Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **495,** 743–757 (2020)
Advance Access publication 2020 May 4

doi:10.1093/mnras/staa1199

The little things matter: relating the abundance of ultrafaint satellites to the hosts' assembly history

Sownak Bose [1]★ Alis J. Deason [2] Vasily Belokurov [3] and Carlos S. Frenk[2]

[1] *Center for Astrophysics | Harvard & Smithsonian, 60 Garden Str, Cambridge, MA 02138, USA*
[2] *Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*
[3] *Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

Accepted 2020 April 25. Received 2020 April 25; in original form 2019 September 9

ABSTRACT

Ultrafaint dwarf galaxies ($M_\star \leq 10^5\,\mathrm{M}_\odot$) are relics of an early phase of galaxy formation. They contain some of the oldest and most metal-poor stars in the Universe which likely formed before the epoch of hydrogen reionization. These galaxies are so faint that they can only be detected as satellites of the Milky Way. They are so small that they are not resolved in current cosmological hydrodynamic simulations. Here, we combine very high-resolution cosmological *N*-body simulations with a semi-analytic model of galaxy formation to study the demographics and spatial distribution of ultrafaint satellites in Milky Way-mass haloes. We show that the abundance of these galaxies is correlated with the assembly history of the host halo: at fixed mass, haloes assembled earlier contain, on average, more ultrafaint satellites today than haloes assembled later. We identify simulated galactic haloes that experience an ancient *Gaia*-Enceladus-Sausage-like and a recent LMC-like accretion event and find that the former occurs in 33 per cent of the sample and the latter in 9 per cent. Only 3 per cent experience both events and these are especially rich in ultrafaint satellites, most acquired during the ancient accretion event. Our models predict that the radial distribution of satellites is more centrally concentrated in early-forming haloes. Accounting for the depletion of satellites by tidal interactions with the central disc, we find a very good match to the observed radial distribution of satellites in the Milky Way over the entire radial range. This agreement is mainly due to the ability of our model to track 'orphan' galaxies after their subhaloes fall below the resolution limit of the simulation.

Key words: methods: numerical – Galaxy: formation – galaxies: dwarf – Local Group – dark ages, reionization, first stars.

1 INTRODUCTION

Dark matter haloes are the fundamental units of cosmic structure. They are the sites where galaxies form and the details of their assembly directly influence the properties of the galaxies within them. In the ΛCDM model, dark matter haloes grow hierarchically through the continuous accretion of mass in the form of discrete lumps and diffuse matter (e.g. Frenk et al. 1985; Wechsler et al. 2002; McBride, Fakhouri & Ma 2009; Fakhouri, Ma & Boylan-Kolchin 2010; Wang et al. 2011; Correa et al. 2015). The faint population of dwarf galaxies that dominate in number are amongst the oldest objects in the cosmos, while, at the other end of the mass range, rare, rich clusters of galaxies assemble later and are still growing at the present day.

Much of what we have learnt about the growth of dark haloes has come from high-resolution numerical simulations. Inferring the assembly of haloes from observational data is considerably more challenging and can only be done by means of indirect tracers. Unsurprisingly, this enterprise has been most successful in the case of our own galaxy, the Milky Way, for which the data are now more complete and comprehensive than ever before. Through the combination of a variety of probes including the kinematics of individual stars, globular clusters, and satellite galaxies, increasingly tight constraints on the mass of the Galaxy's halo – one of the most important parameters in astrophysics – have been obtained (e.g. Deason et al. 2012; Boylan-Kolchin et al. 2013; Patel et al. 2018; Callingham et al. 2019; Watkins et al. 2019). Furthermore, the measurement of high-precision stellar proper motions by the *Gaia* satellite (Gaia Collaboration et al. 2018) has improved our understanding of the assembly of the Galaxy manifold.

There is now increasing evidence that the formation of the Milky Way was punctuated by two distinct events, separated in time by several billion years. The most familiar is the accretion of the small and large magellanic clouds (SMC and LMC), which are thought to have fallen into the potential of the Milky Way's halo 2–3 Gyr

★ E-mail: sownak.bose@cfa.harvard.edu

ago and to be on their first orbit in the Galactic potential (e.g. Kallivayalil, van der Marel & Alcock 2006; Besla et al. 2007; Boylan-Kolchin, Besla & Hernquist 2011; Sales et al. 2011; Besla 2015; Shao et al. 2018). More recently, galactic archaeology based on metal-rich halo stars in *Gaia* data has provided evidence for an ancient dwarf galaxy merger, similar in mass to the LMC, between 8 and 11 Gyr ago, roughly around the time of the formation of the Galactic disc (Belokurov et al. 2018; Helmi et al. 2018; Myeong et al. 2018). The merger remnant, known as the *Gaia*-Enceladus or *Gaia* 'Sausage' – so-named for the elongated appearance of its constituent stars in velocity space – provides a pointer to the early accretion history of the Galactic halo. Numerical simulations are needed to determine the frequency of events of this kind and thus to quantify how typical the Milky Way is compared to the overall population of galaxies (e.g. Fattahi et al. 2019; Mackereth et al. 2019).

Galaxy mergers are important for more than just adding mass to, and disturbing, the central galaxy: they also bring in large populations of satellite galaxies that may survive until the present day (e.g. Deason et al. 2015; Jethwa, Erkal & Belokurov 2016; Dooley et al. 2017; Sales et al. 2017; Shao et al. 2018; Jahn et al. 2019). A complete census of the satellite population of galactic haloes offers a powerful constraint on physical processes associated with galaxy formation (e.g. Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002a; Bovill & Ricotti 2009; Font et al. 2011; Brooks & Zolotov 2014; Sawala et al. 2016; Wetzel et al. 2016; Munshi et al. 2019), as well as on the nature of the dark matter itself (e.g. Macciò & Fontanot 2010; Lovell et al. 2012; Kennedy et al. 2014; Lovell et al. 2016; Newton et al. 2018a). In both instances, the crucial population are the 'ultrafaint' dwarf galaxies ($M_\star \lesssim 10^5\,\mathrm{M}_\odot$, Simon 2019) whose detection, whilst still notoriously difficult, has improved greatly with the advent of the Sloan Digital Sky Survey (SDSS; Adelman-McCarthy et al. 2007; Alam et al. 2015), the Dark Energy Survey (DES; Bechtol et al. 2015; Drlica-Wagner et al. 2015; Kim et al. 2015; Koposov et al. 2015), and Pan-STARRS1 (Laevens et al. 2015; Chambers et al. 2016).

Even at fixed mass, observations find a large degree of scatter in the abundance of satellites surrounding galaxies similar to our own (e.g. Zaritsky et al. 1993; Guo et al. 2011; Wang et al. 2012; Geha et al. 2017; Kondapally et al. 2018; Bennet et al. 2019). The galactic populations we observe today may retain memory of the assembly histories of their hosts. Our aim in this work is to unify these two themes – the assembly history of dark matter haloes and the population of ultrafaint satellites hosted within them – into a single narrative.

The stellar masses of the ultrafaint satellites are so small that they are beyond the reach of the current generation of cosmological hydrodynamical simulations, such as APOSTLE (Fattahi et al. 2016; Sawala et al. 2016), AURIGA (Grand et al. 2017), or FIRE (Hopkins et al. 2014, 2018), which, at best, have a stellar mass resolution of $M_\star \sim 10^3$–$10^4\,\mathrm{M}_\odot$, although higher resolution may be achieved in simulations of *isolated* dwarf galaxies (e.g. Read, Agertz & Collins 2016; Jeon, Besla & Bromm 2017; Corlies, Johnston & Wise 2018; Wheeler et al. 2019; Munshi et al. 2019; Rey et al. 2019; Agertz et al. 2020). Instead, we use the technique of semi-analytic galaxy modelling in which mass resolution is not an issue. We graft our semi-analytic model, GALFORM (Cole et al. 2000; Lacey et al. 2016), on to merger trees constructed from the COCO ΛCDM dark matter simulations (Bose et al. 2016; Hellwing et al. 2016). This introduces a resolution scale in the *dark matter* of $\sim 10^6\,\mathrm{M}_\odot$, which is more than adequate to resolve the haloes of the ultrafaint satellite population. With these simulations, we investigate the diversity of formation histories of galaxies like the Milky Way and explore how the specific accretion events in the Milky Way may have shaped its present-day satellite content. The combination of high-resolution cosmological simulations with observational data from *Gaia* makes this a particularly timely endeavour.

This paper is organized as follows. In Section 2, we describe the simulations and the semi-analytic model of galaxy formation used in this work. Our main results are presented in Section 3, in which we connect the populations of ultrafaint satellites of Milky Way-mass haloes to their assembly histories and predict the radial distribution of these satellites within their hosts. Finally, Section 4 provides a summary.

2 MODELLING TECHNIQUES

This section provides an overview of the numerical set-up used in this work. We describe the *N*-body simulations (Section 2.1) that form the backbone of the semi-analytic model of galaxy formation, GALFORM, which is used to populate dark matter haloes with galaxies (Section 2.2).

2.1 Simulation suite

The *N*-body simulations we analyse are part of the *Copernicus Complexio* (COCO) suite of simulations (Bose et al. 2016; Hellwing et al. 2016), a set of cosmological zoom-in *N*-body simulations that follow the evolution of over 12 billion high-resolution dark matter particles, each of mass $m_p = 1.6 \times 10^5\,\mathrm{M}_\odot$ from ΛCDM initial conditions. The zoom region, which is roughly 24 Mpc in radius, was extracted from a $(100\ \mathrm{Mpc})^3$ periodic volume, the *Copernicus complexio Low Resolution* (COLOUR) simulation, in which the mass of each dark matter particle is $m_p = 8.8 \times 10^6\,\mathrm{M}_\odot$. The simulations were evolved from $z = 127$ to the present day using the P-GADGET-3 code (Springel, Yoshida & White 2001a; Springel 2005). Both COCO and COLOUR assume cosmological parameters derived from *WMAP*-7 data (Komatsu et al. 2011): $\Omega_\mathrm{m} = 0.272$, $\Omega_\Lambda = 0.728$, and $h = 0.704$, where h is related to the present-day Hubble constant, H_0, by $h = H_0/100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$. The spectral index of the primordial power spectrum is $n_\mathrm{s} = 0.967$, and the linear power spectrum is normalized at $z = 0$ taking $\sigma_8 = 0.81$.

Dark matter haloes were identified using the friends-of-friends algorithm (Davis et al. 1985), which connects dark matter particles separated by at most 20 per cent of the mean interparticle separation in each volume. Gravitationally bound substructures within each group are determined using the SUBFIND algorithm (Springel et al. 2001b). To be included in the final halo catalogue, a subhalo is required to contain at least 20 bound dark matter particles, corresponding to a total mass of $3.2 \times 10^6\,\mathrm{M}_\odot$ in COCO and $1.8 \times 10^8\,\mathrm{M}_\odot$ in COLOUR. In what follows, the physical extent of a dark matter halo is defined by r_{200}, the radius within which the mean density of the halo is 200 times the critical density of the Universe. The mass of haloes is quoted in terms of M_{200}, the total mass in dark matter particles enclosed within r_{200}. We identify Milky Way-mass hosts in our simulations based on their M_{200} mass. In particular, we consider haloes in the range $M_{200} = [0.7{-}1.8] \times 10^{12}\,\mathrm{M}_\odot$. This yields a total of 1189 objects in COLOUR and 63 objects in COCO.

Substructures detected by SUBFIND serve as the roots for building merger trees. We establish associations between subhaloes in subsequent output times by identifying objects that share some fraction of their most-bound particles between snapshots; the method is described in detail in Jiang et al. (2014). The (sub)halo merger trees are then traversed to generate galaxy populations using the

GALFORM semi-analytic model of galaxy formation, which we now describe.

2.2 Semi-analytic galaxy formation

Semi-analytic models of galaxy formation provide a flexible and computationally efficient method for generating synthetic galaxy populations. Key advantages of these models over hydrodynamical simulations are their superior resolution and the relative ease with which it is possible to explore the vast parameter space describing the physics of galaxy formation. Semi-analytic modelling is, by now, a mature field and there exist several such models (e.g. Kauffmann, White & Guiderdoni 1993; Cole et al. 1994, 2000; Somerville & Primack 1999; Croton et al. 2006; Benson 2012; Henriques et al. 2015), which differ in the manner in which specific aspects of the galaxy formation process are treated. The general philosophy adopted in most of these models is similar: they follow the properties of (sub)haloes in merger trees and populate them with galaxies by solving sets of coupled differential equations that govern the cooling of gas in haloes, star formation, feedback from stars and black holes, chemical enrichment and the evolution of stellar populations. The free parameters describing these physical processes are calibrated by requiring that the model should reproduce a small selection of properties of the observed local galaxy population (typically at $z = 0$).

