SURFS: Riding the waves with Synthetic UniveRses For Surveys

Pascal J. Elahi, Charlotte Welker, Chris Power, Claudia del P. Lagos, Aaron S. G. Robotham, Rodrigo Cañas and Rhys Poulton

1 International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia
2 ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)

ABSTRACT

We present the Synthetic UniveRses For Surveys (SURFS) simulations, a set of N-body/Hydro simulations of the concordance Λ Cold Dark Matter ($\Lambda$CDM) cosmology. These simulations use Planck cosmology, contain up to 10 billion particles, and sample scales and halo masses down to 1 kpc and $10^8 M_\odot$. We identify and track haloes from $z = 24$ to today using a state-of-the-art 6D halo finder and merger tree builder. We demonstrate that certain properties of haloes merger trees are numerically converged for haloes composed of $\gtrsim 100$ particles. Haloes smoothly grow in mass, $V_{\text{max}}$, with the mass history characterized by $\log M(a) \propto \exp[-(a/\beta)^\alpha]$, where $a$ is the scale factor, $\alpha(M) \approx 0.8$ & $\beta(M) \approx 0.024$, with these parameters decreasing with decreasing halo mass. Subhaloes follow power-law cumulative mass and velocity functions, i.e. $m(> f) \propto f^{-\alpha}$ with $\alpha_M = 0.83 \pm 0.01$ and $\alpha_{V_{\text{max}}} = 2.13 \pm 0.03$ for mass and velocity, respectively, independent of redshift, as seen in previous studies. The halo-to-halo scatter in amplitude is 0.9 dex. The number of subhaloes in a halo weakly correlates with a halo’s concentration $c$ and spin $\lambda$: haloes of high $c$ and low $\lambda$ have 60 per cent more subhaloes than similar mass haloes of low $c$ and high $\lambda$. High cadence tracking shows subhaloes are dynamic residents, with 25 per cent leaving their host halo momentarily, becoming a backsplash subhalo, and another 20 per cent changing hosts entirely, in agreement with previous studies. In general, subhaloes have elliptical orbits, $e \approx 0.6$, with periods of $2.3^{+1.1}_{-1.7}$ Gyr. Subhaloes lose most of their mass at pericentric passage with mass loss rates of $\sim 40$ per cent Gyr$^{-1}$. These catalogues will be made publicly available.

Key words: methods: numerical – dark energy – dark matter.

1 INTRODUCTION

Ongoing and upcoming galaxy surveys, such as ALFALFA (Haynes et al. 2011), GAMA (Driver et al. 2011; Liske et al. 2015), WAVES (Driver et al. 2016; de Jong et al. 2014), and WALLABY (Johnston et al. 2008; Staveley-Smith 2009) will probe galaxy formation and cosmic structure down to stellar galaxy masses of $10^9 M_\odot$ in haloes with maximum circular velocities of 100 km s$^{-1}$, while other large volume surveys, like the Taipan survey (da Cunha et al., in preparation), will focus on sampling millions of galaxies to examine our cosmology. These large observational projects will sample the galaxy population with great statistics, enough to severely test our galaxy formation models. To match these surveys, typically simulation volumes should be of the similar size as the survey volume. However, this is countered by the need to have high enough resolution not only to reliably identify dark matter haloes in cosmological simulations but to robustly follow their evolution. Some cosmological simulations have sacrificed resolution for survey volume (e.g. Angulo et al. 2012; Riebe et al. 2013; Fosalba et al. 2015; Kim et al. 2015) with the goal of populating pure $N$-body runs with galaxies using methods like Halo Occupation Distribution (HOD), combined with SubHalo Abundance Matching (SHAM) (e.g. Zheng et al. 2005; Conroy, Wechsler & Kravtsov 2006; Hearin et al. 2013; Carretero et al. 2015; Skibba et al. 2015; Saito et al. 2016). These methods rely on simple mappings between halo masses and the galaxies that reside in them and lack the predictive power of more physical galaxy formation models such as Semi-Analytic Models (SAMs; e.g. Cole et al. 2000; Baugh 2006; De Lucia & Blaizot 2007; Monaco, Fontanot & Taffoni 2007; Lee & Yi 2013; Henriques et al. 2013; Croton et al. 2016; Lacey et al. 2016) and full hydrodynamical cosmological simulations (e.g. Dubois et al. 2014; Vogelsberger et al. 2014; Schaye et al. 2015). These techniques, however, require high resolution and, in the case of SAMs, accurate reconstruction of the evolution of haloes using high-fidelity halo catalogues coupled with high-cadence merger trees. Higher resolution $N$-body simulations with sufficient resolution (in both halo mass and the temporal resolution of the halo catalogue) (e.g. Springel et al. 2005; Boylan-Kolchin et al. 2009; Klypin, Trujillo-Gomez & Primack 2011) are the only means through which realistic mock surveys capable of
Table 1. Simulation parameters.

<table>
<thead>
<tr>
<th>Name</th>
<th>Box size $L_{\text{box}}$ (h⁻¹ Mpc)</th>
<th>Number of Particles $N_p$</th>
<th>Particle Mass $m_p$ (h⁻¹ M⊙)</th>
<th>Softening Length $\epsilon$ (h⁻¹ ckpc)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>L40N512</td>
<td>40</td>
<td>512²</td>
<td>$4.13 \times 10^7$</td>
<td>2.6</td>
<td>Small-volume, high-resolution test</td>
</tr>
<tr>
<td>L210N512</td>
<td>210</td>
<td>512³</td>
<td>$5.97 \times 10^8$</td>
<td>13.7</td>
<td>Moderate-volume, low-resolution test</td>
</tr>
<tr>
<td>L210N1024</td>
<td>210</td>
<td>1024³</td>
<td>$7.47 \times 10^8$</td>
<td>6.8</td>
<td>Moderate volume, moderate resolution</td>
</tr>
<tr>
<td>L210N1024NR</td>
<td>210</td>
<td>$2 \times 1024^3$</td>
<td>$6.29 \times 10^8$</td>
<td>6.8</td>
<td>Nonradiative (adiabatic gas, no star formation or feedback) analogue to L210N1024.</td>
</tr>
<tr>
<td>L210N1536</td>
<td>210</td>
<td>1536³</td>
<td>$2.21 \times 10^8$</td>
<td>4.5</td>
<td>Moderate volume, current high resolution.</td>
</tr>
<tr>
<td>L900N2048</td>
<td>900</td>
<td>2048³</td>
<td>$7.35 \times 10^9$</td>
<td>14.6</td>
<td>Large volume, low resolution, low cadence for HODs</td>
</tr>
</tbody>
</table>

We present the next stage in these simulations, SURFS (Synthetic UniveRes For Surveys), which consists of a suite of primarily pure N-body simulations spanning a range of cosmological volumes to address both galaxy formation and cosmological surveys. Our simulation volume and resolution choices for our moderate size cosmological runs are primarily motivated by the upcoming WAVES-WIDE survey (Driver et al. 2016), which aims to probe the stellar mass function down to a completeness limit of $M_\star \sim 10^9$ M⊙ in a volume of $\sim 850$ cMpc ($z < 0.2$). The SURFS simulations will produce synthetic analogues of this survey, resolving dark matter haloes down to $10^9$ M⊙, with future simulations resolving haloes down to $10^6$ M⊙, allowing the simulation to probe stellar masses down to $\sim 10^5$ M⊙.¹ These simulations are used to produce high-quality halo catalogues and halo merger trees ideal for SAMs, following in the footsteps of the Millennium simulations (Springel et al. 2005; Boylan-Kolchin et al. 2009) and the MultiDark/Bolshoi series (e.g. Klypin et al. 2011; Riebe et al. 2013). Our large cosmological volume runs sacrifice resolution for larger Gpc volumes and are ideal for populating HODs/SHAM models calibrated using results from SAMs, and/or available observations (see for instance Howlett et al. 2017).

Here in the first of a series of papers, we focus on the properties of dark matter haloes and their evolution over cosmic time. Our goal is to use the precision tracking of cosmic structure evolution to study galaxy formation physics and satellite evolution with SAMs. Additionally, we will provide the community with free access to our halo catalogues and merger trees, useful for producing mock surveys by using these as input to galaxy formation models, a topic we will cover in upcoming papers along with our own mocks. This overview paper covers several topics, highlighting particular results at a variety of cosmological scales.

We begin in Section 2 with an introduction to the simulations, analysis pipeline, data products available, and discuss near-term upcoming simulations that will be released. We then focus on large scales, specifically the cosmic web, the filaments of material connecting knots and surrounding voids in which haloes reside. In Section 3, we show what mass haloes are required to reconstruct the cosmic web and trace the web as defined by gas. We then present the $z = 0$ halo and subhalo population in Section 4, focusing on numerical convergence and properties of the subhalo population. We demonstrate our halo catalogues, show excellent convergence for haloes composed of $\gtrsim 100$ particles, and analyse the subhalo population. We end with the key analysis that can only be done with high fidelity halo merger trees from high-resolution, moderate-volume cosmological simulations: an accurate reconstruction of cosmic growth and the dynamic lives of subhaloes. We show that haloes grow smoothly in dark matter mass and $R_{\text{vir}}$ until they begin to virialize, at which point they continue to grow in mass but become more concentrated and spherical. Subhaloes, despite being dynamic residents of haloes, show smooth internal evolution, gradually losing mass, mostly at pericentric passage. We end in Section 6 and Section 7 with a summary and discussion of SURFS and upcoming results.

2 METHODS

2.1 Simulations

The SURFS suite consists of N-body simulations, most with volumes of 210 h⁻¹ cMpc on a side, and span a range in a particle number, currently up to 8.5 billion particles using a ΛCDM Planck cosmology (Planck Collaboration XIII 2016). The simulation parameters are listed in Table 1. Our simulations are split into moderate-volume, high-resolution simulations focused on galaxy formation for upcoming surveys like WAVES and WALLABY, and larger-volume simulations designed for surveys focused on cosmological parameters like the Taipan survey. Our moderate-volume simulation parameters allow us to resolve the host haloes of galaxies with stellar masses of $10^8$ M⊙ at $z = 0$ with the nominal requirement that the host dark matter haloes of such galaxies be resolved with 100 particles, necessary if we are to build merger trees for coupling to Semi-Analytical Models (SAM) of galaxy formation. Our larger volume simulations parameters allow us to produce several mock surveys using halo catalogues combined with Halo Occupation Distribution (HOD) models. All simulations were run with a memory lean version of the GADGET2 code on the Magnus supercomputer at the Pawsey Supercomputing Centre.

These simulations provide an excellent test-bed for numerical convergence, studies into the growth of haloes and the evolution of subhaloes down to dark matter halo masses of $\sim 10^{10}$ M⊙ (and galaxy stellar masses down to $\sim 10^9$ M⊙). Our paper here will primarily focus on the 210 h⁻¹ Mpc volume (L210) simulations, though we note that our large cosmological volume simulation has already been used to produce mocks to study cosmological parameters (Howlett et al. 2017). We have also run a non-radiative hydrodynamical counterpart to our L210N1024 simulation to examine the effects of gas physics and, more importantly, the rate of cosmic gas accretion, which is an essential piece of information for any SAM (see review by Benson 2010), although from this point we will focus on our DM only simulations.

We produce 200 snapshots and associated halo catalogues in evenly spaced logarithmic intervals in the growth factor starting at

¹ This stellar mass limit is simply based on the stellar mass-to-halo mass relation roughly extrapolated to low mass galaxies (e.g. Behroozi, Conroy & Wechsler 2010; Moster et al. 2010; van Uitert et al. 2016).
$z = 24$ for our L210 and smaller volume simulations. This high cadence, higher than that was used in the Millennium simulations (Springel et al. 2005; Boylan-Kolchin et al. 2009), is necessary for halo merger trees that accurately capture the evolution of dark matter haloes as each snapshot is separated by less than the freefall time of overdensities of 200, i.e. haloes.

