2\sigma$ ) Poisson errors on the mean, which we only show for the LMC suite for visual clarity. The vertical dotted line indicates our conservative convergence limit of  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$ , and the horizontal dashed line indicates perfect agreement between Symphony and GALACTICUS. Right panel: same as the left panel, for the ratio between Symphony and GALACTICUS cumulative radial distributions in units of distance from the host center divided by the host halo virial radius.

dictions in Section 6.4, and areas for future work combining Symphony and UNIVERSEMACHINE in Section 6.5.

### 6.1. UNIVERSEMACHINE Model and Zoom-in Application

UNIVERSEMACHINE (Behroozi et al. 2019) is an empirical galaxy–halo connection model that paints star formation rates (SFRs) onto dark matter halo merger trees. This modeling is performed probabilistically using halos’ maximum circular velocities and accretion rates as a function of redshift. UNIVERSEMACHINE has been calibrated to match luminosity functions, quenched fractions, and auto- and cross-correlation functions of star-forming and quenched galaxies with stellar masses  $M_* \gtrsim 10^8 M_\odot$  for  $0 < z < 10$  (Behroozi et al. 2019). The public version of the UNIVERSEMACHINE DR1 model was calibrated on the cosmological Bolshoi-Planck simulation (Klypin et al. 2016; Rodríguez-Puebla et al. 2016), with a box size of 250 Mpc  $h^{-1}$  and dark matter particle mass of  $m_{\text{part}} \approx 2 \times 10^8 M_\odot$ .<sup>7</sup>

Wang et al. (2021) presented the first application of UNIVERSEMACHINE to zoom-in simulations by using the DR1 model to model the SFHs of halos and subhalos in the original version of Symphony’s Milky Way suite (Mao et al. 2015). These authors developed a technique to join ROCKSTAR halo catalogs from individual simulations to obtain accurate SFH predictions for an entire zoom-in suite. Here, we use the same methodology and DR1 model to obtain SFHs for each Symphony suite’s halo and subhalo populations. In particular, we join halo lists within each suite of Symphony

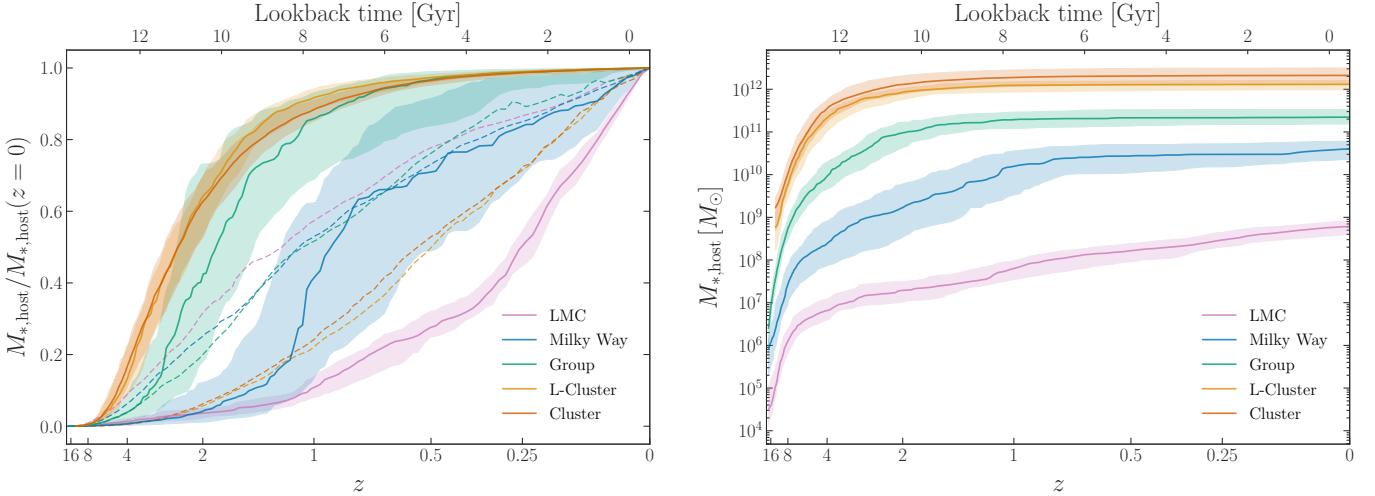
simulations to obtain a statistically robust halo accretion rate distribution for each suite (see Wang et al. 2021 for details). Each distribution is then used to generate SFHs for all simulations within the corresponding suite.

As emphasized by Wang et al. (2021), the UNIVERSEMACHINE DR1 model is *not* constrained for central or satellite galaxies with  $M_* \lesssim 10^8 M_\odot$ , which typically occupy (sub)halos with  $M_{\text{peak,sub}} \lesssim 10^{10} M_\odot$  (e.g., see Wechsler & Tinker 2018). For example, reionization and environmental processes are known to quench galaxies at these stellar masses, while UNIVERSEMACHINE DR1 predicts that nearly all low-mass galaxies actively form stars down to  $z=0$  (Wang et al. 2021). This impacts our predicted SFHs for the galaxies that occupy low-mass (sub)halos in Symphony’s LMC, Milky Way, and Group suites, even though their luminosity functions are consistent with extrapolations of global luminosity function measurements at higher stellar masses (Wang et al. 2021). We also note that, although our L-Cluster and Cluster simulations have similar resolution to the simulation UNIVERSEMACHINE DR1 was calibrated on, the details of galaxy quenching and evolution for low-mass satellites of cluster-mass hosts may not fully be captured by the DR1 model.

### 6.2. Predictions for Symphony Host Halo Star Formation Histories

SFHs for the central galaxies occupying Symphony’s target hosts in each suite are presented in Figure 11. Lower-mass hosts form their stars systematically later than higher-mass hosts, in contrast to their dark matter MAHs (see Figure 5). In particular, Symphony hosts form half of their stars at half-stellar-mass scale factors of  $a_{1/2,\text{host}}^* = 0.8 \pm 0.04$ ,

<sup>7</sup> We use the UNIVERSEMACHINE DR1 version dated 2020-12-19, which can be found at <https://bitbucket.org/pbehroozi/universemachine/src/main/>.



**Figure 11.** Left panel: SFHs of central galaxies that occupy Symphony hosts, stacked over all host halos in each suite. Solid lines show the median evolution of the intrinsic stellar mass (i.e., the predicted stellar mass uncorrected for observational systematics and massive stars’ death over time), normalized to each host’s predicted stellar mass at  $z = 0$ , and shaded regions denote the 68% host-to-host scatter for each suite. Dashed lines show the corresponding mean normalized dark matter MAHs (see the left panel of Figure 5). Right panel: same as the left panel, but for hosts’ un-normalized SFHs.

$0.56 \pm 0.1$ ,  $0.37 \pm 0.09$ ,  $0.3 \pm 0.03$ , and  $0.3 \pm 0.07$  for the LMC, Milky Way, Group, L-Cluster, and Cluster suites, respectively. These findings are consistent with many previous studies (e.g., see Conroy & Wechsler 2009; Wechsler & Tinker 2018) and are expected because higher-mass galaxies are preferentially quenched at early times, while lower-mass galaxies—and particularly our LMC and Milky Way centrals—form stars continuously until low redshifts.

Interestingly, although the dependence of dark matter MAHs on host mass is strongest for the transition between our Group and L-Cluster suites (see Section 3.4), the most noticeable transition among normalized SFHs occurs between the LMC, Milky Way, and Group hosts. This is a robust prediction of the DR1 model, which is well constrained in the stellar mass range of Symphony’s predicted central galaxies. However, note that our LMC, Milky Way, and Group hosts occupy underdense environments that may not be well represented in the UNIVERSEMACHINE DR1 calibration.

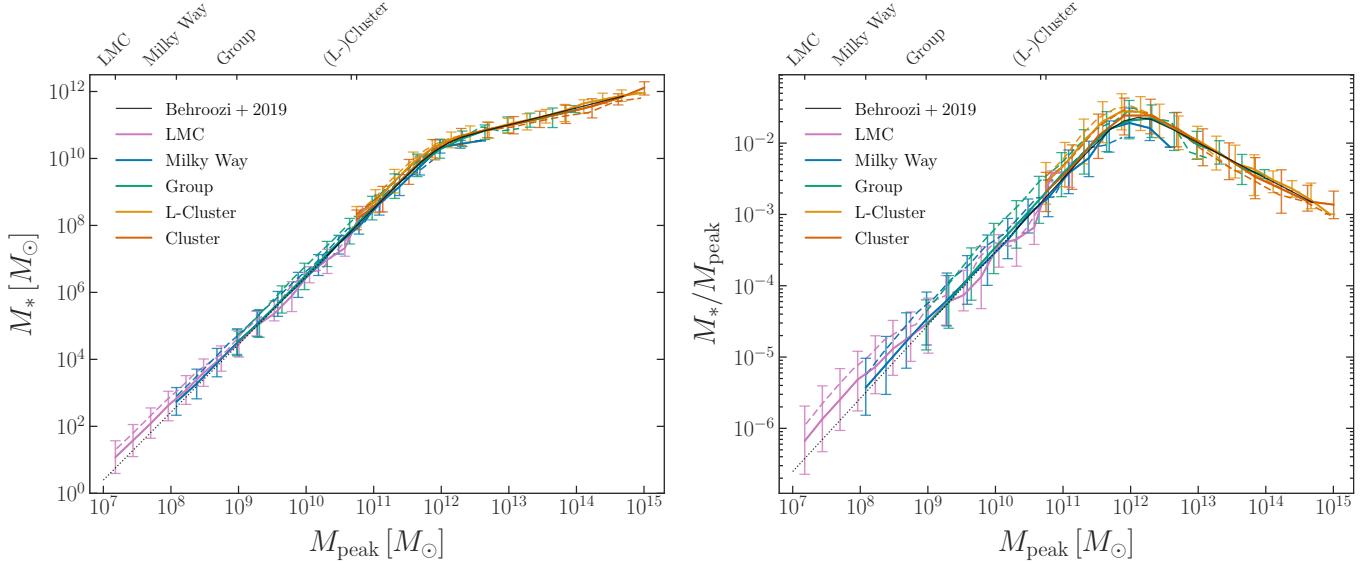
### 6.3. Predictions for the Stellar Mass–Halo Mass Relation Using Symphony

The SMHM relation predicted by UNIVERSEMACHINE at  $z = 0$  for halos and subhalos in the high-resolution regions of our 260 zoom-in simulations is presented in Figure 12, using our fiducial-resolution cut of  $M_{\text{sub}} > 300m_{\text{part}}$ . In particular, solid lines show the SMHM relation for all halos and subhalos within  $5R_{\text{vir},\text{host}}$  of each Symphony host, and dashed lines show the SMHM only for subhalos of the target host in each zoom-in. Note that  $5R_{\text{vir},\text{host}}$  corresponds to the size of the smallest high-resolution regions in Symphony, from the LMC suite, while other Symphony suites’ high-resolution regions extend to  $\approx 10R_{\text{vir},\text{host}}$  (see Section 2.2). We plan to leverage this volume and explore the im-

pact of low-resolution contamination near the high-resolution boundaries in future work.

The SMHM relation, combined over all Symphony suites, spans an immense dynamic range of more than 10 orders of magnitude in stellar mass and 7 orders of magnitude in peak (sub)halo mass. At fixed  $M_{\text{peak},\text{sub}}$ , satellites of Symphony hosts have higher stellar masses than isolated systems, consistent with previous UNIVERSEMACHINE results (Behroozi et al. 2019). Furthermore, the median and the scatter of SMHM relations across different Symphony suites all agree well within the  $1\sigma$  intrinsic and host-to-host scatter where overlap exists. This suggests that our procedure for joining halo lists within each Symphony suite does not significantly bias the resulting SMHM relations. We note that the SMHM scatter is dominated by intrinsic scatter in the galaxy–halo connection, which would persist even when applying UNIVERSEMACHINE to any individual Symphony zoom-in, rather than host-to-host scatter. Exploring the dependence of SFH predictions on (sub)halo environment is an interesting area for future study.

At stellar masses above  $\approx 10^8 M_\odot$ , UNIVERSEMACHINE is calibrated to match various galaxy population statistics, and our SMHM predictions are consistent with fiducial DR1 predictions from Behroozi et al. (2019). This agreement is nontrivial because UNIVERSEMACHINE DR1 was calibrated and run on a cosmological simulation in which host halos at all mass scales are less well resolved than in Symphony. In addition, zooming in on a single massive host halo changes the overall satellite fractions in our simulations relative to those in cosmological volumes, even for suites that do not impose specific isolation criteria (namely, our L-Cluster and Cluster suites). In a follow-up study, we plan to quantify how these effects lead to variations in the faint-end SMHM slope for our LMC, Milky Way, and Group suites and relative



**Figure 12.** Left panel: stellar mass–halo mass (SMHM) relations predicted by applying the UNIVERSEMACHINE DR1 model to all Symphony suites. Solid lines represent the median SMHM relation stacked over all simulations in each Symphony suite using all halos and subhalos above our fiducial-resolution cut of  $M_{\text{sub}} > 300m_{\text{part}}$  and within  $5R_{\text{vir,host}}$  of each Symphony host; dashed lines only include subhalos of the target host in each simulation. Error bars denote the 68% scatter, which is dominated by intrinsic scatter in the galaxy–halo connection rather than host-to-host scatter. The top axis denotes each suite’s resolution limit in terms of  $z = 0$  (sub)halo virial mass (i.e.  $M_{\text{sub}} > 300m_{\text{part}}$ ); however,  $M_{\text{peak,sub}}$  is less well converged than  $M_{\text{sub}}$  (see Appendix B.2), so these ticks should only be interpreted as rough estimates of the resolution limits. The black curve shows the  $z = 0$  SMHM relation from Behroozi et al. (2019) for comparison, and transitions to a dotted power-law extrapolation of the DR1 prediction at stellar masses where it is not constrained. Right panel: same as the left panel, but with stellar mass normalized to each (sub)halo’s peak virial mass.

to fiducial UNIVERSEMACHINE predictions from Behroozi et al. (2019).

At stellar masses below  $\approx 10^8 M_\odot$ , UNIVERSEMACHINE is designed to smoothly extrapolate the global luminosity function even though the DR1 model is not calibrated to the observed properties of such faint galaxies. Thus, our predicted SMHM relation for Symphony’s Milky Way suite is consistent with SMHM relations inferred from Milky Way satellite galaxies with  $10^2 \lesssim M_* \lesssim 10^8 M_\odot$  using abundance matching because these relations are in turn consistent with the global luminosity function (Nadler et al. 2020b; Wang et al. 2021). The predictions at even lower stellar masses and for satellites of lower-mass hosts—particularly for galaxies occupying the lowest-mass (sub)halos in our LMC suite—have not been compared to data or other theoretical predictions. At this extreme faint end of the SMHM, the efficiency and stochasticity of galaxy formation are influenced by both reionization and environmental processes (e.g., Rey et al. 2019; Benitez-Llambay & Frenk 2020; Manwadkar & Kravtsov 2022; Munshi et al. 2021). Thus, as we discuss further in Section 6.5, we plan to revisit our predictions for ultra-faint dwarf galaxy SFHs once UNIVERSEMACHINE is appropriately recalibrated.

#### 6.4. Modeling Observable Galaxy Populations using UNIVERSEMACHINE and Symphony

The results above highlight the dynamic range of Symphony’s well-resolved halo and subhalo populations, which

we have exploited to statistically model the connection between dark matter accretion histories and galaxy SFHs using UNIVERSEMACHINE over an unprecedented range of stellar mass. Here, we argue that our UNIVERSEMACHINE predictions statistically capture the SFHs of systems with luminosities comparable to nearly all currently observable satellite galaxies in the universe, with the exception of the faintest known satellites in nearby clusters. For this to hold, resolved satellites with the lowest stellar masses from our UNIVERSEMACHINE predictions for each suite must be fainter than the dimmest known satellites at the corresponding host halo mass scales, which are generally found in systems at the smallest distances due to selection effects.

