



# Predicting fully self-consistent satellite richness, galaxy growth, and star formation rates from the SStatistical sEmi-Empirical modeL STEEL

Philip J. Grylls<sup>1</sup>, F. Shankar,<sup>1</sup> J. Leja<sup>1,2</sup>, N. Menci,<sup>3</sup> B. Moster<sup>1,4</sup>, P. Behroozi<sup>1,5</sup> and L. Zanisi<sup>1</sup>

<sup>1</sup>Department for Physics and Astronomy, University of Southampton, Highfield SO171BJ, UK

<sup>2</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, USA

<sup>3</sup>INAF – Osservatorio Astronomico di Roma, via di Frascati 33, I-00040 Monte Porzio Catone, Italy

<sup>4</sup>Universitäts-Sternwarte, Ludwig-Maximilians-Universität München, Scheinerstr 1, D-81679 München, Germany

<sup>5</sup>Department of Astronomy and Steward Observatory, University of Arizona, Tucson, AZ 85721, USA

Accepted 2019 October 18. Received 2019 October 14; in original form 2019 May 2

## ABSTRACT

Observational systematics complicate comparisons with theoretical models limiting understanding of galaxy evolution. In particular, different empirical determinations of the stellar mass function imply distinct mappings between the galaxy and halo masses, leading to diverse galaxy evolutionary tracks. Using our state-of-the-art SStatistical sEmi-Empirical modeL, STEEL, we show fully self-consistent models capable of generating galaxy growth histories that simultaneously and closely agree with the latest data on satellite richness and star formation rates at multiple redshifts and environments. Central galaxy histories are generated using the central halo mass tracks from state-of-the-art statistical dark matter accretion histories coupled to abundance matching routines. We show that too flat high-mass slopes in the input stellar mass–halo mass relations as predicted by previous works, imply non-physical stellar mass growth histories weaker than those implied by satellite accretion alone. Our best-fitting models reproduce the satellite distributions at the largest masses and highest redshifts probed, the latest data on star formation rates and its bimodality in the local Universe, and the correct fraction of ellipticals. Our results are important to predict robust and self-consistent stellar mass–halo mass relations and to generate reliable galaxy mock catalogues for the next generations of extragalactic surveys such as Euclid and LSST.

**Key words:** Galaxy: halo – galaxies: abundances – galaxies: clusters: general – galaxies: evolution – galaxies: high-redshift – galaxies: star formation.

## 1 INTRODUCTION

Galaxies are thought to grow and evolve through a combination of ‘*in situ*’ and ‘*ex situ*’ processes. *In situ* processes such as star formation, are thought to be driven by the availability of cold gas in the galaxy. The reserve of cold gas ready to fuel star formation could be regulated by a number of internal and external processes, from stellar and active galactic nuclei feedback to host halo and/or morphological quenching (e.g. Granato et al. 2004; Dekel et al. 2009; Lilly et al. 2013; Schawinski et al. 2014). One important *ex situ* channel affecting galaxy growth is satellite accretion. In particular, in very massive galaxies growth via satellite accretion has been claimed to become progressively more relevant (De Lucia et al. 2006; van Dokkum et al. 2010; Shankar et al. 2013, 2015; Buchan & Shankar 2016; Groenewald et al. 2017; Matharu et al. 2019). Central

galaxies that reside at the centre of massive haloes thus provide a window into the different pathways that have contributed to the mass growth history of galaxies in the local universe. Exploring the way these galaxies build their mass can give insights into the stellar mass–halo mass (SMHM hereafter) relation, the efficiency of the satellite transport from the edge of the cluster to the centre, the balance of the major processes taking place on these satellites, the galaxy merger rate, and the star formation rate (SFR). The characteristic mass at which galaxies transition from being *in situ* to *ex situ* growth dominated has previously been found at  $M_* \sim 10^{11} M_\odot$  (Bernardi et al. 2011; Cattaneo et al. 2011; Shankar et al. 2013).