The set of current simulations will expand to include multiple 8.5 billion particle Gpc scale simulations, along with 64 billion particle 210 $h^{-1}$ Mpc simulations. The moderate-volume, high-resolution simulation will probe haloes down to mass of $3.4 \times 10^9 \mathrm{M}_\odot$, robustly follows the cosmic evolution of $\gtrsim 1.7 \times 10^9 \mathrm{M}_\odot$ haloes. We will produce numerically converged synthetic galaxies down to stellar masses of $\sim 10^8 \mathrm{M}_\odot$, near the completeness limit of the WAVES survey. The end goal is a 500 billion particle 210 $h^{-1}$ Mpc simulation capable of resolving the lives of dwarf galaxies of $\sim 10^6 \mathrm{M}_\odot$.

### 2.2 Halo catalogues

We identify haloes and calculate their properties using VELOCIRAPTOR (a.k.a. STF Elahi, Thacker & Widrow 2011, Elahi et al., in preparation). This code first identifies haloes using a 3DFOF algorithm (3D Friends-of-Friends in configuration space, see Davis et al. 1985) and then identifies substructures using a phase-space FOF algorithm on particles that appear to be dynamically distinct from the mean halo background, i.e. particles which have a local velocity distribution that differs significantly from the mean, i.e. smooth background halo. Since this approach is capable of finding not only subhaloes, but also tidal streams surrounding subhaloes as well as tidal streams from completely disrupted subhaloes (Elahi et al. 2013), for this analysis, we also ensure that a group is roughly self-bound, allowing particles to have potential energy to kinetic energy ratios of at least 0.95.

Like other phase-space finders, such as ROCKSTAR (Behroozi, Wechsler & Wu 2013a), this code is better able to disentangle major mergers than configuration-space based finders (Behroozi et al. 2015) (like SUBFIND, Springel et al. 2001; or AHF, Knollmann & Knebe 2009; see Muldrew, Pearce & Power 2011, for examples of the short-comings of configuration-space halo finders). Specifically, once all deviations from the large-scale, smooth velocity distribution have been identified, a.k.a. subhaloes, the code then searches the remaining background for the cores of merger remnants, i.e. phase-space dense groups, using the velocity dispersion of the halo to scale the velocity-linking lengths. If multiple cores are found, that is the smooth background is characterized by multiple large-scale phase-space distributions, then phase-space dispersion tensors are calculated. Particles in the halo background are then assigned to the closest core in phase space as calculated using the core’s phase-space tensor from the core’s centre-of-mass in phase space. This method is similar to assigning particles based on a Gaussian mixture model, but less time-consuming. For a more thorough discussion of (sub)halo finding, we refer readers to Onions et al. (2012); Knebe et al. (2013) and an upcoming paper on revisions to VELOCIRAPTOR (Elahi et al., in preparation; C. añaes et al., in preparation).

The next step is the construction of a halo merger tree. We use the halo merger tree code that is part of the VELOCIRAPTOR package (see Srisawat et al. 2013, Elahi et al., in preparation, for more details) called TREEFROG. At the simplest level, this code is a particle correlator and relies on particle IDs being continuous across time (or halo catalogues). The cross-matching between catalogue $A \& B$ is done by identifying for each object in catalogue $A$, the object in catalogue $B$ that maximizes the merit function:

$$N_{A,B} = N_{A,B}/(N_A N_B).$$

where $N_{A,B}$ is the number of particles shared between objects $i$ and $j$ and $N_A$ and $N_B$ are the total number of particles in the corresponding object in catalogues $A$ and $B$, respectively. This merit function maximizes the fraction of shared particles in both objects and is generally robust identifying candidate matches. However, there are instances where several possible candidates are identified. This can happen when several similar mass haloes merge at once, as loosely bound particles can be readily exchanged between haloes.

To alleviate these issues, we follow Poole et al. (2017) and use the rank of particles as ordered by their binding energy using

$$S_{A,B} = \sum_i N_{A,B} / R_{A,B},$$

Here the sum is over all shared particles and $R_{A,B}$ is the rank of particle $i$ in halo $A$, with the most bound particle in the halo having $R = 1$. The maximum value of this sum when all particles are shared is $S_{A,B} = \gamma + \ln N_A$, with $\gamma = 0.577 215 6649$ being the Euler–Mascheroni constant.

We combine equation (1) with the normalized version of equation (2), i.e. $\tilde{S}_{A,B} = S_{A,B} / S_{A,B}^\text{max}$, to obtain

$$M_{A,B} = N_{A,B} \tilde{S}_{A,B} = N_{A,B} / S_{A,B}^\text{max},$$

where we calculate the rank ordering in both haloes in question as the rank ordering can be quite different. This would be the case for a subhalo that is completely tidally disrupted in the outskirts of a larger halo. This combined merit maximizes the total shared number of particles while also weighting the match by the number of equally well-bound shared particles.

We produce a tree following haloes forward in time, identifying the optimal links between progenitors and descendants. We rank progenitor(descendant link) as primary and secondary. A primary link is one where the maximum merit for a halo amongst all its candidate descendants points to a descendant which has a maximum merit amongst all its candidate progenitors that points back to the same halo, i.e.: the maximum merit both forward and backward. All other connections are classified as secondary links.

In an ideal case, a halo would only have one descendant and that descendant would only have one progenitor. However, identifying primary links is complicated tidal disruption and by the halo-finding processes, which can lose or join haloes. In the case where a halo has been disrupted and has merged with another halo, the primary progenitor is defined as the link with the best merit in both directions and the tidally disrupted halo is flagged as a secondary progenitor with no primary descendant. If haloes have been artificially merged.

---

2 Freely available https://github.com/pelahi/VELOCIraptor-STF.git

3 We also apply a 6DFOF to each candidate FOF halo using the velocity dispersion of the candidate object to clean the halo catalogue of objects spuriously linked by artificial particle bridges, useful for disentangling early-stage mergers.

4 The mass reconstruction using phase-space tensors for these major merger remnants can be noisy once an object becomes significantly disrupted and its particle distribution is not well characterized by a multi-variate Gaussian with a single global dispersion tensor. This flaw would also be present in full Gaussian mixture models, possibly to a greater extent, and would require generalized distribution functions to assign particles and an evaluation of the number of connections to other particles in the group.
at a given snapshot, it is possible that the merged halo will have several possible descendants at a later snapshot. In this case, the highest merit defines the primary descendant and all other the descendants are flagged has having no primary progenitor.

Objects with no primary progenitor will generate missing links in the tree. This problem occurs for low mass haloes that lie near the particle number threshold used by the halo finder. With fine-scale temporal resolution, these haloes appear to pop in and out of existence, leaving temporally orphan haloes in the tree. Critically, orphan subhaloes can occur at much higher masses as these can be lost by the halo finder as they pass through the dense regions of their host halo.

This problem can be alleviated somewhat by searching multiple snapshots for candidate links (see Behroozi et al. 2013b; Srisawat et al. 2013; Avila et al. 2014; Wang et al. 2016; Poole et al. 2017, for discussions of the pitfalls of tree-building, see for instance). Here we search for primary links. If a halo does not have a primary descendant in the first snapshot, subsequent snapshots are searched till a primary link is identified or the maximum number of snapshots is reached. We typically search up to four snapshots, equivalent to \(\approx 1\) Gyr at late times and \(\approx 30\) Myr at early times or approximately the free-fall time at the virial overdensity, 200\(\sigma\). For a more detailed discussion of tree building, see our upcoming paper, Elahi et al. (in preparation).

These catalogues are freely available on request and will be made available via webserver in the near future and are ideal for input to SAMs and for following the orbits of subhaloes. In an upcoming paper (Lagos et al. in preparation), we will present our mocks produced using an SAM with these catalogues as input.

3 SIMULATION VOLUME

An example of the matter distribution is presented in Fig. 1, where we plot the projected density field of a slice through the simulation volume from our highest resolution simulation. We also plot the upper inset the power spectrum \(P(k) = \langle \delta_k \delta_k^* \rangle\), of our simulations at the initial conditions. A more rigorous examination of the matter power spectrum and biases produced by different tracers will be presented in a later work. The take-home message here is that all our L210 volumes perfectly overlap in the initial conditions, at small \(k\) values where the power spectrum is well sampled below the Nyquist frequency. The upturn in the power at early times and large \(k\) is from the shot noise.

We also plot the haloes identified by VELOCIRAPTOR in the lower insets, where we have applied different cuts on ‘virial’ mass, here defined as \(M_\Delta = 4\pi R_\Delta^3 \Delta \rho_{crit}/3\), with \(\Delta = 200\), \(\rho_{crit}\) is the critical density of the universe, and \(R_\Delta\) is the radius that encloses this mass. Visually, we see the largest haloes appear in the densest regions of the matter field, with smaller haloes residing in a larger variety of environments. Clearly visible in this figure and the lower insets is the cosmic web that is a material network of nodes connected through filaments, at the intersection of walls, themselves segmenting large underdense regions, or voids.

3.1 Cosmic web

The cosmic web naturally arises from the anisotropic gravitational collapse of an initially Gaussian random field of density perturbations (Zel’Dovich 1970; Peebles 1980; Shandarin & Zeldovich 1989; Bond, Kofman & Pogosyan 1996). Haloes form and reside within the overdensities of the cosmic web, accreting smooth material and smaller haloes via filaments (see Bond et al. 1996; for details). Knots, at the intersection of several of the most contrasted filaments, house clusters; the largest virialized objects in the universe. On such scales, the filamentary pattern of the cosmic web is apparent in all large-scale galaxy surveys (e.g. de Lapparent, Geller & Huchra 1986; Doroshkevich et al. 2004; Colless et al. 2003; Alpaslan et al. 2014; Eardley et al. 2015), traced by the galaxy distribution. The classification and study of this anisotropic, multi-scale cosmic density distribution are unsurprisingly a complex task. Numerous methods exist for analysing simulation data and extracting the cosmic web (for an overview of various algorithmic approaches, see Cautun, van de Weygaert & Jones 2013.). Here we do not attempt to compare different schemes or analyse the inferred evolution of the cosmic web, leaving this for future papers. Instead we seek to answer a simple question: what haloes (galaxies) must be sampled to in order to efficiently trace the underlying gaseous structure of the cosmic web, that is the zoology of streams – cold, warm, laminar, turbulent, etc. – funnelled and shaped by the cosmic web on scales that are the most relevant to galaxy formation?

Haloes, and the galaxies that reside in them, are biased spatial tracers of the cosmic velocity and density field. Studies have shown that identifying cosmic structures such as voids depends sensitively on the method and choice of tracer. Using the full, unobservable, density field will give different void regions than using galaxies with different luminosity cuts or haloes of different masses (Paillas et al. 2017). The most luminous galaxies or cluster mass haloes provide information on nexus points of the cosmic web, indicating where the largest filaments terminate. Only by probing smaller halo masses or galaxies can we begin to recover more of the fine-grain features as illustrated in the inset of Fig. 1. Here we show the distribution of haloes for different halo mass cuts. Visually, it does not appear that major density features in this slice, like filaments, change by probing mass scales below \(10^{12} M_\odot\), which would correspond to haloes hosting \(L^*\) galaxies. However, one should bear in mind that in such projections most walls appear as filaments and most filaments as dots, rendering any visual analysis unreliable.