We make use of the Durham semi-analytic model of galaxy formation, GALFORM, first presented in Cole et al. (1994) and refined in Cole et al. (2000). In particular, we use the model developed in Lacey et al. (2016), which unifies several features from previous versions, such as the assumption of a top-heavy initial mass function (IMF) in starbursts, which is required to reproduce the abundance of star-forming sub-millimetre galaxies (Baugh et al. 2005); the model of AGN feedback introduced by Bower et al. (2006), which regulates the growth of massive galaxies; and a star formation law that depends on the molecular gas content of the interstellar medium (Lagos et al. 2011). Broad-band luminosities for the simulated galaxies are computed from the Maraston (2005) stellar population synthesis model. The free parameters in the Lacey et al. (2016) model are calibrated so as to reproduce, at $z = 0$, the optical and near-infrared luminosity functions, the black hole-bulge mass relation, the H I mass function and the fraction of early- and late-type galaxies. For a complete list of model parameters and their calibration, we refer the reader to section 4.2 of Lacey et al. (2016).

Reionization in the early Universe plays a decisive role in the formation of the earliest galaxies. Ionizing UV radiation from the first stars raises the temperature of the ambient hydrogen to $\sim 10^4$ K; this inhibits gas cooling in dark matter haloes whose effective virial temperature, $T_{\rm vir} \lesssim 10^4$ K, halting their growth (e.g. Doroshkevich, Zel'dovich & Novikov 1967; Couchman & Rees 1986; Rees 1986; Efstathiou 1992; Loeb & Barkana 2001; Benson et al. 2002b; Okamoto, Gao & Theuns 2008; Gnedin & Kaurov 2014). These galaxies make up the population of ultrafaint dwarfs seen today as, for example, the smallest satellite galaxies orbiting the Milky Way (Bose, Deason & Frenk 2018). It should be noted, however, that dwarfs with masses similar to the Milky Way's ultrafaints, but which form in the *field*, may continue to form stars long after reionization (Bovill & Ricotti 2011; Weisz et al. 2014). In GALFORM, the effect of the global photoionizing background is approximated by forbidding cooling in haloes of circular velocity, V_c, if $V_c < V_{\rm cut}$ at $z < z_{\rm cut}$. $V_{\rm cut}$ and $z_{\rm cut}$ are input parameters of the model and describe, respectively, the so-called filtering scale (Gnedin 2000; Benson et al. 2002b) and the redshift of reionization. Throughout, we assume $V_{\rm cut} = 30\,{\rm km\,s^{-1}}$ and $z_{\rm cut} = 6$, as recommended by Bose et al. (2018).

Finally, in GALFORM we also keep track of 'orphan galaxies', objects whose dark matter haloes have been disrupted below SUBFIND's detection limit of 20 particles after falling into a more massive halo. In this situation, the galaxies in them are tracked using the methodology described in Simha & Cole (2017): in short, the subhaloes of these galaxies are followed up to the last output time at which they are resolved. Their future orbital evolution is captured by identifying these galaxies with the most-bound particle of the subhalo in which they originally formed.

In our model, an orphan galaxy may be destroyed in one of two ways: (i) by sinking into centre by dynamical friction and merging with the central galaxy or (ii) by being tidally disrupted. The dynamical friction time-scale is determined using the analytic formula derived by Lacey & Cole (1993):

$$T_{\rm df} = \left(\frac{R_{\rm H}}{R_{\rm c}}\right)^{\alpha} \left(\frac{J}{J_{\rm c}}\right)^{\beta} \frac{\tau_{\rm dyn}}{2B(1)\ln\Lambda} \left(\frac{M_{\rm H}}{M_{\rm s}}\right), \quad (1)$$

which assumes a host halo density profile truncated at $R_{\rm H}$. $R_{\rm c}$ is the radius of a circular orbit in the halo that has the same energy as the actual orbit; $J/J_{\rm c}$ is the ratio of the angular momentum of the actual orbit to that of a circular orbit in the halo with the same energy; $\tau_{\rm dyn}$ is the dynamical time of the host halo; $M_{\rm H}$ is its mass; $M_{\rm s}$ is the mass of the satellite; $\ln\Lambda = \ln M_{\rm H}/M_{\rm s}$ is the Coulomb logarithm; $B(x) = {\rm erf(x)} - 2{\rm x}/\sqrt{\pi}\exp(-{\rm x}^2)$. Finally, $\alpha = -1.8$ and $\beta = 0.85$, as determined numerically by Simha & Cole (2017). An orphan galaxy is merged on to the central if the elapsed time between any two steps of the merger tree exceeds $T_{\rm df}$.

A satellite may also be destroyed if it experiences strong enough tidal forces to be considered tidally disrupted. In GALFORM, we determine this as the moment when the distance of the satellite from the centre of the halo falls below the tidal disruption radius, $R_{\rm td}$, defined as

$$R_{\rm td} = r_{\rm sat}\left(\frac{M_{\rm H}(<R_{\rm td})}{M_{\rm s}}\right)^{\frac{1}{3}}; \quad (2)$$

where $r_{\rm sat}$ is a characteristic length scale associated with the satellite (such as the half-light radius of the galaxy). In this definition, the tidal radius is defined as the radius within which the mean density of the host halo exceeds the mean density of the orbiting satellite.

Fig. 1 shows the importance of keeping track of orphans when calculating the satellite population in haloes. Here, we compare the average satellite luminosity function of Milky Way-mass haloes ($M_{200} = 1{-}1.3 \times 10^{12}\,{\rm M_\odot}$; Smith et al. 2007; Deason et al. 2012; Wang et al. 2015; Patel et al. 2018; Callingham et al. 2019; Deason et al. 2019; Grand et al. 2019) with and without the inclusion of orphans. The former shows a significantly larger satellite population, particularly in the regime of the ultrafaint dwarfs ($M_\star \leq 10^5\,{\rm M_\odot}$). The black curve shows the predictions from 100 Monte Carlo realizations of Galactic haloes generated using the extended Press-Schechter (EPS) formalism (Press & Schechter 1974; Bond et al. 1991; Bower 1991): the minimum halo mass in each tree is roughly a factor of 100 smaller than in COCO.[1]

[1]Note that the mass resolution in COCO is enough to resolve the atomic cooling limit ($V_c \sim 17\,{\rm kms^{-1}}$), so that any subhalo that can form stars through atomic cooling can do so. On the other hand, the parent COLOUR simulation just fails to resolve subhaloes near the atomic cooling limit. Thus, satellites in COCO make up a complete sample, whereas the larger volume

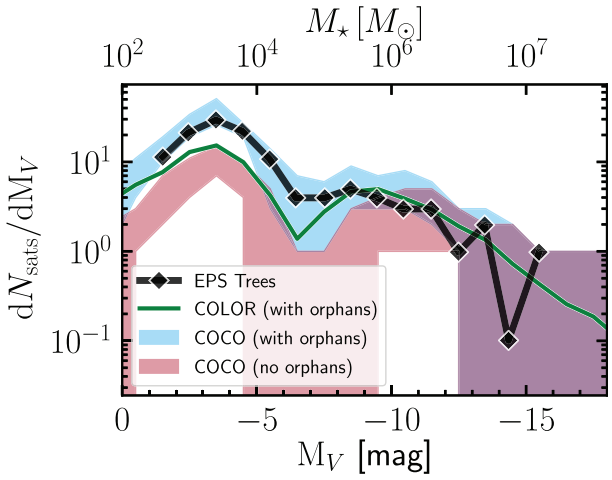

Figure 1. The average luminosity function of satellites in haloes of mass $M_{200} = 1{-}1.3 \times 10^{12}\,\mathrm{M}_\odot$ measured in COCO, with (blue) and without (red) the inclusion of orphan galaxies. Shaded regions encompass the 10th and 90th percentiles. For comparison, the black line shows the mean luminosity function from 100 Monte Carlo merger tree realizations of haloes in the same mass range; the resolution of each merger tree is 100 times better than COCO. The inclusion of orphans significantly boosts the abundance of the smallest galaxies, starting at around $M_\star \sim 10^6\,\mathrm{M}_\odot$. In particular, the calculation including orphans shows excellent agreement with the Monte Carlo trees, which demonstrates that tracking orphans mitigates the effects of limited numerical resolution in our *N*-body simulation. The thin green line shows the luminosity function from COLOUR, including orphans. While COLOUR can resolve ultrafaint galaxies, because of its lower mass resolution it does not produce a complete sample of this population (compared, for example, to our Monte Carlo trees). COLOUR serves to boost the statistics of Milky Way-mass haloes in this work.

The luminosity function including orphans is in excellent agreement with the much higher resolution calculation based on EPS trees. By contrast, a large number of ultrafaints are 'lost' when counting only galaxies in resolved subhaloes at $z = 0$ (see the curve without orphans). Newton et al. (2018b) have shown that including orphans greatly improves convergence in the radial distribution of satellites between simulations of the same halo at different resolution; this is a central aspect of the results presented in Section 3.4.

Tracking orphans also serves to compensate for any artificial disruption of subhaloes that might be caused by numerical effects such as those discussed by van den Bosch et al. (2018) and van den Bosch & Ogiya (2018), who suggest that subhaloes in cosmological simulations undergo excessive (unphysical) disruption, potentially underestimating the 'true' subhalo population. Their observations are related to the findings from idealized, controlled simulations by Peñarrubia et al. (2010) and Errani & Peñarrubia (2020), which suggest that 'cuspy' dark matter subhaloes are, in general, resilient to the effects of tides. Since observed ultrafaints are unlikely to have undergone the type of violent baryonic effects that can produce cores (e.g. Navarro, Eke & Frenk 1996; Read, Walker & Steger 2019), their haloes are more likely to survive and remain cuspy. The inclusion of orphan galaxies mitigates, at least in part, the effects of artificial disruption. Finally, it is important to stress that the orphan galaxy scheme does not add 'new' galaxies to the final catalogue, but simply keeps track of galaxies that had *already formed* in resolved subhaloes at earlier times and are therefore present in previous simulation output times.

Bose et al. (2018) provide a physical interpretation for the general form of the curves in Fig. 1 which they attribute to the way in which the process of reionization shapes the present-day distribution of dwarf galaxies. In particular, the luminosity function of satellites consists of two sub-populations: an ultrafaint component and a bright component separated by a 'valley' at $M_\mathrm{V} \approx -7$. The location of this kink is set by the filtering scale, corresponding to the choice of $V_\mathrm{cut} = 30$ km s^{-1} in the GALFORM model, which is itself motivated by analytic estimates (Rees 1986) and the results of hydrodynamical simulations (Gnedin 2000; Okamoto et al. 2008). The ultrafaint galaxies, located to the left of this valley in Fig. 1, assemble the bulk of their stellar mass prior to reionization and are quenched thereafter; this is consistent with the star formation histories inferred from the ultrafaint satellites orbiting the Milky Way (Brown et al. 2014). Brighter galaxies form later, and their star formation histories bear little memory of reionization (see also results from recent simulations by Simpson et al. 2013; Wheeler et al. 2015; Garrison-Kimmel et al. 2019). Finally, the amplitude of the luminosity function faintwards of the valley depends sensitively on the redshift at which reionization is complete. The bimodality of the satellite luminosity function in Fig. 1 is clearly present in the Milky Way data (Bose et al. 2018). In Section 3, we explore the connection between these satellite galaxies and the assembly histories of their host haloes.

The turnover in the luminosity function at $M_V \sim -4$ is likely a physical outcome of our model, not an artefact of finite resolution (indeed, a turnover at the same scale is also present in the Monte Carlo luminosity function presented in Fig. 1). We interpret the turnover as follows: small-mass haloes have a very short 'window' for forming stars. At early times (i.e. prior to the onset of reionization) the halo mass function builds up at low masses, but few of these haloes have grown massive enough for gas to be able to cool within them. As time proceeds, more haloes fall into the appropriate 'window' for star formation, before reionization inhibits further gas accretion. Thus, the turnover in the satellite luminosity function at the faintest end is likely to be related to this finite timespan for early galaxy formation.