Models of galaxy formation traditionally use the hierarchical growth of dark matter structure as the backbone for galaxy assembly. Hydrodynamical simulations co-evolve the dark matter and baryonic matter allowing for a simultaneous look at the assembly of both components (Vogelsberger et al. 2014; McAlpine et al. 2015). The latter technique, however, requires large computational

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

resources. Less computationally intensive models such as traditional semi-analytic and semi-empirical models use dark matter merger trees from post-processing of dark matter simulations (Guo et al. 2011; Shankar 2013). Dark matter merger trees visualize dark matter assembly as a central trunk and halo mergers happen where branches join. Semi-analytic models initialize gas at high redshift and use a number of physical assumptions and free parameters to tune to observations (De Lucia et al. 2006; Guo et al. 2011). Semi-empirical models use a more direct approach initializing galaxy stellar mass in dark matter haloes most commonly through abundance matching, the association of galaxies to dark matter host haloes via relative abundances (Hopkins et al. 2010b; Zavala et al. 2012; Moster, Naab & White 2013; Shankar et al. 2014; Moster, Naab & White 2018). Both semi-empirical and semi-analytic models follow the merging histories of the underlying dark matter merger trees to track the *in situ* and *ex situ* build-up of galaxy mass. The work of Moster et al. (2018), for example, uses a semi-empirical model to associate the growth of the dark matter halo to the SFR of the host galaxy alongside the build-up of stellar mass from satellites accretion, further strengthening the connection between the dark matter host environment and the build-up of galactic stellar mass.

An average measure of the growth of galaxies can be obtained by comparing stellar mass functions (SMFs), the number densities of galaxies as a function of mass, over multiple epochs. Selecting galaxy populations with the same number density at each epoch, assuming that galaxies maintain rank ordering over cosmic time, allows an estimation of the average growth of galaxies to be made. An estimation of the SFR for each galaxy population can then be computed by taking the time derivative of the stellar mass growth. However, the star formation generated in this way is significantly lower than observational estimates of the SFR (e.g. Leja et al. 2015; Lapi et al. 2017). It is consequently found that if observed SFRs are used in models, they cannot be reconciled with the SMF. In Grylls et al. (2019) it is shown this is also in effect in satellite galaxy distributions, where the predicted number of massive satellites is far too high if satellites evolve using observed SFRs. This is a particular problem for semi-empirical models where one would ideally use the observed SFR as an input. To overcome the inconsistencies between observed SFRs and model predictions it is possible to include SFRs generated by the method above commonly referred to as a continuity SFR. Attributing the stellar mass growth to star formation in this way yields an upper limit to SFR that is consistent with the SMF evolution by design.

To properly constrain the formation of a galaxy one must reproduce the galaxy environment, the distribution of satellites around the central galaxy, at all previous redshifts. Discrepancies with observations of the high-redshift environment will cause modelled satellite stellar mass accretion rates that are either too high or too low. To account for such deficit/surplus modelled *in situ* growth must compensate through other modelling parameters to maintain the evolution of the stellar mass density. Such compensation could, for example, be of the form of suppressed/enhanced SFR or alternatively any number of other physical modelling parameters. Reproducing the number density and distribution of galaxies has, however, proven a challenge for many semi-analytic models (e.g. Asquith et al. 2018). Furthermore, where semi-analytic models have included more physics via an increased number of modelling parameters, this has led to degeneracies that obscure which are the essential physical processes governing galaxy formation (e.g. González et al. 2011; Lapi et al. 2011). Semi-empirical models, due to the direct initialization of galaxies to haloes and smaller

parameter spaces, fare better as they can by design provide more clarity as to which modelling assumptions and related parameters are necessary to fit observations.

In our previous work we presented STEEL (Grylls et al. 2019, hereafter referred to as [Paper I](#)), a Statistical sEmi-Empirical modeL. The basis of STEEL was the shift from discrete merger trees in favour of statistical halo growths and merging histories, which enables to probe galaxy environment unbiased by volume and mass resolution. In [Paper I](#) we presented a detailed study of the richness of the galaxy group and cluster environments in the local Universe. In this work we extend the analysis of satellite richness from STEEL to high redshifts comparing with a large galaxy cluster survey, and state-of-the-art hydrodynamical simulations. Having a clear and well-constrained picture of the building up of satellite population then allows STEEL to create more reliable merger histories for central galaxies across cosmic time. Using STEEL's improved merger histories we are then able to constrain the *ex situ* growth and, by extension, derive more reliable estimates of the *in situ* growth and the implied SFRs of central galaxies.

