



A statistical semi-empirical model: satellite galaxies in groups and clusters

Philip J. Grylls¹, F. Shankar,¹ L. Zanisi¹ and M. Bernardi²

¹Department of Physics and Astronomy, University of Southampton, Highfield SO171BJ, UK

²Department of Physics and Astronomy, University of Pennsylvania, PA 19104, USA

Accepted 2018 November 28. Received 2018 November 27; in original form 2018 July 27

ABSTRACT

We present STEEL, a STatistical sEmi-Empirical modeL, designed to probe the distribution of satellite galaxies in groups and clusters. Our fast statistical methodology relies on tracing the abundances of central and satellite haloes via their mass functions at all cosmic epochs with virtually no limitation on cosmic volume and mass resolution. From mean halo accretion histories and subhalo mass functions, the satellite mass function is progressively built in time via abundance matching techniques constrained by number densities of centrals in the local Universe. By enforcing dynamical merging time-scales as predicted by high-resolution N -body simulations, we obtain satellite distributions as a function of stellar mass and halo mass consistent with current data. We show that stellar stripping, star formation, and quenching play all a secondary role in setting the number densities of massive satellites above $M_* \gtrsim 3 \times 10^{10} M_\odot$. We further show that observed star formation rates used in our empirical model over predict low-mass satellites below $M_* \lesssim 3 \times 10^{10} M_\odot$, whereas star formation rates derived from a continuity equation approach yield the correct abundances similar to previous results for centrals.

Key words: galaxies: clusters: general – galaxies: formation – galaxies: groups: general – galaxies: haloes.

1 INTRODUCTION

Cold dark matter (CDM) predicts a hierarchical assembly of dark matter haloes, which grow in mass through accretion from smaller units to increasingly larger systems. The complex mass assembly of haloes is recorded in the dark matter merger trees. After infall into larger systems, accreted satellite haloes will eventually merge with the parent halo through stripping and orbital decay from dynamical friction. Dark matter structure formation has been studied in numerous numerical works with (e.g. Lemson & Consortium 2006; Klypin et al. 2016) and without (e.g. Vogelsberger et al. 2014; Fattahi et al. 2016) the baryonic component. Dark matter-only simulations have been performed on very large scales, with cosmological volumes up to several Gpc³ (Potter, Stadel & Teyssier 2017). In comparison, due to the computational complexities introduced by hydrodynamics, simulations with baryons have been limited to orders of magnitude smaller volumes, around a hundred Mpc³. The discrepancy in simulated volumes inevitably leads to structures with intrinsically lower number densities limiting their usefulness in studying less common but important galaxy populations, such as massive satellite galaxies residing in relatively small parent haloes. Hydrodynamical simulations have resolution and time constraints imposed by the computational power they can leverage, ultimately being forced to

turn to sub-grid analytic physical recipes below the particle or mesh resolution.

Semi-analytic models (e.g. Granato et al. 2004; Baugh 2006; Monaco, Fontanot & Taffoni 2007; Hirschmann et al. 2012; Shankar et al. 2013) make comprehensive predictions on the evolutionary patterns of galaxies by following the growth and assembly along dark matter merger trees. Semi-analytic models are necessarily characterized by a number of input assumptions and free parameters associated with different formation mechanisms. Semi-analytic models make the admirable effort towards a holistic view of galaxy formation, but they are known to suffer from degeneracies in assumptions and related parameters (e.g. González et al. 2011; Lapi & Cavaliere 2011), and could still be limited by volume effects.

In the past decades, much attention has been devoted to probing the connection between dark matter and galaxies in a more empirical fashion, attempting to build a self-consistent picture of dark matter/galaxy assembly and distribution making use of more observationally driven methods. Halo occupation distribution models were introduced as a first step in this direction (e.g. Berlind & Weinberg 2002; Cooray & Sheth 2002). These models are designed to predict the average number of central and satellite galaxies in haloes by simultaneously matching observed number densities and clustering at a given epoch. Halo occupation models by construction do not account for the full assembly history of the galaxies and may be susceptible to assembly bias effects (e.g. Dalal et al. 2008; Zentner, Hearin & van den Bosch 2014).

* E-mail: p.grylls@soton.ac.uk

Abundance matching between the galaxy luminosity/stellar mass function (SMF) and the halo mass function (HMF) is a semi-empirical technique that has been widely adopted in the last two decades to constrain, on a statistical basis, the mean relation between stellar mass and host halo¹ mass (e.g. Kravtsov et al. 2004; Vale & Ostriker 2004; Yang et al. 2004; Shankar et al. 2006). In its simplest form, abundance matching relies on the basic idea that larger galaxies reside in larger dark matter haloes, and relative abundances are used to create a (mean) stellar mass–halo mass (SMHM) mapping $M_* = \text{SMHM}(M_h, z)$ at any relevant epoch (Yang, Mo & Van den Bosch 2003; Behroozi, Conroy & Wechsler 2010; Moster et al. 2010). This technique has been used as a flexible tool to predict the full mass and structural assembly history of galaxies along dark matter merger trees in a full cosmological context.

Semi-empirical models (e.g. Hopkins et al. 2009b; Cattaneo et al. 2011; Zavala et al. 2012; Shankar et al. 2014a) were conceived as models where galaxies are assigned to haloes at any epoch along dark matter merger trees via abundance matching relations, thus allowing to track the full stellar mass evolution in, for example, the main progenitor branch of the merger tree. Semi-empirical models, although unavoidably more restrained in scope, avoid the often heavy parametrizations incorporated in semi-analytic models, as galaxy stellar masses and other properties (e.g. disc sizes) are empirically assigned to dark matter haloes. In a semi-empirical approach, a minimum number of free parameters are included in the models allowing exploration of specific aspects of galaxy formation minimizing the danger of degeneracies. Despite their flexibility, traditional semi-empirical models can be subject to volume limitation effects due to the finite number of merger trees and/or dark matter haloes. Both semi-analytic and semi-empirical models can therefore suffer from low number statistics especially at high stellar masses when built on top of dark matter simulations.

In this work, we present STEEL, a STastical sEmi-Empirical modeL, built using a novel, fast and flexible methodology to probe the number densities and average properties of galaxies at any given stellar mass, epoch and environment, with negligible computational limitation on volume or mass resolution. This is particularly useful when aiming at simultaneously probing, as in this work, the distribution of massive centrals as well as massive satellites in lower mass haloes. The method employed in this paper starts from tracking backwards in time the mean accretion histories of small bins in halo mass, weighted by the HMF. At each time-step, satellite galaxy number densities are associated with the main halo mass bin via the corresponding time growth in subhalo mass function (SHMF). The main aim of this work is to predict satellite number densities as a function of stellar mass and halo mass. We will show that the dominant parameter controlling the abundance of massive satellites is the dynamical friction time-scale. Nevertheless, the evolution of satellite galaxies, as central ones, is of course a result of a combination of star formation, quenching, and stellar stripping. We will probe the (limited) impact on our results of each one of these processes.

In what follows, we adopt the Planck cosmology with $(\Omega_m, \Omega_\Lambda, \Omega_b, h, n, \sigma_8) = (0.31, 0.69, 0.05, 0.68, 0.97, 0.82)$.² We use M_h and

¹Here and throughout this work, ‘host halo’ refers to the dark matter (sub)halo associated with a given galaxy residing at the (sub)halo’s centre.

²We note that Planck’s best-fitting cosmology has slightly different parameters than those adopted in some of the observational probes included in this work, such as the SMFs ($\Omega_m, h = (0.30, 0.70)$). However, we have checked that correcting the SMF volumes and luminosities to the same Planck’s cosmology yields essentially indistinguishable results in the implied SMHM

M_* to differentiate between halo and stellar mass, respectively; we name $M_{h,\text{cent}}$ and $M_{h,\text{sat}}$ central and satellite, respectively (we direct readers to Appendix C for a complete glossary of the acronyms used throughout this work). Halo masses are defined as virial masses, unless stated otherwise. Wherever relevant, we always assume a Chabrier (2003) stellar initial mass function.

2 DATA

In this work, we take as a reference the Sloan Digital Sky Survey Data Release 7 (SDSS-DR7) from Meert, Vikram & Bernardi (2015, 2016), with improved galaxy photometry. In brief, the data are originally from the SDSS-DR7 spectroscopic sample (Abazajian et al. 2009) containing $\sim 670\,000$ galaxies fitted with a Sérsic+exponential model (PyMorph; Meert et al. 2015, and references therein), with associated halo masses (Yang et al. 2012) updated to the new photometry. The improved photometry provides a number of advantages. For instance, the two-dimensional fit more accurately captures the high-mass end of the SMF (Bernardi et al. 2013, 2017a,b) and uses the mass-to-light ratio from Mendel et al. (2014).

At higher redshifts ($0.3 < z < 3.3$), SMFs are from the COSMOS2015 catalogue (Davidzon et al. 2017). In this survey, masses are derived from multiwavelength spectral energy distribution fitting, including ultra-deep and infrared photometry. We note Davidzon et al. (2017) use Bruzual & Charlot (2003) stellar population synthesis models that are consistent within $\lesssim 0.15$ dex with the mass-to-light ratios adopted in the SDSS catalogue. Both data sets assume a Chabrier (2003) stellar initial mass function.

3 METHODS: A NEW ‘STATISTICAL’ APPROACH TO SEMI-EMPIRICAL MODELLING

Semi-analytic models continue to provide invaluable constraints on the formation and evolution of galaxies. They admirably attempt to analytically model from zero order the full evolution of galaxies since the Big Bang to the local Universe (e.g. Menci et al. 2006). This inevitably requires specific input assumptions and related parameters on, for example, the rate of star formation at different epochs and/or environments.

Semi-empirical models are instead characterized by a ‘bottom-up’ approach, where the least possible assumptions and associated parameters are initially included in the models. Gradually, additional degrees of complexities can be included in the model, wherever needed. Semi-empirical models in this respect are a very powerful complementary tool to semi-analytic models or even cosmological hydrodynamic simulations. Their goal is in fact to provide unique constraints to only specific aspects of galaxy evolution, such as predicting, as in this work, the distribution of satellites in a given dark matter parent halo.³

Semi-empirical models (Hopkins et al. 2009b; Cattaneo et al. 2011; Zavala et al. 2012; Shankar et al. 2014a) are broadly characterized by the following steps:

relation. We will thus ignore such tiny differences in cosmological parameters in what follows. We choose to retain the halo cosmology as the full dark matter backbone of the model presented in this work is built and calibrated against the Planck Cosmology (Planck Collaboration et al. 2016).

³Here and throughout this work, ‘parent halo’ refers to the central halo containing subhaloes/satellite galaxies.

(i) Central galaxies, associated with the main progenitor branch⁴ of the host dark merger tree, are initially defined as gas-rich with disc like profiles at early epochs, as supported by deep *HST* observations in CANDELS (e.g. Huertas-Company et al. (2015, 2016)). All their stellar mass and structural properties are assigned via empirical, time-dependent relations, more significantly the SMHM relation (e.g. Moster, Naab & White 2013).

(ii) Central galaxies are then re-initialized at each time-step during the evolution and can gradually transform their morphology via ‘in situ’ processes, such as (more or less) violent disc instabilities, and/or ‘ex situ’ processes, such as mergers.

(iii) Satellite galaxies are those associated with each dark matter branch merging to the main progenitor. They are assigned all the mass and structural properties of a central galaxy in a typical central halo of the same mass at the time of infall. Satellites may further evolve in stellar mass, given their specific star formation rate (SFR) at infall, possibly accompanied by some stellar stripping.

The typical semi-empirical model described above is a powerful and very competitive tool to explore trends in the expected mass and/or structural evolution of galaxies in a given interval of stellar mass for a given set of parameters in, e.g., amount of stripping and/or orbital energy (Shankar et al. 2015). Based on, by design, closely following the evolution of each dark matter merger tree, the basic semi-empirical approach discussed above is still limited by the resolution and volume of the background dark matter simulation. This could be particularly restraining and computationally expensive especially when modelling very massive galaxies and their accompanying large volumes.

In this paper, we propose a statistically based semi-empirical approach. We explore a large range of stellar masses, with virtually no limitation in sample size, and still retain the high degree of flexibility characteristic of a semi-empirical model. The core steps in building the dark matter backbone of STEEL are summarized as follows (see diagram in Fig. 1):

(i) At the redshift of interest \bar{z} , we start from the HMF to compute the abundances of (parent/central) haloes (panel A).

(ii) Each parent/central halo mass is then followed backwards in time following its average mass growth history, $\langle M_{\text{halo}}(\bar{z}) \rangle$, as expected from extended Press–Schechter (Press & Schechter 1974) and/or numerical simulations (panel B).

(iii) At each time-step, we then calculate the difference in subhalo population between z and $z + dz$ to estimate the expected average number density and masses of subhaloes accreted on to the main progenitor in the redshift interval dz (panel D).

We will show how using analytic HMFs and a ‘statistical accretion history’, in place of halo catalogues and merger trees, we are able to simultaneously examine massive central haloes, and massive satellites in lower mass parents. The latter are usually characterized by low number densities and can be omitted in more traditional approaches that rely on cosmological volumes. For example, massive satellite galaxies with stellar mass $M_* > 10^{11} M_\odot$ residing in relatively smaller parent haloes of mass $M_h \sim 10^{12.75} M_\odot$ have predicted number densities of $\sim 10^{-6} \text{ Mpc}^{-3}$ (see Fig. 11, panel 3), which would correspond to ~ 1 galaxy in a volume of ~ 100

⁴The main progenitor is defined in traditional merger trees as the largest progenitor halo from any given halo merger, usually the largest halo at any time-step. In this work shown in panel B of Fig. 1, we simply have the average growth of this halo but it is convenient to visualize in the traditional sense.