For LMC and Milky Way–mass hosts, the faintest observed satellites are ultra-faint dwarfs orbiting the Milky Way itself (Simon 2019), some of which were originally associated with the LMC (e.g., Kallivayalil et al. 2018; Patel et al. 2020). These galaxies have stellar masses as small as  $\approx 100 M_\odot$  (e.g., Drlica-Wagner et al. 2020) and occupy halos with peak masses down to  $\approx 10^8 M_\odot$ , or smaller (e.g., Read & Erkal 2019; Nadler et al. 2020b). Our LMC and Milky Way–mass suites resolve the abundance and formation histories of subhalos with masses down to  $\approx 10^7 M_\odot$  ( $\approx 10^8 M_\odot$ ), and therefore statistically capture the formation histories of subhalos that host these observed ultra-faints. The details of this comparison depend on the relationship between  $z = 0$  and peak subhalo mass; thus, robustly modeling the ultra-faint end of

the observed Milky Way and/or LMC satellite populations still requires a model for the (potentially premature) disruption of low-mass subhalos (e.g., Nadler et al. 2019b, 2020b).

On the Group scale, we compare our predictions to the faintest known satellites in Centaurus A (Cen A), because it is one of the closest and best-observed systems with a halo mass of  $\approx 10^{13} M_\odot$  (e.g., Karachentsev 2005; Woodley 2006). Cen A’s dimmest satellites have absolute  $V$ -band magnitudes of  $M_V \approx -8$  mag and luminosities of  $\approx 10^5 L_\odot$  (Crnojević et al. 2019), or stellar masses of  $\approx 10^5 M_\odot$  assuming mass-to-light ratios of  $\approx 1$ . Our Group suite resolves the abundance and formation histories of subhalos with masses down to  $\approx 10^9 M_\odot$ , which host satellites with stellar masses of a few times  $10^4 M_\odot$ ; these are dimmer than the limit set by Cen A observations.

Our L-Cluster and Cluster suites have similar resolution; thus, it is sufficient to consider the faintest observed satellites at either host halo mass scale to assess whether our simulations can resolve such galaxies. We assume that the faintest observed satellites of the Virgo cluster—which have absolute  $g$ -band magnitudes of  $M_g \approx -10$  mag (Ferrarese et al. 2016)—set this limit, noting that other nearby clusters have been observed to comparable depths (e.g., the Coma and Fornax clusters; Yamanoi et al. 2012; Venhola et al. 2019). The faintest observed Virgo satellites have stellar masses of  $\approx 10^6 M_\odot$ , which is below our L-Cluster and Cluster suites’ capabilities given that these suites resolve subhalos down to  $\approx 6 \times 10^{10} M_\odot$ , which have stellar masses of  $\approx 10^8 M_\odot$ , on average. However, we note that observations of Virgo (and other nearby clusters) are much more complete at the faint end than typical observations of higher-redshift clusters. Thus, our L-Cluster and Cluster suites’ resolution is sufficient for modeling the majority of observed cluster satellite populations, with the exception of the faintest satellites in several nearby clusters.

Finally, we emphasize that our SFH predictions robustly span a wide range of redshifts because they are tied to dark matter accretion histories that are well resolved at early times. For example, our LMC and MW-mass hosts are well resolved out to  $z \gtrsim 9$  (see Section 3.4); thus, Symphony enables SFH predictions for the progenitors of LMC and MW-mass central galaxies at epochs that are highly relevant given HST observations over the last decade (e.g., Bouwens et al. 2015; Finkelstein et al. 2015; Atek et al. 2018) and ongoing JWST studies. Thus, we anticipate that our predictions will aid the development of empirical models that rely on simulation-based SFH predictions at high redshifts (e.g., Behroozi et al. 2020).

### 6.5. Areas for Future Work

Given the dynamic range of our galaxy–halo connection predictions, we anticipate several interesting observational applications of the UNIVERSEMACHINE–Symphony combination. For example, this combination enables:

- Predictions for the satellite populations of isolated LMC-mass hosts, which will be relevant given upcom-

ing observations of ultra-faint dwarf satellite throughout much of the Local Volume (e.g., Mutlu-Pakdil et al. 2021), using the LMC suite;

- Predictions for the satellite populations of Milky Way analogs, and particularly their quenched fractions, which are in moderate tension with observations of the quenched fraction for satellites of the Milky Way and Andromeda (Geha et al. 2017; Mao et al. 2021), using the Milky Way suite;
- Predictions for the luminosity functions, radial distributions, and SFHs of satellites of early-type galaxy strong lenses, to improve upon current lens modeling techniques and analyses of observed lenses, using the Group suite;
- Predictions for the properties of subhalos and splashback systems around cluster hosts at low redshifts, which have been measured in recent years (e.g., Shin et al. 2019), using the L-Cluster suite;
- Predictions for the satellite populations of intermediate-redshift clusters, some of which display anomalously large quenched fractions (e.g., Noordh et al. 2021), using the Cluster suite.

Once again, we caution that UNIVERSEMACHINE DR1 is not calibrated to dwarf galaxy observations; indeed, DR1 predicts unrealistic SFHs and quenched fractions for  $M_* \lesssim 10^8 M_\odot$  that are inconsistent with Local Group dwarf observations (Wang et al. 2021). Hence, an updated version of UNIVERSEMACHINE that incorporates these constraints is required before using the faintest satellite galaxies predicted for our LMC, Milky Way, and Group suites to interpret observational results. Such updates will be presented in future work (Y. Wang et al. 2023, in preparation) that integrates more realistic satellite quenching models into the current UNIVERSEMACHINE framework; these models will be constrained by Local Volume and SAGA satellite populations (Geha et al. 2017; Mao et al. 2021; Carlsten et al. 2022) and individual Local Group dwarf galaxies SFHs (e.g., Brown et al. 2014; Weisz et al. 2014a,b, 2015; Sacchi et al. 2021). Relevant observables will be studied with the updated model, once UNIVERSEMACHINE is calibrated to match available dwarf galaxy constraints.

## 7. SYMPHONY SCIENCE CASES

In this Section, we discuss potential science cases for the individual Symphony suites (Sections 7.1–7.4) and of the compilation as a whole (Section 7.5). The applications we discuss are simply examples of the science Symphony enables, and are not meant to be comprehensive. Instead, our aim is to highlight the unique characteristics of Symphony that benefit particular science cases at each mass scale, bearing in mind that existing zoom-in suites are also useful for addressing many of these questions.

### 7.1. LMC-mass Suite

Symphony's new LMC suite captures a statistical sample of LMC-mass systems at high resolution and in a narrow host mass range. Thus, this suite allows several basic questions to be addressed for LMC-mass hosts using simulations with comparable resolution to typical Milky Way-mass zoom-ins. For example, we have quantified how LMC-mass hosts' higher concentrations (Figure 3, left panel), earlier formation times (Figure 3, right panel), and earlier subhalo infall times (Appendix C.1) shape their surviving and disrupted subhalo populations relative to the subhalo populations of, e.g., Milky Way-mass systems.

Our LMC-mass suite is timely given recent observational progress in our understanding of the dark matter content and substructure of the LMC and galaxies of similar luminosities. For example, estimates of the LMC's total mass have recently been improved based on its dynamical impact on stellar streams (e.g., Erkal et al. 2019; Shipp et al. 2021), and ultra-faint satellites of the LMC have been identified using new, precise astrometric data (e.g., Kallivayalil et al. 2018; Patel et al. 2020). Meanwhile, brighter satellites of LMC-mass hosts throughout the Local Volume are now commonly being discovered (e.g., see Carlsten et al. 2022 for a recent census). Precise theoretical predictions for the subhalo and satellite populations of LMC-mass hosts in general will be important in order to interpret the LMC's satellite population and observations of satellites around nearby LMC-mass galaxies in a cosmological context. We expect that the combination of Symphony's LMC-mass suite and the population of LMC analogs falling into Milky Way-mass hosts presented in upcoming zoom-in simulations (D. Buch et al. 2023, in preparation) will be particularly helpful for interpreting observations of the actual LMC's luminous substructure. However, we caution that such analyses must account for our LMC suite's selection criteria, which yield LMC hosts in significantly underdense environments relative to typical LMC-mass halos.

Extensions to Symphony's LMC suite may include resimulations with: (1) analytic galaxy potentials or hydrodynamic physics to study the impact of LMC-mass galaxies on their subhalo populations, and (2) nonstandard dark matter physics to study its impact on extremely low-mass halos and subhalos, including those that are not expected to host galaxies. These efforts would build on previous zoom-ins that focus on or include LMC-mass systems (e.g., Jahn et al. 2019; Nadler et al. 2021a; Schwabe & Niemeyer 2022) by leveraging Symphony's large sample of LMC-mass halos to quantify the halo-to-halo scatter and environmental dependence of such baryonic and dark matter physics. We emphasize that our LMC suite robustly resolves the population statistics of halos and subhalos below the galaxy formation threshold (i.e.,  $M_{\text{peak,sub}} \lesssim 10^8 M_\odot$ ; Nadler et al. 2020b; Munshi et al. 2021), and will therefore help facilitate searches for novel dark matter physics using gravitational probes of low-mass halos, including strong gravitational lensing and stellar stream perturbations.

### 7.2. Milky Way-mass Suite

The original version of Symphony's Milky Way-mass suite, from Mao et al. (2015), has already been used to investigate: (1) the dependence of subhalo abundance on host concentration (Mao et al. 2015), (2) the disrupted progenitors of the Milky Way halo (Deason et al. 2016), (3) the Doppler effect on dark matter annihilation signals (Powell et al. 2017); (4) semianalytic models of galaxy formation and feedback physics using the Milky Way satellite population (Lu et al. 2016, 2017), (5) the impact of baryonic physics on subhalo populations (Nadler et al. 2018), (6) the galaxy-halo connection for classical and ultra-faint Milky Way satellite galaxies (Nadler et al. 2019b, 2020b), (7) constraints on dark matter physics beyond the CDM paradigm from Milky Way satellites (Nadler et al. 2019a, 2020a, 2021c,b,a; Mau et al. 2022), (8) the dependence of subhalo populations on secondary host halo properties (Fielder et al. 2019), (9) the impact of subhalos on dark matter density profiles (Fielder et al. 2020), and (10) the connection between dark matter accretion and SFRs in zoom-in simulations (Wang et al. 2021).

The studies listed above illustrate the breadth of science cases Symphony's Milky Way-mass suite enables; in many cases, performing similar studies at different host halo mass scales (e.g., using the LMC and/or Group suites) may yield additional insight. Here, we discuss a few specific examples of new directions for future work given observational and theoretical advances following the publication of the original Milky Way suite in Mao et al. (2015). Observationally, hundreds of new satellites around Milky Way-mass galaxies have been discovered in recent years, e.g., by the SAGA (Geha et al. 2017; Mao et al. 2021) and ELVES (Carlsten et al. 2022) surveys. With appropriate extensions (e.g., by performing resimulations with analytic galaxy potentials) Symphony's Milky Way-mass suite can be used to provide precise predictions for the dependence of subhalo and satellite populations on environment in a narrow host halo mass range, thereby reducing theoretical uncertainties relevant for interpreting these observed satellite systems.

Constrained realizations of Milky Way-mass halos that satisfy additional "Milky Way-like" constraints, including a quiescent merger history following a *Gaia*-Enceladus-like major merger at  $z \approx 2$  (e.g., see Helmi 2020 for a review) and the infall of an LMC analog at late times, are now being developed using zoom-in simulations (D. Buch et al. 2023, in preparation) and have recently been identified in large-volume cosmological simulations (e.g., Evans et al. 2020). Comparing such constrained realizations of Milky Way-like systems to Symphony's Milky Way-mass hosts, which generally do not satisfy specific observational constraints related to the formation history of the Milky Way, will help place the Milky Way's dark matter structure and substructure in a cosmological context. These pursuits are timely given significant ongoing advances in our understanding of the Milky Way's formation history (e.g., Helmi et al. 2018; Naidu et al. 2020) and satellite population (e.g., Drlica-Wagner et al. 2020; Nadler et al. 2020b).

### 7.3. Group Suite

Like the LMC-mass suite, Symphony’s Group suite of strong lens analogs contains a large sample of high-resolution zoom-in simulations in a host halo mass regime that is relatively under-explored (however, see, e.g., Fiacconi et al. 2016, 2017; Despali et al. 2019; Richings et al. 2021). Accurate predictions for the substructure of  $\approx 10^{13} M_{\odot}$  hosts are increasingly important due to recent observations that probe low-mass subhalos of early-type galaxy strong lenses, either statistically or on an individual basis (e.g., Vegetti et al. 2012; Hezaveh et al. 2016; Hsueh et al. 2020; Gilman et al. 2020a,b). Furthermore, the sample of strong lenses available for high-resolution follow-up is expected to increase substantially in the next decade (e.g., Oguri & Marshall 2010; Weiner et al. 2020).

Predictions for the substructure of Group hosts used to facilitate current lensing studies often rely on semianalytic modeling (e.g., using GALACTICUS); however, the precise dependence of subhalo populations on host halo mass, accretion history, environment, etc. has not been characterized in this regime using a statistical sample of high-resolution simulations. We therefore anticipate that Symphony’s Group suite (and extensions thereof) will inform these studies, particularly because the hosts were selected in a narrow mass range at a redshift typical of observed strong lenses ( $z=0.5$ ). As an example, Wagner-Carena et al. (2023) compared their lens substructure model to predictions from Symphony’s Group suite in order to interpret its constraining power.

### 7.4. Low-mass Cluster and Cluster Suites

We discuss the L-Cluster and Cluster suites together because their average host halo masses only differ by a factor of a few. Symphony’s L-Cluster suite was originally presented in Bhattacharyya et al. (2022). These authors also presented a resimulation of the L-Cluster suite in a self-interacting dark matter model to study the impact of self-interactions on cluster subhalos, which is important to quantify given that upcoming cluster weak lensing measurements can potentially distinguish between CDM and self-interacting dark matter models (e.g., Banerjee et al. 2020).

Symphony’s Cluster suite was originally presented in Wu et al. (2013c,b) and has been used to investigate: (1) galaxy cluster virial scaling relations (Wu et al. 2013a), (2) the structural properties and formation history of cluster-size halos (Wu et al. 2013c), (3) tidal stripping of cluster subhalos (Wu et al. 2013b), (4) the connection between hot gas and galaxy mass (Wu et al. 2015), (5) baryonic growth and metal enrichment in clusters (Martizzi et al. 2016), (6) cool cores in galaxy clusters (Hahn et al. 2017), (7) the dark matter velocity distribution in clusters (Mao et al. 2013), (8) assembly bias for cluster halos (Mao et al. 2018), and (9) the impact of subhalos on hosts’ measured dark matter density profiles (Fielder et al. 2020).

As in our discussion of the Milky Way suite, the studies listed above simply illustrate the breadth of potential science cases for Symphony’s L-Cluster and Cluster suites. Similar studies that explicitly compare the L-Cluster and Cluster

suites may yield additional insights (e.g., by quantifying how the results depend on cluster mass). Below, we describe a few new directions for future work using these suites enabled by recent theoretical developments and upcoming cluster substructure measurements.

Observationally, estimates of the cluster mass–richness relation have improved over the last few years (e.g., McClintock et al. 2019; Murata et al. 2019). Interpreting these observations to constrain the galaxy–halo connection or cosmological parameters (e.g., To et al. 2020, 2021) requires accurate predictions for the dark matter content of observed clusters. We anticipate that Symphony’s L-Cluster and Cluster suites, and particularly their combination with UNIVERSEMACHINE, will enable a more complete understanding of the galaxy populations surrounding massive clusters, which directly inform these science cases. These predictions are timely in anticipation of upcoming wide-field optical surveys, including the Rubin Observatory Legacy Survey of Space and Time (Ivezic et al. 2008) and the Nancy Grace Roman Space Telescope High Latitude Wide Area Survey (Eifler et al. 2021), which will simultaneously and accurately measure cluster richness and gravitational lensing signal. Robustly inferring cluster masses from richness-selected samples remains a major systematic uncertainty for these upcoming cluster cosmology analyses. Our high-resolution cluster suites enable self-consistent studies of clusters’ predicted galaxy content and gravitational lensing signatures, and can therefore be used to assess optical cluster selection biases, e.g., related to halo triaxiality and concentration (Wu et al. 2022; Zhang et al. 2022).