The paper is laid out as follows: In Section 2 we present the halo and stellar mass functions we use for our abundance matching as well as the high-redshift clusters we adopt to constrain the model performance. In Section 3 we discuss STEEL, most importantly we provide an overview of the statistical merging history in Section 3.1, the abundance matching in Section 3.3, as well as updates made to the model described in [Paper I](#). We begin the results by testing the high-redshift halo merger rate in Section 4.1, in Section 4.2 we then present the high-redshift satellite galaxy distribution results, and in Section 4.3 the growth of our galaxy population via *in situ* and *ex situ* processes. We then discuss our results in a wider context in Section 5 and summarize in Section 6.

## 2 DATA

In this section we first describe the simulations used for the halo mass functions (HMFs), then the data used to create the SMFs. Together, the HMF and SMF are used to create an SMHM relation, described in Section 3.3, which defines the galaxy–halo connection, essential to STEEL. We then provide details of the cluster data we compare to at high redshifts. All the data presented in this section are converted, wherever necessary, to a Chabrier (2003) stellar initial mass function (IMF). In this work we adopt the Planck cosmology with  $(\Omega_m, \Omega_\Lambda, \Omega_b, h, n, \sigma_8) = (0.31, 0.69, 0.05, 0.68, 0.97, 0.82)$ <sup>1</sup> (Planck Collaboration 2015). Halo masses are defined as virial masses in this cosmology, unless stated otherwise.

### 2.1 Halo mass functions

In this work we use the HMF from the simulations of Despali et al. (2016), generated and converted to appropriate units and cosmology using COLOSSUS (Diemer 2017). The HMF provides the number density of haloes in a given mass bin at a given redshift. We generate the substructure of subhaloes using the subhalo mass functions found in Jiang & van den Bosch (2016). The subhalo mass function provides the number density of subhaloes expected for a parent halo

<sup>1</sup>We note that Planck's best-fitting cosmology has slightly different parameters than those adopted in some of the observations used in this work, such as the SMFs ( $(\Omega_m, h = (0.30, 0.70))$ ). Correcting the SMF volumes and luminosities to the same cosmology yields essentially indistinguishable results.

of a given mass. The (sub) halo mass functions used in this work are all calibrated against the Bolshoi Simulation (Klypin et al. 2016).

## 2.2 Stellar mass functions

In this work we use the SMFs defined below, along with the HMF given above, to constrain the SMHM relationship. The latter in STEEL is constrained first at low redshift  $z = 0.1$ , using SMFs from the Sloan Digital Sky Survey (2.2.1). The evolution of the SMHM relation is then constrained to match the SMF at higher redshifts ( $z > 0.1$ ).

### 2.2.1 Low redshift, $z = 0.1$

At low redshift we use the Sloan Digital Sky Survey Data Release 7 (SDSS-DR7) from Meert, Vikram & Bernardi (2015). The data from the SDSS-DR7 spectroscopic sample (Abazajian et al. 2009) contain  $\sim 670\,000$  galaxies fitted with a Sérsic + exponential model (PYMORPH; Meert et al. 2015) with associated halo masses and central satellite classifications from Yang et al. (2012). The improved photometric analysis by Meert et al. (2015) provides more reliable estimates of the SMF at the high-mass end, which appear more abundant than previous estimates (Bernardi et al. 2016, 2017). In this work we investigate the effect of this enhanced high-mass end on galaxy assembly. We compare to previous determinations of the SMF using as an example the de Vaucoulers (de Vaucouleurs 1948) based cmodel fits from SDSS (Abazajian et al. 2009). The latter definition of galaxy stellar mass has been extensively discussed not to be accurate, partially due to incorrect sky subtraction and adoption of non-ideal light profiles (Bernardi et al. 2013). Bernardi et al. (2017) have clearly shown that the choice of light profile is not a simple matter of ‘semantics’. The single or double Sérsic models perform better in fitting the surface brightness of galaxies independently of the galactic environment (Meert et al. 2015). The performance is thus not related to the inclusion of the intragroup or intracluster light in the fit (Bernardi et al. 2017).

### 2.2.2 High redshift, $z > 0.1$

At higher redshift ( $0.3 < z < 3.3$ ) we use SMFs from the COSMOS2015 catalogue (Davidzon et al. 2017). Here masses are defined using spectral energy distribution fitting, including ultradeep infrared photometry. Davidzon et al. (2017) use Bruzual & Charlot (2003) stellar population synthesis models to estimate stellar masses. As SED fitting is notably different from light profile fitting, one cannot apply the same corrections as in Mendel et al. (2014). Nevertheless, to match the mass-to-light ratios adopted by Mendel et al. (2014), based on the Bell et al. (2003) mass-to-light ratios, we follow Bernardi et al. (2013) and increase the Davidzon et al. (2017) stellar masses, based on Bruzual & Charlot (2003), by  $+0.15$  dex. We note that the resulting  $z = 0.37$  SMF after this correction is in remarkable good agreement with the  $z = 0.1$  SMF by Bernardi et al. (2013). Our result also matches the findings by Bernardi et al. (2016), who showed that, by making use of the BOSS sample, the SMF shows negligible number density evolution up to  $z \sim 0.5$ .