(Mpc/h)³, a typical volume adopted in state-of-the-art cosmological hydrodynamical simulations. On the other hand, even in larger dark matter-only simulations such as the Bolshoi [250 (Mpc/h)³] (Klypin et al. 2016), one would expect at the most ~ 16 of these galaxies. As massive satellite are so rare in other simulations, a robust statistical determination of abundances is therefore vital to fully constrain the connection between galaxy formation and hierarchical assembly predicted by Λ CDM cosmology at all mass ranges (Neistein et al. 2013).

In this section, we detail each step of this statistically based semi-empirical approach. We start by outlining in Section 3.1 the analytic model of dark matter accretion histories characterizing the mean growth of the main progenitor halo and accreted subhaloes including subhalo merging times. We then discuss in Section 3.2 how we populate dark matter haloes with galaxies, using the SMHM relation for central galaxies and its evolution with redshift. Finally, in Section 3.3, we discuss the evolutionary processes that affect the satellite galaxies whilst they reside in their dark matter parent haloes. The parameters used in this section and their origins are given in Table 1.

3.1 Statistical dark matter accretion history

The core principle of our methodology is to treat parent haloes, and satellites galaxies/haloes, as ‘average’ populations avoiding issues with volume and resolution as described previously. Specifically, the goal of this paper is to reproduce the observed distribution of $z = 0.1$ satellite galaxies with mass $M_* > 10^9 M_\odot$. In this subsection, we detail step-by-step construction of the backbone of STEEL complemented with a graphic representation in Fig. 1.

3.1.1 Central haloes

We start by considering a fine grid of central dark matter haloes ranging from $M_h = 10^{11} M_\odot$ to $M_h = 10^{15} M_\odot$ at redshift $z = 0$. Their number densities are given by the HMF, which is obtained using HMFCalc (Murray, Power & Robotham 2013). HMFCalc includes a number of different mass functions. For this work, we use the HMF from Tinker et al. (2010).⁵ The HMF provides the number densities of haloes in a given mass bin (Fig. 1, panel A). The average mass growth histories of all main progenitors with mass in the bin of halo mass ($M_h, M_h + dM_h$) are then calculated using the analytic model from Bosch et al. (2014).⁶ This provides the average ‘main progenitor’ branch of a traditional merger tree for each mass bin at $z = 0$ (Fig. 1, panel B).

3.1.2 Assigning subhaloes to parent haloes

In order to predict the number of satellite galaxies, we must associate to each parent/central halo the number and mass of subhaloes they are expected to contain. To achieve this, we use the SHMF. The SHMF describes the expected distribution of subhaloes of mass $M_{h,\text{sat}}$, in a given parent halo of mass $M_{h,\text{cent}}$, as a function of $M_{h,\text{sat}}/M_{h,\text{cent}}$. Multiple definitions for the SHMF exist depending

⁵We adopt the COLOSSUS PYTHON package (Diemer 2017) for all halo mass conversions required.

⁶This model further improves on the seminal work by Parkinson, Cole & Helly (2008), which was aimed at reproducing numerical merger trees, optimized with small redshift steps minimizing the development of systematic errors at late cosmic epochs.

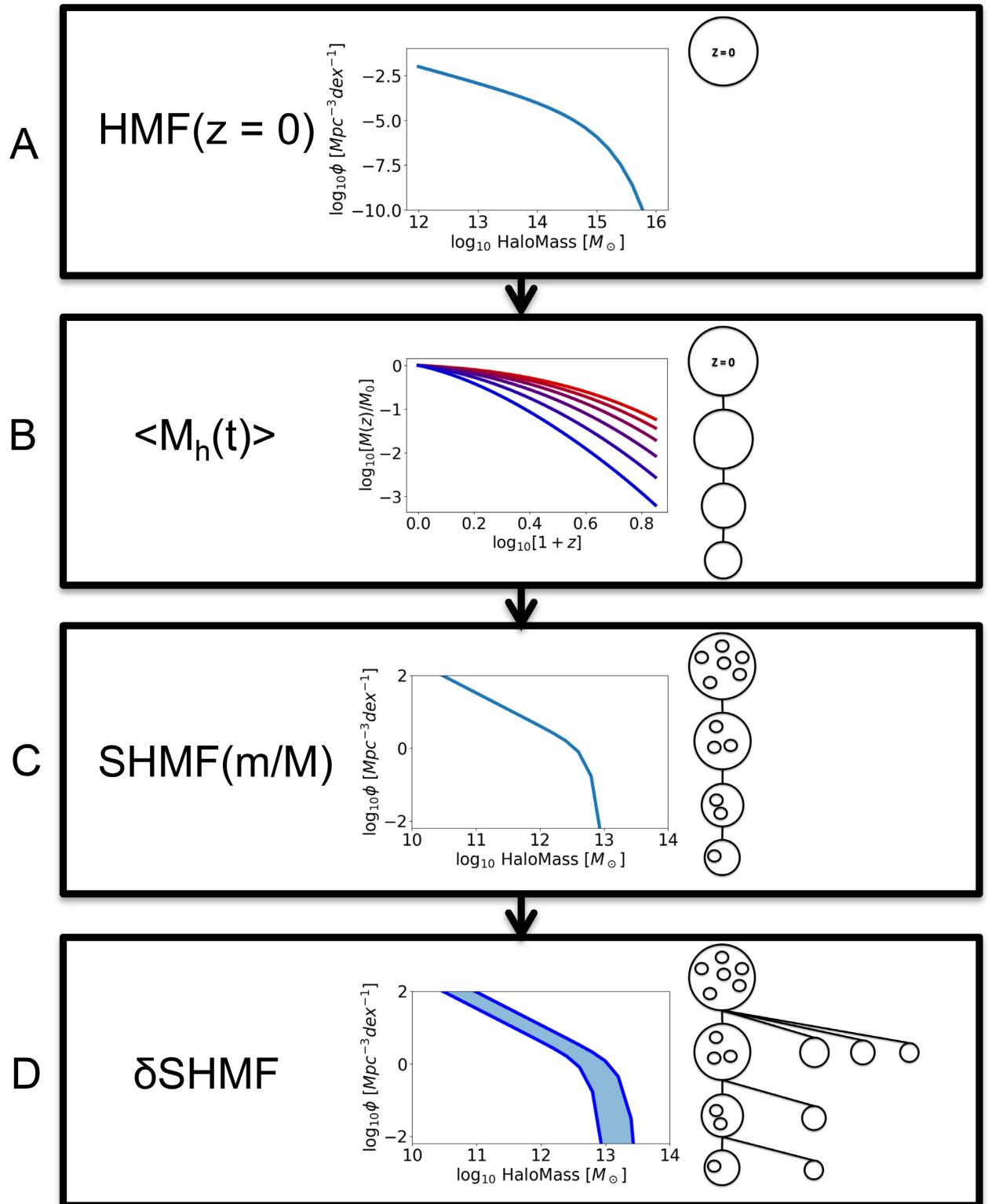


Figure 1. We show the main steps in building the statistical dark matter backbone of STEEL described in Section 3.1. Each panel shows the feature from a traditional merger tree and the statistical function used to replace it. (A) The HMF is used to calculate the number densities of central haloes. (B) Average mass growth histories calculate the size of each mass bin at previous epochs. (C) The SHMF is used to populate each central at each redshift with subhaloes. (D) The average number densities of accreted subhaloes at each epoch are calculated by comparing the SHMF over consecutive redshift steps.

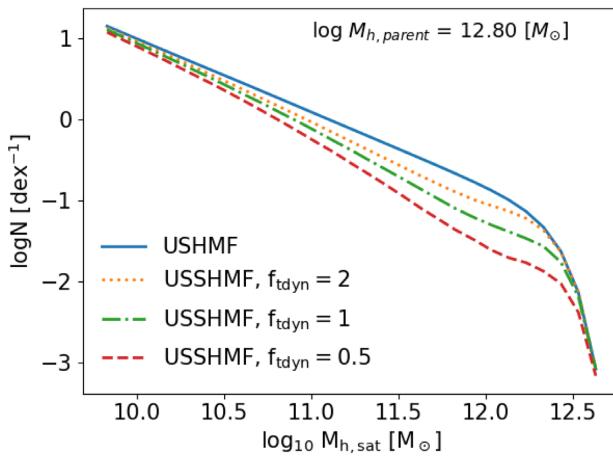


Figure 2. Comparison between the USHMF (solid line) and three USSHMF (dotted lines) for a parent halo of mass $\log M_{\text{h},\text{parent}} [\text{M}_\odot] = 12.80$. The factor f_{dyn} is applied to the merging time-scales of the haloes. Lower factors correspond to lower USSHMF, where more sub-haloes have merged.

on the way a subhalo is defined. In this work, we use two definitions of the SHMF. The first is the unevolved SHMF (USHMF), which describes the total subhaloes accreted over a parent halo's lifetime. In the USHMF, any merging or stripping in the subhaloes occurring after infall are ignored. A number of groups have been able to constrain the USHMF (Giocoli, Pieri & Tormen 2008; Jiang & van den Bosch 2016). In what follows, we use the latest rendition of the USHMF by Jiang & van den Bosch (2016), which is calibrated against the Bolshoi simulation.⁷

The second definition we use in this work is the unevolved ‘surviving’ SHMF (USSHMF). Subhalo masses are assumed frozen at infall, but the subhalo number densities can reduce compared to the USHMF as the USSHMF accounts for subhalo disappearance due to tidal disruption in the parent halo. We show in Fig. 2 for a representative parent halo of mass $\log M_{\text{h},\text{cent}} [\text{M}_\odot] = 12.80$, the USHMF and three USSHMF characterized by different dynamical friction time-scales τ_{dyn} , as described in Section 3.1.4. Larger τ_{dyn} lead to a milder reduction in subhalo number densities as subhaloes take longer to merge with the parent halo. Lower τ_{dyn} are less effective in reducing the number densities of smaller subhaloes that are more likely to have dynamical friction time-scales comparable to or larger than the Hubble time at $z = 0$.

3.1.3 Average subhalo accretion

At each redshift step along the mass growth histories, we calculate the USHMF associated with the parent halo mass. What we create is equivalent to the substructure found in traditional merger trees (Fig. 1, panel C). However, unlike traditional methods, our statistical approach is able to probe ‘rare’ subhaloes without running prohibitively large volumes of merger trees.

For each time-step, we can now calculate a mass function describing the number density of subhaloes accreted on to the population of central haloes in the halo mass bin $[M_{\text{h},\text{cent}}(z), M_{\text{h},\text{cent}}(z) + dM_{\text{h},\text{cent}}(z)]$. The latter is achieved by differentiating the USHMF

⁷We direct the interested reader to Jiang & van den Bosch (2016) for further discussion of the USHMF as well as of other SHMFs, such as the evolved SHMF where the number densities are effected by both subhalo stripping and mergers.

across two neighbouring redshift steps z and $z + dz$, we can calculate the average number density of subhaloes of any given mass $M_{\text{h},\text{sat}}$ that are accreted in the redshift interval dz on to the main progenitor haloes with mass in the bin $[M_{\text{h},\text{cent}}(z), M_{\text{h},\text{cent}}(z) + dM_{\text{h},\text{cent}}(z)]$,

$$\delta \text{USHMF}[z, M_{\text{h},\text{cent}}, M_{\text{h},\text{sat}}] = \text{USHMF}\left(\frac{M_{\text{h},\text{sat}}}{M_{\text{h},\text{cent}}(z)}\right) - \text{USHMF}\left(\frac{M_{\text{h},\text{sat}}}{M_{\text{h},\text{cent}}(z + dz)}\right) \quad (1)$$

In this way, the unevolved subhalo accretion history (δUSHMF) is retrieved for all main progenitor haloes at all redshifts.

3.1.4 Subhalo merging time-scale

From the unevolved subhalo accretion history, we need to isolate which subhalo mass bins survive at each epoch. The sum of all the surviving subhaloes (at each epoch) then yields the USSHMF. A key parameter used to calculate the USSHMF is the ‘observability time-scale’ (or survival time) of each subhalo mass bin $[M_{\text{h},\text{sat}}(z), M_{\text{h},\text{sat}}(z) + dM_{\text{h},\text{sat}}(z)]$ associated with a parent halo mass bin $[M_{\text{h},\text{cent}}(z), M_{\text{h},\text{cent}}(z) + dM_{\text{h},\text{cent}}(z)]$. This time-scale is equivalent to the merger time-scale τ_{merge} of a subhalo of mass $M_{\text{h},\text{sat}}$ in a parent halo mass $M_{\text{h},\text{cent}}$. To calculate τ_{merge} , we use the routines in equation (2), derived from N -body simulations (Boylan-Kolchin, Ma & Quataert 2008)

$$\tau_{\text{merge}} = (f_{\text{dyn}} \tau_{\text{dyn}})^{\frac{A(M_{\text{h},\text{cent}}/M_{\text{h},\text{sat}})^B}{\ln(1 + M_{\text{h},\text{cent}}/M_{\text{h},\text{sat}})}} \exp\left(C \frac{J}{J_c(E)}\right) \left(\frac{r_c(E)}{r_{\text{vir}}}\right)^D, \quad (2)$$

where $A = 0.9$, $B = 1.0$, $C = 0.6$, $D = 0.1$ (McCavana et al. 2012). The factor τ_{dyn} is given by (Jiang & van den Bosch 2016)

$$\tau_{\text{dyn}} = 1.628 h^{-1} \text{Gyr} \left(\frac{\Delta_{\text{vir}}(z)}{178}\right)^{-\frac{1}{2}} \left(\frac{H(z)}{H_0}\right)^{-1}. \quad (3)$$

Our method of considering average halo mass and accretion histories does not allow tracking single orbits and associated orbital energies. We assume instead an average orbit circularity of 0.5 (Khochfar & Burkert 2006), thus reducing the dependence on the angular momentum and radial components, $\frac{J}{J_c(E)}$ and $\frac{r_c(E)}{r_{\text{vir}}}$, to a constant. In other words, this approximation is consistent with the approach of taking the average expected orbits of subhaloes at fixed parent halo mass. The key parameter of our analysis is the factor f_{dyn} included in equation (2). The fudge factor f_{dyn} takes into account the systematic uncertainties induced by numerical resolution effects in N -body simulations that are unable to resolve the full merging time-scales of subhaloes and/or the satellite galaxies they host (Bosch et al. 2018). The parameter f_{dyn} increases or decreases the dynamical times of ‘merging’ satellites enabling an exploration of the effect of dynamical time on the final number density distributions of satellite galaxies at any given epoch.