To properly identify and quantify how well we recover the cosmic web, we use DISPERSE, a topological based filament finder (Sousbie 2011). This algorithm identifies the ridges from a smoothed density field connecting topologically robust saddle points to peaks. DISPERSE measures the robustness of a filament and trims the candidate catalogue in two ways: filament persistence, the ratio of the value at the two critical points in a topologically significant pair of critical points (maximum-saddle, saddle-saddle, or saddle-minimum); and local robustness, the density contrast between the critical points and skeleton segments with respect to background. Removing low-persistence pairs is a multi-scale, non-local method to filter noise/lownoise significance filaments. When applied to point-like distributions of haloes or galaxies, a persistence threshold translates easily into a minimal signal-to-noise ratio, expressed as a number of standard deviations \(\sigma\). This algorithm has not only been used to analyse simulations (e.g. Dubois et al. 2014) but has been successfully applied to real spectroscopic (VIPERS) and photometric (COSMOS) surveys (e.g. Malavasi et al. 2016; Laigle et al. 2018; Malavasi et al. 2017).

We apply this method on the distribution of haloes identified in different cubic sub-volumes of the simulation, with varying persistence threshold and the minimal mass of haloes considered. Let us first consider all the haloes with masses of \(10^{10}–10^{15} M_\odot\) identified in a 75Mpc wide sub-volume of L210N1024NR. Results are presented in a 40Mpc thick projected map in Fig. 2. In the top panel, the persistence threshold is \(1\sigma\) while in the bottom panel
Figure 1. Simulation. We show the density field of our highest resolution simulation, L210N1536. Lower inset shows the cosmic web as outlined by the halo distribution with different mass cuts in a slice \(1/4^{th}\) the box size thick in from the L210N1536 simulation. Upper inset shows the power spectrum, \(P(k)\) with 1\(\sigma\) sampling errors along with the ratio between all simulations and the linear power spectrum, with vertical lines at the Nyquist frequency of each simulation. It is set to 5\(\sigma\). In both panels, haloes are overplotted as circles of varying colour and size depending on \(\log M\). The cosmic web identified with low persistence in the left-hand panel includes numerous short spurious filaments of widely varying directions within one thicker filament and doubled filaments. Although in a sufficiently resolved environment, such a level of precision might be useful to resolve lower density filaments (sometimes dubbed ‘tendrils’) in the vicinity of void galaxies, it fails to provide a simple and smooth characterization of the large-scale filaments on Mpc scale. In the bottom panel, only the most persistent filaments appear. Unlike the top panel, filaments remain coherent over a few segments in between nodes and smoothly flow to nodes, where massive haloes reside.

To determine what halo masses need to be identified in order to capture the Mpc-scale cosmic web, we examine in Fig. 3 the probability density function (PDF) of \(d_\text{fil}\), the distance of haloes to their nearest filament at \(z = 0\). Here we produce a skeleton, shown in the inset of Fig. 3, using haloes with \(M_\Delta > 5 \times 10^{11} \ M_\odot\), keeping filaments that are significant to \(\geq 1\sigma\). This choice is motivated by the need to recover continuous smooth Mpc scale filaments similar to the filaments extracted from the gas density field in simulations such as Horizon-AGN (Dubois et al. 2014), hence with consistent \(d_\text{fil}\) PDFs (see Welker 2015, Chapter 3 for details), and on scales that can be robustly measured in observational studies (Laigle et al. 2018; Malavasi et al. 2017). Such studies can identify filaments with Mpc precision, delineating voids with a typical radius of \(\approx 5 – 30\) Mpc, which is directly comparable to our reconstruction. We plot the PDF for three different halo mass bins: \(M_\Delta < 10^{11} \ M_\odot\), \(10^{11} \ M_\odot < M_\Delta < 10^{12} \ M_\odot\), and \(M_\Delta > 10^{12} \ M_\odot\). We also plot the PDF of \(d_\text{fil}\) for haloes with \(5 \times 10^{11} \ M_\odot < M_\Delta < 10^{12} \ M_\odot\) (dashed magenta...
The cosmic web in SURFS: 40 Mpc thick projected map of a subvolume of the SURFS box, with 75 Mpc on a side. The filaments extracted with DISPERSE from the full distribution of haloes are highlighted in red while haloes in the mass range $10^{10}$–$10^{15} \, M_\odot$ appear as circles of varying size and colours. Top: The persistence threshold for the cosmic web extraction is kept low, with a minimum signal-to-noise ratio of $1\sigma$. A complicated pattern of filaments of various scales and densities is traced by the distribution of haloes. Bottom: The signal-to-noise ratio is raised to $5\sigma$. Now only the most robust filaments appear, highlighting the smooth filamentary pattern of the cosmic web on Mpc scales.

Figure 3. Halo distances to the cosmic web. PDF of $d_{fil}$, the halo distance to the nearest filament, for haloes in three different mass bins with $M_\Delta < 10^{11} \, M_\odot$ (solid green curve), $10^{11} \, M_\odot < M_\Delta < 10^{12} \, M_\odot$ (solid blue curve) and $M_\Delta > 10^{12} \, M_\odot$ (solid red curve). Vertical dashed lines indicate the peak of the distribution. We also plot the PDF of $d_{fil}$ for the sub-sample of haloes with $5 \times 10^{11} \, M_\odot < M_\Delta < 10^{12} \, M_\odot$ (dashed magenta line) to examine the effect of this mass threshold on the PDF. Inset shows the skeleton extracted using haloes with $M > 5 \times 10^{11} \, M_\odot$ and low persistence.

$\lesssim 10^{11} \, M_\odot$ compared to those with masses of $\gtrsim 10^{12} \, M_\odot$. Even the bin containing haloes at the mass threshold used by DISPERSE is not strongly peaked and is similar to the lower mass bins, except from those small deviations and extra noise due to fewer haloes per bin. This result emphasizes the fact that on the scales of interest for this study (Mpc scale filaments along which galaxies drift), the cosmic web is mostly determined by those high mass haloes with $M_\Delta \gtrsim 10^{12} \, M_\odot$. Thus surveys only need to be complete down to halo masses of at least $10^{12} \, M_\odot$, as these haloes are likely to be the nodes of the web, and ideally $5 \times 10^{11} \, M_\odot$ to reliably reconstruct the large-scale cosmic web. Including small galaxies residing in low mass haloes of $< 3 \times 10^{11} \, M_\odot$ necessitates the use of a higher persistence (signal-to-noise) to avoid identifying tendrils.

This demonstrates the ability of surveys like WAVES, which is relatively complete down to halo masses of $\sim 10^{12} \, M_\odot$, to robustly measure the cosmic web.

4 HALO POPULATION

We present results of our $z = 0$ (sub)halo catalogues here. We start with convergence tests and then discuss the subhalo population of our simulations.

4.1 Halo properties and convergence

We start with the simplest comparison, the halo mass and velocity functions, presented in Fig. 4 along with the ratio of the distribution from one simulation to our highest resolution reference simulation in the bottom panels. The distribution of ‘virial’ mass, here defined as $M_\Delta = 4\pi R_\Delta^3 \Delta \rho_{crit}/3$, with $h \Delta = 200$, $\rho_{crit}$ is the critical density of the universe, shows that simulations with the same mass
Figure 4. Halo distributions. We show the $z = 0$ halo mass and maximum velocity functions (left and right). Each plot has the distribution in the upper panel and the residuals relative to our reference simulation in the lower panel. We highlight bins containing fewer than 10 haloes, indicating the number they contain. We also plot the mass/velocity scales of objects composed of 100 particles by vertical lines with the same colour as the corresponding simulation. We also plot four mass functions, Sheth, Mo & Tormen (2001) (solid grey), Angulo et al. (2012) (dotted grey), the modified Angulo et al. (2012) fit keeping only bound particles in the FoF envelope, calculated using IMPCALC (Murray, Power & Robotham 2013). We also plot a fit to the highest resolution simulation (solid magenta) described in the text.

resolution give the same mass function to within $\lesssim 5$ per cent for haloes composed of at least 100 particles. Even with high-resolution, lower volume simulations the variance is well within 10 per cent for mass bins with Poisson fluctuations of $\lesssim 20$ per cent.

The sole systematic differences between simulations are a result of finite volume effects and cosmic variance. For example, larger volume simulations generally have a greater number of large haloes at mass scales above the exponential turnover, as seen by comparing the L900 simulation to the L210 simulations. Cosmic variance is easily seen in the systematic offset between L40N512 and the larger simulations. The large-scale modes present in L40N512 with wavelengths of $40\ h^{-1}$ Mpc, which are below the scale of homogeneity ($150\ h^{-1}$ Mpc based on WiggleZ; Scrimgeour et al. 2012), produce an overall overdensity, enhancing halo formation.

We compare our mass functions to several fitting formulae, which all agree except at the very high mass $\gtrsim 10^{14} M_{\odot}$, where our 210 $h^{-1}$ Mpc boxes are affected by missing power and contain fewer large haloes. We fit the binned differential mass function of our highest resolution simulation at $z = 0$ using EMCEE (Foreman-Mackey et al. 2013). We sample the mass function at the 50 largest haloes and then 50 mass evenly spaced in log $M$, and find parameters broadly in agreement to those found by Watson et al. (2013) (see appendix A for more details).

The $V_{\text{max}}$ distribution (see the right-hand panel of Fig. 4), that is the distribution of the maximum circular velocity defined as $V^2 = GM(r < R)/R$, shows similar convergence. However, the velocity scale where different resolutions diverge by $\gtrsim 10$ per cent occurs for haloes resolved with more than 100 particles, unlike the mass distribution. The divergence occurs for larger haloes because internal properties like $V_{\text{max}}$ require more particles before being resolved. Based on this figure, convergence of $\gtrsim 95$ per cent occurs for haloes composed of $\gtrsim 500$ particles.

The slower $V_{\text{max}}$ convergence indicates that we should be cautious of the internal properties of haloes composed of $\lesssim 500$ particles and is clearly demonstrated in Fig. 5, where we show the radius of this region as a function of $V_{\text{max}}$ for haloes only. Subhaloes are affected by strong tidal fields and hence removed from the analysis here. The simulations all show the same strong correlation between $R_{\text{max}}$ & $V_{\text{max}}$ for well-resolved haloes, with the median and scatter numerically converged for well-resolved haloes, with the distribution in $R_{\text{max}}$ at a given $V_{\text{max}}$ following a Gaussian distribution as shown in the inset. Below velocity scales corresponding to $\sim 500$ particles, haloes deviate away from this correlation. Thus, the internal properties of haloes composed of $\lesssim 500$ particles should be treated with some caution and those composed of $\lesssim 100$ particle generally ignored unless the only property one is interested in is mass. Based on this, we will typically limit our comparison between simulations to haloes composed of $\geq 100$ particles, where we expect differences of at most $\sim 10$ per cent. This limit corresponds to velocity scales of $50\ km\ s^{-1}$ in our current highest resolution L210 simulation.

The mass profiles of dark matter haloes are reasonably well characterized by NFW or Einasto profiles (e.g. Navarro, Frenk & White 1997; Navarro et al. 2004). We do not carefully fit a radial density profile using a maximum likelihood method to each halo to determine its concentration but instead follow Prada et al. (2012) and assume a NFW profile to calculate the concentration...
Figure 5. $V_{\text{max}} - R_{\text{max}}$ relation. We show the median maximum circular velocity radius along with the 16 per cent and 84 per cent quantiles in $V_{\text{max}}$ bins at $z = 0$. Vertical dashed lines indicate the average $V_{\text{max}}$ of haloes composed of 100 particles and shaded region indicates the 2σ region of $V_{\text{max}}$ values for haloes composed of 500 particles. Here we have excluded subhaloes. Inset shows the normalized distribution for haloes composed of 100–2000 particles for each simulation along with the median $V_{\text{max}}$ within this range. Line colours are the same as in Fig. 4.