## 2.3 Clusters

### 2.3.1 Cluster at $z = 2.5$ , Wang et al. 2016

The highest redshift cluster we compare to is an  $M_{\text{vir}} = 10^{13.7} M_{\odot}$  halo containing 15 galaxies with  $M_* > 10^{10} M_{\odot}$  at a redshift of

$z = 2.5$ . This cluster is reported in Wang et al. (2016), and we provide a brief description of the observation and data here. The cluster is observed using IRAM-NOEMA, VLT-KMOS, VLA, XMM-Newton, and Chandra for the spectroscopic observation and redshift determination. The galaxy masses are determined assuming a Salpeter (1955) IMF, which we correct to a Chabrier (2003) IMF, by decreasing the stellar masses by 0.24 dex. The halo mass ( $M_{\text{vir}} \sim 10^{13.93} M_{\odot}$ ) of the cluster is estimated in three different ways, using the total X-ray luminosity, the velocity dispersion of its member galaxies above  $M_* = 10^{10.76} M_{\odot}$ , and the stellar richness of the cluster.<sup>2</sup> Given this object was a targeted cluster, we cannot estimate the cosmic abundance (i.e. the number per cubic megaparsec). For analysis and comparison later in this work we assign this cluster an abundance of  $N(>M_* = 10^{13.93}) = 10^{-7.15} (\text{Mpc}^{-3})$  which is estimated by integrating the HMF in the limits  $[10^{13.93}, \infty]$ , thus providing an upper limit to the number densities associated with clusters of this mass.

### 2.3.2 1959 clusters at $z = 0.7$ – $1.0$ , Wen & Han 2018

We compare to the cluster sample from Wen & Han (2018), which contains 1959 clusters from SDSS-DR14 (Abolfathi et al. 2017) and the WISE survey (Wright et al. 2010). The clusters are identified in the  $WI$  band, and foreground objects are removed using the SDSS photometric data. The cluster mass and richness are estimated using the total  $WI$  band luminosity within 1 Mpc of the central galaxy. As performed above, to each cluster we assign an upper limit to their abundances from the cumulative integration of the HMF.