3.1.5 Surviving subhalo population

At each redshift, we can now use the unevolved subhalo accretion history and the observability time-scale τ_{merge} to calculate the total observable subhalo population associated with any given parent halo mass bin $[M_{\text{h},\text{cent}}(z), M_{\text{h},\text{cent}}(z) + dM_{\text{h},\text{cent}}(z)]$, i.e. the USSHMF (shown by the dashed lines in Fig. 2). To compute the implied total number densities of unmerged subhaloes with mass $[M_{\text{h},\text{sat}}(z), M_{\text{h},\text{sat}}(z) + dM_{\text{h},\text{sat}}(z)]$ at any redshift of interest, we convolve the

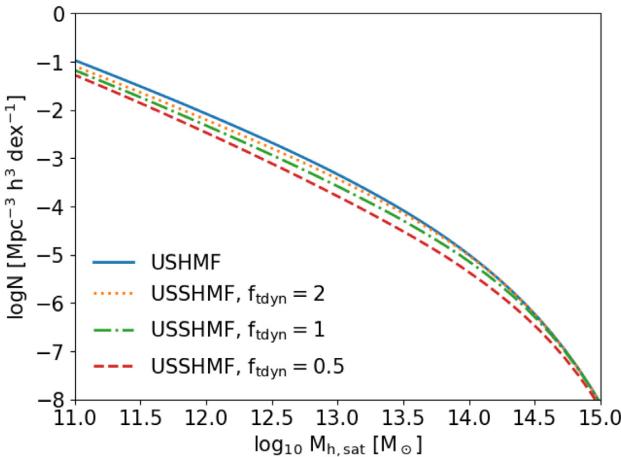


Figure 3. Example of the ‘total’ USHMF (solid line) and three ‘total’ USSHMF (dotted lines) corresponding to three different $f_{t\text{dyn}}$ factors.

USSHMF with the HMF,

$$N(M_{h,\text{sat}}, z) = \int \text{USSHMF}\left(\frac{M_{h,\text{sat}}}{M_{h,\text{cent}}}\right) \text{HMF}(M_{h,\text{cent}}, z) dM_{h,\text{cent}}. \quad (4)$$

Fig. 3 shows the total observable subhalo population for different $f_{t\text{dyn}}$ similar to the SHMF in Fig. 2. Furthermore, via appropriate abundance matching algorithms detailed below, we can assign corresponding satellite galaxies to the unevolved subhalo accretion history and obtain the distribution of satellites in a given parent halo mass bin $[M_{h,\text{cent}}(z), M_{h,\text{cent}}(z) + dM_{h,\text{cent}}(z)]$ by assuming the satellites follow the same merging time-scales as their host subhaloes. Finally, we convolve the latter distributions of satellite galaxies with the HMF to create the total observed satellite stellar mass function (SSMF). Both the distribution of satellite galaxies and the satellite SMF can then be used to constrain $f_{t\text{dyn}}$ as demonstrated in Section 4.

3.2 Abundance matching: Building the local populations of central and satellite galaxies

We now describe how we populate dark matter haloes with galaxies using the abundance matching technique. To compute the distribution of satellites in parent haloes, we populate subhaloes at infall, following the central SMHM relation. For the latter, we adopt a parametric form similar to that proposed by Moster et al. (2010),

$$\begin{aligned} M_*(M_h, z) &= 2M_h N(z) \left[\left(\frac{M_h}{M_n(z)} \right)^{-\beta(z)} + \left(\frac{M_h}{M_n(z)} \right)^{\gamma(z)} \right]^{-1}, \\ N(z) &= N_{0.1} + N_z \left(\frac{z - 0.1}{z + 1} \right), \\ M_n(z) &= M_{n,0.1} + M_{n,z} \left(\frac{z - 0.1}{z + 1} \right), \\ \beta(z) &= \beta_{0.1} + \beta_z \left(\frac{z - 0.1}{z + 1} \right), \\ \gamma(z) &= \gamma_{0.1} + \gamma_z \left(\frac{z - 0.1}{z + 1} \right). \end{aligned} \quad (5)$$

3.2.1 Local Universe $z = 0.1$

To constrain the $z = 0.1$ parameters ($M_n, N, \beta, \gamma, \sigma$), we construct a SMF from the HMF and the SMHM relation with initial parameters close to previous results (Moster et al. 2013; Shankar et al. 2014a; Buchan & Shankar 2016). The HMF provides the number density for each bin of parent halo mass in the range $[M_{h,\text{cent}}(z), M_{h,\text{cent}}(z) + dM_{h,\text{cent}}(z)]$. Using the SMHM relation, each bin of parent halo mass is associated with a Gaussian distribution of stellar masses with width controlled by the scatter parameter (σ). Multiplying by the halo mass number density, we convert this distribution into galaxy number densities, which are then added to the relevant stellar mass bins of the SMF in construction. This operation is repeated over each bin of the HMF. The resulting SMF is then compared to the SDSS central SMF. We vary the parameters on a grid and minimize root mean square between the constructed SMF and the data. In our abundance matching framework, we specifically distinguish between central and SSMFs by taking advantage of the SDSS galaxy-halo catalogue described in Section 2.

For comparison with previous works, we apply the same method as described above to the total galaxy/halo population using the USSHMF ($f_{t\text{dyn}} = 1.0$), the HMF, and the total SMF. Our total result uses single-epoch⁸ abundance matching: we apply the SMHM relation to the total HMF at $z = 0.1$ with no consideration of the infall time of subhaloes.

Table 2 reports the best-fitting parameters using the method above, for central galaxies and total galaxy population. In the left-hand panel of Fig. 4, we show with a solid blue line the $z = 0.1$ SMF for SDSS central galaxies only, abundance-matched with the central HMF (no subhaloes). For comparison we also include the abundance matching fits by Moster et al. (2013), Shankar et al. (2017b), and Behroozi et al. (2018, using the ‘true’ stellar masses, in table J1). The total SMF is reproduced by Moster et al. (2013) and Shankar et al. (2017b) without distinguishing between centrals and satellites galaxies in the SMF. We find our final results for the high mass slope to be steeper than the Moster et al. (2013) but very similar to the ones by Shankar et al. (2017b), as expected given that the latter have tuned their fitting routines to the same galaxy SMF as the one adopted in this paper. At the low mass end our SMHM relation tends to appear slightly steeper than previous works. At least with respect to Shankar et al. (2017b), this is mainly caused by the fact that our results are based on the updated surviving SHMF by Jiang & van den Bosch (2016) that is somewhat steeper than the subhalo mass correction suggested by Behroozi, Wechsler & Conroy (2013) and adopted by Shankar et al. (2017b). Our low-mass slope is, as expected, fully consistent with Behroozi et al. (2018) as their subhaloes have been extracted from the Bolshoi-Planck simulation (Klypin et al. 2016). The right-hand panel of Fig. 4 shows the total, central, and SSMFs as extracted from our SDSS catalogues (black circles, blue triangles, and red squares, respectively) compared with the central SMF generated using the central HMF and the relations from Moster et al. (2013), Behroozi et al. (2018), and our best-fitting relations. The solid black line shows our ‘total’ relationship where the full population of central and satellite galaxies are retrieved with single-epoch abundance matching.

From this comparison, we conclude that for the SMF and HMF used in this work single-epoch abundance matching is acceptable,

⁸A multiepoch approach considers the infall time of subhaloes, and routines are used for stripping, star formation, and quenching of satellite galaxies. The full set of parameters are fit simultaneously (e.g. Moster et al. 2013).

Table 1. Parameters used in the model and how they are constrained.

| Parameter | Description | Equation |
|-------------------------------------|---|----------|
| f_{dyn} | A factor we include to test the impact of shortening or lengthening dynamical times of all satellites. This is a free parameter we use to find the best fit to satellite distributions. | 3 |
| A, B, C, D | Free parameters | |
| M_n, N, β, γ | Dynamical friction fitting parameters given by McCavana et al. (2012) | 2 |
| $M_{n,z}, N_z, \beta_z, \gamma_z$ | Abundance matching parameters, constrained by the central SMF and HMF at $z = 0$ | 5 |
| s, M, α | Abundance matching evolution parameters, constrained using SMF and the HMF at $z > 0$. | 5 |
| $\tau_q, \tau_f, M_{\text{cutoff}}$ | Star formation parameters, constrained by observation given in Tomczak et al. (2016). We also give values to fit our continuity derived SFR in Section 3.3.1 | 6 |
| η_{strip} | Parameters describing the quenching time of satellites. These parameters are all adapted from models in the literature (Wetzel et al. 2013; Fillingham et al. 2016) | 8–10 |
| | Stripping fraction from Cattaneo et al. (2011) | 15 |

Table 2. Parameters for equation (5). The $z = 0.1$ parameters are constrained to 2 d.p. or 2 s.f., whichever is smaller. The $z > 0.1$ parameters are constrained to 1 d.p. or 1 s.f., whichever is smaller.

| | M_n | N | β | γ | σ |
|-----------|-------|-------|---------|----------|----------|
| Central | | | | | |
| $z = 0.1$ | 11.95 | 0.032 | 1.61 | 0.54 | 0.11 |
| Total | | | | | |
| $z = 0.1$ | 11.89 | 0.031 | 1.77 | 0.52 | 0.10 |
| Evolution | | | | | |
| $z > 0.1$ | 0.4 | −0.02 | −0.6 | −0.1 | N/A |

as there is little evolution at late times. However, if the HMF or SMF were to be strongly evolving at low redshift, this approach would yield a result inconsistent with matching the central population. We further argue firmer constraints are set using the central SMHM relation, where there is higher confidence in the populations being matched.

3.2.2 Redshift evolution $z > 0.1$

We assume that at any redshift satellite galaxies at infall follow the same scaling relations as centrals of equal parent halo mass but then may evolve independently afterwards. Therefore, we require an SMHM relation calibrated at all redshifts to initialize satellites in subhaloes at infall.

Keeping fixed the $z = 0.1$ parameters, the redshift-dependent parameters in equation (5) are fit with a similar iterative method. A complication arises as the high redshift data do not distinguish central and satellite SMF. We therefore use a multiepoch⁸ approach without satellite evolution to fit the evolution parameters. We continue to generate central SMF as above then add a satellite SMF and compare to the total SMF from the data minimizing the sum of the root mean squares over eight redshift steps. We generate the satellite SMF⁹ from the model for each set of parameters in our parameter space. In Appendix A, we show the result of our iterative method applied to the Davidzon et al. (2017) and SDSS SMFs.

It is important to realize at this stage that, as demonstrated in Fig. 5, the vast majority (>60–80 per cent) of $z = 0.1$ satellites are accreted below $z < 1.0$ irrespective of the exact dynamical time-

⁹Under the assumption that satellites are ‘frozen’ in stellar mass after infall. We will discuss the impact of relaxing this assumption in Section 3.3.

scale adopted.¹⁰ In this redshift range, the evolution of the SMHM relation has been demonstrated to be rather weak (e.g. Moster et al. 2013; Behroozi et al. 2013; Shankar et al. 2014a; Bernardi et al. 2016, and references thereof). Thus, we expect that our results are robust against possible variations in the redshift dependence of the SMHM relation used to initialize infalling satellites, as demonstrated below.

We will now focus on specifically reproducing the local SSMF (red squares in the right-hand panel of Fig. 4).

3.3 Satellite evolution after infall

After infall, satellites are expected to evolve their stellar mass and structure through several processes such as star formation, stellar stripping, and quenching. In this section, we briefly describe each of these processes and the way we have modelled them in this work. We will then discuss in Section 4.2 how these processes may impact, if at all, the predicted local abundances of satellite galaxies.

3.3.1 Star formation rates

We use the SFR parametrization from Lee et al. (2015) with parameters¹¹ from Tomczak et al. (2016), where s_0 and M_0 have units $\log(M_\odot)$ and M_\odot , respectively,

$$\begin{aligned} \log[\psi(z, M_*)] &= s_0(z) - \log \left[1 + \left(\frac{M_*}{M_0(z)} \right)^{-\alpha(z)} \right], \\ s_0(z) &= 0.195 + 1.157z - 0.143(z^2), \\ \log[M_0(z)] &= 9.244 + 0.753z - 0.090(z^2), \\ \alpha(z) &= Y1.118. \end{aligned} \quad (6)$$

As discussed below, we will also need to explore different recipes for SFRs. To this purpose, following a continuity-equation approach (Leja et al. 2015), we have refitted the parameters in equation (6) to match the SFR distributions expected from self-consistently growing the SSMF in our models (for further details on this, see Ap-

¹⁰We note this does not imply an absence of satellite galaxies at higher redshift, only that these satellites at higher redshift mostly do not survive to be observed at $z = 0.1$.

¹¹These parameters are derived by fitting data from ZFORGE in combination with far-IR imaging from *Spitzer* and *Herschel* in the range of $0.5 < z < 4$. In this work, we extrapolate their fits down to $z = 0$, as this is consistent with the SFR measured by Salim et al. (2007) at lower redshifts.