3 RESULTS

In the following subsections, we present the main results of our study. In Section 3.1, we demonstrate the dependence of the total satellite population on the formation time of host haloes, focussing, in Section 3.2, on assembly histories similar to that of our own Galaxy. In Section 3.3, we study the influence of assembly history on the present-day radial distribution of satellite galaxies. Finally, in Section 3.4 we present a short discussion on corrections to these radial profiles after accounting for the destruction of satellite galaxies by the central galaxy.

3.1 Dependence of the satellite population on the assembly histories of host haloes

We begin our investigation by considering the connection between the present-day dwarf galaxy content of Milky Way-mass haloes and their average mass accretion histories. In particular, we are interested in contrasting the differences, if any, between the mass

COLOUR simulation provides enough statistical power to quantify qualitative trends identified in COCO.

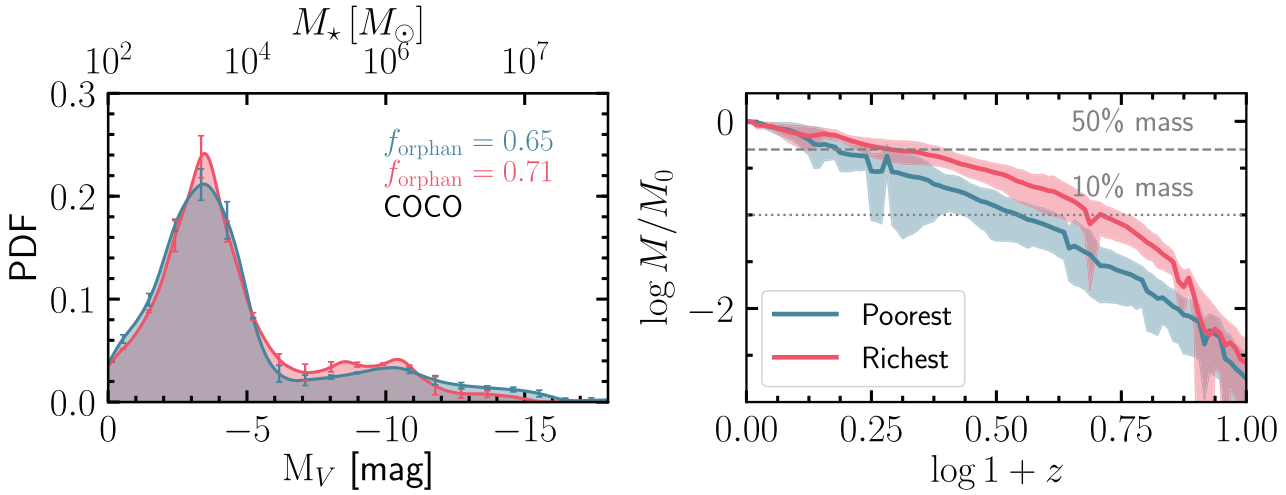

Figure 2. Correspondence between the mean satellite luminosity function (left-hand panel) of Milky Way-mass haloes ($M_{200}^{z=0} \equiv M_0 = 1{-}1.3 \times 10^{12}\,\mathrm{M}_\odot$) and the mean assembly history of their hosts (right-hand panel), obtained from COCO. The figure displays these relations for the 20 per cent 'richest' and 20 per cent 'poorest' haloes in terms of the abundance of their ultrafaint populations. The right-hand panel marks the epoch by which 10 per cent and 50 per cent of the present-day parent halo mass has collapsed. On average, early-forming Milky Way-analogues host more ultrafaint satellites than their later-forming counterparts. We have corrected the mean number of satellites for any differences in the mean final mass of host haloes in each of the 'rich' and 'poor' subcategories. The error bars in the left-hand panel and the shaded regions in the right-hand panel show the 16th and 84th percentile scatter in both quantities. Fig. B1 compares this result with and without the inclusion of orphan galaxies.

assembly of haloes that are particularly 'rich' or 'poor' in their ultrafaint satellite content at $z = 0$.

Predictions based on the COCO simulations are shown in Fig. 2, which focuses on haloes in the mass range $M_{200} = [1{-}1.3] \times 10^{12}\,\mathrm{M}_\odot$. Here, the distinction between the 'richest' and 'poorest' haloes, respectively, refers to the 20 per cent most abundant and the 20 per cent most deficient haloes in satellites of mass $M_\star \leq 10^5\,\mathrm{M}_\odot$ (see the left-hand panel). The mean accretion histories of the haloes, normalized to their present-day mass, are illustrated in the panels on the right. The offset between the two curves shows that, at fixed mass, haloes that form earlier tend to contain a larger population of ultrafaint satellites than their later-forming counterparts. The differences are larger at higher redshift: the 'richest' and 'poorest' haloes differ more in the epoch by which only 10 per cent of the final mass had been assembled than in the time when 50 per cent of the final mass was in place (a quantity often used to denote the redshift of formation of dark matter haloes). On average, early-forming haloes are ∼30 per cent richer in their ultrafaint populations. On the other hand, the abundance of 'classical satellites' ($M_V \leq -8.8$) is similar in the two classes of haloes.

The dependence of the number of ultrafaints on the formation time of haloes may be understood by considering the environmental dependence of halo merger rates at fixed mass. At fixed $z = 0$ mass, early-forming haloes are likely ones originating from Lagrangian patches located in regions of high local overdensity (e.g. Sheth & Tormen 2004; Avila-Reese et al. 2005). Fakhouri & Ma (2009) have shown that at fixed mass, galactic haloes experience, on average, ∼2.5 × more mergers (of all mass ratios) when located in the largest overdensities compared to their counterparts in underdensities. The increased number of mergers may subsequently translate into an increase in the abundance of satellites in early-forming haloes.[2]

Taken at face value, these results contradict the conclusions of previous studies which have found that, in fact, early-forming haloes have *depleted* satellite populations (e.g. Gao et al. 2004; Mao, Williamson & Wechsler 2015; Artale et al. 2018; Zehavi et al. 2018; Bose et al. 2019). However, there are two important distinctions between these studies and the present work: (1) we are concerned with ultrafaint satellites whereas previous studies considered much brighter satellites ($M_\star \gtrsim 10^7\,\mathrm{M}_\odot$), and (2) the contribution of orphans was not taken into account in previous work (i.e. only *resolved* subhaloes were considered). For these reasons, there is no conflict between our results and the conclusions of previous works in the regime where they are all valid (see Appendix B, Fig. B1 for further details).

Fig. 3 shows the correlation between the number of ultrafaint satellites, $N_{\rm uf}$, and the formation epoch of the host halo more quantitatively. The diamonds represent individual haloes extracted from COCO, and have been normalized to the mean number of ultrafaints, $<N_{\rm uf}>$, in the respective mass bin. Each panel corresponds to a different range of halo mass. Normalizing $N_{\rm uf}$ to the mean value allows us to compare the predictions of COCO and COLOUR, despite the difference in mass resolution. This figure shows that there is indeed a positive correlation between the number of ultrafaints hosted by a halo and how early it formed, when the latter is characterized by the half-mass formation time, $z_{\rm form,50\ per\ cent}$.

A simple way to quantify the strength of this correlation is through the Spearman rank correlation coefficient, $r_{\rm s}$, which is quoted in the bottom left corner of each panel. An average score of $r_{\rm s} \approx 0.35$ suggests a weak, positive correlation between $N_{\rm uf}$ and $z_{\rm form,50\ per\ cent}$. We find a similarly significant correlation between $N_{\rm uf}$ and $z_{\rm form,10\ per\ cent}$, the redshift by which 10 per cent of the host halo's final mass was assembled (not shown here). On the other hand, $r_{\rm s} \approx 0$ for the correlation between $N_{\rm uf}$ and $z_{\rm form,90\ per\ cent}$, the

[2]Note, however, that the *total* mass accretion rate, which determines the final-day mass of the halo, is comparable for objects irrespective of their local environment: while accretion through discrete halo mergers is *positively* correlated with environment, the accretion rate of smooth, diffuse material is *negatively* correlated with halo environment (Fakhouri & Ma 2010); the combined environmental dependence of the two modes of mass accretion therefore largely cancels out.

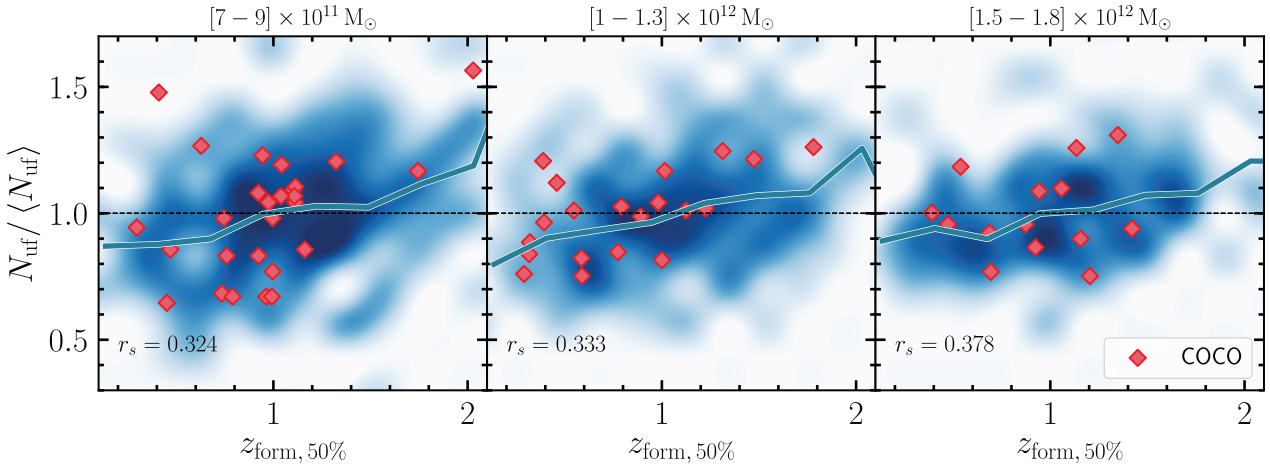

Figure 3. The number of ultrafaint dwarfs ($M_V \geq -6$; $M_\star \leq 10^5\,M_\odot$) as a function of halo formation redshift, z_{form}, defined as the time by which 50 per cent of the halo's final day mass was assembled. Each panel corresponds to a different bin of halo mass: $7-9\times10^{11}\,M_\odot$ (left), $1-1.3\times10^{12}\,M_\odot$ (middle), and $1.5-1.8\times10^{12}\,M_\odot$ (right). Red diamonds represent results from COCO. The blue cloud in the background of shows a 2D histogram of these quantities measured from COLOUR; the mean relation is indicated by the solid blue lines. In each panel, the number of ultrafaints is normalized to the mean number of ultrafaints hosted by haloes in that mass bin. We also give the Spearman rank correlation coefficient, r_s, in the bottom left corner of each panel.

redshift by which 90 per cent of the host halo's final mass was in place (also not shown). This indicates that there is no significant correlation between the number of ultrafaints and the late-time accretion history of haloes.

In summary, these trends indicate that at fixed halo mass, the ultrafaint satellite content of Milky Way-mass haloes is correlated with the early accretion history of the halo, but is largely insensitive to its late-time merger history. We note that by applying a semi-analytic model to a dark matter-only simulation in post-processing, our analysis is unable to capture hydrodynamic/feedback-related processes that may alter the merger history of haloes. For example, Schewtschenko & Macciò (2011) have suggested that in the presence of feedback, host haloes assemble later than their counterparts in dark matter-only simulations. This implies that any mergers (and the satellite galaxies they bring in) are delayed correspondingly. While our semi-analytic treatment is unable to capture this effect, we do not anticipate that this affects the *qualitative* difference in the satellite galaxy census associated with different classes of mergers within a *given* simulation (with or without feedback).

3.2 The satellite content of haloes with Milky Way-like assembly histories

In the previous subsection, we explored the connection between the richness of the ultrafaint satellite population of Milky Way-mass haloes and their accretion history, without specifying details of the events that determine the assembly of the halo. In this subsection, we consider the impact of the major accretion events thought to be relevant for the assembly of our Galaxy.