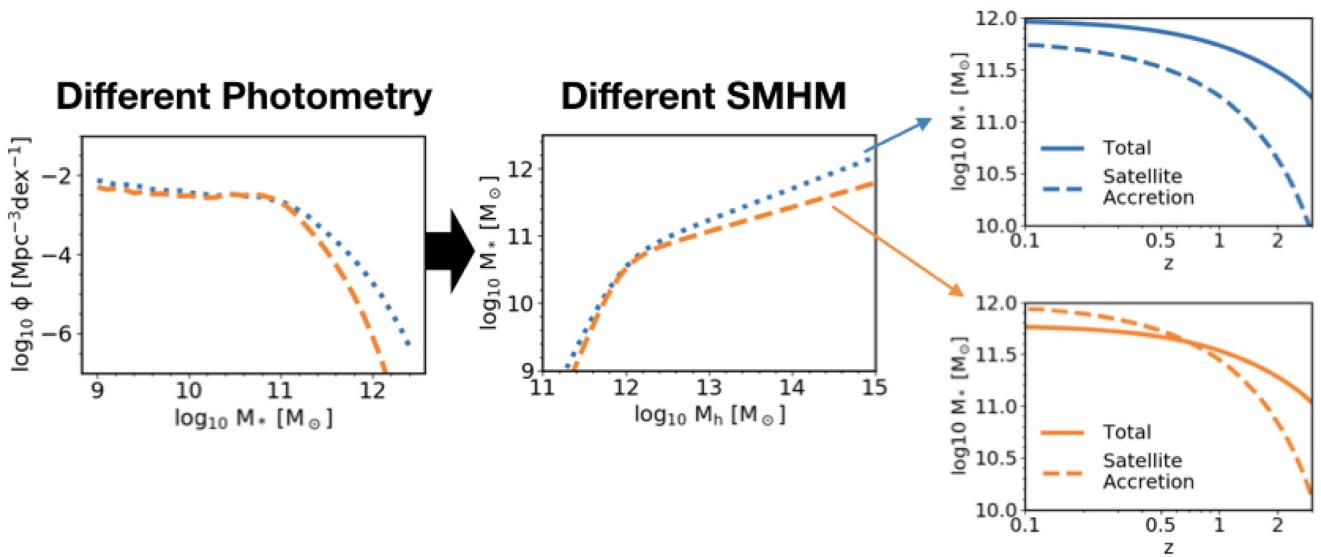
## 3 METHOD: STEEL

Our model STEEL is a STatistical sEmi-Empirical modeL designed to investigate the satellites and subhaloes in groups and clusters. In brief, STEEL removes reliance on discrete dark matter simulations or halo merger trees, commonly used in galaxy simulations, in favour of using mass functions to create a ‘statistical dark matter accretion history’ described in more detail in Section 3.1. This statistical history is then combined with semi-empirical techniques, such as abundance matching (Section 3.3), to create average galaxy population statistics. Whilst a comprehensive description of STEEL can be found in Paper I, we provide in this section highlights of STEEL, including any relevant updates. In this paper we have two objectives. First, we investigate the impact of different SMHM relations on the accretion histories of galaxies. The second aim is to use our semi-empirical accretion histories and galaxy growth histories to derive the SFR/star formation histories of galaxies.

The image in Fig. 1 shows a simple visualization of the process we use to determine the effect of photometry on the accretion histories of galaxy populations. Starting on the left we show two SMFs, the primary difference is the blue (dotted) SMF has a substantially enhanced high-mass end. In the middle panel we show how this high-mass slope changes the SMHM relation, an enhanced high-mass end SMF results in an enhanced high-mass slope. The galaxy growth histories, shown as solid lines, are generated using the SMHM relation used to calculate the average satellite stellar

<sup>2</sup>We note the velocity dispersions and X-ray luminosity estimations give the cluster mass as  $M_{\text{vir}} = 10^{13.73} M_{\odot}$  and the estimate given by mass richness is significantly higher  $M_{\text{vir}} = 10^{14.6} M_{\odot}$ , whilst we used the published average the lower cluster mass excluding the richness estimate is in as-good or better agreement with model results.

## Different Accretion



**Figure 1.** An image showing the steps we follow to connect the differences found in the stellar mass function (left) and the changes in the SMHM relation (SMHM, middle) that propagate into changes in the accretion histories (right). In the right-hand panel dashed lines are mass from satellite accretion and solid lines are total galaxy mass growth. Flatter SMHM relations imply a weaker growth of stellar mass in the central which can be easily overcome by the substantial cumulative growth of merging satellites, rendering the model internally inconsistent.

mass accretion associated with a given central halo mass history. It follows that the galaxy grown using the steeper relation from the enhanced SMF induces more galaxy growth. A flatter high-mass slope induces less growth in the limit where the high-mass slope is flat the central galaxy would not grow with increasing halo mass.