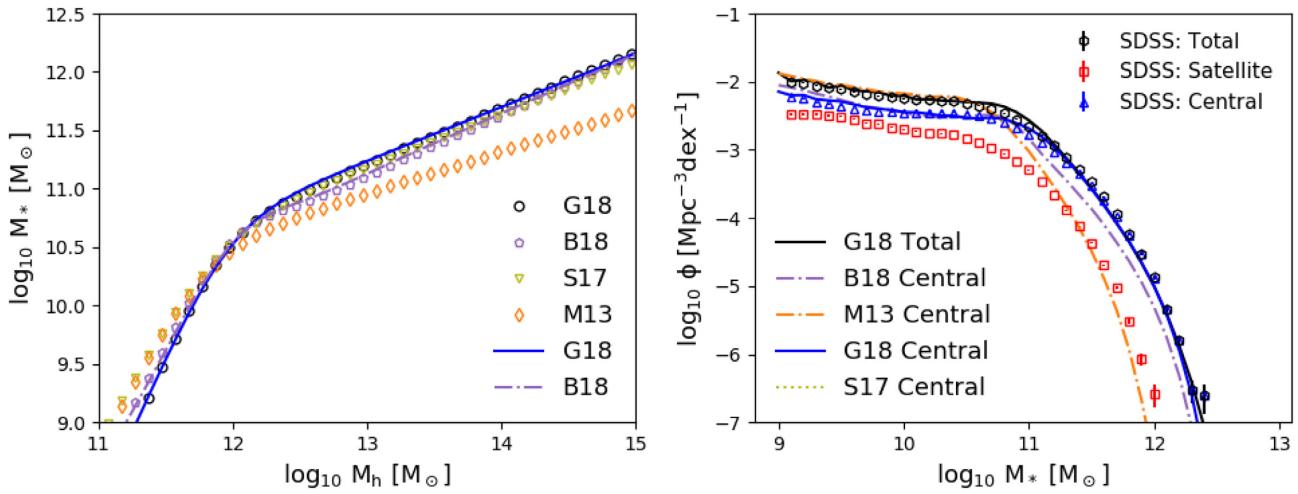


Figure 4. Left-hand panel: The SMHM relation generated from this work (G18, black circles), Behroozi et al. (2018, B18), Shankar et al. (2017a, S17) and Moster et al. (2013, M13). Relations derived from the total SMF/HMF are shown with open symbols, whilst relations derived with the central SMF/HMF are shown as solid lines. Right-hand panel: Central SMF generated from the models using a central only fit, we also show the total SMF from this work (solid black line) from a single-epoch fit. Open symbols show the SDSS central (squares), satellite (triangles), and Total (hexagons) SMFs.

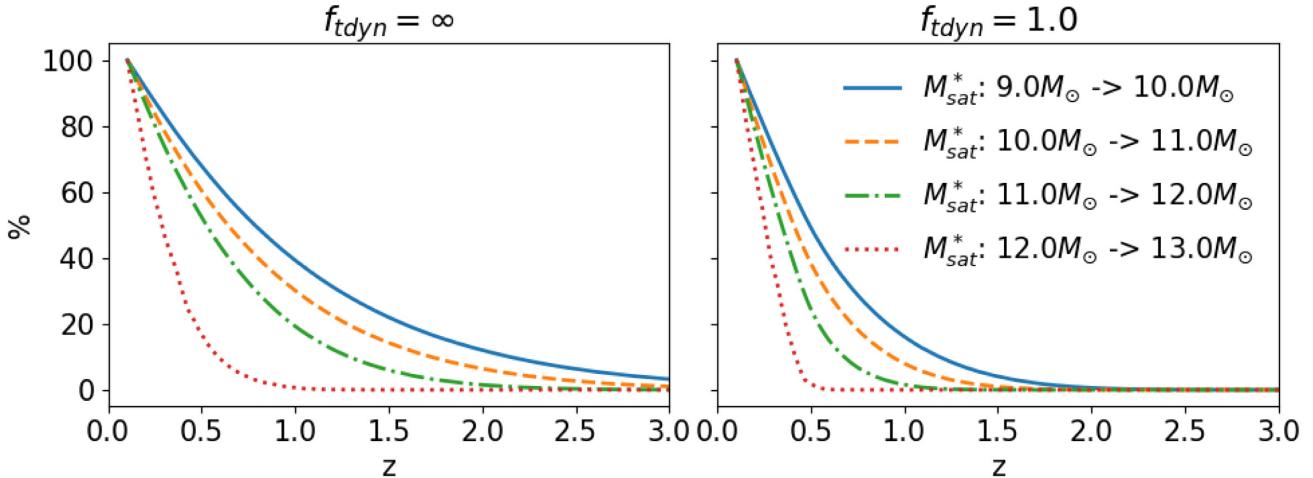


Figure 5. We here show the percentage of satellites observed at a $z = 0.1$ as a function of their redshift of accretion $z > 0$. It can be seen that massive satellites observed at $z = 0.1$ are accreted more recently than smaller satellites. At $z = 0.5$ less than 50 per cent of the total satellites observed at $z = 0.1$ have been accreted, and at $z = 0.1$ this falls to less than 20 per cent.

pendix B). The fit to the resulting SFRs is given by

$$\begin{aligned} \log(\psi(z, M_*)) &= s_0(z) - \log \left[1 + \left(\frac{M_*}{M_0(z)} \right)^{-\alpha(z)} \right], \\ s_0(z) &= 0.6 + 1.22z - 0.2(z^2), \\ \log(M_0(z)) &= 10.3 + 0.753z - 0.15(z^2), \\ \alpha(z) &= 1.3 - 0.1z. \end{aligned} \quad (7)$$

In all cases, the SFR is included in our models with a lognormal scatter of 0.3 dex (Leja et al. 2015).

3.3.2 Quenching

We also include the ability to quench star formation in satellite galaxies after infall. It has been suggested that satellites undergo a ‘delayed-then-rapid’ quenching (Wetzel et al. 2013). The latter model envisions that satellites continue to form stars at the same rate as central galaxies of comparable stellar mass for a time τ_q after

infall, and then quench rapidly over a time-scale τ_f . This quenching is proposed for satellites with stellar mass above $M_* \gtrsim 10^9 M_\odot$, with a minimum $\tau_q \sim 1$ Gyr. For galaxies below $M_* \lesssim 10^9 M_\odot$, we however adopt the more recent results by Fillingham et al. (2016), who put forwards a parent halo mass dependent cut-off,

$$\log(M_{\text{cutoff}}) = 9\log(M_\odot) - (15\log(M_\odot) - \log(M_{\text{h,host}}))/5\log(M_\odot), \quad (8)$$

below which satellite galaxies all share the same quenching time $\tau_q = 2$ Gyr. The rapid quenching time-scale τ_f can be expressed as (Wetzel et al. 2013)

$$\tau_f = -0.5\log(M_{*,\text{sat}}) + 5.7 \text{ Gyr}. \quad (9)$$

We set a minimum τ_f of 0.2 Gyr for all galaxies. This rapid quenching begins at times $t > \tau_q$ after infall, and it is approximated by an exponential decay that is longer for larger satellites, as can be inferred from equation (10). The lookback time at which a galaxy begins fast quenching is then $t_q = t_{\text{infall}} - \tau_q$. After this time the

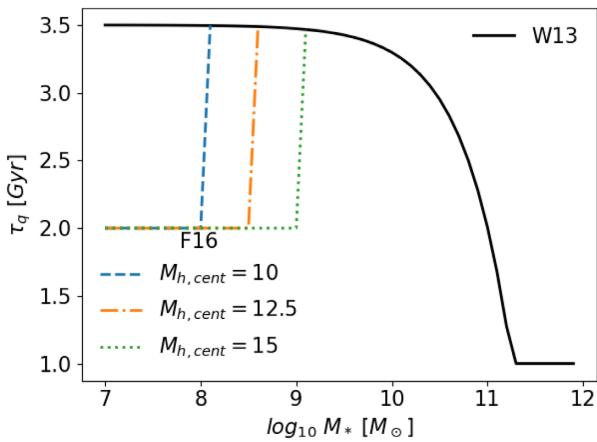


Figure 6. The solid line shows the Wetzel et al. (2013, W13) model for quenching. The dashed lines show the host halo-dependent reduction in quenching time from Fillingham et al. (2016, F16) for three example host masses $\log_{10} M_{\text{h,cent}} = 10, 12.5, 15$ as labelled. Larger hosts are able to reduce the quenching time of larger satellites.

satellite no longer follows the SFR of a typical central galaxy. Fig. 6 illustrates the quenching model, where the dashed/coloured lines demonstrate the halo mass dependence in cut-off mass.

The SFR during the satellite infall is then given by

$$\text{SFR}(t, M_*) = \text{SFR}(t, M_*) \begin{cases} \psi(z(t), M_*), & t > t_q \\ \psi(z(t_q), M_*) e^{-\left[\frac{t_q-t}{\tau_f}\right]}, & t < t_q \end{cases} \quad (10)$$

If at any point a satellite galaxy has an SSFR below $10^{-12} M_\odot \text{ yr}^{-1}$, it is assumed to be fully quenched and assigned an SSFR of $10^{-12} M_\odot \text{ yr}^{-1}$, plus a lognormal scatter of 0.3 dex.

3.3.3 Mass recycling

During infall some of the new stellar mass formed may be returned to the interstellar medium after stellar death. To incorporate this mass recycling process, we use the fractional mass-loss as parametrized in Moster, Naab & White (2018)

$$f(\tau_{\text{ml}}) = 0.05 \ln \left(\frac{\tau_{\text{ml}}}{1.4 \text{ Myr}} + 1 \right). \quad (11)$$

It follows that the mass-loss rate (MLR) at a given time is dependent on the star formation history (SFH), and it is calculated as

$$\text{MLR}(t) = \frac{\sum_{t'=t_{\text{infall}}}^t \text{SFH}(t')(f[t' - (t - \delta t)] - f[t' - t])}{\delta t}. \quad (12)$$

In equation (12), the SFH is the mass of stars formed in each previous time-step. The actual stellar mass growth $\dot{M}_{*,\text{sat}}$ is then given by the difference between the SFR and the MLR

$$\dot{M}_{*,\text{sat}}(t) = \text{SFR}(M_{*,\text{sat}}, t) - \text{MLR}(t). \quad (13)$$

3.3.4 Stripping

The stellar stripping is implemented following the empirically based formalism suggested by Cattaneo et al. (2011). Satellite galaxies strip stellar mass proportionally to the ratio of the host subhalo and the parent halo

$$M_{*,\text{sat}} = M_{*,\text{sat}}(1 - \eta_{\text{strip}})^{\tau_{\text{strip}}}, \quad (14)$$

where τ_{strip} is the dynamical friction time-scale in orbital time. For our model, where the orbital circularity is assumed to take the average value, this becomes the dynamical friction time (Chandrasekhar 1943) times a constant given by Jiang et al. (2008) using the average orbital circularity. Following Cattaneo et al. (2011), the latter can be written

$$\tau_{\text{strip}} = \frac{1.428}{2\pi} \frac{M_{\text{h,cent}}/M_{\text{h,sat}}}{\ln(1 + M_{\text{h,cent}}/M_{\text{h,sat}})}. \quad (15)$$

We set as a reference $\eta_{\text{strip}} = 0.4$ as suggested by Cattaneo et al. (2011) who showed that larger values would be inconsistent with the observational constraints in the local Universe. If at the time of observation, say $z = 0$, a galaxy's full dynamical time is not yet passed, a time-dependent reduction factor is applied to the amount of total stellar stripping given by equation (15).

4 NUMBER DENSITIES AND DISTRIBUTIONS OF MASSIVE SATELLITES

Section 3 has laid out our methodology, which is essentially based on progressively building up the (surviving) satellite number densities over cosmic time based on the input dynamical merger time-scales. Our first goal is now to predict the local ($z = 0.1$) satellite SMF and the distribution of massive satellites as a function of parent halo mass, which we will compare with our refined SDSS galaxy-halo catalogue introduced in Section 2. When computing the distribution of satellites as a function of parent halo mass we will show both full number densities, as well as fractional distributions to better highlight the ‘skewness’ of the predicted distributions with respect to the data. The latter will be simply computed as

$$F(dM_h) = \frac{N(> 10^x)|_{dM_h}}{N(x)}, \quad (16)$$

where $N(> 10^x)$ is the total number (density) of satellites above a threshold stellar mass $x = \log M_*$, and the $N(> 10^x)|_{dM_h}$ is the number of these that reside within the halo mass bin $[M_h, M_h + dM_h]$.

The results discussed in the next sections rely heavily on the concept of mass ratio affecting the merging time-scales of subhaloes and/or satellite galaxies. To ease the discussion of our findings, we provide Fig. 7 as visual guide to how merging times depend on subhalo/satellite masses in three different parent halo masses. In the left-hand panel, we plot dynamical time against subhalo mass for a subhalo accreted at $z = 1.5$. The black dotted line is the time (in Gyr) to $z = 0$. If a subhalo's τ_{merge} is below this value, it will not be observed at redshift $z = 0$. Similarly, for the right-hand panel, we plot the same relation converting the subhalo masses to satellite galaxy masses using the SMHM relation at $z = 1.5$. In both panels, the merging time is increased when the ratio of mass between the subhalo/satellite and the central (parent) halo is smaller.

In Section 4.1, we will first focus on the frozen model. In Section 4.1.1, we will motivate the need for a finite dynamical time and an evolving SMHM relation, and then in Section 4.1.2 set constraints on the dynamical times that well reproduce the number and distribution of SDSS satellites as a function of parent host halo mass. In Section 4.2, we will probe the impact of SFR, stripping and quenching in the predicted satellite distribution in the local Universe.