Meanwhile, a growing body of literature has studied the “splashback” boundaries of dark matter halos using simulations (e.g., Adhikari et al. 2014; Diemer & Kravtsov 2014; More et al. 2015), and the splashback boundary has been observed and used to constrain models of galaxy evolution (e.g., Shin et al. 2019; Adhikari et al. 2021). Because halo populations are well resolved out to many times the host’s virial radius, Symphony can be used to quantify the detailed nature of splashback boundaries and halo populations, including their dependence on host properties and environment. Finally, we note that recent measurements have also been used to set limits on the fraction of “orphan” satellites, whose dark matter subhalos are heavily stripped in the cluster environment (e.g., Kumar et al. 2022). Symphony resolves the detailed tidal stripping of cluster subhalos, which can be used to inform such measurements and empirical models for the contribution of surviving and disrupted satellites to intracluster light (e.g., Behroozi et al. 2013a, 2019).

### 7.5. Science Cases for the Entire Symphony Compilation

To conclude this Section, we highlight a few key questions that the Symphony compilation as a whole can be used to address. This list is not exhaustive, and instead simply demonstrates the breadth of science enabled by the entire Symphony compilation.

1. How do the secondary and environmental properties of host halos above and below the redshift-dependent collapse mass, shape their subhalo populations?
2. To what extent is the evolution of subhalo populations self-similar? How do the normalization and slope of the SHMF evolve, and how does this evolution relate to host halos' formation histories in detail?
3. What are disrupted subhalos' population statistics as a function of host mass, secondary properties, redshift, and environment?
4. How do splashback mass functions and halo correlation statistics in the one to two-halo transition regime depend on host halo mass and environment?
5. What aspects of halo and subhalo populations are not currently captured by semianalytic and/or empirical models of structure and galaxy formation?

Finally, we emphasize that Symphony data products are made publicly available at <http://web.stanford.edu/group/gfc/symphony>, following in the footsteps of previous cosmological simulation data releases (e.g., Lemson & Virgo Consortium 2006; Riebe et al. 2011; Skillman et al. 2014; Heitmann et al. 2019; Ishiyama et al. 2021; Villaescusa-Navarro et al. 2022). Furthermore, Symphony's public data and codebase is designed to enable modular and efficient analysis. For example, our publicly available merger trees are provided in binary, depth-first format, and we provide a number of tutorials that demonstrate Symphony's key functionality. These features differentiate Symphony from high-resolution simulations of large cosmological volumes, some of which contain comparable numbers of Group, L-Cluster, and/or Cluster-mass hosts. Furthermore, to our knowledge, no existing cosmological simulations contain comparable numbers of LMC and/or Milky Way-mass halos that are resolved with  $\approx 10^6$  particles compared to Symphony's 39 LMC-mass hosts and 45 Milky Way-mass hosts.

## 8. CONCLUSIONS

We have presented Symphony, a compilation of 262 cosmological zoom-in simulations spanning host halo masses from  $10^{11}$ – $10^{15} M_\odot$ . The key features of the Symphony compilation include the following:

1. Symphony contains a statistical sample of 262 zoom-in simulations over four decades of host halo mass, from  $10^{11}$ – $10^{15} M_\odot$ ;
2. Symphony includes two new suites of LMC-mass and strong lens analog Group zoom-ins, with 39 and 49 hosts in narrow mass ranges around  $10^{11} M_\odot$  and  $10^{13} M_\odot$ , respectively;
3. Symphony resolves halos and subhalos at similar, high resolution across all suites;

4. Symphony includes extensive convergence tests, including 15 higher-resolution resimulations (five for each of the LMC, Milky Way, and Group suites), and 96 lower-resolution resimulations (one for every Cluster zoom-in);
5. All Symphony halo catalogs and merger trees are publicly available at <http://web.stanford.edu/group/gfc/symphony>.

The key results from the first Symphony analysis presented in this paper include the following:

1. Symphony host halo concentrations are consistent with cosmological simulations after environmental selection effects are accounted for (Figure 3), and their secondary properties correlate with subhalo abundance, with correlation coefficients that peak at the Milky Way host halo mass scale (Figure 4);
2. Subhalo mass functions are approximately self-similar, and scale linearly with host halo mass; at fixed sub-to-host halo mass ratio, higher-mass Symphony hosts' subhalo abundance is approximately two times higher compared to lower-mass hosts (Figure 7);
3. The normalized radial distribution of Symphony subhalos depends weakly on host halo mass, and they are systematically less concentrated than hosts' underlying dark matter density profiles (Figure 8);
4. Semianalytic predictions for subhalo populations from GALACTICUS are consistent with Symphony subhalo mass functions and radial distributions within the  $2\sigma$  Poisson error on the mean for  $M_{\text{sub}}/M_{\text{host}} > 10^{-3}$  across our full range of host halo masses; however, GALACTICUS predicts  $\approx 25\%$  higher subhalo abundances at lower sub-to-host halo mass ratios as well as more concentrated radial distributions (Figure 10);
5. By applying the empirical galaxy–halo connection model UNIVERSEMACHINE to all Symphony simulations, we predict the SFHs of galaxies occupying all halos and subhalos in our zoom-in regions (Figure 12); these predictions statistically capture the correlated dark matter and stellar growth for systems with luminosities comparable to all currently observable satellite galaxies in the universe.

We envision Symphony as a foundation for a range of dark matter science that depends on the detailed evolution of halos and subhalos over a wide dynamic range in a cosmological context. For example, extensions of Symphony suites in alternative dark matter models and/or cosmologies, and including analytic or hydrodynamic treatments of baryonic physics, will be helpful for ongoing efforts to compare substructure observations with theoretical predictions, including in relatively under-explored regimes using Symphony's new LMC-mass and strong lens analog Group suites. More generally,

extending Symphony to other mass scales and cosmic environments while maintaining its statistical and self-consistent ethos will enable a more complete understanding of structure formation.

### ACKNOWLEDGMENTS

We are grateful to Matt Becker for providing the c125-1024 and c125-2048 simulations, Benedikt Diemer for providing data from the Erebus simulations, Alexander Ji for providing the Caterpillar parent box, Ralf Kaehler for assisting with the visualizations in Figure 2, and Bryné Hadnott for helpful discussions and contributions to the public website and notebooks. We thank Tom Abel, Simon Birrer, Deveshi Buch, Fangzhou Jiang, and Shengqi Yang for helpful discussions related to this work.

This research was supported in part by the Kavli Institute for Particle Astrophysics and Cosmology at Stanford University and SLAC National Accelerator Laboratory, and from the U.S. Department of Energy under contract No. DE-AC02-76SF00515 to SLAC National Accelerator Laboratory. This research also received support from the National Science Foundation under grant No. NSF DGE-1656518 through the NSF Graduate Research Fellowship received by E.O.N. A.B. and X.D. acknowledge support from NASA ATP grant 17-ATP17-0120. Y.-Y.M. was supported

by NASA through the NASA Hubble Fellowship grant No. HST-HF2-51441.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. H.-Y.W. is supported by DOE grant DE-SC0021916 and NASA grant 15-WFIRST15-0008. Part of this work was performed at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611.

This research made use of computational resources at SLAC National Accelerator Laboratory, a U.S. Department of Energy Office, and the Sherlock cluster at the Stanford Research Computing Center (SRCC); the authors are thankful for the support of the SLAC and SRCC computing teams. Additional computing resources used in this work were made available by a generous grant from the Ahmanson Foundation.

This research used <https://arXiv.org> and NASA's Astrophysics Data System for bibliographic information.

*Software:* Colossus (Diemer 2018), CONSISTENT-TREES (Behroozi et al. 2013c), GADGET-2 (Springel 2005), Helpers ([bitbucket.org/yymao/helpers/src/master/](https://bitbucket.org/yymao/helpers/src/master/)), Jupyter ([jupyter.org](https://jupyter.org)), Matplotlib (Hunter 2007), NumPy (van der Walt et al. 2011), Pynbody (Pontzen et al. 2013), ROCKSTAR (Behroozi et al. 2013b) SciPy (Virtanen et al. 2020), Seaborn ([seaborn.pydata.org](https://seaborn.pydata.org)).

### REFERENCES

- Abbott, T. M. C., Abdalla, F. B., Alarcon, A., et al. 2018, *PhRvD*, **98**, 043526
- Adhikari, S., Dalal, N., & Chamberlain, R. T. 2014, *JCAP*, **2014**, 019
- Adhikari, S., Shin, T.-h., Jain, B., et al. 2021, *ApJ*, **923**, 37
- Atek, H., Richard, J., Kneib, J.-P., & Schaerer, D. 2018, *MNRAS*, **479**, 5184
- Auger, M. W., Treu, T., Bolton, A. S., et al. 2010, *ApJ*, **724**, 511
- Banerjee, A., Adhikari, S., Dalal, N., More, S., & Kravtsov, A. 2020, *JCAP*, **2020**, 024
- Banerjee, A., Boddy, K. K., Cyr-Racine, F.-Y., et al. 2022, *arXiv e-prints*, arXiv:2203.07049
- Barber, C., Starkenburg, E., Navarro, J. F., McConnachie, A. W., & Fattahi, A. 2014, *MNRAS*, **437**, 959
- Barnes, D. J., Kay, S. T., Bahé, Y. M., et al. 2017, *MNRAS*, **471**, 1088
- Bechtol, K., Birrer, S., Cyr-Racine, F.-Y., et al. 2022, *arXiv e-prints*, arXiv:2203.07354
- Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, **488**, 3143
- Behroozi, P., Conroy, C., Wechsler, R. H., et al. 2020, *MNRAS*, **499**, 5702
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013a, *ApJ*, **770**, 57
- Behroozi, P. S., Wechsler, R. H., Lu, Y., et al. 2014, *ApJ*, **787**, 156
- Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013b, *ApJ*, **762**, 109
- Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., et al. 2013c, *ApJ*, **763**, 18
- Benítez-Llambay, A. & Frenk, C. 2020, *MNRAS*, **498**, 4887
- Benson, A. J. 2012, *NewA*, **17**, 175
- Benson, A. J. 2017, *MNRAS*, **467**, 3454
- Benson, A. J. 2020, *MNRAS*, **493**, 1268
- Benson, A. J. & Du, X. 2022, *MNRAS*, **517**, 1398
- Benson, A. J., Ludlow, A., & Cole, S. 2019, *MNRAS*, **485**, 5010
- Berrier, J. C., Stewart, K. R., Bullock, J. S., et al. 2009, *ApJ*, **690**, 1292
- Bertschinger, E. 2001, *ApJS*, **137**, 1
- Bhattacharyya, S., Adhikari, S., Banerjee, A., et al. 2022, *ApJ*, **932**, 30
- Bland-Hawthorn, J. & Gerhard, O. 2016, *ARA&A*, **54**, 529
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *ApJ*, **803**, 34
- Brown, T. M., Tumlinson, J., Geha, M., et al. 2014, *ApJ*, **796**, 91
- Bryan, G. L. & Norman, M. L. 1998, *ApJ*, **495**, 80
- Callingham, T. M., Cautun, M., Deason, A. J., et al. 2019, *MNRAS*, **484**, 5453
- Campbell, D., van den Bosch, F. C., Padmanabhan, N., et al. 2018, *MNRAS*, **477**, 359

- Carlesi, E., Sorce, J. G., Hoffman, Y., et al. 2016, *MNRAS*, **458**, 900
- Carlsten, S. G., Greene, J. E., Beaton, R. L., Danieli, S., & Greco, J. P. 2022, *ApJ*, **933**, 47
- Chen, L.-H., Magg, M., Hartwig, T., et al. 2022, *MNRAS*, **513**, 934
- Chue, C. Y. R., Dalal, N., & White, M. 2018, *JCAP*, **2018**, 012
- Conroy, C. & Wechsler, R. H. 2009, *ApJ*, **696**, 620
- Correa, C. A., Wyithe, J. S. B., Schaye, J., & Duffy, A. R. 2015, *MNRAS*, **450**, 1514
- Crnojević, D., Sand, D. J., Bennet, P., et al. 2019, *ApJ*, **872**, 80
- Cui, W., Knebe, A., Yepes, G., et al. 2018, *MNRAS*, **480**, 2898
- Deason, A. J., Mao, Y.-Y., & Wechsler, R. H. 2016, *ApJ*, **821**, 5
- Despali, G., Sparre, M., Vegetti, S., et al. 2019, *MNRAS*, **484**, 4563
- Diemand, J., Kuhlen, M., Madau, P., et al. 2008, *Nature*, **454**, 735
- Diemer, B. 2018, *ApJS*, **239**, 35
- Diemer, B. & Joyce, M. 2019, *ApJ*, **871**, 168
- Diemer, B. & Kravtsov, A. V. 2014, *ApJ*, **789**, 1
- Diemer, B. & Kravtsov, A. V. 2015, *ApJ*, **799**, 108
- Dooley, G. A., Peter, A. H. G., Carlin, J. L., et al. 2017, *MNRAS*, **472**, 1060
- Drlica-Wagner, A., Bechtol, K., Mau, S., et al. 2020, *ApJ*, **893**, 47
- D'Souza, R. & Bell, E. F. 2021, *MNRAS*, **504**, 5270
- Dutton, A. A., Macciò, A. V., Frings, J., et al. 2016, *MNRAS*, **457**, L74
- Eifler, T., Miyatake, H., Krause, E., et al. 2021, *MNRAS*, **507**, 1746
- Erkal, D., Belokurov, V., Laporte, C. F. P., et al. 2019, *MNRAS*, **487**, 2685
- Errani, R. & Navarro, J. F. 2021, *MNRAS*, **505**, 18
- Errani, R., Navarro, J. F., Ibata, R., & Peñarrubia, J. 2022, *MNRAS*, **511**, 6001
- Evans, T. A., Fattahi, A., Deason, A. J., & Frenk, C. S. 2020, *MNRAS*, **497**, 4311
- Ferrarese, L., Côté, P., Sánchez-Janssen, R., et al. 2016, *ApJ*, **824**, 10
- Fiacconi, D., Madau, P., Potter, D., & Stadel, J. 2016, *ApJ*, **824**, 144
- Fiacconi, D., Mayer, L., Madau, P., et al. 2017, *MNRAS*, **467**, 4080
- Fielder, C. E., Mao, Y.-Y., Newman, J. A., Zentner, A. R., & Licquia, T. C. 2019, *MNRAS*, **486**, 4545
- Fielder, C. E., Mao, Y.-Y., Zentner, A. R., et al. 2020, *MNRAS*, **499**, 2426
- Finkelstein, S. L., Ryan, Russell E., J., Papovich, C., et al. 2015, *ApJ*, **810**, 71
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, **616**, A1
- Gao, L., Frenk, C. S., Boylan-Kolchin, M., et al. 2011, *MNRAS*, **410**, 2309
- Gao, L., Navarro, J. F., Frenk, C. S., et al. 2012, *MNRAS*, **425**, 2169
- Garrison-Kimmel, S., Boylan-Kolchin, M., Bullock, J. S., & Lee, K. 2014, *MNRAS*, **438**, 2578
- Garrison-Kimmel, S., Hopkins, P. F., Wetzel, A., et al. 2019, *MNRAS*, **487**, 1380
- Gavazzi, R., Treu, T., Rhodes, J. D., et al. 2007, *ApJ*, **667**, 176
- Geha, M., Wechsler, R. H., Mao, Y.-Y., et al. 2017, *ApJ*, **847**, 4
- Gilman, D., Birrer, S., Nierenberg, A., et al. 2020a, *MNRAS*, **491**, 6077
- Gilman, D., Du, X., Benson, A., et al. 2020b, *MNRAS*, **492**, L12
- Giocoli, C., Tormen, G., & van den Bosch, F. C. 2008, *MNRAS*, **386**, 2135
- Graus, A. S., Bullock, J. S., Kelley, T., et al. 2019, *MNRAS*, **488**, 4585
- Green, S. B., van den Bosch, F. C., & Jiang, F. 2021, *MNRAS*, **503**, 4075
- Griffen, B. F., Ji, A. P., Dooley, G. A., et al. 2016, *ApJ*, **818**, 10
- Guedes, J., Callegari, S., Madau, P., & Mayer, L. 2011, *ApJ*, **742**, 76
- Hahn, O. & Abel, T. 2011, *MNRAS*, **415**, 2101
- Hahn, O., Martizzi, D., Wu, H.-Y., et al. 2017, *MNRAS*, **470**, 166
- Hearin, A. P., Chaves-Montero, J., Becker, M. R., & Alarcon, A. 2021, *OJAp*, **4**, 7
- Heitmann, K., Uram, T. D., Finkel, H., et al. 2019, *ApJS*, **244**, 17
- Hellwing, W. A., Frenk, C. S., Cautun, M., et al. 2016, *MNRAS*, **457**, 3492
- Helmi, A. 2020, *ARA&A*, **58**, 205
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Nature*, **563**, 85
- Hezaveh, Y. D., Dalal, N., Marrone, D. P., et al. 2016, *ApJ*, **823**, 37
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, **208**, 19
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, *MNRAS*, **445**, 581
- Hsueh, J. W., Enzi, W., Vegetti, S., et al. 2020, *MNRAS*, **492**, 3047
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Ishiyama, T., Fukushige, T., & Makino, J. 2009, *ApJ*, **696**, 2115
- Ishiyama, T., Rieder, S., Makino, J., et al. 2013, *ApJ*, **767**, 146
- Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, *MNRAS*, **506**, 4210
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2008, *ApJ*, **873**, 111
- Jahn, E. D., Sales, L. V., Wetzel, A., et al. 2019, *MNRAS*, **489**, 5348
- Jenkins, A. 2010, *MNRAS*, **403**, 1859
- Jiang, F. & van den Bosch, F. C. 2016, *MNRAS*, **458**, 2848
- Jing, Y. P. & Suto, Y. 2002, *ApJ*, **574**, 538
- Kaehler, R. 2017, *A&C*, **20**, 68
- Kaehler, R. 2018, *Electronic Imaging*, **2018**, 333
- Kallivayalil, N., Sales, L. V., Zivick, P., et al. 2018, *ApJ*, **867**, 19
- Karachentsev, I. D. 2005, *ApJ*, **129**, 178
- Katz, N. & White, S. D. M. 1993, *ApJ*, **412**, 455