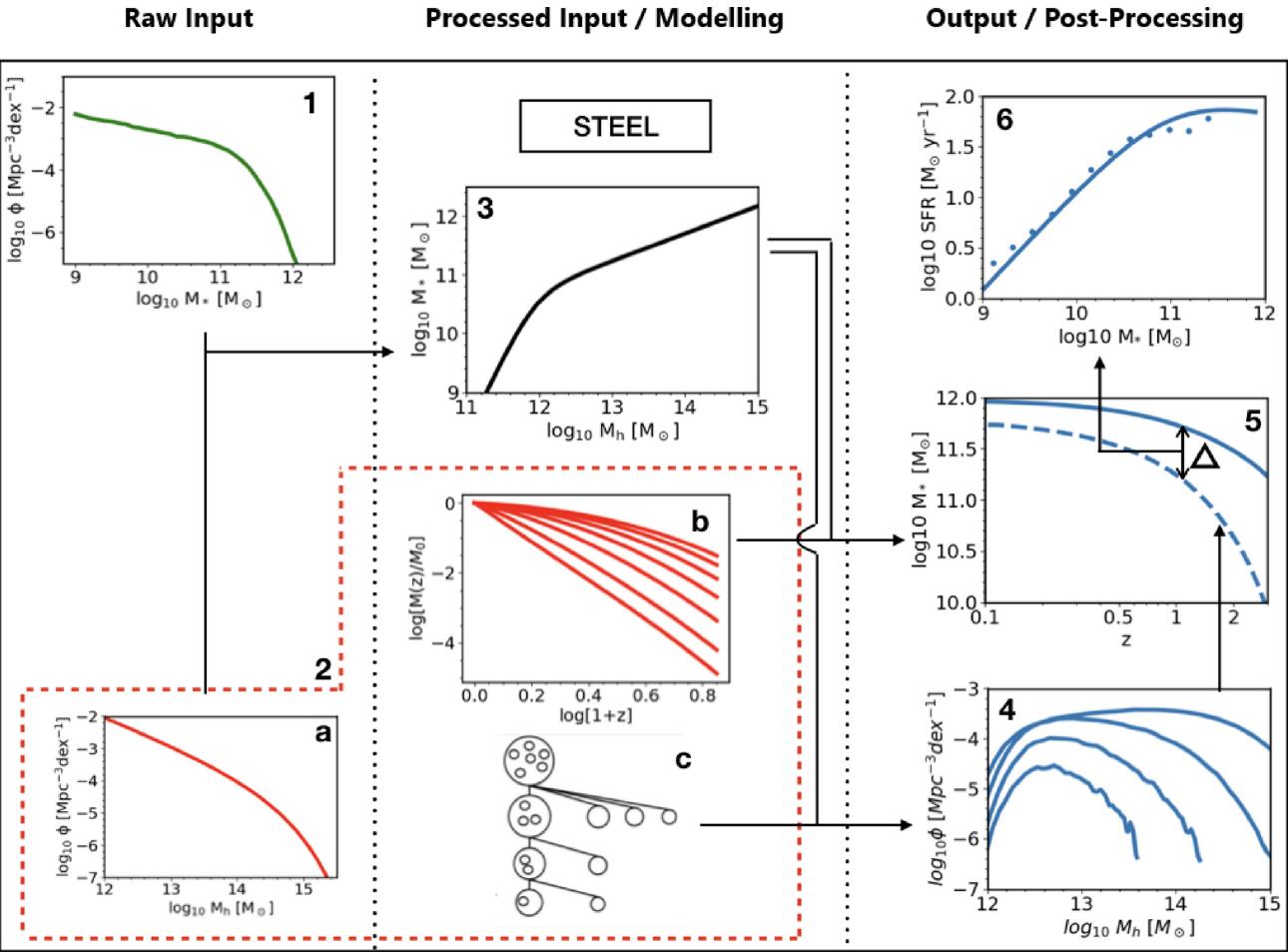
The dashed lines in the rightmost panels are created by calculating the average accretion on to a central galaxy. The majority of accretion is from galaxy mergers that have a low-mass ratio between the central and satellite galaxy. Note, the SMFs show little difference in number density for smaller galaxies (below  $M_* = 10^{11} M_\odot$ ) and thus the low-mass slope of the SMHM relation is unchanged. It follows that the smaller galaxies in a given dark matter assembly history are unchanged and the accretion is similar for both centrals. For the bottom right hand panel showing the accretion history and growth history for a galaxy using the lower SMF it is found that the accretion exceeds the galaxy growth (in this image this is accentuated for clarity). Whereas, for the enhanced SMHM relation the accretion is below that of the galaxy growth. The satellite galaxy accretion history and the central galaxy growth history in a given cosmology are determined by the SMHM relationship and the dark matter halo assembly. In this paper we describe a method that, in a given cosmology, can exclude a set of SMFs over multiple redshift epochs. The evolution of these SMFs is analogous with the growth of the total stellar mass in the Universe over cosmic time. The consistency of the galaxy growth histories and the satellite accretion histories is checked by generating and comparing the ratio of satellite accretion and total mass growth. If the total accretion or rate of accretion is greater than that of the central galaxy mass or galaxy growth rate the set of SMFs is incompatible with the specific  $\Lambda$ CDM cosmology.