The concentration–mass relation of haloes is shown in Fig. 6. Here we have removed subhaloes but have not removed so-called un-relaxed haloes, which are typically not well described by a NFW profile. We see excellent agreement between each simulation above the resolution limit of 500 particles. The simulations reproduce the same relation, in both the mass dependence and the distribution in a given mass bin. The distribution in a given mass bin is reasonably well characterized by a lognormal distribution with a peak that increases with decreasing mass as demonstrated by the inset, where we have plotted the normalized distribution of haloes composed of 1000–2000 particles (0.5 dex in mass), which corresponds to different mass scales in each simulation.

We calculate the shape, using the reduced inertia tensor (Dubinski & Carlberg 1991; Allgood et al. 2006),

$$I_{j,k} = \sum_n m_n x_{j,n}^i y_{k,n}^i \left( r_n^j r_n^k \right)^2 .$$

Here the sum is over particles in the halo, $(r_n^j)^2 = (x_n^j)^2 + (y_n^j/q)^2 + (z_n^j/s)^2$ is the ellipsoidal distance between the halo’s centre-of-mass and the $n$th particle, primed coordinates are in the eigenvector frame of the reduced inertia tensor, and $q$ and $s$ are the semi-major and minor axis ratios, respectively. The shape of haloes, presented in Fig. 7, shows more massive haloes tend to be more triaxial with

5 These unrelaxed haloes are not a major issue due to the 6DFOF algorithm which removes haloes linked by particle bridges and our scheme to separate major mergers.
numerical convergence occurring at the same halo resolution as that seen in Fig. 6. We only show the distribution of \( q \) as the mass trend and numerical convergence for \( s \) is similar, with the difference being \( s \approx 0.7 q \). Here the addition of non-radiative gas makes large haloes more spherical by \( \sim 10\% \). Overall, our simulations show excellent agreement in halo properties with strong numerical convergence for objects composed of \( \gtrsim 100 \) particles. This is true at higher redshifts as well.

### 4.2 Subhaloes

The simulations also have numerous group mass and low cluster mass haloes that are well resolved enough to study the subhalo population. We plot the subhalo mass and circular velocity distributions in Fig. 8 for well-resolved haloes containing at least 50 subhaloes and composed of 50,000 particles in order to have each halo sample the subhalo mass function over a wide range of masses. Our simulations probe host masses from groups of \( \sim 10^{13} M_{\odot} \) up to small clusters of \( \sim 10^{14} M_{\odot} \) (or from velocity scales of \( \sim 300 \text{ km s}^{-1} \) to \( \sim 800 \text{ km s}^{-1} \)) with our highest resolution simulation having over a thousand such haloes, with the median mass of a rich group/small cluster. To stack the distributions, we normalize the mass and \( V_{\text{max}} \) of subhaloes by that of their host halo, i.e. \( f_M = M_{\text{s}}/M_{\text{H}} \) and \( f_V = V_{\text{max,s}}/V_{\text{max,H}} \). Note here that we use the current subhalo masses and velocities, not their peak masses or \( V_{\text{max}} \) prior to accretion, as has been sometimes done in the previous work (see e.g. Rodríguez-Puebla et al. 2016) and normalize by the host halo mass \& \( V_{\text{max}} \) excluding the contribution of subhaloes.

These plots show excellent convergence in the mass and velocity functions, with the median distribution agreeing within the scatter and, critically, the scatter is itself well converged. As has been noted many times, the distribution appears relatively scale free (e.g. Springel et al. 2008; Stadel et al. 2009). However, unlike early studies we do not find that the distribution is characterized by a simple power law with an exponential cut-off at large masses. Instead, we see the presence of major mergers at large mass ratios of \( f_M \gtrsim 10^{-1} \). Haloes typically have at least one long-lived major merger remnant, with typical merger remnants having mass ratios of \( 0.13–0.44 \) with the next largest subhalo typically a factor of 2–20 times smaller. The relative absence of these large subhaloes in previous studies can be attributed to the use of configuration-space-based finders that artificially shrink subhaloes the deeper in previous studies we do not find that the distribution is characterized by a simple power law with an exponential cut-off at large masses. In fact, we see the presence of major mergers at large mass ratios of \( f_M \gtrsim 10^{-1} \). Haloes typically have at least one long-lived major merger remnant, with typical merger remnants having mass ratios of \( 0.13–0.44 \) with the next largest subhalo typically a factor of 2–20 times smaller. The relative absence of these large subhaloes in previous studies can be attributed to the use of configuration-space-based finders that artificially shrink subhaloes the deeper in the parent halo they reside. The result is a mass distribution that is characterized by a double Schechter function as noted in Han et al. (2018), which recovered these merger remnants using a tracking algorithm, HBT+.

We also see a subtle deviation in the abundance of small subhaloes with \( f_M \lesssim 10^{-3.25} \), which corresponds to the average mass ratio of a subhalo composed of \( \lesssim 100 \) particles. The velocity function also shows a deviation away from a simple power law at \( f_V \lesssim 9 \times 10^{-2} \) for all simulations resulting from poorly resolved subhaloes. These small haloes do not have well-converged density profiles, \( V_{\text{max}} \) values, and additionally are susceptible to artificial evaporation (van den Bosch 2017).

The insets show the distribution in the number of subhaloes at a given \( f_M/f_V^{\text{max,H}} \), i.e. the halo-to-halo scatter in the amplitude of the subhalo mass/velocity functions. The halo-to-halo scatter is roughly Gaussian. There does appear to be a shift in the most probable number of subhaloes between simulations probing different host halo mass scales. Cluster mass hosts probed in L900N2048 are richer than the group mass scales probed by our L210 simulations.

![Figure 8. Subhalo distributions. We show the \( z = 0 \) subhalo mass and maximum velocity functions (top and bottom). We plot the median distribution (thick line) and the 16 per cent and 84 per cent quantiles (shaded region) using all haloes composed of more than 50,000 particles containing more than 100 subhaloes to ensure a well-sampled mass and \( V_{\text{max}} \) function for each halo. We also indicate the number of host haloes used and the median and 16 per cent and 84 per cent quantiles of the host halo mass or \( V_{\text{max}} \). Colour, marker, and line styles are the same as in Fig. 4. In the insets, we show the distribution of \( n \) at a given \( f_M \) for simulations with \( \geq 100 \) host haloes.](image)

We fit the differential subhalo mass and velocity functions with a power-law functions using for all haloes simultaneously using EMCEE with the log likelihood given by

\[
\ln L = \sum_j \left( -1/2 \sum_i \left( \frac{n_i - n_{\text{model},i}}{\sigma_i} \right)^2 - \sum_i \ln 2\pi \sigma_i \right),
\]  

(6)
where \( j \) is the sum over haloes, \( i \) is the sum over the bins in the binned differential mass function of halo \( j \), \( \sigma_i = \sqrt{\sigma_2 + \frac{1}{2}} \) is the associated modified Poisson error, and
\[
 n_{\text{model}} = \int \frac{dn}{df} df = \int_0^\infty f^{\alpha + \beta} (A f^{-\alpha} + B f^{-\beta}) df, \tag{7}
\]
is the integral of the differential mass function over the bin. For the purposes of fitting, we do not fit the double power law simultaneously, instead we first limit the fit to the steeper index \( \alpha \), corresponding to substructures with well-defined dynamical masses as opposed to objects with masses of \( \gtrsim 5 \) percent of host halo’s mass that have been flagged by VELOCIraptor as possible merger remnants. We only fit well-resolved haloes with large subhalo populations and limit the fit to subhaloes composed of at least 100 particles. We also try fitting each halo individually to assess the halo-to-halo scatter, which is much larger than the scatter on the median parameters arising from fitting all haloes at once.

We find the average power law for the mass function to be \( \alpha_M = 0.83 \pm 0.01 \) for our highest resolution simulation. Fitting each halo individually gives \( \alpha_M = 0.77 \pm 0.26 \), in agreement with the fit to the average, to other SURFS simulations, and previous estimates (e.g. Madau, Diemand & Kuhlen 2008; Springel et al. 2008; Stadel et al. 2009; Gao et al. 2012; Onions et al. 2012; Rodríguez-Puebla et al. 2016; Han et al. 2018). The amplitude from fitting each halo individually is \( A_M = 0.11^{+0.82}_{-0.02} \) indicating the halo-to-halo scatter of \( \sim 0.9 \) dex, although the overall scatter inferred from fitting each halo individually is likely an overestimate as each halo has few subhaloes, so the scatter in each fit is dominated by the poor constraining power of each halo. We also note that the amplitude is highly correlated with \( \alpha \) (higher amplitudes, lower \( \alpha \) values).

The flatter high-mass fraction tail, corresponding to minor and major merger remnants, is less well defined because these remnants are comparatively rarer than subhalo accretion, resulting in poor sampling. Only on average does this region look like a power law, with a poorly constrained slope of \( \beta = 0.42^{\pm0.05}_{-0.32} \). This region is better characterized by a (possibly skewed) Gaussian distribution in \( \log f \), representing the distribution of ratios of merger events. By fitting the average current subhalo mass function, we are estimating the average mass ratios of long-lived mergers. We find \( \log f_{\text{merger}} = -0.83 \pm 0.01 \), with the dispersion being \( \sigma_{\log f_{\text{merger}}} = 0.312 \pm 0.003 \), i.e. the typical merger remnant has a mass ratio of 7–30 percent.

For the velocity function, limiting the fit to subhaloes composed of \( \gtrsim 100 \) particles constrains it to \( f_{\text{vmax}} \gtrsim 10^{-1} \). We find that the average distribution has a steep slope of \( \alpha_V = 2.13 \pm 0.03 \), while the fit to each halo independently gives \( \alpha_V = 2.09 \pm 0.86 \).

The middle panel splits haloes according to spin. Here we see the reverse trend (dashed lines above solid lines): haloes with low spin have more subhaloes than those with high spin, having typically \( \sim 30 \) percent more subhaloes. Again here the systematic bias is generally within the scatter of typical haloes.

The amount of substructure shows significant halo-to-halo scatter, but does it correlate with other bulk halo quantities? Does the subhalo abundance show some form of assembly bias? We find a correlation between the amount of substructure and halo’s concentration and spin. This dependence is presented in Fig. 9, where we split the host halo population by \( c \) and \( \lambda \) and compare the subhalo

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure9.png}
\caption{Subhalo distribution dependence. We show the \( z = 0 \) subhalo maximum velocity function ratio split according to concentration \( c \), spin \( \lambda \), and both (top, middle, and bottom, respectively). In all panels we compare the subhalo distribution in the lower (solid thin line) and upper (dashed thick line) 25 per cent to the median distribution within 25 per cent – 75 per cent. Gray lines indicate the scatter relative to the mean within the 25 per cent – 75 per cent quantile region. Colours are the same as in Fig. 4.}
\end{figure}

\begin{footnotesize}
\begin{itemize}
\item Han et al. (2018) also found a wide range of slope values depending on the mass used, the peak or current mass, mass enclosing 200 times the critical density, or 200 times the mean density.
\end{itemize}
\end{footnotesize}
5 COSMIC GROWTH

The simple picture of cosmic mass growth is one in which haloes grow in mass via smooth mass accretion and through the tidal disruption of subhaloes, and subhaloes slowly lose mass as they are pulled towards the centre of their host via dynamical friction till they are completely disrupted. However, this neglects tidal fields produced by nearby haloes, subhaloes leaving their host halo as they near apocentre, and major mergers, which can excite particles resulting in some mass (and angular momentum) loss. We analyse the growth of haloes and the evolution of subhaloes in this section but begin with convergence tests.