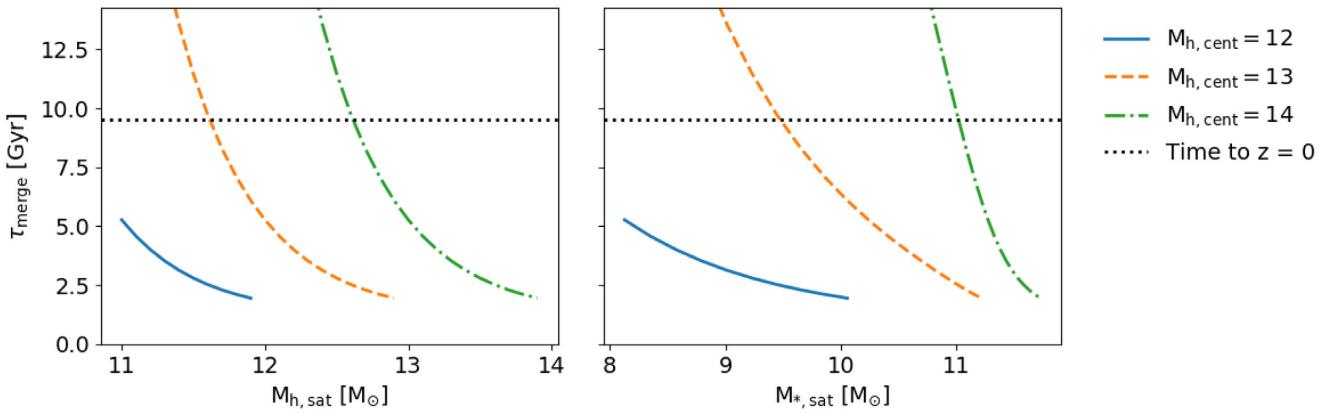


Figure 7. Range of merging time-scales for a range of subhalo (left) and satellite (right) masses when accreted on to three different host masses: $\log M_{h,\text{cen}} [\text{M}_\odot] = 12$ (blue solid), 13 (orange dashed), and 14 (green dot-dashed). The dotted black line shows the time to redshift $z = 0$, i.e. the minimum amount of time a satellite would need to survive to be observable in the local Universe.

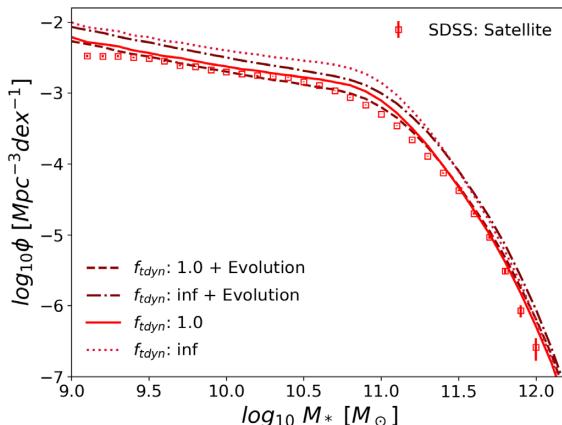


Figure 8. SSMFs generated by the model compared to SDSS data (open squares). The dashed and solid lines refer to models with a reference $f_{\text{dyn}} = 1.0$ (as suggested by N-body simulations), respectively, with and without redshift evolution in the input SMHM relation. The dot-dashed and dotted lines have infinite f_{dyn} (i.e. no mergers) with and without redshift evolution in the input SMHM relation.

4.1 Frozen model

We begin by showing the model predictions of satellite number densities and distributions in parent halo mass in a frozen model, where satellites do not evolve in stellar mass after infall.

4.1.1 Redshift evolution and mergers

We start by exploring a basic model where τ_{dyn} is infinitely long (no satellite mergers) and the SMHM relation does not evolve in cosmic time, i.e. keeping the same $z = 0$ parameters at all redshifts. We choose to start with the simplest assumption of an unevolving SMHM, as also suggested by some empirical evidence at least up to $z \sim 1$ (e.g. Shankar et al. 2006, 2014b; Moster et al. 2013; Buchan & Shankar 2016; Tinker et al. 2017; Guo et al. 2018). We then compare the outputs of this model with those from models in which we either allow for galaxy–satellite mergers (f_{dyn} is finite), and/or the SMHM relation evolves with redshift. Fig. 8 shows how the four combinations of these assumptions affect the satellite SMF. For both the models characterized by an infinite dynamical time ($f_{\text{dyn}} = \infty$, dot-dashed and dotted lines), the number densities

of satellites are too high compared to the observed satellite SMF (red squares). When setting $f_{\text{dyn}} = 1.0$, i.e. the merger time-scales are exactly those parametrized by McCavanaugh et al. (2012), both the evolving and unevolving SMHM models produce very similar number densities and well consistent with the observed satellite SMF. The model with an evolving SMHM relation is particularly successful.

We now test in Fig. 9 the four models against the SDSS distributions of the satellites in groups and clusters. The top row shows the comoving number density per halo mass bin. The bottom row shows the fractional distributions (equation 16) that better highlights the differences induced by varying model inputs. In both rows, the SDSS data are shown as grey bands. As expected, the largest difference is seen between the models with $f_{\text{dyn}} = \infty$ and $f_{\text{dyn}} = 1.0$. When τ_{dyn} is infinite,¹² the distribution of satellites skews such that high-mass satellites are found more frequently in relatively lower mass parent haloes. We explain the origin of this effect when discussing dynamical time in more detail in Section 4.1.2. The evolution¹³ of the SMHM relation has a low impact on the number densities. This is in part due to the evolution of the SMHM relation being weak especially in the high-mass end, and in part due to most (>60–80 per cent $z = 0.1$) satellites being accreted at relatively low redshift ($z < 1.0$). We thus support the view that local satellite distributions are largely independent of the specific input of the evolution parameters of the SMHM relation.

4.1.2 Dynamical time

In this section, we show the predictions of models with different merging time-scales, $f_{\text{dyn}} = 0.5, 1.0, 2.5$, to probe the effects of dynamical time on the satellite population. In Fig. 10, we find, as expected, that longer dynamical times tend to increase satellite number densities especially towards lower stellar masses. The number densities of higher-mass satellites are more resilient to increase with dynamical time. In fact, when $f_{\text{dyn}} \gtrsim 1$ the number densities of massive satellites become already very close to the theoretical maximum number density (Fig. 8) set by $f_{\text{dyn}} = \infty$. It follows that massive satellites are on average a recently accreted population. In

¹²In PYTHON in this case, we set this variable using ‘numpy.inf’.

¹³We direct readers to Appendix A for further discussion and motivation for the redshift evolution parameters used in this work.

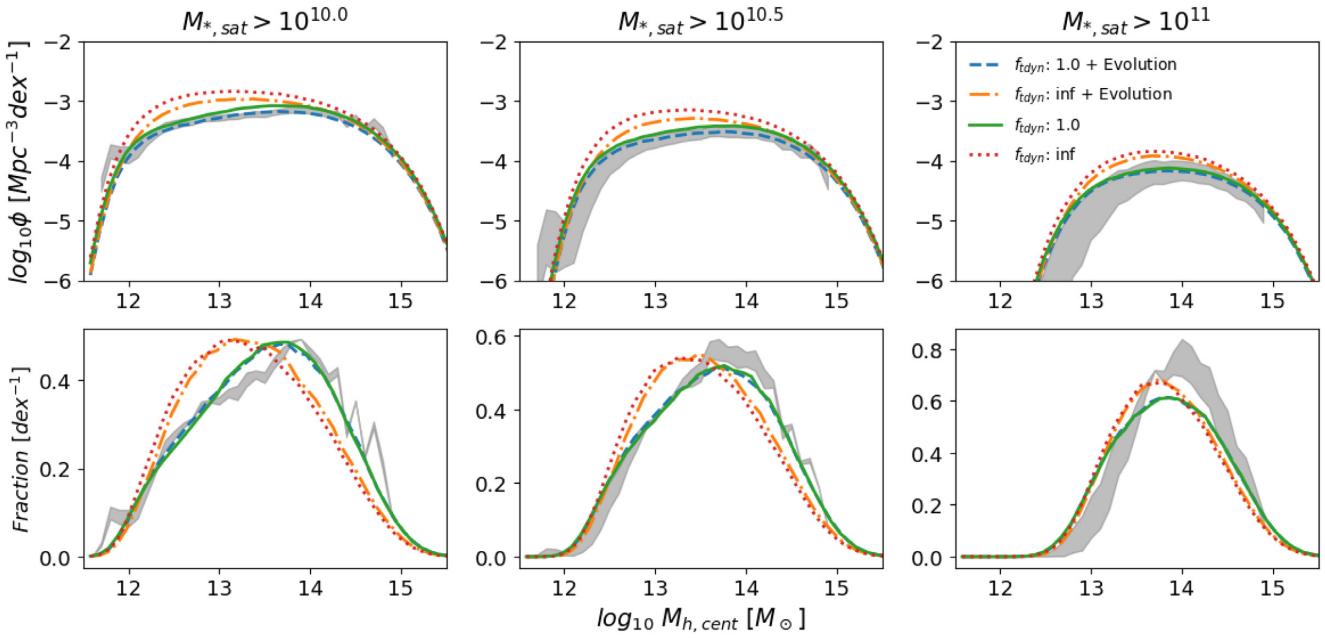


Figure 9. Satellite distributions in parent haloes generated from the model are compared to those observed in SDSS (grey band). Columns from left to right show increasing satellite stellar mass cuts as labelled. The top row shows the number density of satellites expected to be found in each parent halo mass. The bottom row shows the fractional distribution described by equation (16). The solid and dashed lines show a reference $f_{\text{dyn}} = 1.0$ with and without evolution in the SMHM relation, respectively. The dot-dashed and dotted lines show $f_{\text{dyn}} = \infty$ (i.e. no mergers) with and without evolution in the SMHM relation, respectively. The width of the grey band corresponds to a 10 per cent uncertainty in satellite stellar masses.

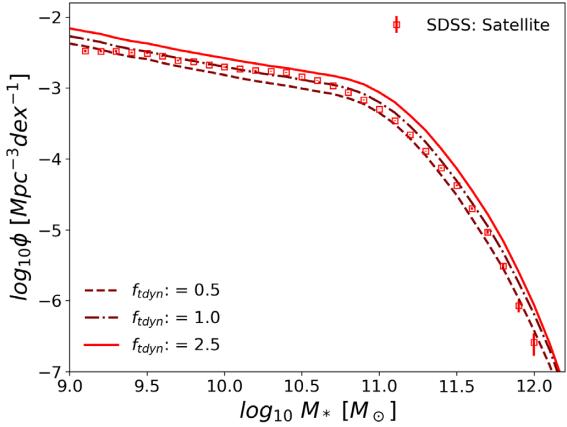


Figure 10. SSMFs generated by the model compared to SDSS data (open squares). The solid, dot-dashed, and dashed lines show $f_{\text{dyn}} = 0.5, 1.0$, and 2.5 , respectively.

other words, there are only a few high mass satellites that had enough time to merge when $f_{\text{dyn}} \gtrsim 1$. In contrast, lower mass satellites have not yet reached their theoretical limit, and thus they still have room to increase their number densities with increasing dynamical time.

Similarly to Fig. 9, Fig. 11 shows how the distribution of satellites as a function of parent halo mass is affected by f_{dyn} . The number density distribution (top row) shows results similar to the satellite SMF where increasing dynamical time increases the number densities. However, there is also an apparent steepening effect for which lower mass host haloes end up containing relatively less satellites with respect to models with longer dynamical times. The fractional plot (bottom row) accentuates this change in the number density distributions shown in the top row: shorter dynamical times shift

the peak of the distribution to the right as relatively more satellites are observed in high-mass host haloes.

This steepening of the satellite distribution as a function of halo mass, as well as the shift mentioned in Section 4.1.1 where infinite dynamical times move satellites preferentially to lower halo masses, are both caused by the amount of time satellites survive in their hosts. Massive satellites are far more common in massive hosts, as can be inferred from the SHMF. Therefore, irrespective of the chosen merging time-scale, there will always be a high number density of surviving massive satellites in higher mass parent haloes. However, when merging time-scales are increased, the lower number densities of massive satellites in moderately-sized haloes are also increased. Given that lower mass parent haloes are more abundant, the reduction of merging time-scales tends to shift the peak of the fractional distribution of galaxies to lower mass parents. Otherwise put, the merging time-scales in clusters ($\log_{10} M_{\text{h,cent}} [\text{M}_\odot] > 13.75$) are so long that even a factor of 5 reduction in merger time-scale still does not give the satellite galaxies sufficient time to merge with the central galaxy. This effect can be inferred from Fig. 7, where we see steeper gradients for a given subhalo/satellite mass with increasing parent halo mass. Merging is more efficient for the lower mass satellites, as can be seen by the steepness of the $f_{\text{dyn}} = 0.5$ model (dashed lines) in the top row of Fig. 11.