Fig. 4 shows the dependence of $N_{uf}/ < N_{uf} >$ on $z_{form,50\ per\ cent}$ for haloes in the mass range $M_{200} = [1-1.3]\times10^{12}\,M_\odot$, where we have highlighted objects that have undergone distinct past accretion events in different colours. In particular, we show haloes with an LMC-like accretion event ($M_{halo}^{infall} \sim 10^{11}\,M_\odot$ around 2 Gyr ago, with the condition that at least one satellite as bright as the LMC is present at $z = 0$; Besla 2015; Peñarrubia et al. 2016) in yellow, a *Gaia*-Enceladus-Sausage-like (GES-like) event ($M_{halo}^{infall} \sim 10^{11}\,M_\odot$ between 8 and 10 Gyr ago; Belokurov et al. 2018; Helmi et al. 2018) as a grey line (which shows the median relation trend with $z_{form,50\ per\ cent}$); haloes that have experienced both types of accretion events are coloured in red. Note that we use symbols to represent individual haloes that fall in the LMC-like and LMC+GES-like categories, as the statistics are not good enough to construct a meaningful median relation. We also show results from COLOUR only, as the number of candidate host haloes in COCO is too small to be able separate them into the three merger categories. In this figure, M_{halo}^{infall} is the dark matter halo mass of the accreted satellite just prior to infall. The fraction of haloes that fall into each category is given at the top of this panel; the corresponding collapsed histograms of N_{uf} are shown in the right-hand panel.

It is interesting to contrast the relative frequency of each type of merger. While an early GES-like accretion event is relatively common (roughly one in three haloes experience an event of this kind; see also Fattahi et al. 2019), a late-time LMC-like accretion event is less common (only around one in ten haloes experience this). The latter fraction is reduced by requiring that a satellite as bright as the LMC should survive to $z = 0$ (the fraction increases to 28 per cent if this condition is relaxed). Assembly histories like that of our own Galaxy, where *both* types of events have occurred, are exceedingly rare: only 3 per cent of the haloes in this mass range in COLOUR fall in this category.

Fig. 4 displays variations in the number of ultrafaint satellites in haloes that have experienced each type of accretion event. There is a roughly equal split of haloes above and below the mean population abundance in the category of LMC-like accretion events, while there is a marginal preference of an 'ultrafaint excess' in haloes that have experienced an early GES-like accretion i.e. 56 per cent of haloes in this category lie above the one-to-one line, as evidenced by the extended tail >1 in the grey histogram. Interestingly, the distribution for haloes that have experienced both early and late-time accretion events (red diamonds) shows a tendency towards an excess of ultrafaint satellites relative to the mean population. This is especially true for lower values of $z_{form,50\ per\ cent}$.

The top panel in Fig. 4 shows histograms of formation time. The median formation time of haloes in our chosen mass bin is marked by the vertical dashed line. We find that haloes that have experienced GES-like mergers are weighted more heavily towards the right side of this line ('early-forming'), while the yellow histogram denoting

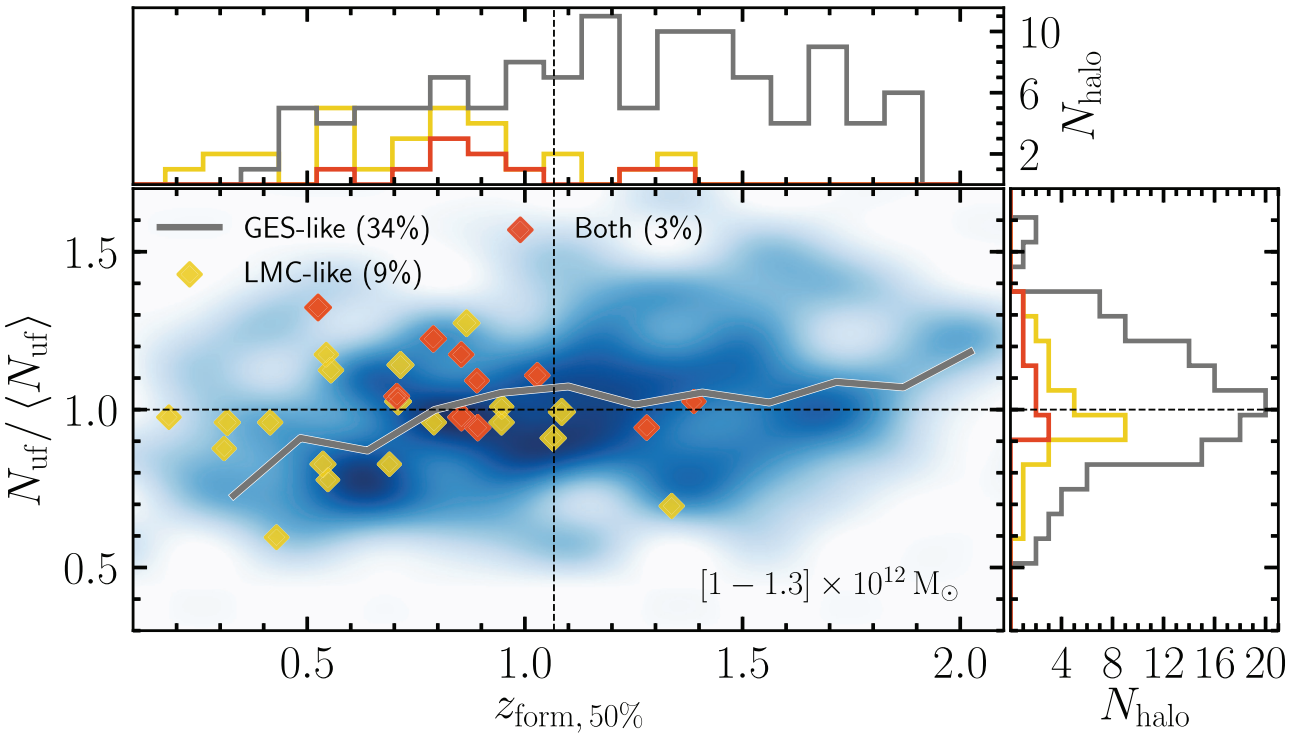

Figure 4. As Fig. 3, but we now highlight subpopulations of haloes that have undergone accretion events characteristic of our own Galaxy, specifically, an early GES-like accretion event (grey line, which shows the median trend with $z_{\mathrm{form,50\ per\ cent}}$), a late-time LMC-like accretion event (yellow) and both types of accretion event (red). The dashed vertical line marks the median value of $z_{\mathrm{form,50\ per\ cent}}$ for haloes in the chosen host halo mass bin. Results are shown for COLOUR only, as the size of the sample in COCO is too small to segregate into these three distinct categories.

LMC-like accretion events is weighted preferentially to the left ('late-forming'). Haloes that have experienced both types of events are those that were initially in the 'early-forming' category, but were reclassified as 'late-forming' after the accretion of the LMC-mass satellite. Haloes in this category have the double benefit of experiencing a significant, early accretion event in the form of the GES-like merger, which brings in its own ultrafaints, as well as the late time LMC-like event, which also brings in its own ultrafaints. This may explain, at least in part, why the red diamonds lie preferentially above the horizontal dashed line.

The observations presented in Figs 2–4 paint a consistent picture: early-forming haloes contain, on average, a larger than average number of ultrafaint satellites, and an 'ancient' GES-like accretion event is more influential in shaping the present-day satellite luminosity function than a late-time LMC-like accretion event.

3.3 Dependence of the radial distribution of satellites on the assembly histories of host haloes

The radial distribution of satellites contains information about the formation and accretion history of the host halo and their dynamical evolution after infall. From a practical point of view, predictions from cosmological simulations for the radial distribution of satellites serve as an important prior when correcting for incompleteness in estimating the Milky Way's total satellite population from the partial observed set (e.g. Koposov et al. 2008; Tollerud et al. 2008; Belokurov 2013; Hargis, Willman & Peter 2014; Jethwa et al. 2016; Kim, Peter & Hargis 2018; Newton et al. 2018b).

The radial distribution of satellite galaxies after infall is the net result of a number of physical processes that interact non-trivially. Gravitational processes such as dynamical friction and tidal disruption are clearly important in determining the bias of the subhalo population relative to the underlying dark matter profile (e.g. Ghigna et al. 2000; Nagai & Kravtsov 2005; Diemand, Kuhlen & Madau 2007; Springel et al. 2008; Ludlow et al. 2009; Han et al. 2016; Sawala et al. 2016), while the spatial bias of the galaxies relative to the smooth dark matter distribution and to subhaloes depends on the details of how galaxies occupy dark matter subhaloes (e.g. Frenk et al. 1996; Gao et al. 2004; van den Bosch et al. 2005; Conroy, Wechsler & Kravtsov 2006; Macciò et al. 2010; Budzynski et al. 2012; Reddick et al. 2013). In this subsection, we explore how the satellite distributions reflect differences in the assembly histories of haloes.

Fig. 5 shows the radial profiles of satellites in hosts of mass $M_{200} = [1-1.3]\times10^{12}\,\mathrm{M}_\odot$ in COCO (dashed lines) and COLOUR (solid lines). The black curves show the average radial profile, while the red lines show the average profile of the 20 per cent earliest-forming, and the blue line of the 20 per cent latest-forming, haloes in this mass bin, defined by their values of $z_{\mathrm{form,50\ per\ cent}}$. The left-hand panel shows the mean number of satellites identified within distance r from the centre of the halo, while the right-hand panel shows the fraction of the total satellite population located within this radius. The dashed vertical lines mark the mean virial radii of haloes in each category.

Fig. 5 shows clear departures from the average profile when splitting haloes by their assembly time. Although early-forming haloes typically have lower masses (and smaller virial radii), we see that they tend to contain more satellites within a fixed physical halocentric radius than their later-forming counterparts. This is especially true within the innermost 100 kpc or so. That late-forming haloes have a more spatially extended satellite population is seen

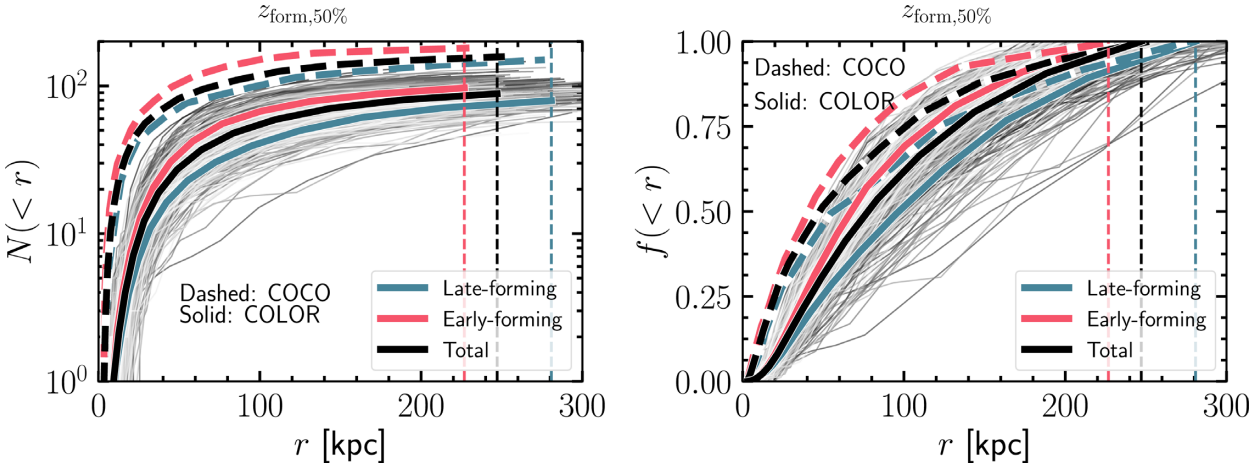

Figure 5. The radial distribution of satellites in galactic haloes ($M_{200} \sim 1{-}1.3 \times 10^{12}\,{\rm M}_\odot$) identified in the COCO and COLOUR simulations. Left-hand panel: The total number of satellites within distance, r, from the halo centre. Right-hand panel: The fraction of the total satellite population within this radius. In each panel, the grey curves show the radial profiles of individual haloes in this mass bin; the black curve is the corresponding mean profile. We distinguish between the profiles of the 20 per cent earliest- (red) and 20 per cent latest-forming (blue) haloes as defined by the redshift at which 50 per cent of the present-day halo mass was in place. The dashed vertical lines mark the corresponding virial radii. We note that the systematic differences observed in this figure are preserved when the radial profile is rescaled by the virial radius (i.e. when the x-axis is expressed as r/r_{200}). Neglecting orphan galaxies changes this radial distribution considerably; this comparison is presented in Fig. B2.