We note that the last four snapshots have not been fully corrected as we have not evolved our simulations in the future. However, this only spans \( z = 0 \) to \( z = 0.028 \), a small fraction of cosmic time. Moreover, only \( \approx 1.5 \) per cent of haloes composed of \( \lesssim 100 \) particles are affected and this percentage drops to \( \lesssim 0.1 \) per cent for haloes composed of \( \geq 500 \) particles (see the following section).

5.1 Numerical convergence

We examine how well we recover (sub)halo evolution using several diagnostics. We first start with comparing the properties of a (sub)halo to its immediate progenitor. The time difference between halo and progenitor is a minimum of \( \sim 200 \) Myr, but can be up to 1 Gyr for haloes which have an optimal progenitor found four snapshots in the past. This is rare and only occurs for \( \sim 0.3 \) per cent of haloes at all cosmic times. Typically only \( \sim 1 - 2 \) per cent of haloes have progenitors found more than a single snapshot in the past or not found at all. We show the dependence on halo particle number in Fig. 10. There is a strong resolution dependence on the fractions of haloes missing immediate progenitors, present in all our simulations, regardless of mass resolution. There may be a subtle halo mass dependence on the fractions themselves, with simulations probing lower mass scales having larger fractions of less than ideal progenitor links (going from L210N512 given by the green curve to L210N1536 given by the orange curve, for example).

These less than ideal links or missing links drop to \( \lesssim 1 \) per cent for haloes composed of \( \geq 100 \) particles. Large haloes that do not have ideal progenitors occur in multi-merging systems.

Subhaloes, shown in the lower panel, display different behaviour due to the highly non-linear, tidally disruptive environment in which they live. Here, no resolution dependence is seen and 2 per cent of all subhaloes have less than ideal links, evenly split between finding a progenitor two snapshots in the past and not finding an ideal progenitor within four snapshots. The subhaloes with missing progenitors are typically major merger events, and occasionally multi-mergers. In the upcoming work, we describe further improvements to tree building combined with additional particle tracking that is necessary to fix these rare cases (Elahi et al. in preparation; Poulton et al. in preparation).

Next we investigate the progenitor’s properties compared to the current (sub)halo in Fig. 11, specifically the mass and \( V_{\text{max}} \) ratio between a (sub)halo and its progenitor. The thick lines show the median ratio of progenitor to descendant mass and \( V_{\text{max}} \) for each simulation, with the shaded regions showing the scatter. For all mass scales, most objects do not evolve significantly over this period, with the median being \( \approx 1 \) to within \( \lesssim 1 \) per cent. If we separate subhaloes from haloes, we find that the median ratio is \( \approx 0.99 \), whereas for subhaloes the median ratio is \( \approx 1.01 \): unsurprisingly haloes grow and subhaloes shrink on average. The median and, more importantly, the scatter in this evolution are the same for all simulations for haloes composed of \( \geq 500 \) particles. Clearly the...
merger history of objects is only strongly numerically converged for objects above this particle limit. Even haloes composed of 100 particles have a 1σ scatter ∼10 per cent larger than better resolved haloes, showing this numerical effect. The scatter in the mass ratio is larger than the scatter in \(V_{\text{max}}\), \(\Delta M_{\text{prog}}/M \approx 5\) per cent compared to \(\Delta V_{\text{max,prog}}/V_{\text{max}} \approx 2\) per cent. The larger scatter is a result of the mass loss and growth mechanisms such as accretion of mass from subhaloes and subhaloes leaving their host having a larger impact on the outer regions of halo relative to the central regions defined by \(R_{\text{max}}\), despite the longer dynamical times. The asymmetry in the 2σ scatter is a result of major mergers, which can drastically increase the mass of a halo.

Changes in the mass accretion rate between two consecutive steps are also informative. We follow Contreras, Padilla & Lagos (2017) and define this change as

\[
\delta \Gamma = \frac{[\Delta i M_{\Delta} + \Delta i+1 M_{\Delta}]}{\Delta i+1 M_{\Delta} + [\Delta i+1, i+2 M_{\Delta}]} \eqno(8)
\]

where we average over steps \(i, i + 1, \& i + 2\). This ratio is only \(\leq 1\) in instances where the halo undergoes a mass decrease (increase) followed by an increase (decrease). Contreras et al. (2017) found it necessary to clean Millennium (Bolshoi) catalogues of those haloes with \(\delta \Gamma \leq 0.1\) (\(\delta \Gamma \geq 0.3\)), haloes which experienced significant and likely artificial fluctuations in the accretion rate. The Millennium catalogues were particularly affected across all halo masses according to Contreras et al. (2017), even for haloes composed of 1000s of particles.

We find only haloes above \(\sim 200\) particles show little mass dependence in \(\delta \Gamma\), with 6 per cent of haloes with \(\delta \Gamma \leq 0.1\). This mass accretion rate change for haloes composed of fewer particles show a strong dependence on particle number, increasing to 12 per cent for haloes composed of fewer than 50 particles, indicating mass accretion convergence for haloes composed of \(\geq 200\) particles. Moreover, this test indicates SURFS halo catalogue is not as affected by spurious mass accretion changes as the Millennium catalogues.

The results are the same at higher redshifts, indicating convergence in reconstruction of merger histories.

### 5.2 Halo evolution

#### 5.2.1 Formation time

We start with the formation time of haloes, which is itself a useful numerical convergence diagnostic. We show the formation time, here defined as the redshift \(z_{\text{form}}\) at which a halo’s \(M_{\Delta}(z)\) has 25 per cent of its current-day \(M_{\Delta}\). The formation time monotonically decreases with increasing present-day halo mass. For all well-resolved haloes composed of \(\geq 200\) particles, the formation time becomes biased to lower redshifts and later times. The \(z_{\text{form}} - M_{\Delta}\) relation is well characterized by \(z_{\text{form}} = \alpha \log M_{\Delta} + \beta\), with \(\alpha = -0.5 \pm 0.001\) & \(\beta = 8.15 \pm 0.01\) when fitting the full halo population (not the binned data plotted here), similar to the results of Power, Knebe & Knollmann (2012) although they looked at the formation time for 50 per cent of the current-day mass.

#### 5.2.2 Dark matter growth

The average evolution of haloes is presented in Fig. 13 for the L210N1536, where we have removed subhaloes, which have very different evolutionary paths. Here we split haloes into \(z = 0\) mass bins a decade in size from \(10^9 M_\odot\) to \(10^{15} M_\odot\) and for completeness we have included haloes down to 20 particles at \(z = 0\). For each mass bin we calculate the median evolutionary track along with 1σ quantiles.

Fig. 13 clearly shows smooth mass evolution with the largest haloes having accreted 10 per cent of their \(z = 0\) mass by \(z = 2\). The mass bin that differs in evolution, the red curve, corresponds to haloes composed of 20–90 particles, hence the lack of mass growth due to progenitors being below the resolution limit. If we examine the same mass bin in higher mass resolution simulations such as L40N512 where this bin contains haloes composed of 163–1634 particles, we find that the mass growth is the same as haloes of larger mass (see Fig. B1). The 1σ scatter in the evolution is \(\approx 0.3\) dex. The initial high mass growth at early times followed by a turnover at late times has been noted in numerous studies (e.g. van den Bosch 2002; Wechsler et al. 2002; Tasitsiomi et al. 2004; McBride, Fakhouri & Ma 2009; Rodríguez-Puebla et al. 2016).

Several parametrizations of halo mass growth have been used, from single parameter exponential growth (Wechsler et al. 2002) to two-parameter models (van den Bosch 2002; Tasitsiomi et al. 2004). We use a three-parameter model, a simpler functional form than that proposed by Rodríguez-Puebla et al. (2016) to characterize the average growth of haloes of mass \(M_{\Delta, o}(\alpha = 1)\),

\[
\log M_{\Delta}(\alpha) = A(M_{\Delta, o}) \exp \left[-(\alpha/c(M_{\Delta, o}))^{-\alpha(M_{\Delta, o})}\right] \eqno(9)
\]

where optimal fit parameters, \(A, \alpha, \& c\) depend on the final mass of the halo in question. Our fits are listed in Table 2. In general we find, parameters decrease with decreasing halo mass, although the statistical significance of this trend is low due to the scatter in the average evolution. Additionally, the smaller mass bins have artificially flatten growth rate as smaller progenitors lie below the

---

Figure 12. Halo Formation Time. We show the halo formation time as a function of mass dependence. Like Fig. 6, we bin haloes in mass, determine medians and quantiles and show the distribution at a particular mass range in the inset. Vertical lines mark the mass scale of haloes composed of 100 particles. Colour, marker, and line styles are the same as in Fig. 5. We also show a fit to the distribution by a dashed grey line.
resolution threshold. The growth rate, $dm/da = \alpha f^{\alpha+1} a^{-(\alpha+1)}$, decreases with increasing $a$.

The $V_{\text{max}}$ curves trace the mass growth, monotonically increasing with time save for mass bins of poorly resolved haloes. Comparing this figure to Fig. B1, we find that only mass bins of haloes composed of $\gtrsim 500$ particles show numerically converged evolution. The average scatter in evolutionary paths across cosmic time is $\approx 0.1$ dex. Again, the growth can be fit by a similar function to the mass growth and the fits are also listed in Table 2. The inferred growth rate of $V_{\text{max}}$ is smaller than the mass growth rate.

The comoving size of a halo as defined by $R_{V_{\text{max}}}$ also shows smooth evolution with little scatter. However, unlike mass and $V_{\text{max}}$, the comoving size does not grow continuously but instead peaks at mass-dependent redshifts and gradually decreases over time. This is not a result of the use of comoving gravitational softening lengths, nor is the turnover resolution dependent except for poorly resolved haloes as can be seen by again comparing this figure to Fig. B1 in Section B. The redshift at which this turnover occurs depends on mass with smaller mass haloes found at $z = 0$ turning over at higher redshift. The turnover corresponds to when the mass variance $\sigma(M) \approx 1.0$ at the average progenitor mass that is when haloes are common non-linear regions. As these scales become more non-linear, the haloes tend to virialize, becoming more concentrated and the comoving size shrinks. For halo masses of $\lesssim 10^{12} \, M_\odot$, $R_{V_{\text{max}}}$ decreases by $\sim 50$ per cent from its peak size.

The median evolution for concentration, spin, and shape is generally a simple function of $z$. However, these quantities show significant scatter in evolutionary tracks, on the order of 0.4 dex. The abrupt change in the median concentration seen at high redshift for large masses is a result of haloes being on average resolved enough for physically meaningful concentrations using maximum circular velocities to be calculated ($\sim 50$ particles). The upturn in $c$ and $\lambda$ occurs at roughly the same time as the downturn in $R_{V_{\text{max}}}$.

In general, key trends are that haloes become more concentrated at late times, with the comoving size shrinking below $z = 3$ with large haloes identified today contracting later than smaller haloes. The angular momentum does not evolve significantly, although the internal angular momentum of large haloes steadily decreases with time. Current-day haloes at a given mass are less triaxial and more compact that their high $z$ counterparts at similar halo mass.