The least-square residuals to the SMF, number density distribution and fractional distributions are given in Table 3. There is no model that simultaneously fully matches all the observations in all mass ranges. The trend in both the number density distribution and the fractional distribution is that slightly longer dynamical times ($f_{\text{dyn}} = 1.2$) are favoured by the less massive satellites. Longer dynamical time-scales better match the halo mass distributions (number density distribution/fractional distributions) for lower mass satellites, and vice versa for higher mass satellites with $\log M_{\text{*}}/\text{M}_\odot > 11$. Nevertheless, the simple combination of abundance matching

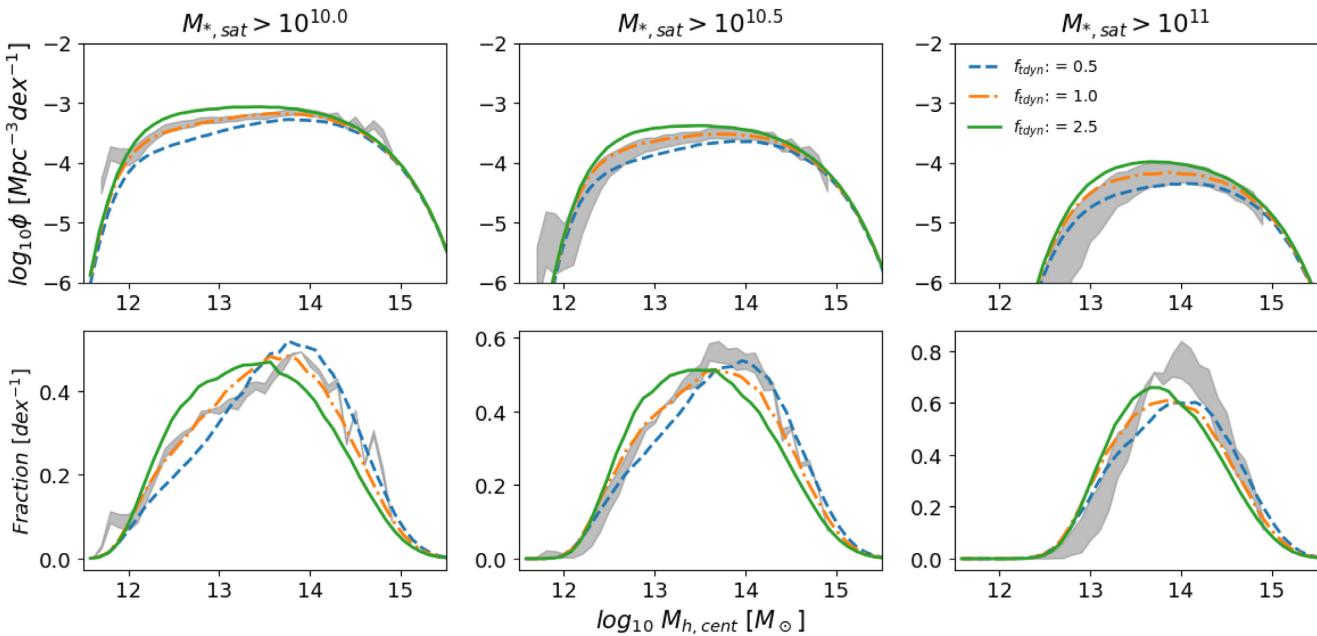


Figure 11. Satellite distributions in parent haloes generated from the model are compared to those observed in SDSS (grey band). Columns from left to right show increasing satellite stellar mass cuts as labelled. The top row shows the number density of satellites expected to be found in each parent halo mass. The bottom row shows the fractional distribution described by equation (16). The dashed, dot-dashed, and solid lines show $f_{t\text{dyn}} = 0.5, 1.0$, and 2.5 , respectively. The width of the grey band corresponds to a 10 per cent uncertainty in satellite stellar masses.

Table 3. We show the sum of the squared residuals between the SDSS and our model. The satellite SMF is calculated between 9.1 and $12.0 M_*$. The SDF fit is calculated between 12 and $14.9 M_h$ for the >10 and >10.5 plots, and between 12.5 and $14.9 M_h$ for >11 . The fractional plot fit is calculated between 11.6 and $14.9 M_h$.

| $f_{t\text{dyn}}$ | SSMF (Fig. 10) | | SDF | | | Fractional distribution | | |
|-------------------|----------------|-------|-------------------|---------|--------|-------------------------|---------|-------|
| | (Fig. 10) | >10 | (Top row Fig. 11) | >10.5 | >11 | >10 | >10.5 | >11 |
| 0.5 | 0.022 | 0.19 | 0.55 | 0.073 | 0.0042 | 0.0047 | 0.0078 | |
| 0.8 | 0.025 | 0.13 | 0.51 | 0.089 | 0.0020 | 0.0017 | 0.0054 | |
| 1.0 | 0.034 | 0.12 | 0.56 | 0.10 | 0.0015 | 0.0011 | 0.0050 | |
| 1.2 | 0.043 | 0.12 | 0.53 | 0.11 | 0.0015 | 0.00094 | 0.0046 | |
| 1.5 | 0.054 | 0.12 | 0.52 | 0.13 | 0.0017 | 0.0010 | 0.0045 | |

and dynamical merging time-scales as suggested by pure N -body simulations ($f_{t\text{dyn}} = 1.0$) tends to provide overall good agreement to both the satellite SMF and the satellite distributions, without the need to invoke additional physics in the (late) evolution of satellite galaxies after infall.

4.2 ‘Non-frozen’ evolutionary models

We showed in the Section 4.1 that the number and distribution of satellite galaxies are primarily driven by the accretion history and dynamical friction time-scale, and depend weakly on the exact redshift evolution in the input SMHM relation.

However, as discussed in Section 3.3, satellites may not be a strictly frozen population. After infall a number of processes such as star formation, quenching and/or stellar stripping may all well affect the final stellar masses of satellite galaxies. In this section, we explore the impact that each one of these processes could have on the predicted local population of satellite galaxies. We will prove, in particular, that additional, late stellar stripping, star-formation, and quenching, whilst not necessarily playing a major role in varying integrated number densities, are still all necessary ingredients to

fully reproduce SDSS data in terms of, for example, specific SFR distributions.

In what follows, we explore the two SFR models with the quenching model described in Sections 3.3.1 and 3.3.2. The first, purely observationally-based, SFR model strictly follows the SFR parametrization by Tomczak et al. (2016, T16 hereafter), given in equation (6). The second SFR model (which we label as ‘CE’) is instead based on a continuity equation approach similar to Leja et al. (2015). In essence, the latter is based on first assuming number conservation of galaxies, and then deriving the implied SFRs as a function of stellar mass from the time growth in any stellar mass bin implied by the redshift evolution of the SMF (corrected for gas loss from dying stars assumed to be an average of 40 per cent). The novelty in the latter model with respect to previous work is that we do not tune the resulting SFR on the total but rather only on the SMF of *central* galaxies, which is in turn iteratively constrained by matching the local SMF of SDSS centrals. This has the main effect of lowering the implied average SFR at fixed stellar mass at any given epoch (see equations 6 and 7).

We compare the satellite SMFs produced by the two star formation+quenching models addressed above to our SDSS SMF

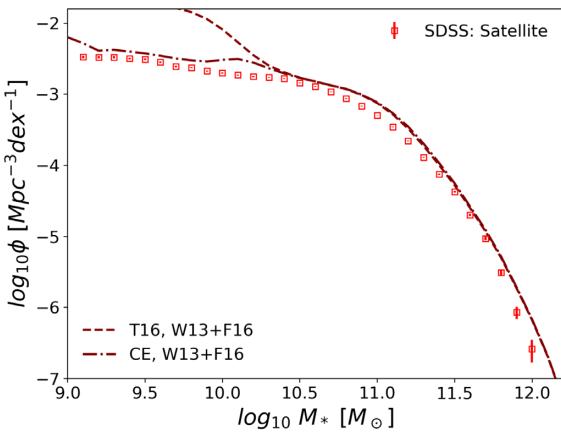


Figure 12. SSMFs generated from the model using both the Tomczak et al. (2016) (dashed) and continuity (dot-dashed) SFRs compared to the SDSS SSMF (open squares).

of satellites in Fig. 12. It is apparent that using the observed SFR by T16 (dashed line), even inclusive of the best recipes for quenching, still substantially overproduces the number density of galaxies below $M_* \lesssim 3 \times 10^{10} M_\odot$. This is a well-known problem affecting the full (dominated by central) galaxy population (e.g. Leja et al. 2015): the integrated (observed) SFR is not consistent with the moderate growth over time of the SMF causing an overproduction of galaxies becoming gradually more severe at lower stellar masses. Our results point to a similar problem affecting the satellite population, on the assumption that the latter at infall share the same SFR distribution as a typical central galaxy of the same stellar mass.

To further constrain our adopted continuity equation-based SFR distributions, we compare our best models with the bimodality in (satellite) SSFR measured in the SDSS data (grey shading in Fig. 13). The data show a clear bimodality in SSFR at least for galaxies below $M_* \lesssim 10^{11} M_\odot$, at variance with the SSFR distribution of centrals (black thick long dashed lines), which point to a much larger population of star-forming galaxies. This difference in SSFR distributions has been often ascribed to environmental quenching (e.g. Peng et al. 2010), either in the form of stripping or strangulation. Both the SFR reference models predict very similar distributions in SSFR for galaxies above $M_* \gtrsim 3 \times 10^{10} M_\odot$, as expected from the invariance in the predicted number densities above this stellar mass. When using the continuity equation-based method (CE, dot-dashed line) the overproduction is greatly reduced and the SSMF is very close to the SSMF with no evolution in the satellite population.

We now move on showing the relative impact of SFR and stellar stripping on the SSMF. Figs 14 and 15 show SMF and host halo mass distributions for the $f_{\text{dyn}} = 1.0$ reference model with neither stripping nor star formation, the CE star formation model, and the CE star formation model with stripping (long-dashed, dot-dashed, and solid lines, respectively). We see from both figures that the reference and star formation model are almost indistinguishable. The stripping, at least at the level implemented in this work, also has a rather minor effect, at the most reducing the number densities of the most massive satellites ($> 10^{11} M_\odot$) by $\lesssim 0.2$ dex. Note that that the *cumulative* effect of stripping on the stellar mass growth of central galaxies could be much more prominent, as emphasized by a number of groups (e.g. Cattaneo et al. 2017).

Table 4 shows the sum of the square residuals to the SMF, SDF,

and fractional distributions for the same models discussed above, our reference frozen one, and the one with evolution of satellites after infall (stellar stripping and continuity equation-based star formation). Table 4 shows that the satellite late evolution has little effect on the fractional distribution. In the number density distribution, we see an improved fit for galaxies in the $M_* > 10^{11} M_\odot$ range, mainly induced by the stripping which slightly reduces the number density of massive galaxies. Table 4 shows the sum of square residuals for the dynamical time with $f_{\text{dyn}} = 1$, for the frozen and evolved models.

In summary, our results point to the fact that direct abundance matching between centrals and satellite galaxies and the halo plus SHMFs in the local Universe (and at higher redshifts) is a good approximation to the mean SMHM relation for central galaxies, irrespective of the precise late evolution of satellites after infall.

5 DISCUSSION

5.1 Strengths of the statistical approach with respect to the state of the art

The statistical semi-empirical model introduced in this work presents a highly complementary approach to the existing suit of cosmological galaxy evolution models. The latter models, either being fully analytic or fully numerical, inevitably rely on large boxes to simulate both low- and high-mass galaxies with sufficient statistics. Even traditional semi-empirical models, based on applying abundance matching techniques to vast catalogues of dark matter merger trees (e.g. Behroozi et al. 2018, and references therein), still require large computational resources. Our approach, based on statistical mean dark matter accretion histories, allows to explore the full range of galaxy stellar and halo masses without the need to simulate large volumes or even weight dark matter merger trees. Whilst we lose the ability to track an individual galaxy through cosmic time, we are able to rapidly predict the statistics of any subpopulation of chosen galaxies at any redshift and environment.

There is vast literature on the modelling of satellite galaxies. Here, we recall just a couple of examples of semi-empirical and semi-analytic models to highlight some of the key similarities and key differences. Neistein et al. (2013), with an approach similar to ours, separated the central and satellite populations in an attempt to better define the galaxy halo connection. By allowing in an N -body simulation the stellar mass of satellite halaxies to depend on both the host subhalo mass and on the parent halo mass, Neistein et al. (2013) find that the local satellite SMHM is substantially less well defined than the one for central galaxies. In our model, satellites instead strictly follow the SMHM relation of centrals at infall. In this way, we find the resulting satellite distributions to be well reproduced. Our semi-empirical statistical model was able to reproduce multiple observables such as the stellar and parent halo mass distributions, with essentially only one parameter, f_{dyn} . By working with minimal assumptions and related free parameters, our approach is thus less prone to possible degeneracies affecting more traditional, multiparameter techniques.

Another key difference with respect to previous models concerns ‘orphan galaxies’. In N -body or merger tree-based simulations, when a subhalo goes below the resolution limit, an orphan galaxy is created (e.g. De Lucia et al. 2011; Guo et al. 2011). It is then necessary to make an assumption on how much longer that subhalo (and hosted satellite galaxy) will survive. In our model, we avoid this complication by self-consistently assigning to all satellites a (full) observability time-scale at infall.

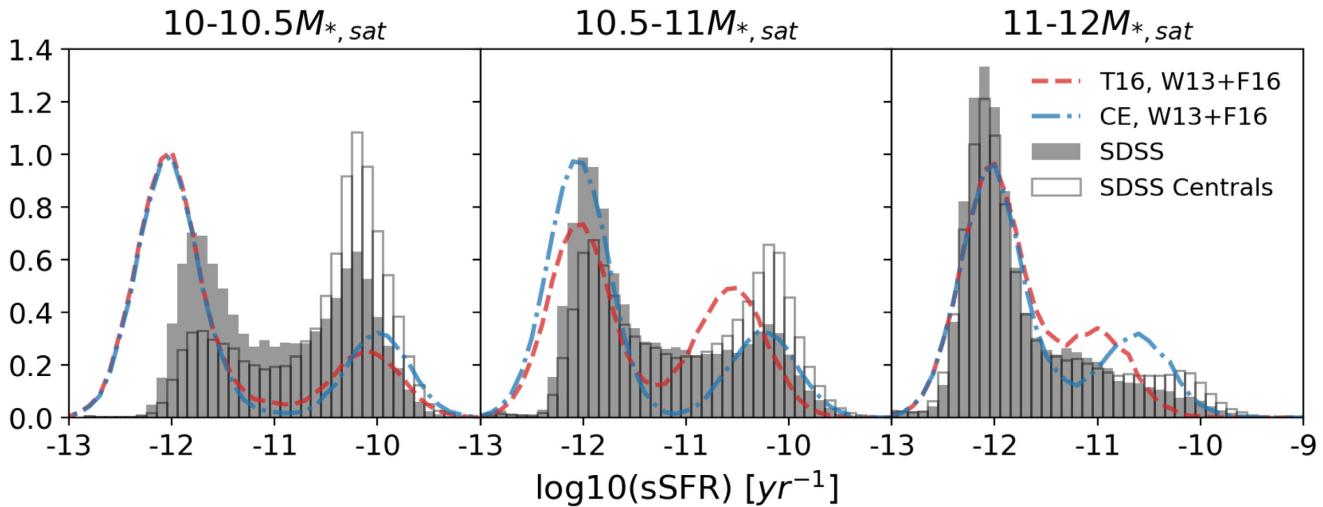


Figure 13. Specific SFR distributions in three mass ranges (as labelled) generated from the model using both the Tomczak et al. (2016) (dashed) and continuity (dot-dashed) SFRs. Data shown are the distribution of specific SFR in SDSS satellites (shaded bars) and SDSS centrals (unfilled bars).