- Kelley, T., Bullock, J. S., Garrison-Kimmel, S., et al. 2019, *MNRAS*, **487**, 4409
- Klypin, A., Gottlöber, S., Kravtsov, A. V., & Khokhlov, A. M. 1999, *ApJ*, **516**, 530
- Klypin, A., Yepes, G., Gottlöber, S., Prada, F., & Heß, S. 2016, *MNRAS*, **457**, 4340
- Kravtsov, A. 2010, *AdAst*, **2010**, 281913
- Kravtsov, A. & Manwadkar, V. 2022, *MNRAS*, **514**, 2667
- Kumar, A., More, S., & Rana, D. 2022, *MNRAS*, **517**, 4389
- Lemson, G. & Virgo Consortium, t. 2006, arXiv e-prints, astro
- Lu, Y., Benson, A., Mao, Y.-Y., et al. 2016, *ApJ*, **830**, 59
- Lu, Y., Benson, A., Wetzel, A., et al. 2017, *ApJ*, **846**, 66
- Ludlow, A. D., Bose, S., Angulo, R. E., et al. 2016, *MNRAS*, **460**, 1214
- Ludlow, A. D., Schaye, J., & Bower, R. 2019, *MNRAS*, **488**, 3663
- Ludlow, A. D., Navarro, J. F., Boylan-Kolchin, M., et al. 2013, *MNRAS*, **432**, 1103
- Mansfield, P. & Avestruz, C. 2021, *MNRAS*, **500**, 3309
- Mansfield, P. & Kravtsov, A. V. 2020, *MNRAS*, **493**, 4763
- Manwadkar, V. & Kravtsov, A. V. 2022, *MNRAS*, **516**, 3944
- Mao, Y.-Y., Geha, M., Wechsler, R. H., et al. 2021, *ApJ*, **907**, 85
- Mao, Y.-Y., Strigari, L. E., Wechsler, R. H., Wu, H.-Y., & Hahn, O. 2013, *ApJ*, **764**, 35
- Mao, Y.-Y., Williamson, M., & Wechsler, R. H. 2015, *ApJ*, **810**, 21
- Mao, Y.-Y., Zentner, A. R., & Wechsler, R. H. 2018, *MNRAS*, **474**, 5143
- Martizzi, D., Hahn, O., Wu, H.-Y., et al. 2016, *MNRAS*, **459**, 4408
- Mau, S., Nadler, E. O., Wechsler, R. H., et al. 2022, *ApJ*, **932**, 128
- McAlpine, S., Helly, J. C., Schaller, M., et al. 2022, *MNRAS*, **512**, 5823
- McBride, C., Berlind, A., Scoccimarro, R., et al. 2009a, in AAS Meeting Abstracts, Vol. 213, AAS Meeting Abstracts #213, 425.06
- McBride, J., Fakhouri, O., & Ma, C.-P. 2009b, *MNRAS*, **398**, 1858
- McCall, M. L. 2014, *MNRAS*, **440**, 405
- McClintock, T., Varga, T. N., Gruen, D., et al. 2019, *MNRAS*, **482**, 1352
- Meneghetti, M., Davoli, G., Bergamini, P., et al. 2020, *Science*, **369**, 1347
- Moliné, Á., Sánchez-Conde, M. A., Palomares-Ruiz, S., & Prada, F. 2017, *MNRAS*, **466**, 4974
- Moliné, Á., Sánchez-Conde, M. A., Aguirre-Santaella, A., et al. 2023, *MNRAS*, **518**, 157
- More, S., Diemer, B., & Kravtsov, A. V. 2015, *ApJ*, **810**, 36
- Munshi, F., Brooks, A. M., Applebaum, E., et al. 2021, *ApJ*, **923**, 35
- Murata, R., Oguri, M., Nishimichi, T., et al. 2019, *PASJ*, **71**, 107
- Mutlu-Pakdil, B., Sand, D. J., Crnojević, D., et al. 2021, *ApJ*, **918**, 88
- Nadler, E. O., Banerjee, A., Adhikari, S., Mao, Y.-Y., & Wechsler, R. H. 2020a, *ApJ*, **896**, 112
- Nadler, E. O., Banerjee, A., Adhikari, S., Mao, Y.-Y., & Wechsler, R. H. 2021a, *ApJL*, **920**, L11
- Nadler, E. O., Birrer, S., Gilman, D., et al. 2021b, *ApJ*, **917**, 7
- Nadler, E. O., Gluscevic, V., Boddy, K. K., & Wechsler, R. H. 2019a, *ApJL*, **878**, L32
- Nadler, E. O., Mao, Y.-Y., Green, G. M., & Wechsler, R. H. 2019b, *ApJ*, **873**, 34
- Nadler, E. O., Mao, Y.-Y., Wechsler, R. H., Garrison-Kimmel, S., & Wetzel, A. 2018, *ApJ*, **859**, 129
- Nadler, E. O., Wechsler, R. H., Bechtol, K., et al. 2020b, *ApJ*, **893**, 48
- Nadler, E. O., Drlica-Wagner, A., Bechtol, K., et al. 2021c, *PhRvL*, **126**, 091101
- Nagai, D. & Kravtsov, A. V. 2005, *ApJ*, **618**, 557
- Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, *ApJ*, **901**, 48
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, **490**, 493
- Navarro, J. F., Hayashi, E., Power, C., et al. 2004, *MNRAS*, **349**, 1039
- Navarro, J. F., Ludlow, A., Springel, V., et al. 2010, *MNRAS*, **402**, 21
- 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
- Nierenberg, A. M., Gilman, D., Treu, T., et al. 2020, *MNRAS*, **492**, 5314
- Noordeh, E., Canning, R. E. A., Willis, J. P., et al. 2021, *MNRAS*, **507**, 5272
- Oñorbe, J., Garrison-Kimmel, S., Maller, A. H., et al. 2014, *MNRAS*, **437**, 1894
- Ogiya, G., van den Bosch, F. C., Hahn, O., et al. 2019, *MNRAS*, **485**, 189
- Oguri, M. & Marshall, P. J. 2010, *MNRAS*, **405**, 2579
- Patel, E., Kallivayalil, N., Garavito-Camargo, N., et al. 2020, *ApJ*, **893**, 121
- Pontzen, A., Roškar, R., Stinson, G., & Woods, R. 2013, pynbody: N-Body/SPH analysis for python, Astrophysics Source Code Library, record ascl:1305.002
- Poole-McKenzie, R., Font, A. S., Boxer, B., et al. 2020, *JCAP*, **2020**, 016
- Powell, D., Laha, R., Ng, K. C. Y., & Abel, T. 2017, *PhRvD*, **95**, 063012
- Pullen, A. R., Benson, A. J., & Moustakas, L. A. 2014, *ApJ*, **792**, 24
- Read, J. I. & Erkal, D. 2019, *MNRAS*, **487**, 5799
- Reddick, R. M., Wechsler, R. H., Tinker, J. L., & Behroozi, P. S. 2013, *ApJ*, **771**, 30
- Rey, M. P., Pontzen, A., Agertz, O., et al. 2019, *ApJL*, **886**, L3

- Richings, J., Frenk, C., Jenkins, A., Robertson, A., & Schaller, M. 2021, *MNRAS*, 501, 4657
- Richings, J., Frenk, C., Jenkins, A., et al. 2020, *MNRAS*, 492, 5780
- Riebe, K., Partl, A. M., Enke, H., et al. 2011, arXiv e-prints, arXiv:1109.0003
- Rodríguez-Puebla, A., Behroozi, P., Primack, J., et al. 2016, *MNRAS*, 462, 893
- Sacchi, E., Richstein, H., Kallivayalil, N., et al. 2021, *ApJL*, 920, L19
- Samuel, J., Wetzel, A., Tollerud, E., et al. 2020, *MNRAS*, 491, 1471
- Sawala, T., Frenk, C. S., Fattahi, A., et al. 2016, *MNRAS*, 457, 1931
- Schwabe, B. & Niemeyer, J. C. 2022, *PhRvL*, 128, 181301
- Sheth, R. K. & Tormen, G. 1999, *MNRAS*, 308, 119
- Shin, T., Adhikari, S., Baxter, E. J., et al. 2019, *MNRAS*, 487, 2900
- Shipp, N., Erkal, D., Drlica-Wagner, A., et al. 2021, *ApJ*, 923, 149
- Simon, J. D. 2019, *ARA&A*, 57, 375
- Skillman, S. W., Warren, M. S., Turk, M. J., et al. 2014, arXiv e-prints, arXiv:1407.2600
- Springel, V. 2005, *MNRAS*, 364, 1105
- Springel, V., Wang, J., Vogelsberger, M., et al. 2008, *MNRAS*, 391, 1685
- Starkenburg, E., Helmi, A., De Lucia, G., et al. 2013, *MNRAS*, 429, 725
- Tasitsiomi, A., Kravtsov, A. V., Gottlöber, S., & Klypin, A. A. 2004, *ApJ*, 607, 125
- To, C., Krause, E., Rozo, E., et al. 2021, *PhRvL*, 126, 141301
- To, C.-H., Reddick, R. M., Rozo, E., Rykoff, E., & Wechsler, R. H. 2020, *ApJ*, 897, 15
- van den Bosch, F. C. 2002, *MNRAS*, 331, 98
- van den Bosch, F. C. & Jiang, F. 2016, *MNRAS*, 458, 2870
- van den Bosch, F. C., Jiang, F., Hearin, A., et al. 2014, *MNRAS*, 445, 1713
- 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 der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, 13, 22
- Vegetti, S., Lagattuta, D. J., McKean, J. P., et al. 2012, *Nature*, 481, 341
- Venholo, A., Peletier, R., Laurikainen, E., et al. 2019, *A&A*, 625, A143
- Villaescusa-Navarro, F., Genel, S., Anglés-Alcázar, D., et al. 2022, arXiv e-prints, arXiv:2201.01300
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261
- Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2020, *NatRP*, 2, 42
- Wagner-Carena, S., Aalbers, J., Birrer, S., et al. 2023, *ApJ*, 942, 75
- Wang, J., Bose, S., Frenk, C. S., et al. 2020a, *Nature*, 585, 39
- Wang, K., Mao, Y.-Y., Zentner, A. R., et al. 2020b, *MNRAS*, 498, 4450
- Wang, L., Dutton, A. A., Stinson, G. S., et al. 2015, *MNRAS*, 454, 83
- Wang, Y., Nadler, E. O., Mao, Y.-Y., et al. 2021, *ApJ*, 915, 116
- Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, *ApJ*, 568, 52
- Wechsler, R. H. & Tinker, J. L. 2018, *ARAA*, 56, 435
- Weiner, C., Serjeant, S., & Sedgwick, C. 2020, *RNAAS*, 4, 190
- Weisz, D. R., Dolphin, A. E., Skillman, E. D., et al. 2014a, *ApJ*, 789, 147
- Weisz, D. R., Dolphin, A. E., Skillman, E. D., et al. 2014b, *ApJ*, 789, 148
- Weisz, D. R., Dolphin, A. E., Skillman, E. D., et al. 2015, *ApJ*, 804, 136
- Wetzel, A. R., Deason, A. J., & Garrison-Kimmel, S. 2015, *ApJ*, 807, 49
- Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, *ApJL*, 827, L23
- Woodley, K. A. 2006, *AJ*, 132, 2424
- Wu, H.-Y., Evrard, A. E., Hahn, O., et al. 2015, *MNRAS*, 452, 1982
- Wu, H.-Y., Hahn, O., Evrard, A. E., Wechsler, R. H., & Dolag, K. 2013a, *MNRAS*, 436, 460
- Wu, H.-Y., Hahn, O., Wechsler, R. H., Behroozi, P. S., & Mao, Y.-Y. 2013b, *ApJ*, 767, 23
- Wu, H.-Y., Hahn, O., Wechsler, R. H., Mao, Y.-Y., & Behroozi, P. S. 2013c, *ApJ*, 763, 70
- Wu, H.-Y., Costanzi, M., To, C.-H., et al. 2022, *MNRAS*, 515, 4471
- Yamanoi, H., Komiyama, Y., Yagi, M., et al. 2012, *AJ*, 144, 40
- Yang, S., Du, X., Benson, A. J., Pullen, A. R., & Peter, A. H. G. 2020, *MNRAS*, 498, 3902
- Zavala, J. & Frenk, C. S. 2019, *Galaxies*, 7, 81
- Zentner, A. R., Berlind, A. A., Bullock, J. S., Kravtsov, A. V., & Wechsler, R. H. 2005, *ApJ*, 624, 505
- Zhang, Z., Wu, H.-Y., Zhang, Y., et al. 2022, arXiv e-prints, arXiv:2202.08211

## APPENDIX

## A. FIGURE 1 REFERENCES

Apart from Symphony, Figure 1 shows the following cosmological zoom-in simulations: NIHAO (Wang et al. 2015; Dutton et al. 2016), ELVIS (Garrison-Kimmel et al. 2014, 2019), Artemis (Poole-Mckenzie et al. 2020), Caterpillar (Griffen et al. 2016), Via Lactea (Diemand et al. 2008), Aquarius (Springel et al. 2008), Latte (Wetzel et al. 2016; Samuel et al. 2020), APOSTLE (Sawala et al. 2016), Ponos (Fiacconi et al. 2016, 2017), Phoenix (Gao et al. 2012), the Three Hundred Project (Cui et al. 2018), C-EAGLE (Barnes et al. 2017), and simulations from Despali et al. (2019) and Richings et al. (2020). This collection is not comprehensive, and does not include: resimulation suites (e.g., PhatELVIS; Kelley et al. 2019), zoom-ins of cosmological volumes that do not target specific hosts (e.g., Copernicus Complexio; Hellwing et al. 2016), hydrodynamic simulations of lower-mass galaxies (e.g., from FIRE; Hopkins et al. 2014), or simulations only presented in a hydrodynamic context (e.g., ERIS; Guedes et al. 2011).

## B. CONVERGENCE TESTS

This Appendix describes Symphony resimulation suites used to test for numerical convergence and the resulting convergence properties. Appendix B.1 describes our suite of zoom-in resimulations run with varying resolution, Appendix B.2 describes their convergence properties, Appendix B.3 describes resimulations with varying softening, and Appendix B.4 describes resimulations with varying time stepping.