### 5.3 Subhalo evolution

#### 5.3.1 Do subhaloes leave home?

It has been known for a while that subhaloes live interesting lives. They can momentarily leave their host (so-called backsplash galaxies), exchange hosts, are subject to the tidal field of their host, and encounter other subhaloes (e.g. Knebe et al. 2011). van den Bosch (2017) recently outline 12 evolutionary paths for subhaloes, most physical, others due to the limitations of the temporal resolution at

<table>
<thead>
<tr>
<th>Mass (M_\odot)</th>
<th>$A$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1.7 \times 10^{12}$</td>
<td>$8.1^{+0.07}_{-0.02}$</td>
<td>$0.80^{+0.08}_{-0.09}$</td>
<td>$0.028^{+0.02}_{-0.03}$</td>
</tr>
<tr>
<td>$1.7 \times 10^{13}$</td>
<td>$8.01 \pm 0.05$</td>
<td>$0.73 \pm 0.09$</td>
<td>$0.020^{+0.04}_{-0.02}$</td>
</tr>
<tr>
<td>$1.8 \times 10^{12}$</td>
<td>$7.87^{+0.21}_{-0.07}$</td>
<td>$0.58^{+0.13}_{-0.23}$</td>
<td>$0.013^{+0.21}_{-0.05}$</td>
</tr>
<tr>
<td>$1.7 \times 10^{11}$</td>
<td>$9.65^{+0.11}_{-1.75}$</td>
<td>$0.062^{+0.21}_{-0.01}$</td>
<td>$0.320^{+0.47}_{-0.315}$</td>
</tr>
</tbody>
</table>

$^a$Affected by resolution. $^b$Severely affected by resolution.
which subhaloes are identified and the biases of the halo finder. The dominant evolutionary channel is one in which a subhalo continues to be a subhalo from one output to the next, followed by those that momentarily leave their host. But how often does this happen over the lifetime of a subhalo? Are subhaloes sedentary, rarely leaving a host halo once accreted?

We examine the lives of subhaloes by following subhaloes identified at $z = 0.4$ (4 Gyr ago, to allow subhalo lifetimes to be measured) back to their accretion and then following their lives afterwards. We only examine subhaloes that were composed of $\geq 100$ when first accreted and limit our analysis to host haloes that have a subhalo population of $\geq 10$, that is moderately resolved host haloes and do not split results based on host halo mass or examine any dependence on cosmic time, leaving a more detailed analysis for a future study. These selection criteria mean that different simulations explore different mass scales.

We define flybys as subhaloes that were subhaloes for at most four snapshots (or less than the freefall time at $R_c$) over the course of their life. Haloes that swap hosts are, naturally, subhaloes that have changed host at least once since being first accreted. Here we define back splash subhaloes as those that momentarily leave their host halo,7 excluding those subhaloes that have also swapped hosts. We define preprocessed subhaloes here as subhaloes that were at earlier times subhaloes of a different host. The fractions of these subhaloes are listed in Table 3.

Clearly, subhaloes are dynamic, with $\sim 40$–50 per cent either leaving their host halo momentarily, switching hosts entirely, or just momentarily becoming a subhalo. About 1/4 of all subhaloes also leave the virial radius of their host halo before being re-accreted, with another $\sim 2$ per cent leaving their host entirely. Approximately 25 per cent of all long-lived subhaloes truly leave one host and enter another one. The preprocessing of subhaloes, that is where a subhalo’s host halo is itself accreted, is a natural outcome of hierarchical structure formation. We do not find this to be a dominant channel of accretion, though it is still significant at 30 per cent. Given the resolution limits, this fraction is likely an underestimate.

These populations may have host mass-scale dependence. For example, L40N512, which resolves lower mass host haloes and subhaloes than the L210 simulations, has a higher fraction of subhaloes that change hosts or that are flybys, though it has fewer subhaloes that have been preprocessed. We will explore the host dependence of subhalo orbits in great detail in the upcoming work.

5.3.2 How subhaloes orbit their host

We examine the orbit of subhaloes about their host halo, interpolating their radial and tangential positions and velocities between snapshots and identify peri/apocentric passages using changes in the radial motion. An example of an orbit is shown in Fig. 14 for a subhalo that was accreted at a look-back time of 5 Gyr. Note that this subhalo is not disrupted, the curves halt $\approx 5$ Gyr after accretion as this is the end of the simulation, i.e. $z = 0$.

The pericentre/apocentres are easily identifiable, giving an orbital period of 3.5 Gyr. This subhalo typically experiences gradual mass loss over its orbits, which starts principally after first pericentric passage. Although there is a minor fluctuation in the virial mass when the subhalo is near apocentre, this fluctuation is a result of the close passage of another subhalo, artificially skewing the recovered mass. There is also a kink in the mass associated from being identified by the FOF algorithm as a field halo to being identified as a subhalo by the velocity/phase-space subhalo algorithm, as FOF masses are inclusive, including substructure within the object whereas substructures have exclusive masses that do not include internal substructure. There is also a drop in mass of 10 per cent from the loss of high angular momentum material in the FOF

### Table 3. Subhalo population

We list the number of subhaloes $N_f$ used to estimate the following statistics: Backsplash Fraction $f_{BS}$; fraction that swap hosts at least once $f_{BH}$; the preprocessed fraction $f_{PP}$; and the flyby fraction $f_{FB}$. Simulations limited to those with numerous well-resolved haloes across cosmic time.

<table>
<thead>
<tr>
<th>Name</th>
<th>$N_f$</th>
<th>$f_{BS}$</th>
<th>$f_{BH}$</th>
<th>$f_{PP}$</th>
<th>$f_{FB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>L40N512</td>
<td>2355</td>
<td>0.234</td>
<td>0.282</td>
<td>0.291</td>
<td>0.033</td>
</tr>
<tr>
<td>L210N1024</td>
<td>14774</td>
<td>0.257</td>
<td>0.193</td>
<td>0.368</td>
<td>0.028</td>
</tr>
<tr>
<td>L210N1536</td>
<td>46861</td>
<td>0.246</td>
<td>0.235</td>
<td>0.291</td>
<td>0.040</td>
</tr>
</tbody>
</table>

---

7 Here we mean by ‘leave a halo’ as a subhalo that is no longer in the 6DFOF envelop defining a halo. The exact definition is not critical, but it is worth noting that there are several definitions of the edge of a halo and what constitutes a subhalo. The typical halo edge is commonly delineated by the virial radius, which has several definitions in the literature. For example, More, Diemer & Kravtsov (2015) discuss some of these definitions, present some of their drawbacks, and advocate using the ‘splashback’ radius, the mean apocentre of particles belonging to the halo as a more physical boundary to a halo. This radius is related to $R_c$ and more critically, this first caustic surface is likely related to the 6DFOF envelop used to define our haloes, hence we feel justified here to use our definition of subhalo.
envelop upon accretion. In general, the evolution of the object is well recovered in the dense environment of a halo.\(^8\)

The resulting probability distribution functions (PDFs) of orbital properties for subhaloes are shown in Fig. 15. Here we focus on the orbital period, identified between two pericentric passages, the radius of pericentre (normalized by the virial radius of the host), the orbital ellipticity \(e = \frac{\delta r}{\Delta 1}/ \Delta 1\), and three evolutionary properties: the sum of the change in the radial velocity, the mass change over an orbit, and the fractional change in the orbital angular momentum over an orbit. These evolutionary properties are defined as

\[
\begin{align*}
\delta M_{\rm {r,peri}} &= \frac{1}{T_{\rm orb}} \sum_{i,\rm{orb}} \left( \frac{M(t_{i+1}) - M(t_i)}{M(t_i)} \right), \\
\delta M_{\rm m,peri} &= \frac{1}{T_{\rm orb}} \sum_{i,\rm{orb}} \left( M(t_{i+1}) - M(t_i) - 0.5\delta t \right), \\
\delta j_{\rm orb} &= \frac{1}{T_{\rm orb}} \sum_{i,\rm{orb}} \left( j(t_{i+1}) - j(t_i) \right).
\end{align*}
\]

(10) Here \(j \equiv |\mathbf{r} \times \mathbf{v}|\) is the specific orbital angular momentum, \(\delta t\) is the time between snapshots, and the sums are for times over the orbit.

Overall, our simulations give similar PDFs for all these orbital properties. Typical orbital times are \(\sim 2\) Gyr or more, so that first pericentric passage occurs roughly a Gyr after accretion. The distribution shows little variation between simulations. The few subhaloes with very short periods (<0.2 Gyr) result from the simple identification of peri/apocentric passage. We identify these points along an orbit, using changes in the sign of \(v_r\), the radial velocity relative to the host halo. At large radii outside the host halo where \(v_r\) is small, encounters with other haloes can alter the infall velocity leading to spurious identification of peri/apocentric passage, which happens for \(\sim 3\) per cent of all pericentres measured (these typically also \(e \sim 0\)). Very long orbital times of \(\gtrsim 8\) Gyr are also a result of misidentified apocentres and are more representative of infall times.

The pericentre PDF shows typical pericentres of \(\lesssim 0.3R_\Delta\). The few subhaloes with pericentres identified at large radii \(r_{\rm peri}/R_\Delta > 2.0\) are due to spurious identification of pericentric passage, and occurs less than 1 per cent of the time.

The orbital ellipticity PDFs appear to be composed of two distinct populations. One at very low \(e\), which drops quickly and another which is much broader than peaks at \(e \sim 0.7\). In general, most subhaloes are on quite elliptical orbits, with 50 per cent having \(e \gtrsim 0.59\). The origin of these populations is simple: orbits with \(e \lesssim 0.2\) arise from subhaloes that have already completed two pericentric passages, whereas \(e \gtrsim 0.2\) arise from infalling orbits, where the subhalo has completed a passage from infall to pericentre and out to apocentre. The dense environment within a halo circularizes orbits. There is also a small fraction (\(\lesssim 3\) per cent) with very elliptical \((e \gtrsim 0.99)\) radially plunging orbits, where an apocentre has been incorrectly identified, and is actually on first infall with long orbit times.

All the evolutionary properties show that most subhaloes do not experience large changes in either their mass or their orbit. Averaged over an orbit, a subhalo does not change mass significantly, \(\delta M_{\rm m,peri} \sim 0\). This is not too surprising given the dynamic lives of subhaloes, where they can spend significant amounts of time outside the virial radius of their host, and where they could re-accrete some material. The figure shows that subhaloes experience mass loss at pericentric passage, with median mass loss rates of \(\sim 40\) per cent per Gyr, though there can be significant variations in the mass loss rate.

---

\(^8\) Configuration-space halo finders exhibit systematic artificial decrease in assigned subhalo mass with decreasing radius (Muldrew et al. 2011), along with larger fluctuations than the factor of 3 change in mass seen here, with these fluctuations occurring more often.

---

**Figure 15. Orbital properties.** We show the median distribution for orbital periods, pericentric passage, and ellipticity of the orbit in the top row and the mass change over pericentric passage, mass change over an orbit, and the change in orbital angular momentum over an orbit in the bottom row. We show the median normalized histograms (binmed pdfs) (solid lines) along with bootstrap estimated errors (shaded regions) for each property. We also show the median value for the property along with the 16 per cent, 84 per cent quantiles. Colour, marker, and line styles are the same as in Fig. 4.
The orbital angular momentum is generally conserved with some change in \( j \) accompanying the mass loss. The positive angular momentum change can circularise an orbit. This can be seen from relating \( j_{\text{tab},} e, r_{\text{peri}}, \) and \( v_{\text{apo}}. \) If \( j_{\text{tab}} > 0 \) then as \( j = r_{\text{apo}}v_{\text{apo}} \), both \( r_{\text{apo}} > 0 \) & \( v_{\text{apo}} > 0 \) with the same also true at pericentre. With a little bit of algebra, one can show that \( j_{\text{tab}} > 0 \) is satisfied for

\[
1 - e^2 \frac{r_{\text{peri}}}{r_{\text{peri}}} < \frac{1 - e^2 v_{\text{peri}}}{v_{\text{peri}}}. \tag{13}
\]

This range spans \( e < 0 \), corresponding to circularizing orbits.