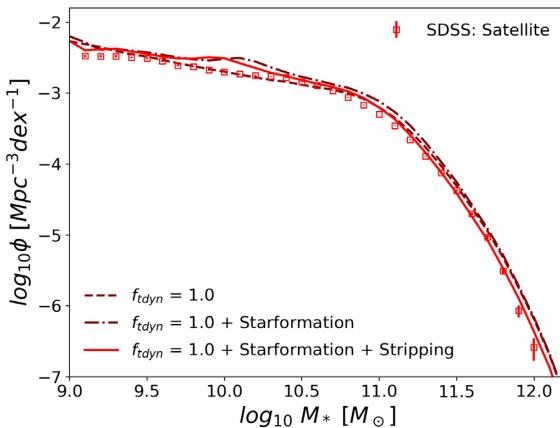


Figure 14. SSMFs generated from the model compared to SDSS satellites (open squares). The models shown all have $f_{\text{dyn}} = 1.0$ and are the reference ‘frozen model’ (dashed line), starformation (CE model) only (dot-dashed line) and star formation and stripping (solid line).

5.2 Future applications of the statistical semi-empirical model

We have shown here the ability of the statistical semi-empirical model to reproduce the observed group and cluster richness at $z = 0.1$ using simply the dynamical friction time-scale from N -body simulations. In future work, we will also compare our model outputs with the richness of groups and clusters at high redshift thus providing additional important constraints to models of structure formation. For example, we will use our statistical modelling to fit extreme objects such as the high redshift ($z = 2.5$) massive cluster reported in Wang et al. (2016), which has been claimed to be challenging in other Λ CDM cosmology-based models. By predicting the (correct) distributions of satellite galaxies predicted over many epochs, more accurate merger rates and pair fractions can be calculated and compared to observations found in surveys such as those reported in Mundy et al. (2017). Via more robust estimates of galaxy merger rate, we will be able to set upper limits on the impact of mergers in shaping galaxy morphology (Hopkins et al. 2009a), build galactic bulges (Cole et al. 2000; Hopkins et al. 2010; Shankar et al. 2013), and trigger AGN activity (Villforth et al. 2017).

Our technique being fast and flexible is an extremely valuable resource in view of the next large and deep extra-galactic surveys for example *Euclid* (Refregier et al. 2010). Our model can be both predictive, working out specific expectations for such surveys, and then reactive, using the model outputs as constraints on existing formation models.

6 CONCLUSIONS

In this paper, we have presented ‘STEEL’ a *Statistical sEmi-Empirical modeL* that replaces the traditional merger tree dark matter backbone with a ‘statistical accretion history’. In essence, we first trace backwards the mean accretion history of dark matter haloes in a given bin of halo mass at the redshift of interest z . We then gradually build in time their satellite population by integration of the SHMF, and at each time-step assign stellar masses to subhaloes via abundance matching techniques. Our approach is extremely fast and accurate, allowing to probe the galaxy number densities and average properties within the full range of stellar masses probed by large extragalactic surveys, with virtually no limitation on volume size or mass resolution.

We use the statistical semi-empirical model to predict the satellite galaxy population at stellar masses above $M_* \gtrsim 10^9 M_\odot$. We find that irrespective of the exact input SMHM relation, the main driver shaping the local SMF of satellite galaxies is by far the input dynamical friction time-scale. In particular, adopting a merging time-scale very close to the one calibrated from high-resolution N -body dark matter simulations provides an excellent match to the satellites SMF as inferred from SDSS. Shorter dynamical friction time-scales not only reduce the number densities of satellites (as more satellites disappear merging with central galaxies) but also shift the abundances of satellites into higher mass parent haloes than actually observed. Conversely, longer dynamical friction time-scales create an overproduction in the SSMF and induce too many satellites in lower mass parents. Thus, a traditional hierarchical dark matter cosmology naturally provides the right abundances of satellites in the local Universe without any additional fine-tuning.

We then incorporate popular recipes for star formation, stellar stripping, and quenching in satellite galaxies after infall. We find

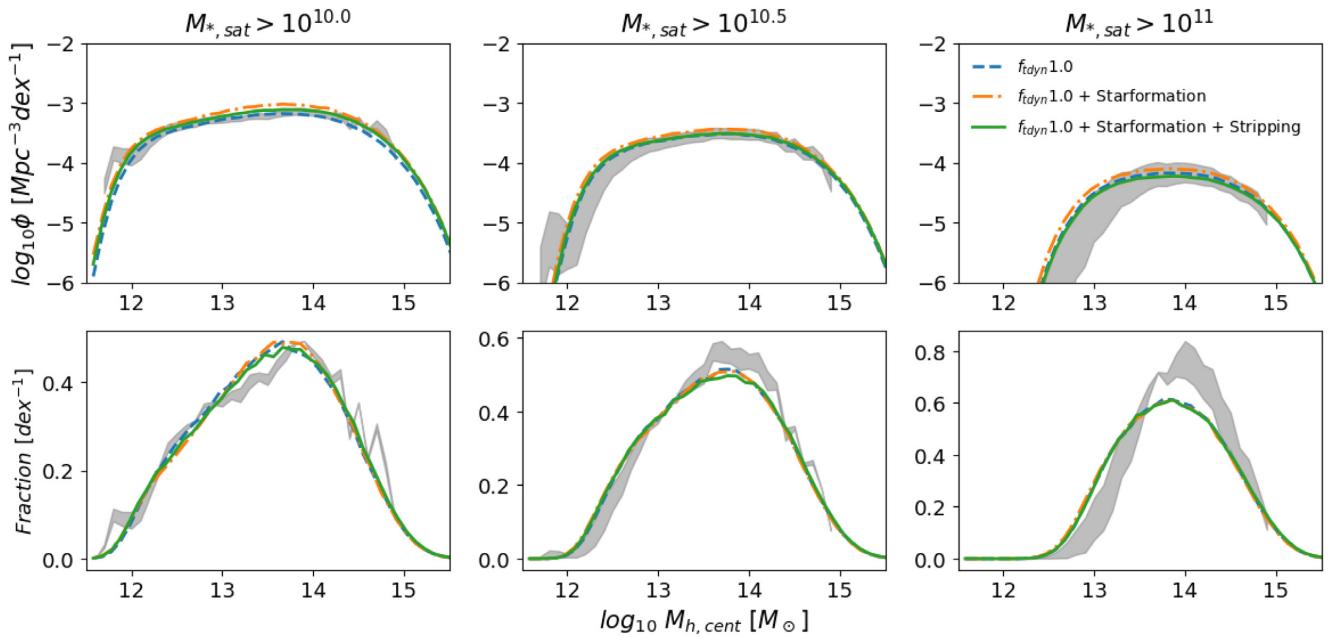


Figure 15. Satellite distributions in parent haloes generated from the model are compared to those observed in SDSS (grey band). Columns from left to right show increasing satellite stellar mass cuts as labelled. The top row shows the number density of satellites expected to be found in each parent halo mass. The bottom row shows the fractional distribution described by equation (16). The models shown all have $f_{\text{dyn}} = 1.0$ and are the reference ‘frozen model’ (dashed line), star formation only (dot-dashed line) and star formation and stripping (solid line). The width of the grey band corresponds to a 10 per cent uncertainty in satellite stellar masses.

Table 4. We show the sum of the squared residuals between the SDSS and our model as in Table 3 with the same mass ranges for the fitting. All models have $f_{\text{dyn}} = 1.0$ from top to bottom, we then have the reference frozen model, the model with star formation, and the model with stripping and star formation.

| f_{dyn} | SSMF (Fig. 14) | SDF (Top row Fig. 15) | | | Fractional distribution (Bottom row Fig. 15) | | |
|-----------------------------------|-------------------|--------------------------|-------|-------|---|--------|--------|
| | | >10 | >10.5 | >11 | >10 | >10.5 | >11 |
| 1.0 | 0.034 | 0.12 | 0.53 | 0.10 | 0.0015 | 0.0011 | 0.0049 |
| 1.0 | 0.049 | 0.077 | 0.43 | 0.14 | 0.0018 | 0.0012 | 0.0059 |
| With star formation | | | | | | | |
| 1.0 | 0.021 | 0.087 | 0.47 | 0.088 | 0.0016 | 0.0015 | 0.0056 |
| With stripping and star formation | | | | | | | |

that, within reasonable variations of the input parameters, all these processes play a secondary role. At fixed merging time-scale the resulting SMF and halo mass distribution of local satellites end up being very similar. All in all, our results highly suggest that at any given epoch, a direct abundance matching between the total galaxy SMF and halo plus SHMFs provides a very good approximation to the mean SMHM relation of central galaxies, irrespective of the exact (late) evolution of satellites after infall. In other words, our results confirm the common assumption that at any given epoch satellites can be abundance matched as centrals at infall, irrespective of their redshift of infall.

We find that, similarly to what previously noticed for centrals, adopting the latest empirical determinations of the SFR as a function of stellar mass and redshift significantly overproduces the abundance of galaxies below $M_* \lesssim 3 \times 10^{10} M_\odot$, irrespective of the quenching, stripping, or merger time-scales adopted. An SFR as predicted by a continuity equation approach well matches both the local abundances of satellites and the bimodality in specific SFRs for all galaxies above $M_* \gtrsim 10^{11} M_\odot$ as measured in the SDSS.

STEEL is an ideal tool to rapidly compute accurate galaxy merger rates as a function of stellar mass, time, and environment. It can

reveal the impact of mergers in shaping the structural and dynamical evolution of galaxies, and eventually their central supermassive black holes. We will look into some of these topics in more detail in future work.

ACKNOWLEDGEMENTS

We warmly thank B. Moster and J. Leja for detailed discussions on the SMHM relation and the SFR, respectively. We also thank C. Conselice, N. Menci, and C. Marsden for useful discussions on the project. We also acknowledge extensive use of the PYTHON libraries ASTROPY, MATPLOTLIB, NUMPY, PANDAS, and SCIPY. PJG acknowledges support from the STFC for funding this PhD.

REFERENCES

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543
- Baugh C. M., 2006, A Primer on Hierarchical Galaxy Formation: The Semi-analytical Approach. Available at: <http://iopscience.iop.org/article/10.1086/379875/pdf>; <http://stacks.iop.org/0004-637X/600/i=2/a=580>
- Behroozi P., Wechsler R., Hearn A., Conroy C., 2018, preprint ([arXiv:1806.07893](https://arxiv.org/abs/1806.07893))

- Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, 717, 379
 Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
 Berlind A. A., Weinberg D. H., 2002, *ApJ*, 575, 587
 Bernardi M., Fischer J. L., Sheth R. K., Meert A., Huertas-Company M., Shankar F., Vikram V., 2017b, *MNRAS*, 468, 2569
 Bernardi M., Meert A., Sheth R. K., Fischer J. L., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2017a, *MNRAS*, 467, 2217
 Bernardi M., Meert A., Sheth R. K., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2016, *MNRAS*, 455, 4122
 Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, *MNRAS*, 436, 697
 Bosch F. C. v. d., Jiang F., Hearin A., Campbell D., Watson D., Padmanabhan N., 2014, *MNRAS*, 445, 1
 Bosch F. C. v. d., Ogiya G., Hahn O., Burkert A., 2018, *MNRAS*, 474, 3043
 Boylan-Kolchin M., Ma C. P., Quataert E., 2008, *MNRAS*, 383, 93
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Buchan S., Shankar F., 2016, *MNRAS*, 462, 2001
 Cattaneo A., Mamon G. A., Warnick K., Knebe A., 2011, *A&A*, 533
 Cattaneo A. et al., 2017, *MNRAS*, 471, 1401
 Chabrier G., 2003, *PASP*, 115, 763
 Chandrasekhar S., 1943, *ApJ*, 97, 255
 Cole S., Lacey C., Baugh C., Frenk C., 2000, *MNRAS*, 319, 168
 Cooray A., Sheth R., 2002, *Phys. Rep.*, 372, 1
 Dalal N., White M., Bond J. R., Shirokov A., 2008, *ApJ*, 687, 12
 Davidzon I. et al., 2017, *A&A*, 605, A70
 De Lucia G., Fontanot F., Wilman D., Monaco P., 2011, *MNRAS*, 414, 1439
 Diemer B., 2017, Technical report, COLOSSUS: A Python Toolkit for Cosmology, Large-Scale Structure, and Dark Matter Halos, Harvard-Smithsonian Center for Astrophysics. Available at: <http://arxiv.org/abs/1712.04512>
 Fattahi A. et al., 2016, *MNRAS*, 457, 844
 Fillingham S. P., Cooper M. C., Pace A. B., Boylan-Kolchin M., Bullock J. S., Garrison-Kimmel S., Wheeler C., 2016, *MNRAS*, 463, 1916 (F16)
 Giocoli C., Pieri L., Tormen G., 2008, *MNRAS*, 387, 689
 González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, *ApJ*, 735, L34
 Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, *ApJ*, 600, 580
 Guo H. et al., 2018, preprint ([arXiv:1810.05318](https://arxiv.org/abs/1810.05318))
 Guo Q. et al., 2011, *MNRAS*, 413, 101
 Hirschmann M., Naab T., Somerville R. S., Burkert A., Oser L., 2012, *MNRAS*, 419, 3200
 Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009b, *ApJ*, 691, 1168
 Hopkins P. F. et al., 2009a, *MNRAS*, 397, 802
 Hopkins P. F. et al., 2010, *ApJ*, 715, 202
 Huertas-Company M., et al., 2015, *ApJ*, 809, 95
 Huertas-Company M., et al., 2016, *MNRAS*, 462, 4495
 Jiang C. Y., Jing Y. P., Faltenbacher A., Lin W. P., Li C., 2008, *ApJ*, 675, 1095
 Jiang F., van den Bosch F. C., 2016, *MNRAS*, 458, 2848
 Khochfar S., Burkert A., 2006, *A&A*, 445, 403
 Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, *MNRAS*, 457, 4340
 Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, *ApJ*, 609, 35
 Lapi A., Cavaliere A., 2011, *Adv. Astron.*, 2011, 1
 Lee N. et al., 2015, *ApJ*, 801, 80
 Leja J., Van Dokkum P. G., Franx M., Whitaker K. E., 2015, *ApJ*, 798, 115
 Lemson G., Consortium T. V., 2006, Technical Report, Halo and Galaxy Formation Histories from the Millennium Simulation: Public Release of a VO-oriented and SQL-queryable Database for Studying the Evolution of Galaxies in the LambdaCDM Cosmogony. Max Planck Institutes for Astrophysics and for Extraterrestrial Physics
 McCavanna T., Micic M., Lewis G. F., Sinha M., Sharma S., Holley-Bockelmann K., Bland-Hawthorn J., 2012, *MNRAS*, 424, 361
 Meert A., Vikram V., Bernardi M., 2015, *MNRAS*, 446, 3943
 Meert A., Vikram V., Bernardi M., 2016, *MNRAS*, 455, 2440
 Menci N., Fontana A., Giallongo E., Grazian A., Salimbeni S., 2006, *ApJ*, 647, 753
 Mendel J. T., Simard L., Palmer M., Ellison S. L., Patton D. R., 2014, *ApJS*, 210, 3
 Monaco P., Fontanot F., Taffoni G., 2007, *MNRAS*, 375, 1189
 Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, 428, 3121
 Moster B. P., Naab T., White S. D. M., 2018, *MNRAS*, 477, 1822
 Moster B. P., Somerville R. S., Maulbetsch C., Van Den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903
 Mundy C. J., Conselice C. J., Duncan K. J., Almaini O., Häußler B., Hartley W. G., 2017, *MNRAS*, 470, 3507
 Murray S., Power C., Robotham A., 2013, *Astron. Comput.*, 3-4, 23
 Neistein E., Li C., Khochfar S., Weinmann S. M., Shankar F., Boylan-kolchin M., 2013, *MNRAS*, 15, 1
 Parkinson H., Cole S., Helly J., 2008, *MNRAS*, 383, 557
 Peng Y. J. et al., 2010, *ApJ*, 721, 193
 Planck Collaboration XIII et al., 2016, *A&A*, 594, A13
 Potter D., Stadel J., Teyssier R., 2017, *Comput. Astrophys. Cosmol.*, 4, 2
 Press W. H., Schechter P., 1974, *ApJ*, 187, 425
 Refregier A., Amara A., Kitching T. D., Rassat A., Scaramella R., Weller J., Euclid Imaging Consortium F., 2010, preprint ([arXiv:1001.0061](https://arxiv.org/abs/1001.0061))
 Salim S. et al., 2007, *ApJS*, 173, 267
 Shankar F., Lapi A., Salucci P., De Zotti G., Danese L., 2006, *ApJ*, 643, 14
 Shankar F., Marulli F., Bernardi M., Mei S., Meert A., Vikram V., 2013, *MNRAS*, 128, 109
 Shankar F. et al., 2014a, *MNRAS*, 439, 3189
 Shankar F. et al., 2014b, *ApJ*, 797, L27
 Shankar F. et al., 2015, *ApJ*, 802, 73
 Shankar F. et al., 2017a, *MNRAS*, 466, 4029
 Shankar F. et al., 2017b, *ApJ*, 840, 34
 Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878
 Tinker J. L. et al., 2017, *ApJ*, 839, 121
 Tomczak A. R. et al., 2016, *ApJ*, 817, 118
 Vale A., Ostriker J. P., 2004, *MNRAS*, 353, 189
 Villforth C. et al., 2017, *MNRAS*, 466, 812
 Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
 Wang T. et al., 2016, *ApJ*, 828, 56
 Wetzel A. R., Tinker J. L., Conroy C., Van Den Bosch F. C., 2013, *MNRAS*, 432, 336 (W13)
 Yang X., Mo H. J., Jing Y. P., van den Bosch F. C., Chu Y., 2004, *MNRAS*, 350, 1153
 Yang X., Mo H. J., Van den Bosch F. C., 2003, *MNRAS*, 339, 1057
 Yang X., Mo H. J., Van Den Bosch F. C., Zhang Y., Han J., 2012, *ApJ*, 752
 Zavala J., Avila-Reese V., Firmani C., Boylan-Kolchin M., 2012, *MNRAS*, 427, 1503
 Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, *MNRAS*, 443, 3044

APPENDIX A: ABUNDANCE MATCHING

In Fig. A1, we show the central and SSMFs from the evolving abundance matching relation described in Section 3.2. We show nine different redshifts, as labelled, using the frozen model with a dynamical-time factor $f_{\text{dyn}} = 1.0$, which best matches the local data from SDSS. The $z > 0.3$ data are from Davidzon et al. (2017).

A limitation of the present methodology is that it lacks, at present, robust observational estimates of the high-redshift central and SSMFs to better anchor our fitting procedure. Nevertheless, we see in Figs 9 and 8, the satellite SMF, the number density distribution or the fractional distribution do not necessarily require any redshift evolution in the input parameters of the SMHM relation to fit the local data. This is mainly due to the fact that most (> 60 per cent) of the accretion of the satellites observed at $z = 0$ occurs at relatively recent times, $z < 0.5 - 1$. Redshift evolution is however required for a good fit to the data when $z > 1$, as demonstrated by the grey lines in Fig. A1.

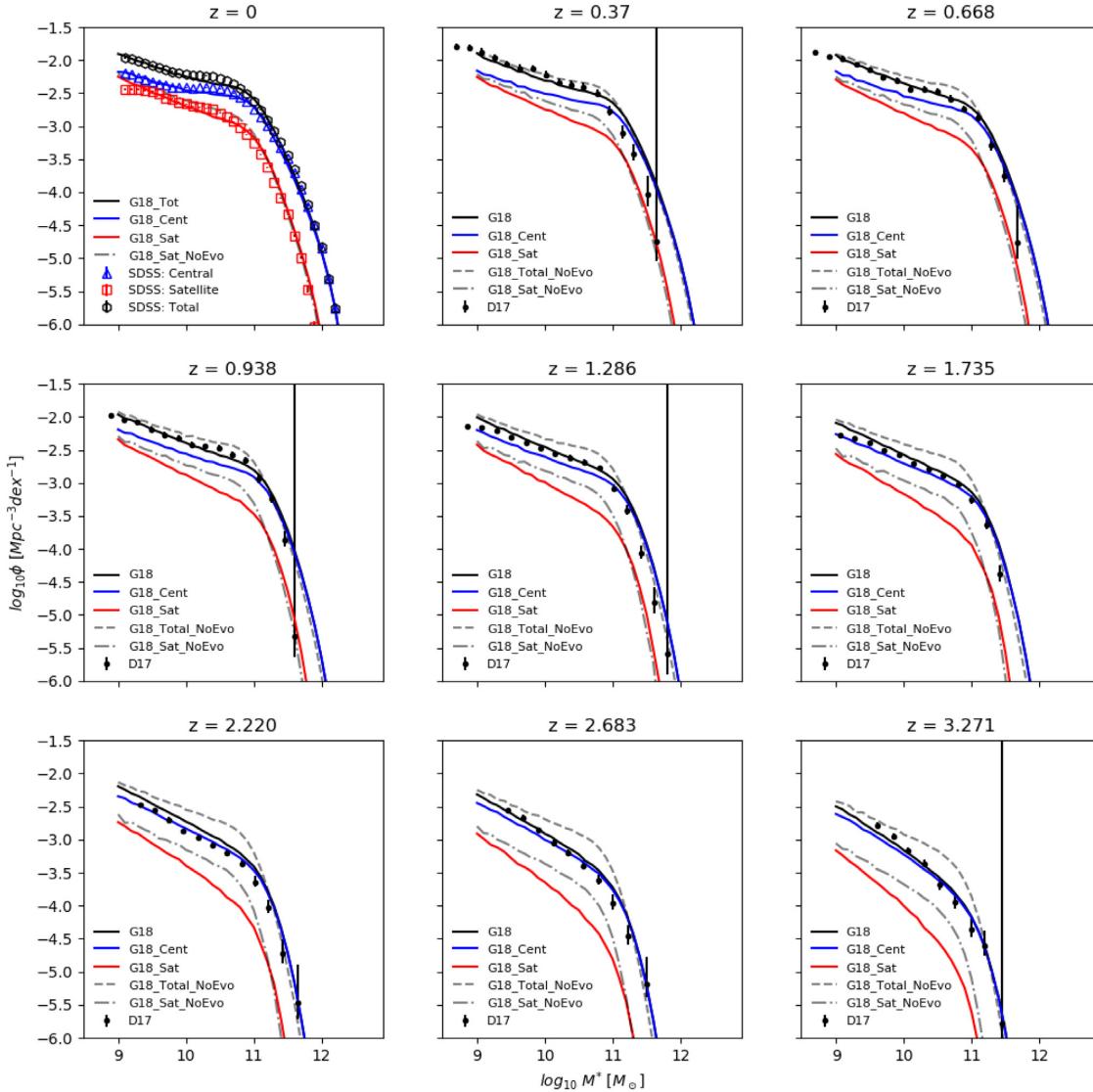


Figure A1. We show the results of the redshift dependent abundance matching at eight different epochs above redshift $z = 0$. The SDSS data are shown in the redshift $z = 0.1$ panel with crosses, higher redshift panels show COSMOS2015 (D17) data as green dots. The final result is shown in each panel as green blue and red lines for total, central, and satellite galaxies, respectively. The no-evolution is shown in grey in all panels. In this plot, G18 refers to the central SMHM fit from this work.

APPENDIX B: SFR FROM A CONTINUITY EQUATION APPROACH

A general SFR continuity equation model considers the change in the SMF over time step δt . Assuming galaxies maintain rank order the model calculates the average SFR a galaxy would have to maintain the implied growth in stellar mass. In this work, using a continuity method, as described above, we generate a SFR that is self-consistent with the HMF and abundance matching used in this work. Usually, the input SMF used is an observed quantity. However, we use as input for our continuity equation the central SMF generated by the SMHM relation from Section 3.2 and the central HMF from Tinker et al. (2010).¹⁴ When considering the mass

growth we must consider mass-loss, else the SFR is massively underpredicted. We use an instantaneous loss fraction of 40 per cent. We also neglect the mass gained from mergers. Mergers would decrease the SFR calculated (and therefore provide a better fit though out this work) so SFR from this method should be considered as an upper limit. We show in Section 3.3 our SFR has limited effect on the results of this work. We alter the parameters in equation (6) to fit our continuity equation-driven M_* –SFR relations and the new fit is given in equation (7). During this fit, we ignore the SFR inferred from galaxies that sit above the knee of the SMHM relation where the growth becomes merger dominated (Tomczak et al. 2016). We show the difference between the continuity equation approach and the Tomczak SFR in Fig. B1.

¹⁴We also note that given the release of resolved centrals at high redshift, improved abundance matching, and/or improved local SMF the SFR (as

well as the rest of the model) is simply adjusted self-consistently as is the power of a semi-empirical model.

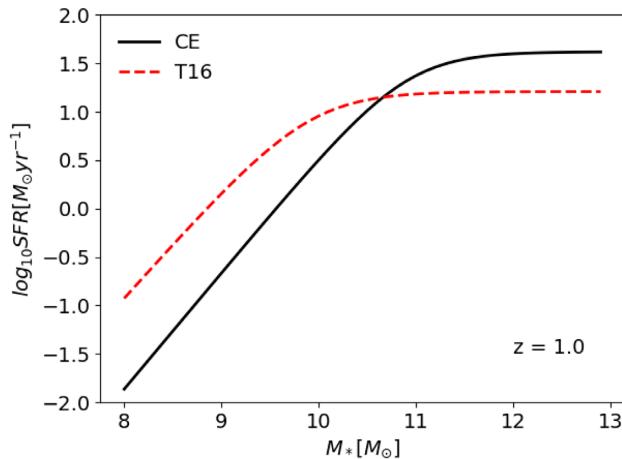


Figure B1. The Tomczak (red dashed) and continuity equation approach (black solid) SFR–stellar mass relations as given in Section 3.3.1 at redshift $z = 1$.

APPENDIX C: GLOSSARY OF TECHNICAL ACRONYMS

Throughout this work, we use several technical acronyms. We provide the reader with a glossary of these in Table C1 to aid reading of the paper and equations therein.

Table C1. A list of technical acronyms used throughout this paper.

| Acronym | Definition |
|----------------|---|
| STEEL | The name of our model: STatistical sEmi-Empirical modeL |
| HMF | Halo mass function |
| SHMF | Subhalo mass function |
| USHMF | Unevolved subhalo mass function |
| USSHMF | Unevolved surviving subhalo mass function |
| δ USHMF | Unevolved subhalo mass function accretion |
| SMF | Stellar mass function |
| SSMF | Satellite stellar mass function |
| SFR | Star formation rate |
| SFH | Star formation history |
| MLR | Mass-loss rate |

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