## B.1. Resimulations with Varying Resolution

We perform high-resolution (HR) resimulations for five zoom-ins from each of our LMC, Milky Way, and Group suites, for a total of 15 resimulations. These resimulations use one additional MUSIC refinement level relative to each suite’s fiducial initial conditions; thus, the HR LMC, Milky Way, and Group resimulations use six, five, and four refinement regions, yielding an equivalent of 32,768, 16,384, and 8192 particles per side in the most refined regions, respectively. The corresponding particle mass in the highest-resolution regions of these resimulations is a factor of 8 lower than the fiducial value for each suite, yielding  $m_{\text{part,HR}}/M_{\odot} = 6.3 \times 10^3$ ,  $5.0 \times 10^4$ , and  $4.1 \times 10^5$  in the highest-resolution regions for the LMC, Milky Way, and Group HR runs, respectively. The softening lengths in the highest-resolution regions are reduced by a factor of 2 relative to the fiducial values, yielding values of  $\epsilon_{\text{HR}}/(\text{pc } h^{-1}) = 40, 80$ , and  $170$ , respectively. All other simulation parameters are fixed at their fiducial values.

Furthermore, the entire Cluster suite was originally resimulated at one resolution level lower than the fiducial Cluster simulations we present (Wu et al. 2013c,b). These low-resolution (LR) resimulations were performed with a parti-

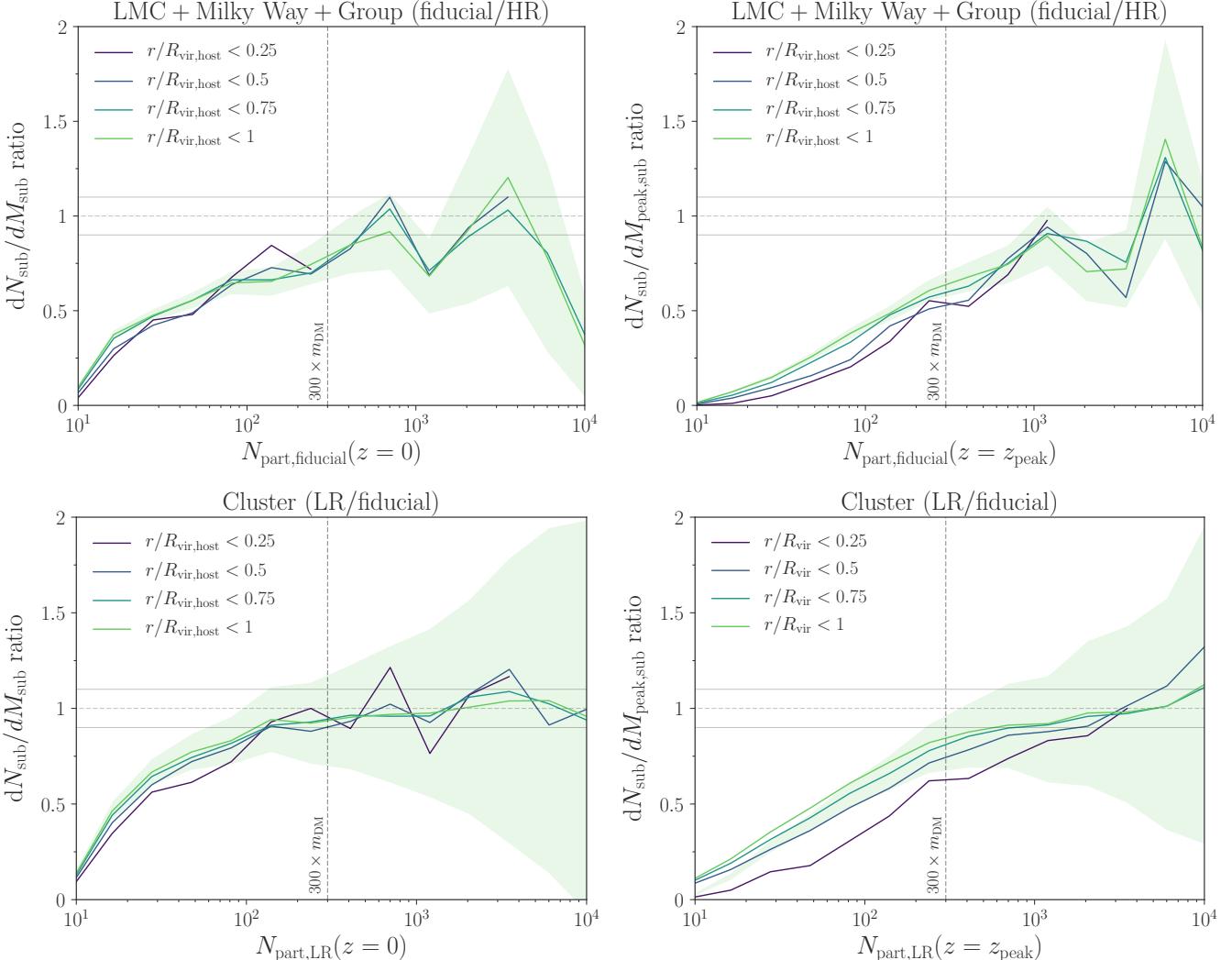
cle mass of  $m_{\text{part,LR}} = 1.4 \times 10^9 M_{\odot}$  (eight times higher than the fiducial value) and a softening of  $\epsilon_{\text{LR}} = 6687 \text{ pc } h^{-1}$  ( $\approx$  two times higher than the fiducial value). Note that, because our fiducial Cluster suite has relatively high mass resolution compared to our other suites (see Figure 1), host halo particle counts in these “low-resolution” runs only differ from those in the fiducial versions of our other suites by factors of a few.

Finally, we note that Bhattacharyya et al. (2022) resimulated one host from the L-Cluster suite at higher resolution, using a particle mass of  $m_{\text{part,HR}} = 2.8 \times 10^7 M_{\odot}$  (eight times lower than the fiducial value) and a softening of  $\epsilon_{\text{HR}} = 600 \text{ pc } h^{-1}$  (two times lower than the fiducial value). These authors find that the  $V_{\text{peak,sub}}$  function of surviving subhalos is enhanced at low  $V_{\text{peak,sub}}$  in the high-resolution resimulation, although with only a single zoom-in, it is difficult to draw strong conclusions from this test. As shown in Appendix B.2, the agreement between the convergence properties of our LMC, Milky Way, and Group resimulations compared to our Cluster convergence tests suggests that there is no significant mass trend in our zoom-ins’ convergence properties. We therefore assume that the convergence properties of the L-Cluster suite are similar to our other suites, deferring additional resimulations to future work.

## B.2. Convergence Properties with Varying Resolution

We compare simulations at varying resolution against one another in Figure 13. The top-left panel shows the ratio of differential SHMFs for our 15 HR resimulations compared against the SHMFs for those halos at fiducial resolutions. The ratio of the SHMFs is computed separately for each suite and then averaged into a combined set of ratios. The top-right panel shows the same for differential  $M_{\text{peak,sub}}$  functions, where  $M_{\text{peak,sub}}$  is the maximum virial mass that a subhalo ever achieved along its main branch; this quantity is often used in analyses where one wants to minimize the impact of subhalo disruption (e.g., see Campbell et al. 2018). The bottom panels show the same quantities measured for our LR vs. fiducial Cluster simulations. Figure 13 also shows the dependence of convergence behavior on subhalo distance from the host center,  $r/R_{\text{vir,host}}$ . Previous convergence tests have shown that convergence properties of subhalos depend on distance in idealized configurations (e.g., van den Bosch et al. 2018; van den Bosch & Ogiya 2018) and cosmological settings (e.g., Mansfield & Avestruz 2021). This is because tidal forces and subhalo mass-loss rates increase at smaller distances, leading to more demanding numerical requirements.

SHMFs are converged at the 10% level for  $N_{\text{part}} = M_{\text{sub}}/m_{\text{part}} > 300$  at  $z = 0$ . Trends with subhalo distance from the host center are weak and only reach statistical significance for low-particle count subhalos that are below the convergence limit. The trend with  $r/R_{\text{vir,host}}$  is not well characterized in the three low-mass suites due to the relatively small number of hosts tested, but the Cluster test suggests that con-



**Figure 13.** Top-left panel: ratio of the mean differential subhalo mass function, evaluated using subhalo virial mass at  $z = 0$ , in 15 fiducial-resolution simulations (five from each of the LMC, Milky Way, and Group suites), normalized to that in high-resolution (HR) resimulations of these halos. Results are plotted in terms of equivalent particle number in the fiducial-resolution runs, i.e.,  $N_{\text{part,fiducial}}(z = 0) = M_{\text{sub}}/m_{\text{part,fiducial}}$ . Ratios are shown for subhalos within 0.25 (purple), 0.5 (blue), 0.75 (blue-green), and 1.0 (green) times the host halo virial radius in each fiducial or HR simulation. The green band indicates  $1\sigma$  Poisson error on the mean for the result using all subhalos within the virial radius. Thin solid horizontal lines indicate ratios of  $\pm 10\%$ , and curves are truncated when the error on the mean exceeds 100%. The vertical dashed line corresponds to our fiducial convergence limit of  $300m_{\text{part}}$ . Top-right panel: same as the top-left panel, but for differential  $peak$  subhalo mass functions. Note that a more stringent particle number threshold of  $N_{\text{part,LR}}(z = z_{\text{peak}}) \gtrsim 1000$  is required to achieve a similar level of convergence in terms of  $M_{\text{peak,sub}}$  for subhalos at large distances from the host center. Furthermore, the convergence properties of the  $M_{\text{peak,sub}}$  function depend strongly on distance (see Appendix B.2 for details). Bottom-left panel: same as the top-left panel, but using subhalo mass functions from low-resolution (LR) resimulations of all 96 Cluster halos, normalized to those in our fiducial-resolution Cluster simulations and stacked over all hosts. Results are plotted in terms of equivalent particle number in the LR runs, i.e.,  $N_{\text{part,LR}}(z = 0) = M_{\text{sub}}/m_{\text{part,LR}}$ . Bottom-right panel: same as the bottom-left panel, but for differential peak subhalo mass functions.

vergence behavior at distances down to at least  $\approx 0.25R_{\text{vir,host}}$  is consistent with that for more distant subhalos.

SHMFs evaluated using peak virial mass show weaker convergence trends, requiring between  $\approx 700$  and  $1000$  particles for convergence at larger distances from the host center. Unlike SHMFs evaluated using  $z = 0$  masses, the convergence behavior is strongly dependent on  $r/R_{\text{vir,host}}$ , even at high

particle counts, such that subhalos closer to the host center require more particles for convergence than distant subhalos. We note that the apparent convergence of SHMFs within  $r/R_{\text{vir,host}} < 0.25$  is consistent with being a statistical fluctuation, and it is possible that a larger suite of multiresolution halos would measure stricter convergence criteria.

While entirely consistent with existing convergence tests, this radial dependence has important implications for studies of satellite radial profiles that rely on abundance matching using subhalos’ peak properties (e.g., Reddick et al. 2013) or efforts to use radial profiles to constrain the galaxy–halo connection (e.g., Graus et al. 2019). The effect of this lack of convergence will be to preferentially suppress subhalo number density profiles in the inner regions of host halos. We note that the true level of bias in subhalo profiles may be larger than indicated by Figure 13, which shows the ratio of lower-to-higher-resolution runs. This is possible because it is possible that—until strict convergence is achieved—inner radial profiles in the higher-resolution runs are also biased low.

At first it may seem counterintuitive that SHMFs evaluated using peak mass are more poorly converged than those using  $z = 0$  mass, given that  $M_{\text{peak,sub}}$  is unchanged by the exact mass-loss rate of a subhalo after infall as long as that subhalo can still be tracked by the halo finder. However, the  $M_{\text{peak,sub}}$  function is effectively a convolution of the  $M_{\text{sub}}$  function with the  $M_{\text{peak,sub}}/M_{\text{sub}}$  distribution (see Appendix C.2), and is therefore sensitive to subhalos below the  $M_{\text{sub}}$  convergence limit, even for subhalos with large peak particle counts. Furthermore, the radial dependence of convergence properties—which are only significant for low-particle-count halos when using  $z = 0$  masses—fluence the  $M_{\text{peak,sub}}$  function at particle counts where mass functions appear converged for all subhalos within  $R_{\text{vir,host}}$ . The radial dependence of  $M_{\text{peak,sub}}$  convergence is unlikely to be a problem specific to our zoom-in suites; this behavior has been noted in earlier convergence studies (van den Bosch et al. 2018; van den Bosch & Ogiya 2018; Mansfield & Avestruz 2021), and, as explained above, is expected from first principles.

### B.3. Resimulations with Varying Softening

We perform additional resimulations for three host halos in the Milky Way suite using the fiducial initial conditions and particle mass but varying the softening length. In particular, these resimulations adopt 0.25, 0.5, 2, and 4 times the fiducial softening length of  $\epsilon_{\text{fiducial}} = 170 \text{ pc } h^{-1}$  while holding all other parameters fixed. The top-left panel of Figure 14 demonstrates that SHMFs evaluated using  $z = 0$  subhalo mass in these runs are consistent with our fiducial results at the  $\approx 10\%$  level down to the convergence limit of  $300m_{\text{part}}$ .

These results are broadly consistent with similar tests performed in Mansfield & Avestruz (2021) that studied the impact of force softening on the 50 largest clusters in c125-1024. The average mass of this sample was  $\approx 10^{14} M_{\odot}$ , leading to  $N_{\text{part}} \approx 7 \times 10^5$  and resolved sub-to-host halo mass ratios of  $M_{\text{sub}}/M_{\text{host}} \approx 4 \times 10^{-4}$ , on average. Thus, our tests have better mass resolution but much less statistical power. They also probe force softening scales that are four times smaller relative to the mean interparticle spacing,  $L/N$ , than the tests in Mansfield & Avestruz (2021). These authors showed that variations in  $\epsilon$  only begin to suppress SHMFs significantly at very large  $\epsilon \approx 0.12 \times L/N$ , similar to softenings that lead to the historical “overmerging” problem in

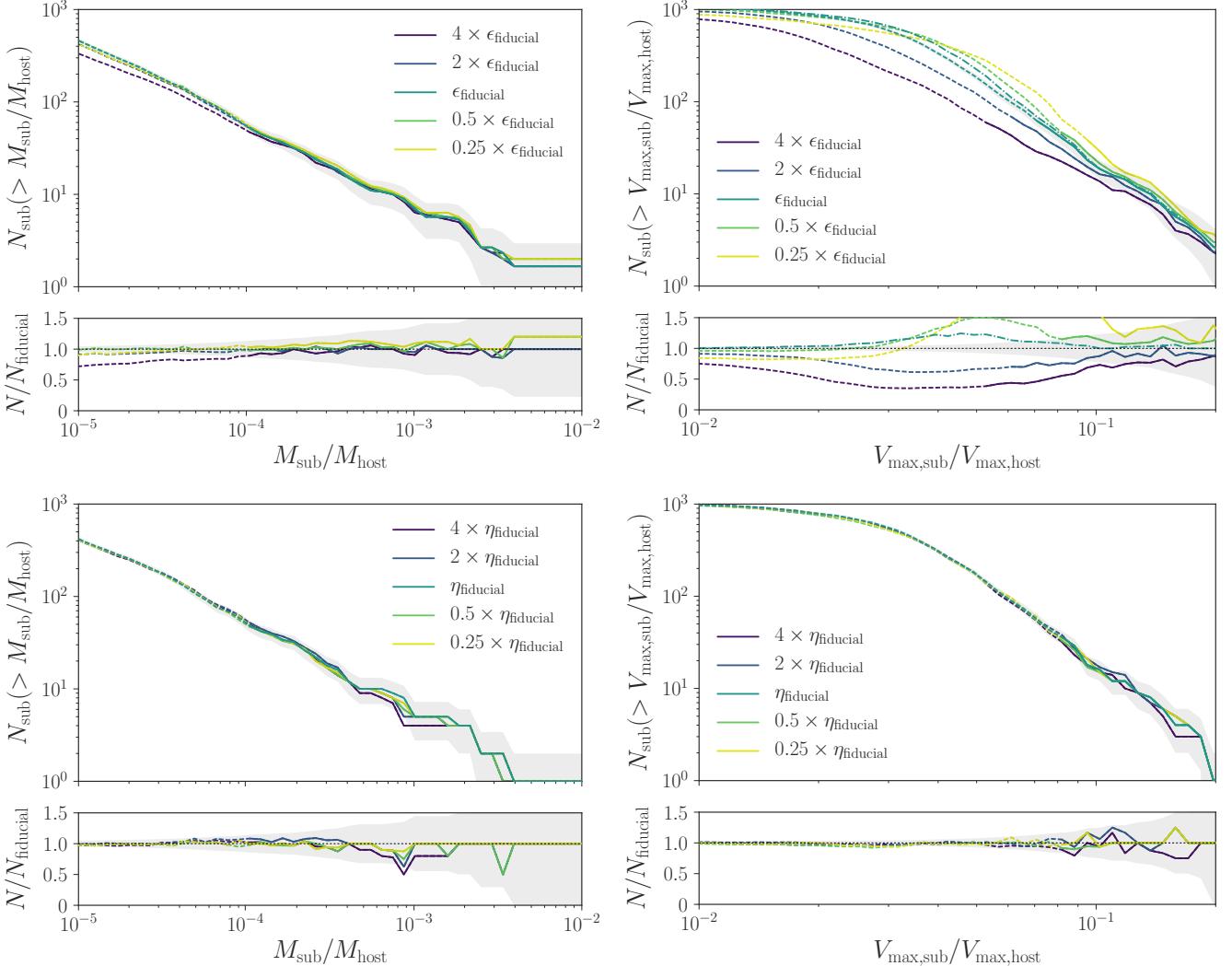
$N$ -body simulations (see discussion in, e.g., Klypin et al. 1999). Our suites’ softening values are well below this value, with  $0.01 < \epsilon_{\text{fiducial}}/(L/N) < 0.012$  for the LMC, Milky Way, Group, and L-Cluster suites, and  $\epsilon_{\text{fiducial}} = 0.027 \times L/N$  for the Cluster suite (see Section 2.2). The lack of impact on our SHMFs due to this effect is therefore expected.