### 5.3.3 Life after accretion

We examine the lives of subhaloes after accretion in more detail in Fig. 16, again by following subhaloes identified at \( z = 0.4 \). Here we only keep those objects that spend most of their life as a subhalo after accretion (removing flybys but keeping back splash and host swapping subhaloes\(^9\)). We only show results from L210N1536 as results from other simulations are similar. We calculate the fractional change of a variety of properties relative to the accretion value as a function of time, such as mass \( \delta M_{\text{acc}} \equiv M(t)/M_{\text{acc}} \), and plot the median and 1 \( \sigma \) variation given by the 16 per cent and 84 per cent quantiles. We split the evolution of subhaloes based on the mass ratio between the subhalo and its host prior to accretion going from blue to red in decreasing mass ratios of \( (\geq 1, 1: 10, 1: 10^2, 1: 10^3, 1: 10^4, \leq 10^{-3}) \). Note that mass ratios indicative of mergers are given by dark blue curves for subhaloes which are more massive than their host prior to accretion and blue curves for major/minor mergers (mass ratio of 1:1 to 1:10). We should also point out that the subhalo sample presented here is biased towards larger mass ratios as we include all host haloes composed of \( \geq 5000 \) particles, for which the smallest mass ratio that can be identified given the 20 particle limit is \( 10^{-2.5} \). Subhaloes with smaller mass fractions can only reside in the largest host haloes with \( M_{\Lambda} \geq 2 \times 10^{12} \odot \).

The survival fraction depends on mass ratio between the subhalo and its host. For very small mass ratios of \( \lesssim 10^{-3} \) (yellow and red curves) there are resolution effects at play. Subhaloes can only be identified when composed of 20 particles, thus when a subhalo drops below this threshold, the subhalo is considered to be tidally disrupted. Of greater importance is artificially shortened lifetimes due to artificial evaporation. van den Bosch (2017) found \( \approx 80 \) per cent of subhaloes in the Bolshoi+ROCKSTAR (sub)halo catalogue are artificially disrupted, although this typically occurs for moderately resolved (sub)haloes composed of \( \lesssim 1000 \) particles after they have lost \( \geq 90 \) per cent of their accretion mass. The artificial evaporation rate is higher for less well-resolved subhaloes. The mass ratio bins with \( \lesssim 10^{-2} \) (green to red) are likely to have artificially reduced survival fractions as these bins are dominated by subhaloes composed of \( \leq 100 \) particles.

Considering these caveats, the lifetime of most objects is quite long, \( \gtrsim 5 \) Gyr. It does appear that of the subhaloes identified at \( z = 0.4, \approx 60 \) per cent are on long enough orbits with pericentres far enough away to avoid disruption. We will examine the orbits and the associated survival rates in an upcoming paper.

The median evolution in mass relative to the accretion \( M_{\text{acc}} \) indicates subhaloes initially do not significantly change mass upon accretion, with a delay of \( \approx 1 \) Gyr, before typically losing mass at

\[^9\] A more detailed analysis of mass loss resulting from preprocessing similar to Joshi, Wadsley & Parker (2017) is saved for future work.
an almost linear rate. This slight delay is roughly 1/2 the typical orbit time (see Fig. 15 in Section 5.3.2). Even major mergers (blue and dark blue curves) typically do not suffer much mass loss, gradually shrinking to $1/10^9$ their mass over the course of Gyrs. For subhaloes with small mass ratios, accretion can remove some of the loosely bound particles assigned to the object by the FOF algorithm as seen by the small dip in mass just after accretion, particularly for poorly resolved subhaloes with artificially shallow potential wells.

The scatter in the evolution indicates how complex the lives of subhaloes can be. The upper quantile for major mergers indicates mass growth, though in such instances one could argue that neither object should be thought of as a host halo and the other as a subhalo. These mass growth instances typically occur when the subhalo/host halo tag switches between haloes and the subhalo mass reconstruction is less certain or in instances where the merger is initially a glancing one, both objects becoming field haloes for a time before re-coalescing. Even minor mergers can experience some mass growth, accreting some of the more loosely bound outer regions of a host halo. In general, the scatter is $\approx 0.3$ dex save for poorly resolved subhaloes (red and yellow curves) where it can be $\geq 0.5$ dex, with the upper quantile occasionally indicating mass growth. Here, these instances of mass growth occur when a subhalo has momentarily left its host halo and is able to accrete mass or has merged with another subhalo. The high 84 per cent quantile lines where $\delta M_{\text{acc}} > 1.2$ at late times for poorly resolved haloes is principally due to subhaloes that have become haloes. These former subhaloes have longer lifetimes than similar mass subhaloes that remain with their host, dominating the median and upper quantiles at late times, resulting in the ever-increasing upper quantile and upturn in the median seen.

We fit the mass change with a simple exponential decay,

$$\delta M_{\text{acc}}(t) = \begin{cases} 1, & t < t_0, \\ \exp [-\alpha (t - t_0)^\tau], & t \geq t_0, \end{cases} \tag{14}$$

where $t_0$ is the time delay between accretion and mass loss, $\tau$ is the time-scale, and $\alpha$ describes the loss rate. For mass accretion ratios of $10^{-3} < M_{\text{acc}}/M_{\text{fit}} < 10^{-1}$, we find delay times of $t_0 = 0.68^{+0.22}_{-0.20}$ Gyr, times scales of $\tau = 4.9 \pm 1.0$ Gyr and $\alpha = 0.80^{+0.59}_{-0.31}$. It should be noted that $\alpha$ and $\tau$ are strongly anti-correlated. The large scatter between the lower and upper quantiles is reflected in the uncertainties in the parameters. Nevertheless we find haloes with larger mass ratios have longer delay times and time-scales to $\sigma$ significance. However, given that these differences are likely driven not only by accretion mass ratios but orbital parameters, we leave a more detailed analysis for later. In general, the delay times are consistent with the time-scales for first pericentric approach, whereas the time-scales for significant mass loss indicate that several pericentric passages are required to significantly alter the mass.

The evolution in $V_{\text{max}}$ follows the mass: objects gradually become smaller over time. The evolution appears almost linear with time but to compare with the time-scales of mass evolution, we fit $\delta V_{\text{max}, \text{acc}}$ with the same exponential function as mass. We find a similar though slightly smaller delay time, $t_0 = 0.76 \pm 0.11$ Gyr and similar power-law $\alpha = 0.48^{+0.93}_{-0.33}$. The key difference is the longer time-scales,

$$\tau = 17.6 \pm 2.0 \text{ Gyr. The evolution is still not quite linear with time, } \delta V_{\text{max}, \text{acc}} \approx 1 - x + x^2/2 - x^3/6, x = ((t - t_0)/\tau)^\alpha. \text{ There is also less scatter, } \approx 0.1 \text{ dex, in the evolutionary paths, indicating that } V_{\text{max}} \text{ is a better tracer of subhalo evolution than total mass.}$$

We should note that these fits and associated time-scales are biased towards long-lived subhaloes. Due to artificially enhanced mass loss for poorly resolved subhaloes, these fits are more representative of subhaloes with mass ratios of $\gtrsim 10^{-3}$. The true mass loss and evolution of smaller subhaloes require even higher resolution simulations. A more complete follow-up of the dependence of lifetimes and mass loss on orbital parameters will be discussed in future work.

As typical subhaloes lose mass and decrease in $V_{\text{max}}$, well-resolved subhaloes shrink as seen by the decrease in $R_{\text{max}}$ in the fourth panel in Fig. 16. As subhaloes shrink, they typically become more concentrated (for well-resolved haloes at least). Due to the numerical limitations, low mass ratio subhaloes have systematically enlarged $R_{\text{max}}$ (see Fig. 5), affecting the inferred evolution, hence it is difficult to draw physically meaningful conclusions from the average population here. The scatter is large in both $R_{\text{max}}$ and $\sigma$, approximately 50 per cent.

For mass ratios of $\lesssim 10^{-2}$, subhaloes initially lose loosely bound, high angular momentum material upon accretion as indicated by the decrease in spin, within both $R_{\text{max}}$ and $V_{\text{max}}$ Afterwards, on average these subhaloes spin up due to tidal torques, though the scatter in $\delta V_{\text{acc}}$ is quite large, around 0.4 dex, with the upper region of the upper quantiles steadily increasing with time, till the object either becomes poorly resolved or has undergone several orbits. The increase in global spin is not tracked by the spin within $R_{\text{max}}$, which on average remains unchanged, in qualitative agreement with the tidal torques spinning up the outskirts of subhaloes.

The single feature that does not change is the average shape as indicated by $\delta q_{\text{acc}}$. This remains around 1 with a scatter of 0.2 dex.

In general, once accreted and on bound orbits, subhaloes gradually decrease in $M$, $V_{\text{max}}$, $R_{\text{max}}$, become more concentrated. With each orbit, the outskirts of subhaloes are spun-up, with this high angular momentum material stripped first, while the central angular momentum changes little.

### 6 Discussion

We have given an overview of the Synthetic Universes For Surveys (SURFS) simulation suite, consisting of both pure N-Body and non-radiative hydrodynamical simulations. These simulations have large enough volumes to be useful for surveys and studying the statistical properties of haloes and high enough resolution to resolve the internal properties of dark matter haloes and subhaloes, a requirement for any investigation into cosmic structure formation and galaxy formation physics. Our high fidelity halo catalogues and merger trees allow us to follow the orbits of subhaloes and their dynamical lives to a precision never before reached in synthetic survey simulations. Plus, our simulation volumes are large enough to study the cosmic web, and examine how the life of a halo is affected by where it resides in this web.

We have shown numerical convergence between our various volumes and resolutions, clearly demonstrating that internal halo properties and merger histories should be limited to halo composed of $\gtrsim 100$ particles. Ideally, any analysis that takes halo merger trees from simulations as input should limit it to haloes with numerically converged accretion histories. The limit for a strongly numerically converged temporal halo catalogue is even more conservative, only haloes composed of $\gtrsim 500$ particles should be used. This limit
affects any SAM, whereas less physically meaningful HOD models can use haloes composed of $\gtrsim 20$ particles. Thus, our catalogues can be used as input to SAMs to produce numerically converged galaxy population residing in haloes down to masses of $10^{11} \, M_\odot$ and study the orbits of galaxies in small groups up to cluster environments. Our overview highlights a selection of interesting results from haloes and subhaloes.

An analysis of large-scale structure shows:

(i) The cosmic web is best reconstructed with haloes of $M_\Delta \gtrsim 10^{11.5} \, M_\odot$ and requires a survey to be complete to at least $M_\Delta \gtrsim 10^{12} \, M_\odot$. Only surveys like WAVES will produce robust cosmic web reconstructions.

(ii) Large haloes typically occupy knots and filaments, whereas smaller haloes reside in a larger variety of environments, spanning a broad range of distances from the nearest cosmic web filament.