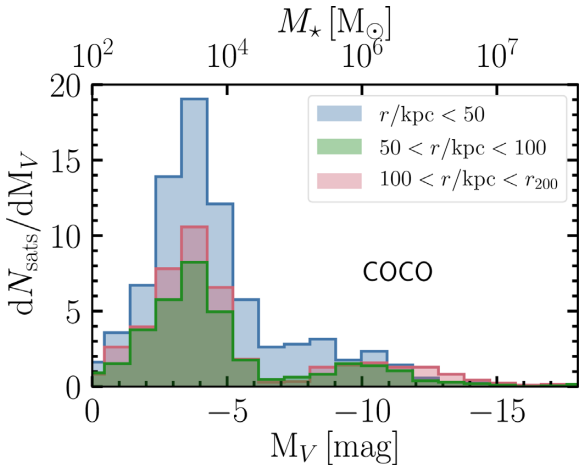

Figure 6. The luminosity function of satellites in COCO split into radial shells. The halo mass range adopted here corresponds to $M_{200} = 1{-}1.3 \times 10^{12}\,{\rm M}_\odot$. Whereas the classical satellites are found in roughly equal number throughout the halo, the ultrafaint population is found predominantly in the innermost 50 kpc. The results for COLOUR (not shown in this diagram) are qualitatively similar, albeit with different number counts. The effect of excluding the orphan galaxy population is shown in Fig. B3.

clearly in the right-hand panel of Fig. 5, while 75 per cent of the total satellite content of early-forming haloes is contained within 100 kpc of halo centre, only about 50 per cent of the late-forming population lies in this region. This is perhaps unsurprising: later-forming haloes are more likely to have undergone a more recent merger, in which case any satellites that have been brought in during this event are located preferentially in the outskirts of the host halo as dynamical friction has not operated for long enough to bring them to the centre.

Simulations also show that satellites accreted later tend to have larger apocentres (see fig. 7 in Deason et al. 2013). This is likely the determining factor in the extended distribution of low-mass dwarf satellites for which dynamical friction is not very effective in any case. These qualitative trends change when orphan galaxies are not included, as shown in Fig. B2. We also note that the same qualitative dependence of the radial profile on assembly history is seen in both COCO and COLOUR, although the number of galaxies found within any given halocentric radius is different because of the different mass resolutions of the two simulations. Although not shown here, we find that the radial distribution of satellite galaxies in GALFORM is, in general, more centrally concentrated than that of the dark matter subhaloes, with a slope that is comparable to that of the total dark matter density profile (see also Springel et al. 2008; Bose et al. 2019).

Fig. 6 splits the luminosity function of satellites into radial shells to highlight the mass range of satellites that dominates the population at a given radius. The classical dwarfs ($M_\star \geq 10^6\,{\rm M}_\odot$) are distributed relatively evenly throughout the halo, although the brightest satellites are located preferentially in the outer parts. The ultrafaints, on the other hand, are more centrally concentrated. We note that it is in the inner regions of haloes where the effects of finite resolution are most important, and where the orphan galaxy tracking scheme is essential. In Fig. B3, we show how the radial occupation of satellites of a given mass changes significantly when orphans are not followed explicitly.

3.4 The destruction of satellites by the central disc

In the previous subsections, we have described how the assembly histories of haloes give rise to differences in their present-day ultrafaint dwarf galaxy populations, both in number and in spatial distribution. We have especially highlighted the importance of tracking orphan galaxies when making these comparisons.

A well-known process that we now need to consider is the destruction of satellite galaxies resulting from interactions with the central baryonic disc. Both idealized and cosmological hydrodynamical simulations have shown that the amount of destruction is significant, particularly for satellites moving in radial orbits. The destruction is more severe near the centre and is enhanced around more massive central galaxies (e.g. Zentner et al. 2005; D'Onghia et al. 2010;

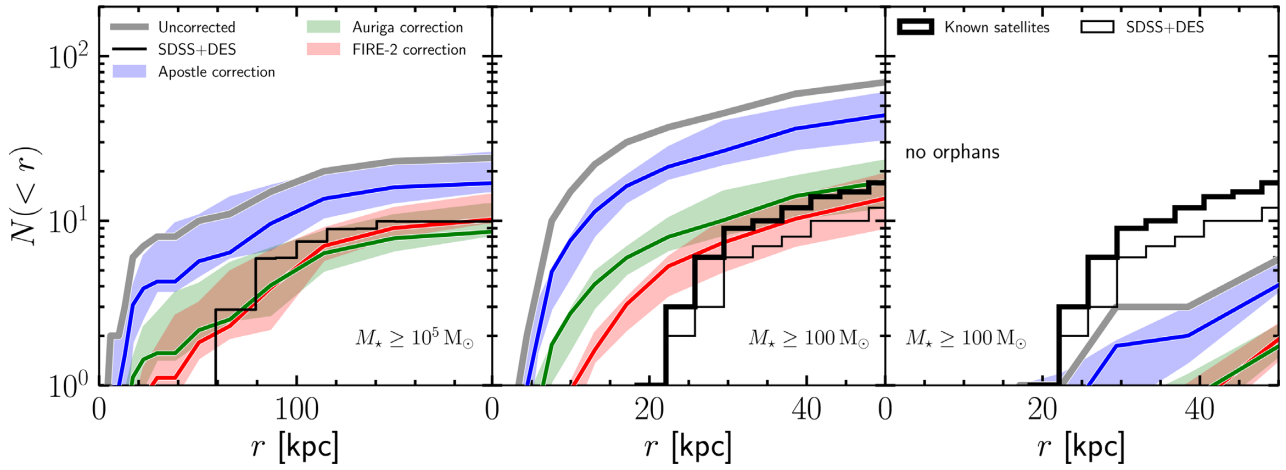

Figure 7. The radial distribution of satellites in galactic haloes from the COCO simulation. Left-hand panel: The mean radial distribution of satellites more massive than $M_\star \geq 10^5\,\mathrm{M}_\odot$, before (grey line) and after (coloured bands) correcting for disc destruction. The size of the depletion correction varies from simulation to simulation: blue and green bands are the results found by Richings et al. (2020) in the APOSTLE and AURIGA simulations, respectively, and red bands are those found in FIRE-2 by Samuel et al. (2020). The shaded regions encompass the 5th–95th percentile scatter in the profiles measured in COCO. The black histogram shows the radial distribution of spectroscopically confirmed satellites in SDSS and DES that fall in this mass range. Middle panel: A zoom into the radial distribution within the innermost 50 kpc, with linestyles and colours identical to those in the left-hand panel. The figure now shows the radial distributions including all satellites more massive than $M_\star \geq 100\,\mathrm{M}_\odot$. The thick black histogram is the distribution of *all-known* satellites within 50 kpc, including those that have not been confirmed spectroscopically while the thin black histogram includes only those that have been confirmed. In both panels, the observational histograms have not been corrected for survey incompleteness, and should be treated as lower limits. In general, there is good agreement between the observed radial distribution and the simulated profiles after correcting for the effects of disc destruction. Right-hand panel: As in the middle panel, but neglecting 'orphan' galaxies in the simulations. The exclusion of the orphan population reduces the satellite population dramatically; this is exacerbated when the effects of disc destruction are taken into account.

Zhu et al. 2016; Garrison-Kimmel et al. 2017; Sawala et al. 2017; Simpson et al. 2018; Buck et al. 2019; Richings et al. 2020).

Graus et al. (2019) have suggested that when disc disruption is accounted for, ΛCDM simulations are inconsistent with the abundance of satellites observed in the inner parts of the Milky Way. They find that under the assumption of 'standard reionization quenching' – in which galaxies with peak circular velocity $V_{\rm peak} \leq 20$ km s^{-1} no longer accrete gas after reionization (Okamoto et al. 2008, 2010) there are far fewer subhaloes within 50 kpc of the centre in the dark matter-only simulations than the observed number of satellites in the Milky Way. For their model to produce enough satellites in the inner galaxy, subhaloes as small as $V_{\rm peak} \simeq 7$ km s^{-1} need to be populated with galaxies. As we shall see below the inclusion of orphan galaxies is enough to explain the Milky Way data within the conventional model of reionization.

While the merger of satellites with the central baryonic disc is accounted for in GALFORM, the explicit destruction of satellites by interaction with the disc is not. We therefore correct the radial profiles in GALFORM using the fitting functions provided by Samuel et al. (2020). These authors compare dark matter-only simulations with hydrodynamical simulations using the FIRE-2 model (Hopkins et al. 2018) and find that the relative difference, $f(d)$, in the corresponding radial profiles is well fit by the functional form:

$$f(d) = \begin{cases} 0\,, & 0 \leq d \leq d_0 \\ \alpha \left[1 - \exp\left(-\frac{d-d_0}{d_1}\right)\right], & d \geq d_0 \end{cases} \quad (3)$$

where d is the distance from the host galaxy and α, d_0, and d_1 are free parameters that are fixed by fitting this function to the simulated radial profiles. This effect has also been quantified by Richings et al. (2020), who performed a comprehensive analysis of the effects of disc destruction in the APOSTLE (Sawala et al. 2016; Fattahi et al. 2016) and AURIGA (Grand et al. 2017) cosmological, hydrodynamical simulations. We find that the functional form in equation 3 provides an equally good fit to the results of these simulations. The best-fitting values of α, d_0, and d_1 for FIRE-2, APOSTLE, and AURIGA are listed in Table A1.

The effects of disc destruction are illustrated in Fig. 7. The mean radial profile before applying any correction is shown by the solid grey line; the radial profiles after applying the corrections for disc destruction (using equation 3) are shown by the coloured bands, with each colour corresponding to the correction inferred from a different simulation: FIRE-2 (red), APOSTLE (blue), and AURIGA (green). Note that the radial distributions in this figure come from the high-resolution COCO simulation (rather than the lower resolution COLOUR simulation used for Fig. 5). The higher resolution of COCO is required to follow the very small ($M_\star \simeq 100\,\mathrm{M}_\odot$) satellites shown in the right-hand panel of Fig. 7. We recall that the luminosity function of satellites in COCO gives a very good match to the 'resolution-free' Monte Carlo calculation (see Fig. 1).

Accounting for disc destruction clearly reduces the abundance of satellites in all cases. The reduction is particularly large within the innermost 80 kpc or so, especially for the FIRE-2 model. The average number of satellites with $M_\star \geq 10^5\,\mathrm{M}_\odot$ within the virial radius is reduced, respectively, by 18 per cent, 56 per cent, and 62 per cent when using the APOSTLE, AURIGA, and FIRE-2-based corrections. The differences amongst the different simulations arise primarily from the mass of the disc which is smaller in APOSTLE than in AURIGA and FIRE-2. Further details of the reasons behind these differences are given in Appendix A. In general, the profiles predicted by GALFORM reproduce the rather flat distribution at large radii characteristic of the Milky Way's radial satellite profile, extending out to ∼300 kpc.

The middle panel in Fig. 7 highlights the severity of disc disruption close to the galaxy by zooming into the innermost 50 kpc of the haloes. The black histograms show the radial distribution

of observed satellites distinguishing the subset that have been spectroscopically confirmed as satellite galaxies in the SDSS and DES catalogues (thin black histogram) from the entire population of candidate satellites (thick black histogram). If disc destruction is neglected, our model overpredicts the abundance of satellites within 50 kpc. However, once disc destruction is accounted for, the predicted radial profiles are in excellent agreement with the observations, at least when the corrections based on AURIGA and FIRE-2 are applied. The histograms in the figure show that there are no observed satellites within ∼20 kpc from the Galactic Centre; however, these data have not been corrected for sky coverage or survey incompleteness, and should therefore be treated as lower bounds on the 'true' distribution. The AURIGA and FIRE-2 models predict $\sim 7 \pm 2$ and 4 ± 2 ultrafaint satellites within 20 kpc, respectively.