In the top-right panel of Figure 14, we study the impact of force softening on the subhalo maximum circular velocity functions. We find that subhalos contract and that the  $V_{\text{max,sub}}$  function increases with decreasing  $\epsilon$ . We do not find that the subhalo  $V_{\text{max,sub}}$  function converges with decreasing  $\epsilon$ . This raises the question of whether convergence in the  $V_{\text{max,sub}}$  function requires even smaller force softening scales than tested here. This would have serious implications, because virtually all cosmological and zoom-in simulations use force softening scales larger than our simulations’  $0.25\epsilon_{\text{fiducial}}$  when expressed in terms of the mean interparticle spacing.

One critical question is whether the lack of convergence in the small- and large- $\epsilon$  regimes is caused by the same effect. Mansfield & Avestruz (2021) argued that, in the large epsilon regime ( $\epsilon \gtrsim \epsilon_{\text{fiducial}}$ ), contraction with decreasing  $\epsilon$  is due to the reduction of centripetal forces. This suppression can occur far beyond the formal force softening scale because lower centripetal forces lead to lower densities, which in turn cause lower (physical) centripetal forces. Mansfield & Avestruz (2021) showed that a model based on this effect predicts convergence limits across a wide range of cosmological boxes in the low- $\epsilon$  regime and quantitatively predicts the exact biases in  $V_{\text{max,sub}}$  for controlled test simulations. However, this model predicts that the effect would be weak at  $\epsilon_{\text{fiducial}}$  and nonexistent at  $0.5\epsilon_{\text{fiducial}}$ , suggesting that the lack of convergence at small  $\epsilon$  in Figure 14 has a separate cause.

To illustrate this point, we apply the  $\epsilon$ -debiasing model from Mansfield & Avestruz (2021) to our fiducial run and show it as the dotted–dashed curve in the top-right panel of Figure 14. This model was shown to exactly recreate the  $\epsilon$ -dependent  $V_{\text{max}}$  bias in resimulations of c125-1024, the parent simulation of these zoom-ins. The debiasing model predicts that  $V_{\text{max}}$  suppression is negligible at fiducial force softening scales. We note that we assume an NFW profile for subhalos when applying the Mansfield & Avestruz (2021), which will overestimate the density of subhalos in their outskirts (e.g. Ogiya et al. 2019; Errani et al. 2022), thus leading to larger  $R_{\text{max}}/R_{\text{vir}}$  at a fixed value of  $V_{\text{max}}/V_{\text{vir}}$ . The level of bias in the Mansfield & Avestruz (2021) model decreases with increasing  $R_{\text{max}}/\epsilon$ , meaning that the use of NFW profiles causes us to *overestimate* the level of biasing. Thus, the dotted–dashed line is a conservative upper limit.

The analysis in Ludlow et al. (2019) offers an alternative explanation for the small- $\epsilon$  behavior. The authors tested the impact of a very wide range of force softening scales on subhalo profiles and found that small, isolated halos experience substantial contraction when force softening scales are smaller than 0.0018 to  $0.003 \times (L/N)$  (see their Figure 2, and note that its upper x-axis is in units of  $10^{-2}\epsilon/(L/N)$ ), which corresponds to  $\epsilon \lesssim 0.25\epsilon_{\text{fiducial}}$  for Symphony suites. This effect only occurs for simula-

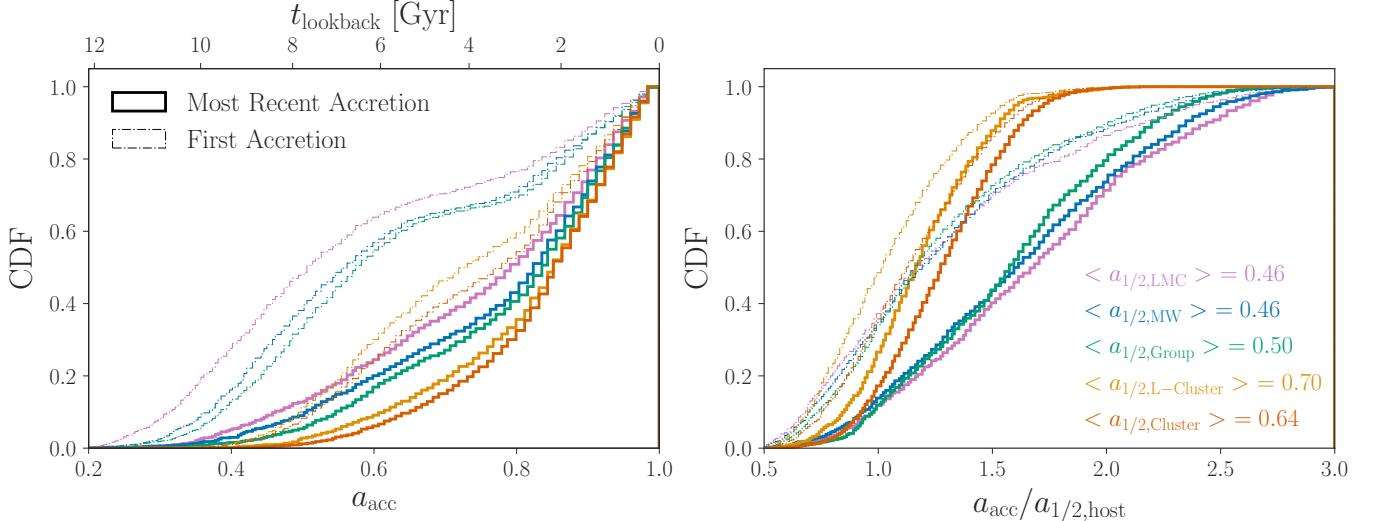


**Figure 14.** Top-left panel: subhalo mass functions evaluated using sub-to-host halo virial mass ratio at  $z = 0$ , stacked over three hosts from the Milky Way suite resimulated with Plummer-equivalent gravitational softening lengths of 4.0 (purple), 2.0 (blue), 0.5 (green), and 0.25 (yellow) times the fiducial value of  $\epsilon_{\text{fiducial}} = 170 \text{ pc h}^{-1}$  (blue-green). Solid lines show the mean cumulative mass function stacked over these three hosts, and transition to dashed at the largest value of  $M_{\text{sub}}/M_{\text{host}}$  among these hosts set by our convergence limit of  $300m_{\text{part}}$ . The dotted-dashed line shows the fiducial  $V_{\text{max,sub}}$  function, corrected for low- $\epsilon$  suppression of centripetal forces using the Mansfield & Avestruz (2021) model. The shaded region shows the  $1\sigma$  Poisson error on the mean, and the bottom panel shows the ratio of the runs with varying softening length relative to the fiducial setting. These subhalo mass functions are consistent with our fiducial results at the  $\approx 10\%$  level down to the convergence limit. Top-right panel: same as the top-left panel, but for normalized maximum circular velocity functions at  $z = 0$ . Note that the resimulations transition from solid to dashed at slightly different values of  $V_{\text{max,sub}}/V_{\text{max,host}}$  because there is not a one-to-one relationship between particle count and  $V_{\text{max,sub}}$ . The systematic variations in  $V_{\text{max,sub}}$  as a function of softening length are discussed in Appendix B.3. Bottom panels: same as the top panels, but for varying time stepping.

tions run using the default GADGET-2 time-stepping criterion (`ErrTolIntAccuracy`),  $\eta = 0.025$ , which is similar to the value of  $\eta = 0.01$  used in our simulations. For finer time stepping, no contraction is found in the Ludlow et al. (2019) tests. This suggests that the nonconvergence shown in the top-right panels of Figure 14 in the small- $\epsilon$  limit is caused by time-stepping errors.

#### B.4. Resimulations with Varying Time Stepping

We resimulate one host halo in the Milky Way suite for several values of the GADGET-2 time stepping criterion,  $\eta$  (`ErrTolIntAccuracy`). Specifically, these resimulations use 0.25, 0.5, 2, and 4 times our simulations’ fiducial setting of  $\eta = 0.01$ ; the resulting SHMFs and maximum circular velocity functions compared to our fiducial run are shown in the bottom panels of Figure 14.



**Figure 15.** Left panel: cumulative distributions of Symphony subhalos’ most recent (solid, thick lines) and first (thin, dot-dashed lines) accretion scale factors into *any* larger halo, stacked over all subhalos from each suite with  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$ . Top ticks indicate lookback time assuming the cosmological parameters for our LMC, Milky Way, and Group suites. Right panel: same as the left panel, but with accretion scale factors normalized by the half-mass scale factor of each Symphony host. The legend indicates the mean half-mass scale factors for the LMC (pink), Milky Way (blue), Group (green), L-Cluster (gold), and Cluster (red) suites (see Table 1).

We find that SHMFs at  $z = 0$  are robust to changes in the time-stepping parameter at the  $\approx 1\%$  level down to our convergence limit. Maximum circular velocity functions are similarly well converged. Moreover, these deviations do not systematically increase as a function of the time-stepping criterion, suggesting that they are partly due to noise associated with the resimulations of the particular halo we selected. Thus, we expect that including resimulations of additional Symphony zoom-ins with varying time stepping would further reduce this systematic.

In summary, in Appendix B.3 we showed that large- $\epsilon$  suppression does not impact our  $V_{\text{max},\text{sub}}$  functions at our fiducial  $\epsilon$ , but that our subhalos continue to contract as  $\epsilon$  is decreased below the fiducial level. Ludlow et al. (2019) showed that, as  $\epsilon$  is decreased below our fiducial level, insufficient time stepping can also lead to subhalo contraction. However, in this section we showed that such time-stepping-dependent effects are not important *at* our fiducial  $\epsilon$ . Therefore, our  $\epsilon_{\text{fiducial}}$  sits in an island of stability, in which it is not large enough to suppress rotation curves out to  $V_{\text{max}}$  for resolved subhalos, but not small enough to lead to time-stepping errors.

## C. AUXILIARY SUBHALO POPULATION STATISTICS

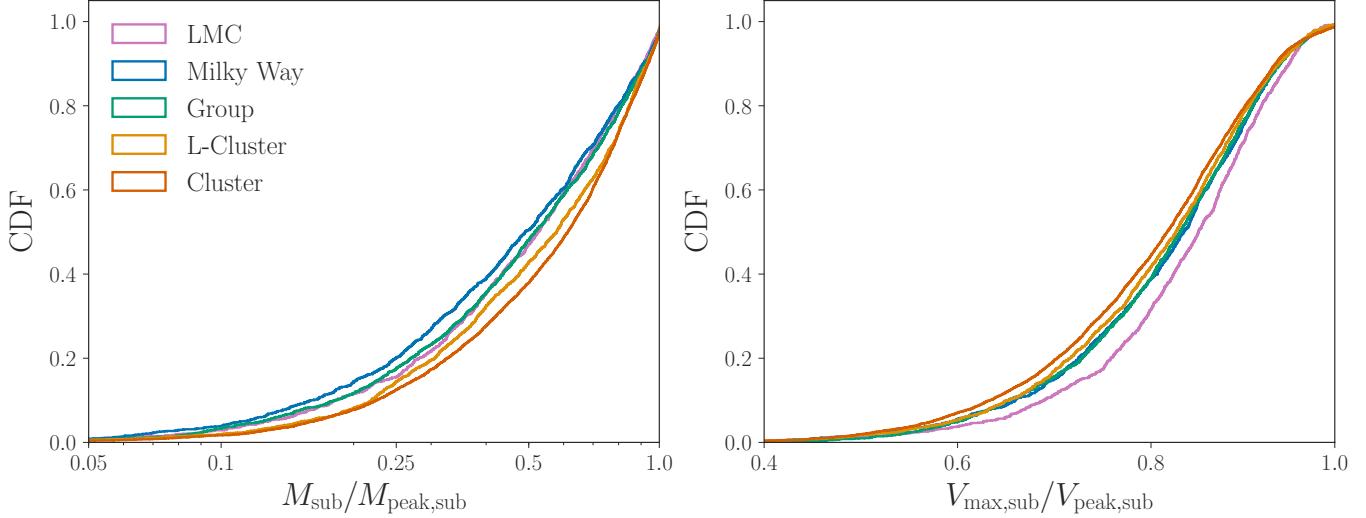
This Appendix studies characteristics of Symphony subhalo populations beyond those discussed in the main text, including subhalo infall time distributions (Appendix C.1) and the amount of tidal stripping subhalos above our resolution limit experience (Appendix C.2).

### C.1. Subhalo Infall Time Distributions

First, Figure 15 shows cumulative infall time distributions for subhalos above our conservative resolution limit of  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$ , stacked over all hosts

within each Symphony suite. Infall times are measured using the `Acc_Scale` (solid, thick distributions) and `First_Acc_Scale` (thin, dashed distributions) output by CONSISTENT-TREES. Note that these variables record the scale factor at which a subhalo most recently or first accreted into *any* larger halo, which may not be the Symphony host halo it resides in at  $z = 0$ . Thus, these distributions include “pre-processing” effects, in which subhalos first orbited within another halo before falling into the  $z = 0$  host. Furthermore, even for objects that have not been subhalos of any halo other than the Symphony target host, the first and most recent accretion scale factors differ for splashback subhalos with apocenters outside of the host’s virial radius.

As expected, subhalos of lower-mass hosts accrete systematically earlier. Interestingly, the difference between first and most recent infall time distributions grows with decreasing host mass. Because the fraction of pre-processed subhalos is expected to be fairly small at all host halo mass scales we consider (e.g., Berrier et al. 2009; Wetzel et al. 2015), we mainly attribute this trend to the varying fraction of splashback subhalos as a function of host mass. In particular, the fraction of splashback subhalos is larger for lower-mass hosts because they are older and accrete more slowly than higher-mass hosts. Thus, we expect a larger difference between the first and most recent accretion scale factors for subhalos of lower-mass hosts, which is consistent with both panels of Figure 15. For the LMC, Milky Way, and Group suites, the flattening of the first accretion scale factor cumulative distributions in the left panel of Figure 15 starts near  $t_{\text{lookback}} \approx 2$  Gyr, roughly one orbital timescale ago, and extends to  $t_{\text{lookback}} \approx 5$  Gyr. This is consistent with the expectation that many subhalos of lower-mass hosts accreted within



**Figure 16.** Left panel: cumulative distributions of Symphony subhalos’  $z = 0$  virial mass,  $M_{\text{sub}}$ , divided by the largest virial mass achieved along each subhalo’s main branch,  $M_{\text{peak,sub}}$ , stacked over all subhalos from each suite with  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$ . Right panel: same as the left panel, for the ratio of  $z = 0$  maximum circular velocity,  $V_{\text{max,sub}}$ , to the largest  $V_{\text{max,sub}}$  achieved along each subhalo’s main branch,  $V_{\text{peak,sub}}$ .

the last few orbital timescales are splashback objects at  $z = 0$  (e.g., Barber et al. 2014).