By tracing haloes at high cadence across cosmic time, we find:

(i) Haloes smoothly grow in mass and $V_{\text{max}}$, in agreement with the previous work (e.g. van den Bosch 2002; Wechsler et al. 2002; McBride et al. 2009; Rodríguez-Puebla et al. 2016). The mass history of a halo identified at $z = 0$ with mass $M_{\Delta, 0}$ is well characterized by $\log M(a) = -\log M_{\Delta, 0} \exp\left[-(a/\beta)\sigma\right]$, with parameters that depend weakly on the $z = 0$ mass. This functional form is simpler than the recent one proposed by Rodríguez-Puebla et al. (2016). The typical cosmic evolution of a $1.6 \times 10^{12} \, M_\odot$ halo has $\alpha = 0.84^{+0.09}_{-0.11}$ & $\beta = 0.022^{+0.002}_{-0.003}$. The velocity scale of a halo is also well characterized by this function with $1.6 \times 10^{12} \, M_\odot$ haloes having $\alpha_{\text{max}} = 1.12^{+0.18}_{-0.26}$ & $\beta_{\text{max}} = 0.028^{+0.03}_{-0.04}$. Mass grows faster than $V_{\text{max}}$.

(ii) Haloes also grow in comoving $R_{\text{max}}$ till they begin to virialize, which on average occurs when these haloes are common non-linear density peaks, i.e. when the mass variance $\sigma(M) \approx 1$. Haloes continue to grow in mass as they virialize but contract in comoving $R_{\text{max}}$, becoming more concentrated and spherical. This turnover point depends on mass. The exact relationship between contraction, the virial state, mass accretion rate, merger rate, and the mass scale will be explored in a future paper.

Our simulations have enough well-resolved haloes for a statistical analysis of the subhalo population. Our simulations show:

(i) Subhaloes follow power-law mass and velocity functions, $n(> M) \propto f^{-\gamma}$ with indices of $\alpha_M = 0.83 \pm 0.01$ and $\alpha_V = 2.13 \pm 0.03$, in agreement with previous studies (e.g. Onions et al. 2012; Rodríguez-Puebla et al. 2016), roughly independent of redshift. The halo-to-halo scatter in the mass function is $\sigma_{\log M} \approx 0.26$. The scatter in $V_{\text{max}}$ is not as well constrained, given the limited dynamic range, giving an overestimated scatter of $\sigma_{\log V_{\text{max}}} \approx 0.86$. The amplitude shows $0.9$ dex scatter.

(ii) The number of subhaloes residing within a host halo is weakly correlated with the host halo’s concentration $c$ and spin $\lambda$. A halo with $(c, \lambda) \pm \sigma$ (above, below) the mean values will have $60$ per cent more subhaloes than similar mass halo with $c$ and $\lambda \pm \sigma$ below the mean.

(iii) Subhaloes are dynamic residents. Approximately $25$ per cent leave their host halo momentarily, becoming a backsplash subhalo, and another $\sim 25$ per cent changing hosts entirely. This is in rough agreement with Warnick, Knebe & Power (2008); Knebe et al. (2011); van den Bosch (2017), with differences in the fraction of dynamic subhaloes depending on the exact definition of a halo’s boundary, be it defined by critical density, mean density, or some other definition. We find that a moderate fraction of subhaloes, $\sim 30$ per cent, are preprocessed, having been a subhalo of a host which itself becomes a subhalo.

(iv) Two distributions of ellipticities are observed. Subhaloes that have complete several orbits can have low $e$ values due to dynamical friction and angular momentum exchange circularizing orbits. These subhaloes comprise the smaller fraction of subhaloes with only $16$ per cent of the population having $e \lesssim 0.15$. In general, $e$ peaks at $\sim 0.75$ and has a broad distribution. Similarly, orbital periods show two populations, short orbital times for subhaloes that have complete several orbits and a broader distribution peaked at $2.4$ Gyr.

(v) Evolution during an orbit is on average smooth, gradually losing mass and decreasing in $V_{\text{max}}$, following a delay of $\sim 0.7$ Gyr, roughly close to first pericentric passage. Most of the mass loss occurs at pericentric passage, where subhaloes can lose on average $40$ per cent Gyr$^{-1}$. The time-scale for exponential mass loss is $4.9 \pm 1.0$ Gyr, that is several pericentric passages.

(vi) The central regions of long-lived subhaloes are less perturbed by tidal fields, with $V_{\text{max}}$ decreasing on longer time-scales of $17.6 \pm 2.0$ Gyr. Given the steady evolution of $V_{\text{max}}$, which shows less scatter than mass, this quantity is a better tracer of subhalo evolution.

7 WHAT’S NEXT?

These results are just the tip of the proverbial iceberg: much more remains to be examined, both from the dark matter halo point of view and from the perspective of galaxy formation physics. Even here we have focused on just dark matter haloes, synthetic galaxy catalogues using SAMs will follow. Future work will explore a variety of topics. For instance, the improved subhalo tracking will necessitate updates.
to SAMs, particularly in dynamical friction schemes and orbital angular momentum evolution. We have hundreds of group mass haloes resolved with several hundred thousand particles containing hundreds of subhaloes and tens of low mass clusters composed of millions of particles. These objects are ideal for testing the internal structure and evolution of groups across cosmic time.

We are even able to follow the evolution of Milky Way+Magellanic cloud like systems, as shown in Fig. 17, one of the key science goals of the WAVES survey.

As work progresses, we plan to make the halo catalogues, trees, and eventually semi-analytic galaxy catalogues available to the public in a searchable database and raw data, allowing the community to use SURFS data for their own research purposes.

ACKNOWLEDGEMENTS

We would like to thank the anonymous referee for their comments. PJE is supported by the Australia Research Council (ARC) Discovery Project Grant DP160102235. CW is supported by the Jim Buckee Fellowship. CP is supported by ARC Future Fellowship FT130100041. CL is funded by a Discovery Early Career Researcher Award DE150100618. RC also thanks the MERAC Foundation for a Postdoctoral Research Award. AR acknowledges the support of ARC Discovery Project grant DP140100395. RC is supported by the SIRF awarded by the University of Western Australia Scholarships Committee, and the Consejo Nacional de Ciencia y Tecnología (CONACYT) scholarship No. 438594 and the MERAC Foundation. RP is supported by a University of Western Australia Scholarship. Parts of this research were conducted by the ARC Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020, and supported by the ARC Discovery Project DP160102235. This research was undertaken on Magnus at the Pawsey Supercomputing Centre in Perth, Australia, and on Raijin, the NCI National Facility in Canberra, Australia, which is supported by the Australian Commonwealth Government. Parts of this research were conducted by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through the project number CE170100013.

The authors contributed to this paper in the following ways: PJE ran simulations and analysed the data, made the plots, and wrote the bulk of the paper. CW analysed the data and wrote sections of the paper. CP ran simulations. RC and RP assisted with the development of the software used to produce the data, VELOCIRAPTOR & TREEFROG respectively. PJE, AR, CP, and CL lead the SURFS project. All authors have read and commented on the paper.

Facilities

Magnus (Pawsey Supercomputing Centre), Raijin (NCI National Facility).

Software

PYTHIA, MATPLOTLIB (Hunter 2007), SCIPY (Jones et al. 2001@), EMCEE (Foreman-Mackey et al. 2013), SCIKIT (Pedregosa et al. 2011), GADGET (Springel 2005), VELOCIRAPTOR (Elahi et al. 2011, Elahi et al. in preparation), and TREEFROG (Elahi et al. in preparation).

REFERENCES

Johnston S. et al., 2008, Exp. Astron., 22, 151
Johnston S. et al., 2008, Exp. Astron., 22, 151
Jones E., Oliphant T., Peterson P. et al., 2001, SciPy: Open source scientific tools for Python. [Online; accessed <today>]
APPENDIX A: MASS FUNCTION

Following numerous studies (e.g. Watson et al. 2013; Poole et al. 2016), we fit the halo mass function

\[ \frac{dn}{d \ln M_A} = \frac{\rho_{\text{tot}}}{\pi \sigma_i} f(\sigma, z) \left( \frac{d\ln \sigma}{d\ln M} \right) , \]  

(A1)

where \( f(\sigma, z) \) is the scaled mass function, which here we assume a universal (i.e. redshift independent) functional form of

\[ f(\sigma, z) = A \left( \frac{\beta}{\sigma + 1} \right)^{\gamma/\sigma^2} . \]  

(A2)

Here \( A, \beta \) are normalization parameters and \( \sigma, \gamma \) are shape parameters. We determine the PDFs of these model parameters using EMCEE with the log likelihood given by

\[ \ln L = -1/2 \sum_i \left( \frac{n_i - n_{\text{model}}}{\sigma_i} \right)^2 - \sum_i \ln \sqrt{2\pi} \sigma_i , \]  

(A3)

the integral of the differential mass function over the bin.

Typically, studies fit the binned mass function derived from the entire simulation volume, correcting for missing power on scales larger than the simulation volume. Though this might give the most precise fit to simulation data, it is not an accurate representation of the true HMF.\(^{11}\) Baryon physics can significantly alter dark matter haloes. Moreover, the resulting PDFs of the parameters will be missing the scatter coming from cosmic variance (and missing large-scale power). Given these issues, we do not attempt to account for missing large-scale power or cosmic variance. We would argue that fitting the HMF from a pure N-Body simulation to within 1 per cent is not a meaningful task unless compared to other simulations in a careful check of softening lengths, hydrodynamical parameters, and subgrid physics (e.g. Gnedin et al. 2004; Cui et al. 2012; Zolotov et al. 2012; Chan et al. 2015; Bosquet et al. 2016; Despali & Vegotti 2017; Garrison-Kimmel et al. 2017), are just a few of the studies that show the significant impact baryons have on dark matter, from dark matter concentration to dark matter mass).

Instead we simply fit the binned differential mass function sampled at the 100 different mass scales, the first 50 largest haloes to capture the exponential turn-over, and the next equally spaced in log \( M \). The results of the fitting procedure from the EMCEE package (Foreman-Mackey et al. 2013) using uninformed priors are shown in Fig. A1. Our parameter PDFs are qualitatively the same as any other work; there are degeneracies between normalization parameters, \( \log A \) and \( \beta \), between the normalization parameters and the shape parameter \( \sigma \). The exponential shape parameter \( \gamma \) is weakly correlated with other parameters. The normalization parameter \( \beta \) has a long tail due to the number of haloes constraining the turn-over from power law to exponential, and the most likely value of \( \beta = 4.42 \). Our results are broadly in agreement with Watson et al.

\(^{11}\) Additionally, some studies fit the FOF mass, others the mass enclosing densities of \( \rho = \Delta \rho_{\text{crit}} \), or \( \rho = \Delta \rho_{\text{S}} \), with \( \Delta = 200 \) to \( \Delta = 500 \), a function of redshift based on spherical collapse.
APPENDIX B: COSMIC GROWTH AND RESOLUTION

We show the growth of cosmic structure of different simulations here, specifically our higher mass resolution, smaller volume simulation (Fig. B1), and our lower resolution L210N1024 simulation (Fig. B2). The median growth of haloes observed in L210N1536 is reproduced by L40N512 for well-resolved haloes in both simulations, that is mass bins of $\gtrsim 10^{12} \, M_\odot$. The halo mass bin of $10^{10}-10^{11} \, M_\odot$ (red curve), which contains haloes composed of $\sim 162-1619$ particles in L40N512, reproduces the evolutionary behaviour of large haloes, which is not seen in L210N1536 due to resolution effects.

A similar impact of reducing our resolution is seen in L210N1024. Here the $10^{11}-10^{12} \, M_\odot$ (yellow curve) shows little evolution beyond a redshift of 2, and the $10^{10}-10^{11} \, M_\odot$ (red curve) shows no evolution. In contrast L210N1536 $10^{11}-10^{12} \, M_\odot$ haloes begin to evolve from $z \sim 4$ onwards.

Figure B1. Halo Evolution of L40N512. The average halo evolution similar to Fig. 13 but for L40N512.
Figure B2. Halo Evolution of L210N1024. The average halo evolution similar to Fig. 13 but for L210N1024.

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