The main conclusion drawn from Fig. 7 is at odds with the results of Graus et al. (2019) who, after including the effects of disc destruction and assuming the standard model of reionization, found about 10 times fewer satellites within 50 kpc of the halo centre than the observed number in this region of the Milky Way. The reason why GALFORM predicts far more ultrafaint satellites in the inner regions of galactic haloes is simply the inclusion of 'orphan' galaxies, that is galaxies whose dark matter haloes have been lost due to numerical resolution effects as the number of particles in the subhalo drops below a certain level (20 particles in the case of SUBFIND subhaloes). The resolution of the dark matter-only simulations used by Graus et al. (2019) is $3 \times 10^4\,\mathrm{M}_\odot$. This is about a factor of 3 lower resolution than the 'Level-2' Aquarius simulations of galactic haloes (Springel et al. 2008) for which Newton et al. (2018a) have shown orphan galaxies are important, particularly in the central regions. (They are important even at Aquarius 'Level 1' resolution of $\sim 10^3\,\mathrm{M}_\odot$.)

That neglecting orphan galaxies leads to a significant underestimate of the number of faint satellites can be seen in the rightmost panel of Fig. 7. The radial profiles that we find in this case are consistent with those of Graus et al. (2019). Including the orphans obviates the need to populate extremely small subhaloes, as Graus et al. (2019) had to do in their simulations to obtain as many ultrafaint satellites as observed. In fact, in GALFORM the smallest haloes that ever form a galaxy have $V_{\rm peak} \sim 16.8\,\mathrm{km\,s^{-1}}$. At $z = 7$, when, according to GALFORM, these satellites have, on average, formed around 50 per cent of their $z = 0$ stellar mass, this corresponds to a *halo mass* of $\sim 1.4 \times 10^8\,\mathrm{M}_\odot$. It is important to emphasize that the GALFORM model used here was not tuned to reproduce the observed radial distribution of satellites. The importance of tracking subresolution orphan galaxies, particularly when considering the statistics of satellite populations in the vicinity of the Milky Way disc, has also been recently pointed out by Nadler et al. (2018, 2019) and Li, Gao & Wang (2019).

This discussion underlines the importance of limited numerical resolution in *N*-body simulations and the need to account for it, for example, using the orphan galaxy tracking scheme implemented in GALFORM. Including or ignoring these objects leads to very different inferences for the galaxy-halo connection in the dwarf galaxy regime.

4 CONCLUSIONS

We have used *N*-body simulations in the ΛCDM cosmology to explore the connection between the assembly history of Milky Way-mass dark matter haloes and their satellite population at $z = 0$, including 'ultrafaint' satellites and focusing on the abundance and spatial distribution of satellites.

Our sample of Milky Way-mass haloes was extracted from the dark matter-only *Copernicus Complexio Low Resolution* (COLOUR) and *Copernicus Complexio* (COCO) *N*-body simulations (Bose et al. 2016; Hellwing et al. 2016; Sawala et al. 2016) of structure formation in a ΛCDM universe. The large computational volume (10^6 Mpc^3) of the former provides a rich statistical sample of haloes of mass comparable to that of the Milky Way, whereas the 60 times higher mass resolution of COCO probes the smallest galaxies expected to form in haloes above the atomic gas cooling limit ($V_c \sim 17\,\mathrm{kms^{-1}}$). To embed galaxies within these simulations, we made use of the Durham semi-analytic model of galaxy formation, GALFORM (Cole et al. 2000; Lacey et al. 2016), which calculates the physical processes involved in galaxy formation along halo merger trees built from each simulation. GALFORM includes a detailed treatment of early hydrogen reionization, gas cooling, star and black hole formation, feedback from stars and AGN, metal production, the synthesis of stellar populations, etc. It provides a flexible and computationally inexpensive environment to explore the parameter space of galaxy formation.

This work focuses on the demographics of the ultrafaint satellites of the Milky Way ($M_\star \leq 10^5\,\mathrm{M}_\odot$, $M_V \geq -7$), and asks what their present-day abundance and radial distribution tells us about the assembly history of the host dark matter halo. These galaxies typically form in low-mass dark matter subhaloes, often near the detection threshold of the substructure finding algorithms that are applied to *N*-body simulations. To ensure that these galaxies are accounted for, even after their dark matter (sub)halo falls below the nominal resolution limit, we employ the technique of 'orphan galaxy tracking'. This method uses information from the last epoch at which the subhalo was resolved and follows its subsequent evolution by tagging its galaxy with the most-bound particle of the subhalo. Analytic prescriptions are then used to assess their survival under the influence of dynamical friction and tidal disruption (e.g. Simha & Cole 2017). The orphan method results in a dramatic improvement in the agreement between simulations at different resolution, as well as with Monte Carlo realizations of the galactic population that are not limited by numerical resolution (Guo et al. 2011 and e.g. fig. 1 of Newton et al. 2018b). Our main results are summarized as follows:

(i) At fixed halo mass, that galactic dark matter haloes that form earlier than the average population of that mass contain more satellite galaxies than their later-forming counterparts (Fig. 2).

(ii) The disparity in the total number of satellite galaxies between early and late-forming haloes is due primarily to ultrafaint satellites. The difference is manifest only when selecting haloes according to their early formation history (i.e. by the time when 10 per cent or 50 per cent of the present-day mass was in place). We find no correlation between the number of ultrafaint satellites and variations in the recent merger history of the host haloes (Fig. 3).

(iii) Translated into the language of the Milky Way's assembly, our models predict that an ancient GES-like accretion event is more likely to have brought in a large number of ultrafaint galaxies than the more recent accretion of an LMC-like system. One plausible explanation for this is that an appreciable number of ultrafaint dwarfs may have already been destroyed inside the late-accreted host, diminishing its contribution to the Milky Way's satellite population. A full investigation would require tracking the fate of every satellite galaxy (and its progenitors) that ever falls into an LMC-mass halo; we leave this to future work. We also find that

systems that have undergone *both* GES-like and LMC-like accretion events in their history are exceedingly rare: only ∼3 per cent of the haloes in our sample have experienced GES and LMC-like mergers (Fig. 4).

(iv) The radial distribution of satellites is more centrally concentrated in early-forming haloes. Haloes of a given present-day mass that assemble early have more centrally concentrated matter density profiles, and the radial distribution of their satellite populations is similarly more concentrated. A (perhaps secondary) effect is that massive structures accreted early on (such as the GES progenitor), along with their satellites, are brought close to the central regions of the host by dynamical friction. The majority of these satellites are ultrafaint (Fig. 5). A sizeable fraction of these galaxies are identified as orphans (Fig. B2).

(v) We accounted for the destruction of satellites by a central disc using the results of Samuel et al. (2020) for the FIRE-2, and those Richings et al. (2020) for the APOSTLE and AURIGA hydrodynamical simulations (see Appendix A). When orphans are not included, satellites are entirely missing within the inner 50 kpc of the halo, in severe tension with the observed radial distribution which has about 20 satellites in this region (the right-hand panel of Fig. 7). Including orphans, however, bring the theoretical prediction into good agreement with the data (the middle panel of Fig. 7).

In this work, we have explored the intimate connection between the present-day satellite population of galactic dark matter haloes and their formation history. The synergy of observational facilities such as the SDSS, DES, and *Gaia* have transformed our understanding of the build-up of our own Galaxy. As we continue to extend the census of ultrafaint dwarfs around galaxies other than our own, the potential for developing a general picture of the assembly of galaxies becomes even greater. Theoretical models need to keep pace with the data; it is clear that the extreme ultrafaint dwarf galaxy regime continues to push the boundaries of state-of-the-art simulations, in resolution, quality of the galaxy formation physics models and subhalo tracking methods. Although this task is challenging, the wealth of information provided by these low-mass denizens, for both cosmology and galaxy formation, provides strong motivation to continue to strive for further advancements in observations, cosmological simulations, and numerical methods.

ACKNOWLEDGEMENTS

We thank the referee for suggesting a number of changes that have improved this paper significantly. SB is supported by Harvard University through the ITC Fellowship. AJD is supported by a Royal Society University Research Fellowship, and she and CSF by the STFC Consolidated Grant for Astronomy at Durham (ST/L00075X/1). CSF is also supported by ERC Advanced Investigator grant, DMIDAS [GA 786910]. This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grants ST/H008519/1 and ST/K00087X/1, STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. This research was carried out with the support of the HPC Infrastructure for Grand Challenges of Science and Engineering Project, co-financed by the European Regional Development Fund under the Innovative Economy Operational Programme. The work of SB was performed in part at Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611. The data analysed in this paper can be made available upon request to the author.

REFERENCES

Adelman-McCarthy J. K. et al., 2007, ApJS, 172, 634
Agertz O. et al., 2020, MNRAS, 491, 1656
Alam S. et al., 2015, ApJS, 219, 12
Artale M. C., Zehavi I., Contreras S., Norberg P., 2018, MNRAS, 480, 3978
Avila-Reese V., Colín P., Gottlöber S., Firmani C., Maulbetsch C., 2005, ApJ, 634, 51
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Bechtol K. et al., 2015, ApJ, 807, 50
Belokurov V., 2013, NewAR, 57, 100
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Bennet P., Sand D. J., Crnojević D., Spekkens K., Karunakaran A., Zaritsky D., Mutlu-Pakdil B., 2019, ApJ, 885, 153
Benson A. J., 2012, New Astron., 17, 175
Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002a, MNRAS, 333, 156
Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002b, MNRAS, 333, 177
Besla G., 2015, preprint (arXiv:1511.03346)
Besla G., Kallivayalil N., Hernquist L., Robertson B., Cox T. J., van der Marel R. P., Alcock C., 2007, ApJ, 668, 949
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bose S., Hellwing W. A., Frenk C. S., Jenkins A., Lovell M. R., Helly J. C., Li B., 2016, MNRAS, 455, 318
Bose S., Deason A. J., Frenk C. S., 2018, ApJ, 863, 123
Bose S., Eisenstein D. J., Hernquist L., Pillepich A., Nelson D., Marinacci F., Springel V., Vogelsberger M., 2019, MNRAS, 490, 5693
Bovill M. S., Ricotti M., 2009, ApJ, 693, 1859
Bovill M. S., Ricotti M., 2011, ApJ, 741, 18
Bower R. G., 1991, MNRAS, 248, 332
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Boylan-Kolchin M., Besla G., Hernquist L., 2011, MNRAS, 414, 1560
Boylan-Kolchin M., Bullock J. S., Sohn S. T., Besla G., van der Marel R. P., 2013, ApJ, 768, 140
Brooks A. M., Zolotov A., 2014, ApJ, 786, 87
Brown T. M. et al., 2014, ApJ, 796, 91
Buck T., Macciò A. V., Dutton A. A., Obreja A., Frings J., 2019, MNRAS, 483, 1314
Budzynski J. M., Koposov S. E., McCarthy I. G., McGee S. L., Belokurov V., 2012, MNRAS, 423, 104
Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517
Callingham T. M. et al., 2019, MNRAS, 484, 5453
Chambers K. C. et al., 2016, preprint (arXiv:1612.05560)
Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Corlies L., Johnston K. V., Wise J. H., 2018, MNRAS, 475, 4868
Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, MNRAS, 450, 1514
Couchman H. M. P., Rees M. J., 1986, MNRAS, 221, 53
Croton D. J. et al., 2006, MNRAS, 365, 11
D'Onghia E., Springel V., Hernquist L., Keres D., 2010, ApJ, 709, 1138
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Deason A. J. et al., 2012, MNRAS, 425, 2840
Deason A. J., Belokurov V., Evans N. W., Johnston K. V., 2013, ApJ, 763, 113
Deason A. J., Wetzel A. R., Garrison-Kimmel S., Belokurov V., 2015, MNRAS, 453, 3568