As shown in the right panel of Figure 15, the dependence of infall time on host mass is reversed when measured in units of the host’s formation time. In particular, relative to their hosts’ half-mass scale factors, subhalos of higher-mass hosts accrete *earlier* than subhalos of lower-mass hosts. This results from a combination of two effects: (1)  $a_{1/2,\text{host}}$  increases for higher-mass hosts, and (2) there is a “survivor bias” due to the  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$  resolution cut imposed when calculating the cumulative distributions. Specifically, subhalos with earlier infall times, comparable to the half-mass scale factors of the LMC, Milky Way, and Group hosts, are likely to be tidally stripped below this resolution limit by  $z = 0$ . Exploring the detailed dependence of subhalo population statistics on accretion time (e.g., following Green et al. 2021) is an interesting area for future work.

### C.2. Subhalo Tidal Stripping

Figure 16 shows cumulative distributions of  $M_{\text{sub}}/M_{\text{peak,sub}}$  and  $V_{\text{max,sub}}/V_{\text{peak,sub}}$  for subhalos above our conservative sub-to-host halo mass resolution limit of  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$ , stacked over all hosts within each Symphony suite. Subhalos above this limit typically retain  $\approx 50\%$  of their peak mass and  $\approx 80\%$  of their peak maximum circular velocity. Very few subhalos with  $M_{\text{sub}}/M_{\text{peak,sub}} \lesssim 0.05$  or  $V_{\text{max,sub}}/V_{\text{peak,sub}} \lesssim 0.4$  remain above the mass resolution limit, regardless of whether an  $M_{\text{sub}} > 300m_{\text{part}}$  cut is applied.

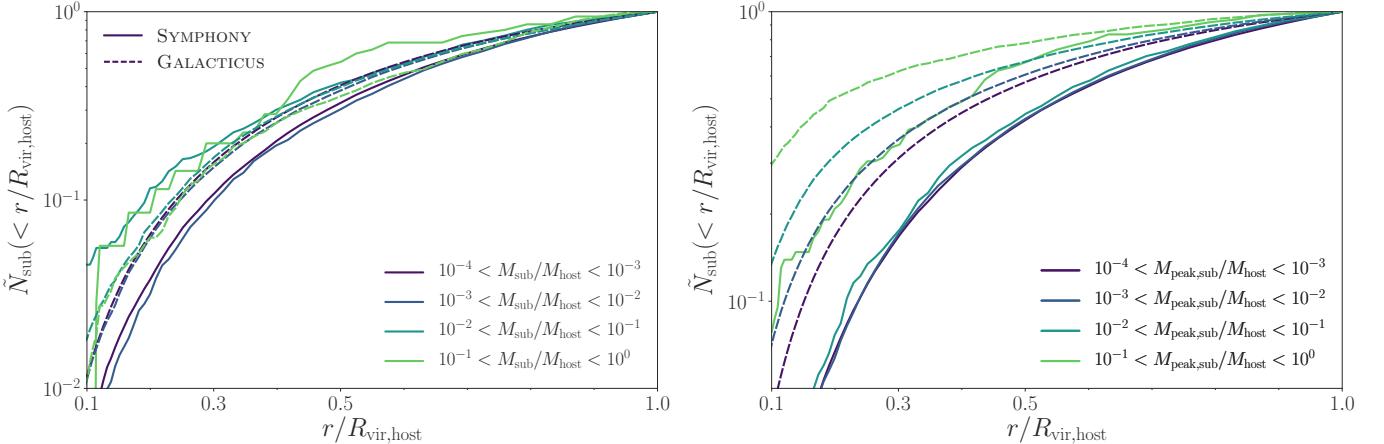
Mass loss is slightly enhanced for subhalos of lower-mass hosts, likely due to a combination of their earlier infall times (Figure 15) and their hosts’ higher concentrations (Figure 3). Meanwhile, subhalos of lower-mass hosts retain slightly *more* of their peak maximum circular velocity than subhalos of higher-mass hosts. This may result from the higher concentrations of lower-mass subhalos (e.g., Moliné et al. 2017,

2023), which makes their inner regions probed by  $V_{\text{max,sub}}$  more resilient to stripping despite earlier infall times.

## D. ADDITIONAL GALACTICUS COMPARISONS

This Appendix provides comparisons between Symphony and GALACTICUS predictions beyond those discussed in the main text. First, Figure 17 compares Symphony and GALACTICUS normalized radial distributions stacked over all suites, binned by  $M_{\text{sub}}/M_{\text{host}}$  (left panel) and  $M_{\text{peak,sub}}/M_{\text{host}}$ . As for Figure 8, note that the lowest-mass bin extends slightly below our fiducial  $z = 0$  sub-to-host mass ratio convergence limit (Equation 2), while the radial distributions binned by peak mass are not formally converged at small radii given our convergence tests (see Appendix B.2), and should be interpreted with caution. Interestingly, GALACTICUS radial profiles are largely insensitive to  $M_{\text{sub}}/M_{\text{host}}$  and become systematically more concentrated with increasing  $M_{\text{peak,sub}}/M_{\text{host}}$ . On the other hand, Symphony radial profiles are nearly identical for all  $M_{\text{peak,sub}}/M_{\text{host}} \lesssim 10^{-2}$  and  $M_{\text{peak,sub}}/M_{\text{host}} \lesssim 10^{-1}$ , while the subhalos with the largest sub-to-host halo mass ratios in our simulations (in terms of both  $z = 0$  and peak mass) are more centrally concentrated due to dynamical friction.

Next, Figure 18 shows cumulative SHMF comparisons between Symphony and GALACTICUS for each suite individually. These plots demonstrate that Symphony and GALACTICUS SHMFs are consistent within the host-to-host scatter, although GALACTICUS displays slightly but systematically steeper SHMF slopes. Finally, Figure 19 shows the corresponding comparisons between Symphony and GALACTICUS cumulative normalized radial distributions. Again, these plots demonstrate that Symphony and GALACTICUS radial profiles are consistent within the host-to-host scatter, although GALACTICUS displays slightly more concentrated radial profiles in the very inner regions ( $r/R_{\text{vir,host}} \lesssim 0.2$ ) down



**Figure 17.** Left panel: normalized cumulative subhalo radial distributions, stacked over all suites, in bins of  $z = 0$  sub-to-host halo mass ratio for Symphony (solid) and GALACTICUS (dashed). Right panel: same as the left panel for bins of peak sub-to-host halo mass ratio.

to a fixed sub-to-host halo mass ratio. In turn, the GALACTICUS predictions slightly exceed Symphony at larger radii because the radial profiles we plot are normalized to unity at  $r/R_{\text{vir,host}} = 1$ .

### E. COMPARISONS BETWEEN SYMPHONY MILKY WAY, CATERPILLAR, AND PARENT BOXES

In this Appendix, we compare the Symphony Milky Way suite SHMF to (1) its parent simulation, c125-1024 ( $m_{\text{part}} = 2.1 \times 10^8 M_{\odot}$ ) and a higher-resolution version of the same box, c125-2048 ( $m_{\text{part}} = 2.6 \times 10^7 M_{\odot}$ ), and (2) the LX14 resolution level of the Caterpillar Milky Way–mass zoom-in simulations (Griffen et al. 2016;  $m_{\text{part}} = 3.0 \times 10^4 M_{\odot}$ ). We also compare Caterpillar SHMFs to those measured in its parent simulation, “CaterpillarParent,” (described in Griffen et al. 2016;  $m_{\text{part}} = 1.2 \times 10^8 M_{\odot}$ ), and we compare both the Symphony and Caterpillar parent box SHMFs to the cosmological simulation VSMDPL (Klypin et al. 2016;  $m_{\text{part}} = 8.8 \times 10^6 M_{\odot}$ ) as an independent check.

For the following comparisons, all SHMFs are measured using the ratio of  $z = 0$  subhalo virial mass to host halo mass,  $M_{\text{sub}}/M_{\text{host}}$ . For plotting purposes, we assume that SHMFs are converged for  $M_{\text{sub}} > 300m_{\text{part}}$  in each zoom-in or cosmological simulation, following Symphony convergence results in Appendix B; this assumption does not affect our conclusions. Furthermore, based on the consistency among the cosmological box SHMFs described below, we assume that differences in cosmological parameters between these simulations result in subdominant differences between their SHMFs compared to the selection effects that we discuss. We refer the reader to the respective references for additional details about the numerical and cosmological parameters for each simulation.

#### E.1. Subhalo Mass Function Measurements

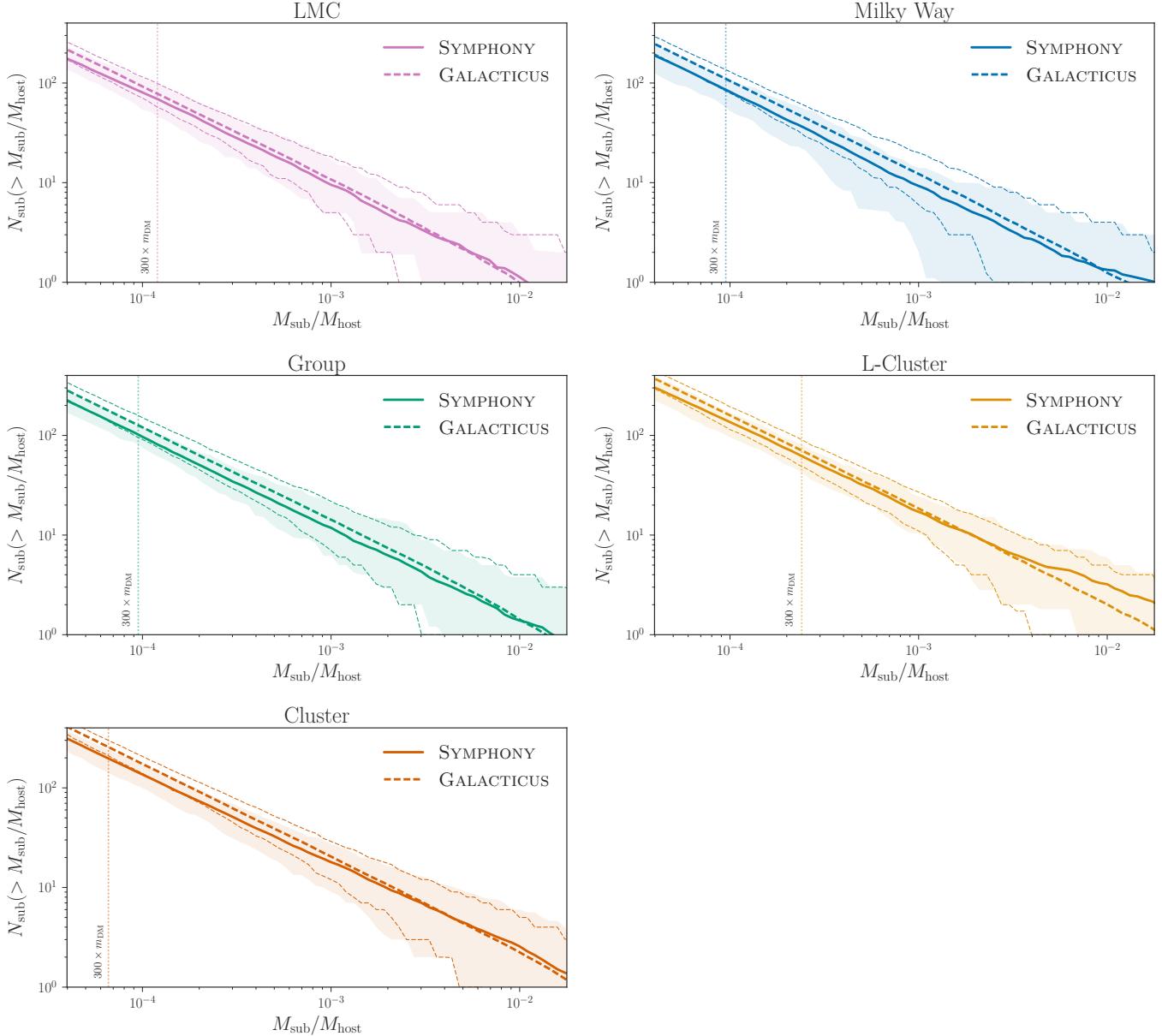
Figure 20 shows the mean SHMFs, in units of  $z = 0$  sub-to-host halo mass ratios, for the 45 Symphony Milky Way hosts, the 35 Caterpillar zoom-ins hosts, and samples of hosts from c125-1024, c125-2048, VSMDPL, and the Caterpillar

parent box that are chosen to either satisfy combinations of (1) the “primary” Symphony Milky Way mass and environmental cuts described in Section 2.2.2, (2) the exact Symphony Milky Way selection, i.e., the primary mass and environmental cuts plus the “secondary” selection of which halos to zoom in on, and (3) the corresponding primary and secondary Caterpillar selection cuts.<sup>8</sup> Each SHMF extends down to a sub-to-host halo mass ratio of 300 particles for the lowest-mass host in each host sample from the corresponding zoom-in or cosmological simulations. When interpreting the differences between cosmological simulations, note that the Caterpillar parent box resolution is slightly better than c125-1024, while VSMDPL’s resolution is roughly a factor of 3 better than c125-2048.

#### E.2. Comparing Symphony to Its Parent Box

First, we study the SHMF of Symphony Milky Way hosts in the parent boxes at different resolution levels relative to all hosts in c125-2048 that pass Symphony’s primary mass and environmental cuts. Specifically, the solid black line in Figure 20 shows the SHMF for 450 hosts that pass the Symphony Milky Way environmental cuts described in Section 2.2.2, where 10 hosts with masses closest to each Symphony Milky Way host’s  $z = 0$  virial mass are selected from all halos that pass the environmental cuts in order to match the host mass distributions. Meanwhile, the dashed black and gray lines show SHMFs for the 45 Symphony target halos in c125-2048 and c125-1024, respectively. The identities of these hosts in c125-1024 are known, because they are used to generate the Symphony initial conditions, and we find their matches in c125-2048 by minimizing a metric based on differences in  $z = 0$  mass and phase-space distance. We have verified that all of the c125-1024 and c125-2048 matches’ MAHs match each Symphony zoom-in at the percent level or better.

<sup>8</sup> The first 24 Caterpillar hosts were presented in Griffen et al. (2016); we use the 35 available at <https://www.caterpillarproject.org/>.



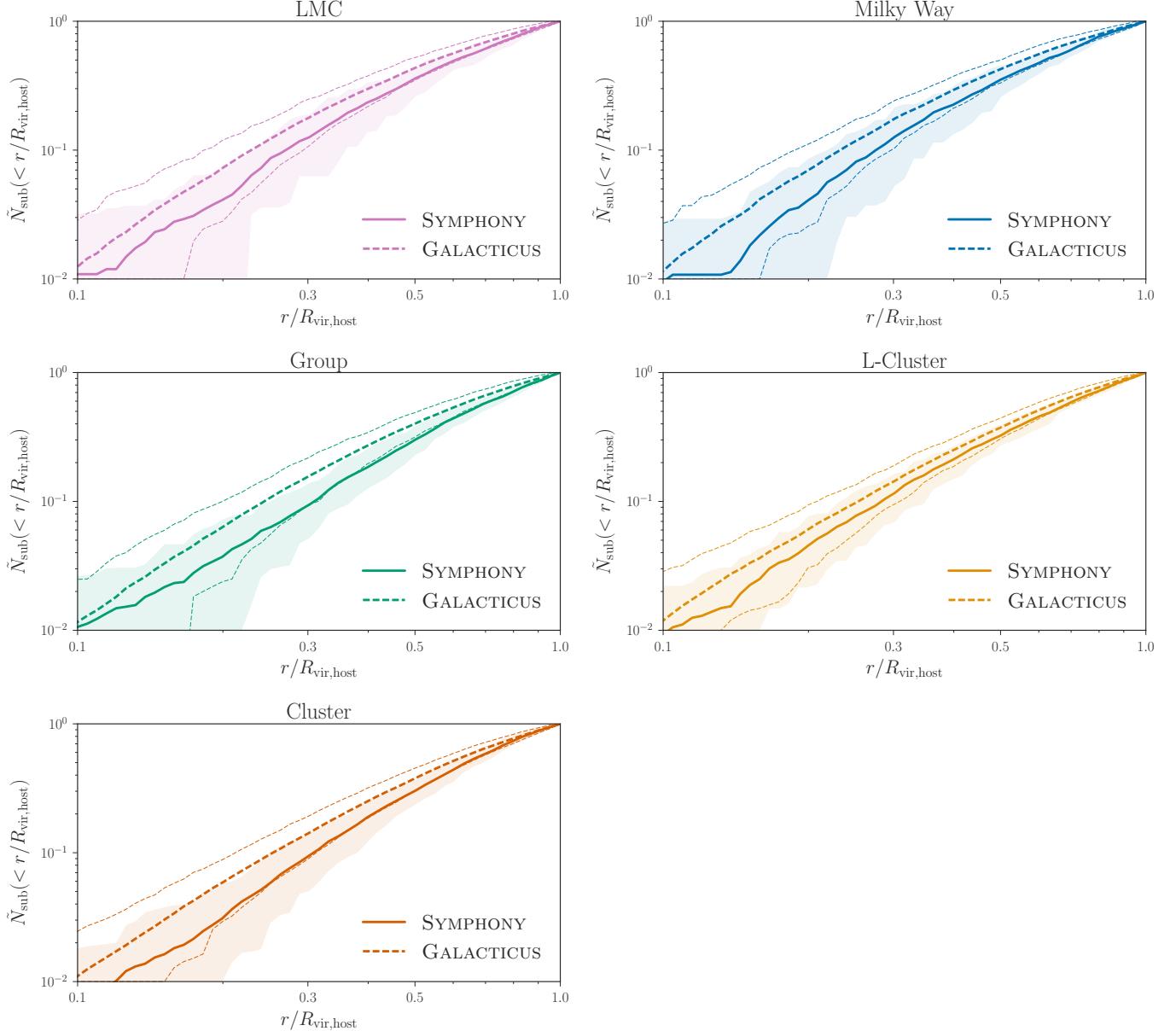
**Figure 18.** Comparison of subhalo mass functions in our five zoom-in simulation suites with GALACTICUS predictions. In each panel, the solid line shows the mean cumulative Symphony SHMF stacked over all simulations in each suite, the shaded band shows the 2.5th–97.5th percentile range, and dashed lines show the corresponding GALACTICUS predictions and percentiles. Vertical dashed lines show the conservative subhalo mass resolution limit for each suite, corresponding to 300 times the highest-resolution dark matter particle mass.

As described in Section 5.3, comparing these SHMFs reveals that, in both the zoom-ins and parent boxes, Symphony hosts’ mean SHMF is  $\approx 30\%$  higher than the mass and environmentally selected sample’s mean SHMF at high sub-to-host halo mass ratios. Because this effect appears in both the parent and zoom-in simulations, we attribute it to either a random fluctuation or the effects of secondary selection cuts—i.e., a consequence of choices among which host halos to zoom in on after the primary mass and environmental cuts are applied—rather than an artifact (or feature) of our zoom-in simulation procedure or analysis. Meanwhile, at sub-to-

host halo mass ratios near c125-2048’s resolution limit of  $M_{\text{sub}}/M_{\text{host}} \approx 10^{-2}$ , the mean SHMFs of both the Symphony zoom-in suite and the matched sample in c125-2048 are consistent with the mass and environmentally selected host sample’s SHMF in both c125-2048 and VSMDPL.

We therefore conclude that:

1. For  $M_{\text{sub}}/M_{\text{host}} \gtrsim 10^{-2}$ , the mean Symphony Milky Way SHMF is  $\approx 30\%$  high relative to all Milky Way-mass hosts in c125-1024 and c125-2048 that satisfy the Symphony mass and environmental cuts described in Section 2.2.2; the SHMF of the 45 Symphony hosts



**Figure 19.** Comparison of normalized radial distributions in our five zoom-in simulation suites with GALACTICUS predictions. In each panel, the solid line shows the mean cumulative normalized radial profile, stacked over all simulations in each Symphony suite, and the shaded band shows the corresponding 2.5th–97.5th percentile range. Dashed lines show the corresponding GALACTICUS predictions and percentiles. Only subhalos above our conservative sub-to-host halo resolution threshold of  $M_{\text{sub}}/M_{\text{host}} > 2.7 \times 10^{-4}$  are included.

in the fiducial and high-resolution versions of the parent box displays the same offset, which is therefore caused by the selection of the 45 Symphony hosts, and may either be a random fluctuation or a selection effect related to these hosts' properties.

2. For  $M_{\text{sub}}/M_{\text{host}} \lesssim 10^{-2}$ , the mean Symphony Milky Way SHMF is consistent with the SHMF of a corresponding mass and environmentally selected sample in both c125-2048 and VSMDPL.

We note that, for computational efficiency, Symphony Milky Way hosts were typically selected to have small Lagrangian volumes relative to all hosts in the parent box that satisfy the relevant mass and environmental cuts (Mao et al. 2015). However, we only observe a weak correlation (with significant scatter) between Lagrangian volume and the amplitude of the normalized SHMF at any sub-to-host halo mass ratio, so Lagrangian volume bias does not necessarily explain the SHMF bias at high sub-to-host halo mass ratios.

### E.3. Comparing Caterpillar to its Parent Box

Next, we compare Caterpillar zoom-in and parent simulation SHMFs. The mean SHMF of Caterpillar hosts in the Caterpillar parent box, which includes both primary mass and environmental cuts and the effects of secondary selection for the 35 Caterpillar hosts, is shown by the dashed red line in Figure 20. At high sub-to-host halo mass ratios, the Caterpillar sample’s SHMF in both the parent box and zoom-ins is  $\approx 30\%$  lower than for a mass-matched sample of hosts in the Caterpillar parent box that only pass the primary Caterpillar cuts. Thus, selection of the specific Caterpillar host sample also biases its zoom-ins’ SHMF at high sub-to-host halo mass ratios, but in a different direction than in Symphony.

Meanwhile, at lower sub-to-host halo mass ratios, the mean SHMF of the Caterpillar sample is  $\approx 20\%$  high relative to the host samples with Symphony’s mass and environmental cuts in c125-2048 and VSMDPL. We have verified that this difference persists if Caterpillar’s mass and environmental cuts are applied instead. Nonetheless, it is difficult to interpret this difference without a higher-resolution resimulation of Caterpillar’s parent box for comparison. In general, we expect that cosmological and zoom-in resimulations with varying resolution are needed to robustly interpret SHMF biases at intermediate sub-to-host halo mass ratios.

Thus, we conclude that:

1. For  $M_{\text{sub}}/M_{\text{host}} \gtrsim 10^{-2}$ , the mean Caterpillar SHMF is  $\approx 30\%$  low relative to all hosts that pass the Caterpillar mass and environmental cuts (which are similar to the Symphony Milky Way cuts, but differ in detail; see Griffen et al. 2016) in the Caterpillar parent box; the SHMF of the 35 hosts in the Caterpillar parent box that were resimulated displays the same offset, which is therefore caused by the selection of the 35 Caterpillar hosts, and may either be a random fluctuation or a selection effect related to these hosts’ properties.
2. For  $M_{\text{sub}}/M_{\text{host}} \lesssim 10^{-2}$ , the mean Caterpillar SHMF is  $\approx 25\%$  high relative to all hosts that pass the Symphony mass and environmental cuts in c125-2048 and VSMDPL. Our tests using these cosmological simulations suggest that this difference is not caused by Caterpillar’s mass and environmental cuts, and is instead either a random fluctuation due to the selection of the specific sample of 35 Caterpillar hosts (which would also appear for a matched sample of hosts in a higher-resolution version of Caterpillar’s parent box), or a result of Caterpillar’s zoom-in procedure or analysis (which would only appear in the zoom-in results).

### E.4. Comparing Symphony to Caterpillar

Finally, we compare Symphony and Caterpillar SHMFs in the context of our previous findings. Before doing so directly, we first compare their respective parent simulations’ SHMFs to each other and to an independent measurement from VSMDPL. We reiterate that Caterpillar’s primary mass and environmental cuts (which differ in detail from those used in

Symphony’s Milky Way suite) do not appear to drive the differences between Symphony and Caterpillar SHMFs. Thus, we interpret discrepancies between Symphony and Caterpillar through the comparisons to matched samples in their parent cosmological simulations described above.

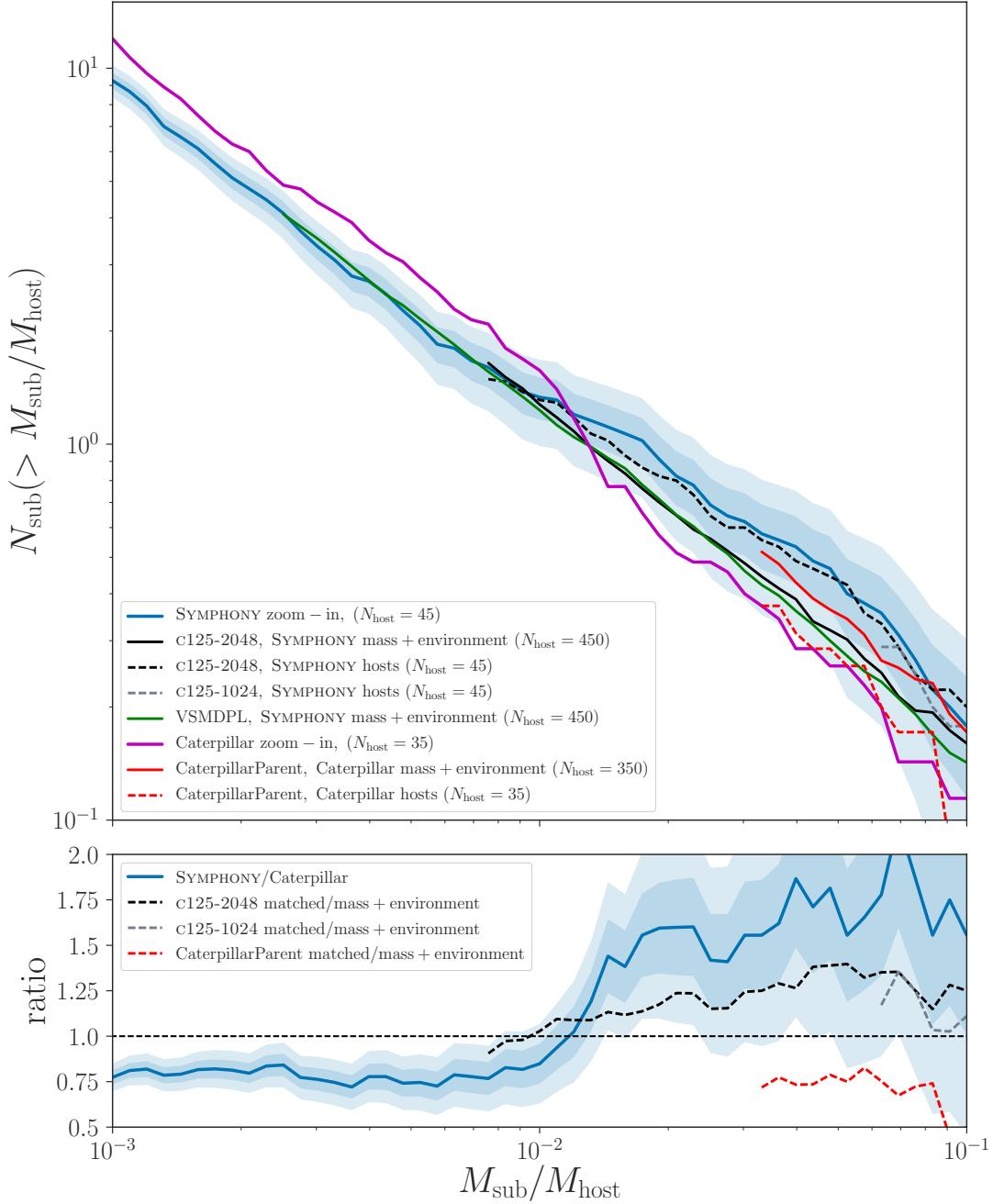
A comparison between Symphony and Caterpillar SHMFs indicates that:

1. At high sub-to-host halo mass ratios,  $M_{\text{sub}}/M_{\text{host}} \gtrsim 10^{-2}$ , the upward fluctuation in the mean Symphony SHMF combines with the downward fluctuation in the mean Caterpillar SHMF to yield a  $\approx 60\%$  (or  $\approx 2\sigma$ ) discrepancy between the zoom-in suites. This difference is driven by the selection of the specific 45 Symphony hosts and 35 Caterpillar hosts that were resimulated, and may either be a random fluctuation or a selection effect related to these hosts’ properties.<sup>9</sup>
2. At intermediate sub-to-host halo mass ratios,  $10^{-3} \lesssim M_{\text{sub}}/M_{\text{host}} \lesssim 10^{-2}$ , the overabundance in the mean Caterpillar SHMF relative to cosmological simulations yields a  $\approx 25\%$  (or  $\approx 2\sigma$ ) discrepancy relative to Symphony. This difference may be caused by a fluctuation due to the selection of the specific Caterpillar host sample after mass and environmental cuts are applied, or may be caused by Caterpillar’s zoom-in procedure and/or analysis. Resimulations of the Caterpillar parent box at varying resolution would help to assess these potential explanations.
3. At low sub-to-host halo mass ratios,  $M_{\text{sub}}/M_{\text{host}} \lesssim 10^{-3}$ , the Symphony and Caterpillar SHMF slopes do not significantly differ. Thus, the SHMF overabundance in Caterpillar relative to Symphony at intermediate subhalo masses propagates to very low sub-to-host halo mass ratios, resulting in the  $\approx 25\%$  discrepancy at the lowest resolved masses shown in Figure 10.

We emphasize that Caterpillar’s LX14 resolution is better by a factor of roughly 8 than the Symphony Milky Way suite’s fiducial resolution, and that future comparisons will need to account for resolution differences carefully to draw firm conclusions about the physical or numerical origins of SHMF discrepancies at low sub-to-host halo mass ratios.

To conclude, we note that our direct comparison between Symphony and Caterpillar SHMFs is consistent with the comparison between Symphony’s Milky Way suite and GALACTICUS predictions discussed in Section 5.3. This is expected, because Yang et al. (2020) calibrated the GALACTICUS model we use to reproduce Caterpillar SHMFs, and is confirmed by our tests.

<sup>9</sup> Curiously, our tests on Symphony Milky Way hosts using the modified version of ROCKSTAR described in Griffen et al. (2016) and used to analyze the Caterpillar zoom-in simulations indicate that even more high-mass subhalos are recovered with the Caterpillar analysis tools; thus, the discrepancy at the high-mass end of the SHMF may be even larger when using a unified analysis pipeline.



**Figure 20.** Top panel: mean SHMFs, in units of  $z = 0$  sub-to-host halo mass ratio, in the Symphony Milky Way suite (blue), the c125-2048 box (a higher-resolution version of Symphony Milky Way’s parent simulation) using (1) a mass-matched sample of hosts that pass the Symphony Milky Way mass and environmental cuts (solid black), and (2) the sample of Symphony target halos (dashed black), the c125-1024 box (Symphony Milky Way’s parent simulation) using the sample of 45 Symphony target hosts (dashed gray), the VSMDPL box using Symphony Milky Way’s mass and environmental cuts (green), the Caterpillar zoom-in suite (magenta), and the Caterpillar parent box (“CaterpillarParent) using a mass-matched sample of hosts that pass the Caterpillar mass and environmental cuts (solid red) and the 35 Caterpillar target host halos (dashed red). Dark (light) blue bands indicate  $1\sigma$  ( $2\sigma$ ) Poisson uncertainty on the mean Symphony zoom-in SHMF; these uncertainties are representative of other SHMFs shown that use the same (or a similar) number of hosts. Each SHMF extends down to a sub-to-host halo mass ratio of 300 particles for the lowest-mass host in each sample. Both the Symphony and Caterpillar zoom-in resolution limits are significantly below  $M_{\text{sub}}/M_{\text{host}} = 10^{-3}$ . Bottom panel: same as the top panel, but for ratios of mean SHMF in Symphony relative to Caterpillar (blue), the specific Symphony host sample in c125-2048 and c125-1024 relative to the sample with only mass and environmental cuts applied (dashed black and gray, respectively), and the specific Caterpillar sample in the Caterpillar parent box relative to the sample with only mass and environmental cuts applied (dashed red).