Deason A. J., Fattahi A., Belokurov V., Evans N. W., Grand R. J. J., Marinacci F., Pakmor R., 2019, MNRAS, 485, 3514
Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859
Dooley G. A., Peter A. H. G., Carlin J. L., Frebel A., Bechtol K., Willman B., 2017, MNRAS, 472, 1060
Doroshkevich A. G., Zel'dovich Y. B., Novikov I. D., 1967, SvA, 11, 233
Drlica-Wagner A. et al., 2015, ApJ, 813, 109
Efstathiou G., 1992, MNRAS, 256, 43P
Errani R., Peñarrubia J., 2020, MNRAS, 491, 4591
Errani R., Peñarrubia J., Laporte C. F. P., Gómez F. A., 2017, MNRAS, 465, L59
Fakhouri O., Ma C.-P., 2009, MNRAS, 394, 1825
Fakhouri O., Ma C.-P., 2010, MNRAS, 401, 2245
Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, MNRAS, 406, 2267
Fattahi A. et al., 2016, MNRAS, 457, 844
Fattahi A. et al., 2019, MNRAS, 484, 4471
Font A. S. et al., 2011, MNRAS, 417, 1260
Frenk C. S., White S. D. M., Efstathiou G., Davis M., 1985, Nature, 317, 595
Frenk C. S., Evrard A. E., White S. D. M., Summers F. J., 1996, ApJ, 472, 460
Gaia Collaboration et al., 2018, A&A, 616, A1
Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004, MNRAS, 355, 819
Garrison-Kimmel S. et al., 2019, MNRAS, 489, 4574
Garrison-Kimmel S., Bullock J. S., Boylan-Kolchin M., Bardwell E., 2017, MNRAS, 464, 3108
Geha M. et al., 2017, ApJ, 847, 4
Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 2000, ApJ, 544, 616
Gnedin N. Y., 2000, ApJ, 535, 530
Gnedin N. Y., Kaurov A. A., 2014, ApJ, 793, 30
Grand R. J. J., Deason A. J., White S. D. M., Simpson C. M., Gómez F. A., Marinacci F., Pakmor R., 2019, MNRAS, 487, L72
Grand R. J. J. et al., 2017, MNRAS, 467, 179
Graus A. S., Bullock J. S., Kelley T., Boylan-Kolchin M., Garrison-Kimmel S., Qi Y., 2019, MNRAS, 488, 4585
Guo Q. et al., 2011, MNRAS, 413, 101
Han J., Cole S., Frenk C. S., Jing Y., 2016, MNRAS, 457, 1208
Hargis J. R., Willman B., Peter A. H. G., 2014, ApJ, 795, L13
Hellwing W. A., Frenk C. S., Cautun M., Bose S., Helly J., Jenkins A., Sawala T., Cytowski M., 2016, MNRAS, 457, 3492
Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663
Hopkins P. F. et al., 2018, MNRAS, 480, 800
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Jahn E. D., Sales L. V., Wetzel A., Boylan-Kolchin M., Chan T. K., El-Badry K., Lazar A., Bullock J. S., 2019, MNRAS, 489, 5348
Jeon M., Besla G., Bromm V., 2017, ApJ, 848, 85
Jethwa P., Erkal D., Belokurov V., 2016, MNRAS, 461, 2212
Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, MNRAS, 440, 2115
Kallivayalil N., van der Marel R. P., Alcock C., 2006, ApJ, 652, 1213
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Kennedy R., Frenk C., Cole S., Benson A., 2014, MNRAS, 442, 2487
Kim D., Jerjen H., Mackey D., Da Costa G. S., Milone A. P., 2015, ApJ, 804, L44
Kim S. Y., Peter A. H. G., Hargis J. R., 2018, Phys. Rev. Lett., 121, 211302
Komatsu E. et al., 2011, ApJS, 192, 18
Kondapally R., Russell G. A., Conselice C. J., Penny S. J., 2018, MNRAS, 481, 1759
Koposov S. et al., 2008, ApJ, 686, 279
Koposov S. E., Belokurov V., Torrealba G., Evans N. W., 2015, ApJ, 805, 130
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lacey C. G. et al., 2016, MNRAS, 462, 3854
Laevens B. P. M. et al., 2015, ApJ, 802, L18
Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, MNRAS, 418, 1649
Li M., Gao L., Wang J., 2019, MNRAS, 483, 2000
Loeb A., Barkana R., 2001, ARA&A, 39, 19
Lovell M. R. et al., 2012, MNRAS, 420, 2318
Lovell M. R. et al., 2016, MNRAS, 461, 60
Ludlow A. D., Navarro J. F., Springel V., Jenkins A., Frenk C. S., Helmi A., 2009, ApJ, 692, 931
Macciò A. V., Fontanot F., 2010, MNRAS, 404, L16
Macciò A. V., Kang X., Fontanot F., Somerville R. S., Koposov S., Monaco P., 2010, MNRAS, 402, 1995
Mackereth J. T. et al., 2019, MNRAS, 482, 3426
Mao Y.-Y., Williamson M., Wechsler R. H., 2015, ApJ, 810, 21
Maraston C., 2005, MNRAS, 362, 799
McBride J., Fakhouri O., Ma C.-P., 2009, MNRAS, 398, 1858
Munshi F., Brooks A. M., Christensen C., Applebaum E., Holley-Bockelmann K., Quinn T. R., Wadsley J., 2019, ApJ, 874, 40
Myeong G. C., Evans N. W., Belokurov V., Sand ers J. L., Koposov S. E., 2018, ApJ, 863, L28
Nadler E. O., Mao Y.-Y., Wechsler R. H., Garrison-Kimmel S., Wetzel A., 2018, ApJ, 859, 129
Nadler E. O., Gluscevic V., Boddy K. K., Wechsler R. H., 2019, ApJ, 878, L32
Nagai D., Kravtsov A. V., 2005, ApJ, 618, 557
Navarro J. F., Eke V. R., Frenk C. S., 1996, MNRAS, 283, L72
Newton O., Cautun M., Jenkins A., Frenk C. S., Helly J. C., 2018a, Proc. IAU Symp. 344, Dwarf Galaxies: From the Deep Universe to the Present. p. 109
Newton O., Cautun M., Jenkins A., Frenk C. S., Helly J. C., 2018b, MNRAS, 479, 2853
Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920
Okamoto T., Frenk C. S., Jenkins A., Theuns T., 2010, MNRAS, 406, 208
Patel E., Besla G., Mandel K., Sohn S. T., 2018, ApJ, 857, 78
Peñarrubia J., Benson A. J., Walker M. G., Gilmore G., McConnachie A. W., Mayer L., 2010, MNRAS, 406, 1290
Peñarrubia J., Gómez F. A., Besla G., Erkal D., Ma Y.-Z., 2016, MNRAS, 456, L54
Press W. H., Schechter P., 1974, ApJ, 187, 425
Read J. I., Agertz O., Collins M. L. M., 2016, MNRAS, 459, 2573
Read J. I., Walker M. G., Steger P., 2019, MNRAS, 484, 1401
Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30
Rees M. J., 1986, MNRAS, 218, 25P
Rey M. P., Pontzen A., Agertz O., Orkney M. D. A., Read J. I., Saintonge A., Pedersen C., 2019, ApJ, 886, L3
Richings J.et al., 2020, MNRAS, 492, 5780
Sales L. V., Navarro J. F., Cooper A. P., White S. D. M., Frenk C. S., Helmi A., 2011, MNRAS, 418, 648
Sales L. V., Navarro J. F., Kallivayalil N., Frenk C. S., 2017, MNRAS, 465, 1879
Samuel J. et al., 2020, MNRAS, 491, 1471
Sawala T. et al., 2016, MNRAS, 457, 1931
Sawala T., Pihajoki P., Johansson P. H., Frenk C. S., Navarro J. F., Oman K. A., White S. D. M., 2017, MNRAS, 467, 4383
Schewtschenko J. A., Macciò A. V., 2011, MNRAS, 413, 878
Shao S., Cautun M., Deason A. J., Frenk C. S., Theuns T., 2018, MNRAS, 479, 284
Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385
Simha V., Cole S., 2017, MNRAS, 472, 1392
Simon J. D., 2019, ARA&A, 57, 375
Simpson C. M., Bryan G. L., Johnston K. V., Smith B. D., Mac Low M.-M., Sharma S., Tumlinson J., 2013, MNRAS, 432, 1989
Simpson C. M., Grand R. J. J., Gómez F. A., Marinacci F., Pakmor R., Springel V., Campbell D. J. R., Frenk C. S., 2018, MNRAS, 478, 548
Smith M. C. et al., 2007, MNRAS, 379, 755
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Springel V. et al., 2008, MNRAS, 391, 1685

Springel V., 2005, MNRAS, 364, 1105
Springel V., Yoshida N., White S. D. M., 2001a, New Astron., 6, 79
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001b, MNRAS, 328, 726
Tollerud E. J., Bullock J. S., Strigari L. E., Willman B., 2008, ApJ, 688, 277
van den Bosch F. C., Ogiya G., 2018, MNRAS, 475, 4066
van den Bosch F. C., Yang X., Mo H. J., Norberg P., 2005, MNRAS, 356, 1233
van den Bosch F. C., Ogiya G., Hahn O., Burkert A., 2018, MNRAS, 474, 3043
Wang J. et al., 2011, MNRAS, 413, 1373
Wang J., Frenk C. S., Navarro J. F., Gao L., Sawala T., 2012, MNRAS, 424, 2715
Wang W., Han J., Cooper A. P., Cole S., Frenk C., Lowing B., 2015, MNRAS, 453, 377
Watkins L. L., van der Marel R. P., Sohn S. T., Evans N. W., 2019, ApJ, 873, 118
Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52
Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Gilbert K. M., Dalcanton J. J., Williams B. F., 2014, ApJ, 789, 148
Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, ApJ, 827, L23
Wheeler C. et al., 2019, MNRAS, 490, 4447
Wheeler C., Oñorbe J., Bullock J. S., Boylan-Kolchin M., Elbert O. D., Garrison-Kimmel S., Hopkins P. F., Kereš D., 2015, MNRAS, 453, 1305
Zaritsky D., Smith R., Frenk C., White S. D. M., 1993, ApJ, 405, 464
Zehavi I., Contreras S., Padilla N., Smith N. J., Baugh C. M., Norberg P., 2018, ApJ, 853, 84
Zentner A. R., Berlind A. A., Bullock J. S., Kravtsov A. V., Wechsler R. H., 2005, ApJ, 624, 505
Zhu Q., Marinacci F., Maji M., Li Y., Springel V., Hernquist L., 2016, MNRAS, 458, 1559

APPENDIX A: DISC DISRUPTION IN HYDRODYNAMICAL SIMULATIONS

In Section 3.4, we saw that the destruction of satellite galaxies by the central galaxy reduces the size of the population substantially, particularly in the vicinity of the central disc. In this Appendix, we compare the extent of this depletion in the FIRE-2, APOSTLE, and AURIGA simulations.

Fig. A1 shows the reduction in the number density of subhaloes as a function of radius in hydrodynamical versions of each simulation, compared to their dark matter-only counterparts. In all three cases, there is a general trend for the depletion to become more pronounced near the centre of the halo (see also Errani et al. 2017). This is to be expected as the tidal forces experienced by subhaloes are largest in the vicinity of the central galaxy. The extent of the depletion at any given radius, however, is markedly different in each of the three simulations.

It is clear that the degree of subhalo depletion is least severe in APOSTLE. Richings et al. (2020) ascribe this to the fact that galactic discs in the highest resolution APOSTLE simulations are 2–3 times less massive than the Milky Way disc. In AURIGA, on the other hand, the central galaxy mass is perhaps too large by a factor of 1.5–2. The presence of more massive central galaxies in AURIGA and FIRE-2 compared to APOSTLE causes subhaloes to experience stronger tidal forces during pericentric passages and thus to be destroyed more easily. This disruption extends well beyond the virial radius of the halo, and is particularly strong for subhaloes on radial orbits. Richings et al. (2020) provides a detailed explanation of the radial dependence of the depletion in the APOSTLE and AURIGA simulations.

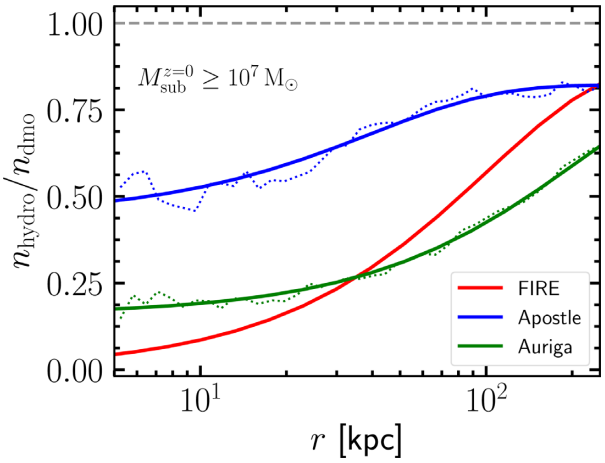

Figure A1. Ratio of the radial number density of resolved subhaloes with mass $M_{\rm sub}^{z=0} \geq 10^7\,{\rm M}_\odot$ in hydrodynamical and dark matter-only versions of the FIRE-2, APOSTLE, and AURIGA simulations. The thin dotted lines show the median ratios measured by Richings et al. (2020), while the corresponding solid curves are fits to these ratios using the functional form suggested by Samuel et al. (2020, equation 3). For each of these models, we list the best-fitting values of the free parameters α, d_0, and d_1 in Table A1.

Table A1. Best-fitting values for the parameters describing the disc destruction correction function, $f(d)$ equation 3), in each of the models listed in the first column. For the FIRE-2 model, the values in parentheses correspond to satellites more massive than $M_\star \geq 10^5\,{\rm M}_\odot$, as given by Samuel et al. (2020).

Model	α	d_0 (kpc)	d_1 (kpc)	Reference
FIRE-2	0.9 (0.8)	0 (8)	100 (78)	Samuel et al. (2020)
APOSTLE	0.82	0	40	Richings et al. (2020)
AURIGA	0.85	0	207	Richings et al. (2020)

Another interesting feature of Fig. A1 is that while the depletion near the halo centre is stronger in FIRE-2 than in AURIGA, the trend reverses in the outer halo ($r > 30$ kpc). The difference in the outer parts is perhaps due to the fact that the central galaxies in AURIGA are more massive than those in FIRE-2. However, at $r < 30$ kpc, the depletion is less in AURIGA despite its more massive central galaxies. This may be due, at least in part, to the specific way in which the subhalo orbits are interpolated between snapshots in the two simulations. Due to the rarity of satellites in this region an accurate interpolation is crucial for estimating the average depletion, particularly near the halo centre. We refer the reader to Richings et al. (2020) for a comprehensive treatment of this topic.

APPENDIX B: THE IMPORTANCE OF ORPHAN GALAXIES

A key aspect of our analysis is the consistent tracking of 'orphan' galaxies, a technique used to follow the dynamical evolution of GALFORM galaxies even after the dark matter subhalo in which they are hosted is disrupted below the nominal resolution limit of the COCO simulation ($\sim 3 \times 10^6\,{\rm M}_\odot$ in halo mass; see Section 2.2 for details). The results presented in the main body of the paper always include the orphan population; the purpose of this Appendix is to demonstrate how ignoring this population can lead to qualitatively different results.

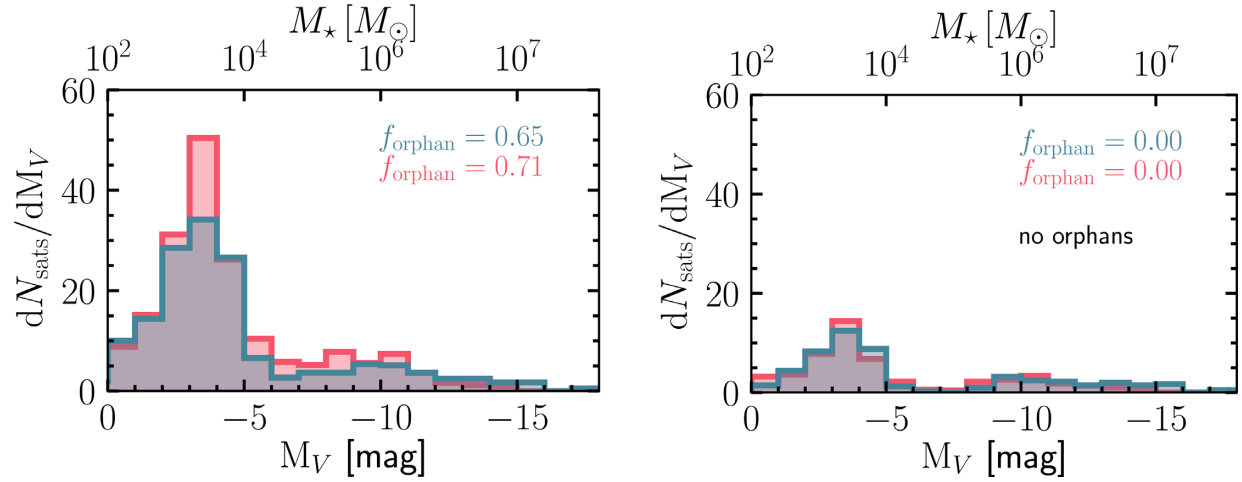

Figure B1. As Fig. 2, but now showing the comparison between the luminosity functions with (left) and without (right) orphans. The fraction of ultrafaint satellites identified today as orphans is given in the legend. The fraction is large, 60–70 per cent; these galaxies would be missed entirely in current hydrodynamical simulations but can be followed in our semi-analytic model coupled to a very high resolution *N*-body simulation. When the orphan population is neglected, the luminosity functions are nearly identical for early (red) and late-forming (blue) haloes. The host halo mass range in this and subsequent figures is $M_{200} = 1{-}1.3 \times 10^{12}\,{\rm M}_\odot$.

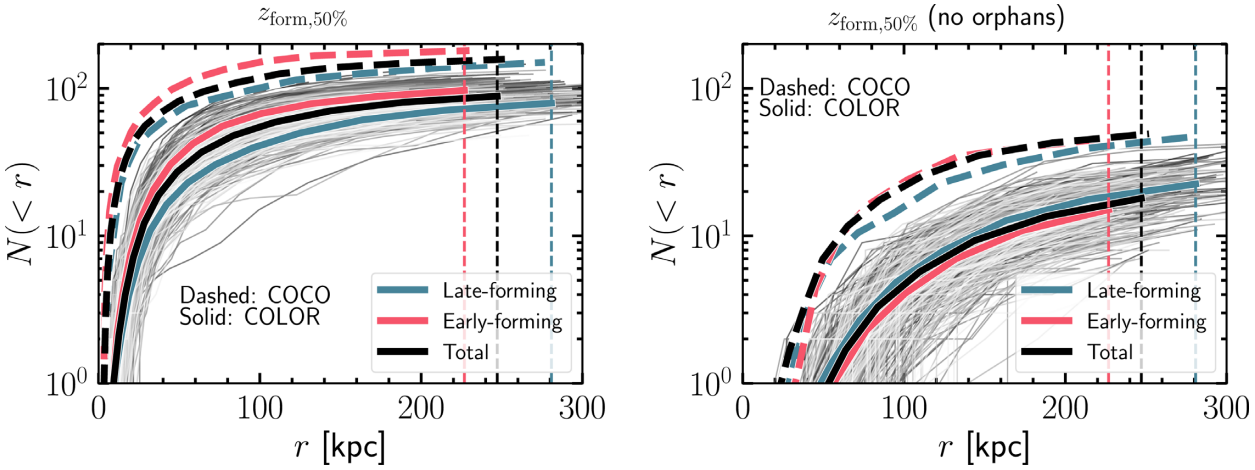

Figure B2. As Fig. 5, but now showing the comparison between radial profiles with (left) and without (right) orphans galaxies. Orphan galaxies make up a significant fraction of the total population at all radii and are by far the dominant population in the innermost 50 kpc.

Fig. B1 revisits the connection between the number of ultrafaint satellites and the assembly history of the host halo, this time comparing the results with (left) and without (middle) orphan galaxies. The fraction of the total satellite population that is identified as orphans is given in each panel. This number, which is made up predominantly of ultrafaint satellites, is sizeable: around 65 per cent in the later-forming haloes, rising to over 70 per cent in early-forming haloes. Qualitatively, this dependence on halo assembly is expected. Early-forming haloes are likely to accrete satellites earlier than their late-forming counterparts. These satellites orbit and lose mass through tidal stripping, in many cases bringing their subhaloes below the SUBFIND detection limit, resulting in the galaxies being tagged as orphans. On the other hand, satellites which fall into their hosts later experience fewer pericentric passages and thus a smaller fraction of them are likely to be tagged as orphans. Finally, the right-hand panel in Fig. B1 shows that neglecting the orphan population results in little to no difference in the satellite abundance in early-forming *versus* late-forming haloes.

Fig. B2 shows that there are also significant differences in the radial profiles of satellite galaxies when orphans are not included. While the total number of galaxies within a fixed radius, *r*, is, of course, significantly lower in all cases, there is also a change in the relative behaviour of early and late-forming haloes when ignoring orphan satellites. Indeed, the trends are inverted: at all *r*, late-forming haloes are now predicted to host more satellites than early-forming haloes, reproducing a common (but incorrect) conclusion often found in previous literature (see the discussion in Section 3.1), which is contrary to the result presented in Fig. 5 (reproduced in the left-hand panel of Fig. B2).

The radial profiles predicted by our semi-analytic model further highlight the importance of tracking orphan galaxies. These objects, which are largely ultrafaint galaxies, are distributed differently within the host halo relative to their more massive counterparts. The different spatial distributions are clearly illustrated in Fig. B3, which shows the satellite luminosity function from the high resolution COCO simulation, split into radial shells. When orphans are not

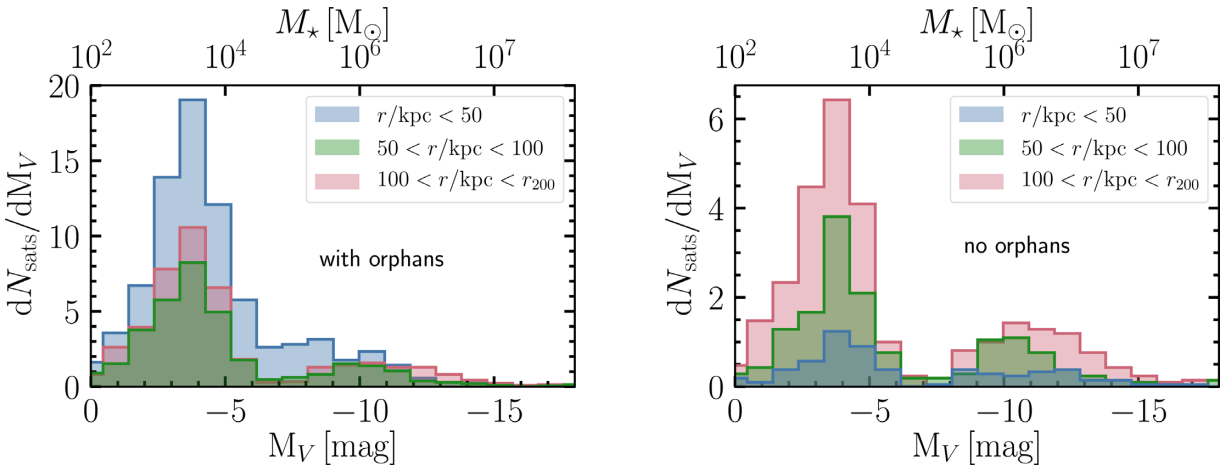

Figure B3. As Fig. 6, showing the luminosity function of satellites split into radial shells, with (left) and without (right) the inclusion of orphan galaxies (note the different y-axis limits in the two panels). The orphan satellites are distributed throughout the halo, with the majority of their contribution at the centre. Neglecting orphans result in a near absence of galaxies within the innermost 50 kpc.

included (right-hand panel), there is a near absence of satellites of any mass within the inner 50 kpc of the halo centre. Beyond this radius, we retrieve the familiar bimodal satellite distribution discussed in Section 2.2. In COCO, we find, on average, only around 6 satellites (across all masses) within the innermost 50 kpc when ignoring orphans; but when orphans are counted, there are on average ∼70 satellites within this same radius. As shown in Fig. B3, the inner population is primarily made up of satellites with $M_\star \leq 10^6\,\mathrm{M}_\odot$, although there is a small contribution from more massive satellites as well. Galaxies near the centres of haloes are more likely to be identified as orphans where tidal forces are more severe, and more likely to strip subhaloes below the resolution limit of the simulation. Furthermore, substructure finding algorithms often 'lose' subhaloes near halo centres where the relatively low density contrast complicates the identification of substructures.

This paper has been typeset from a TEX/LATEX file prepared by the author.