The image in Fig. 2 shows a simplification of the processes we follow to derive the SFR by following galaxy populations along their halo mass histories. The plot labelled 1 (green) is the input SMF. The box in red is the statistical dark matter accretion history described in Section 3.1, including the HMF (2a), the central growth

histories (2b), and the halo substructure (2c) shown here as a discrete merger tree for visualization purposes. Using the abundance matching routines described in Section 3.3, the SMF (1) and the HMF (2a) are used to create the SMHM relationship (3, black). In Paper I we showed how the dark matter accretion histories (2) and abundance matching (3) can be used to generate distributions of satellites for any central halo. In this work we generate satellite distributions for central haloes at multiple redshifts (4) then test them against simulations and observations in Section 4.2. For each central halo mass track (2b) the average number density of satellites that reach the centre of the halo and merge with the central galaxy, is calculated thus generating the average satellite accretion history (dashed line, 5). Using the central halo growth histories (2b) and the SMHM relation (3), we can generate the average central galaxy growth history (solid line, 5). These two quantities can be compared to check for self-consistency, as described above and shown in Fig. 1. Where a self-consistent central growth and accretion history is found, any deficit between the accreted mass and the growth history is attributed to SFR (delta, 5). The derived SFR for central galaxies (solid line, 6) is compared to observational data (points, 6). Where the SFR prediction generated from the model is found to be consistent with the observed SFR this is a good indication that the model is correct. Additional observational constraints not shown in Fig. 2 can be added to improve the analysis, such as the specific SFR (ssSFR) distribution which is discussed in Section 4.4. In future work other constraints such as the pair fraction of galaxies and the intracluster light generated from dynamical process during satellite accretion will also be considered.

### 3.1 Statistical merging history

Common modelling techniques, such as hydrodynamical, semi-analytic, and traditional semi-empirical models rely on a discrete set of haloes within a simulation ‘volume’. These discrete haloes come



**Figure 2.** An image showing the constituent steps of the method to generate SFRs in this paper. In brief, the three columns from left to right are raw inputs, derived inputs/modelling, and output/post-processing. The subplots are: (1) The SMF, (2a) The HMF, (2b) Halo mass growth histories, (2c) Accretion histories/Merger tree, (3) The SMHM relation, (4) Group/Cluster satellite richness, (5) Central growth histories/satellite accretion histories, (6) SFR. The SFRs are derived from the difference between the total growth in stellar mass and that from satellite accretion (panel 5).

in three forms: an  $N$ -body cosmological box, post-processed merger trees, or catalogue of haloes.<sup>3</sup> Due to the inevitably restrained simulated cosmic volumes, all the aforementioned models are biased towards smaller haloes and galaxies, largely missing a statistically comprehensive description of the most massive central and satellite galaxies.

STEEL is instead designed to model all haloes (and subhaloes) within the simulation range without volume or resolution constraints. We remove the dependence on discrete halo sets through the use of a *statistical* dark matter backbone, and simulate all haloes and all mass-ratio mergers<sup>4</sup> with equal weights regardless of their number density. The following steps represent the core method for creating the statistical dark matter accretion histories.

<sup>3</sup>We note that in many cases merger trees and halo catalogues are extracted from the cosmological box of a dark matter only  $N$ -body simulation. The alternatives being Press-Schechter analytically derived merger trees (e.g. Press & Schechter 1974; Parkinson, Cole & Helly 2008) and halo catalogues obtained by sampling the HMF.

<sup>4</sup>Within the simulation mass range, however, this can be set arbitrarily wide as long as the choice of HMF and empirical techniques are valid in the proposed mass range.

(i) At the redshift of interest  $\bar{z}$  we start from the HMF to compute the abundances of (parent/central) haloes for any given central halo bin  $[M_{\text{h,cent}}(\bar{z}), M_{\text{h,cent}}(\bar{z}) + dM_{\text{h,cent}}(\bar{z})]$ .

(ii) Each parent/central halo mass is then followed backwards in time following its average mass growth history,  $\langle M_{\text{halo}}(\bar{z}) \rangle$ , calculated using the routines from van den Bosch et al. (2014).

(iii) An unevolved subhalo mass function<sup>5</sup> (Jiang & van den Bosch 2016) is then assigned to each central halo mass bin  $[M_{\text{h,cent}}(\bar{z}), M_{\text{h,cent}}(\bar{z}) + dM_{\text{h,cent}}(\bar{z})]$  at each redshift epoch.

(iv) At each time-step we then calculate the difference in subhalo population between  $z$  and  $z + dz$  to estimate the average number density and masses of subhaloes accreted on to the main progenitor in the redshift interval  $dz$ , which we call the ‘accreted’ subhalo mass function.

(v) Each bin  $[M_{\text{h,sub}}, M_{\text{h,sub}} + dM_{\text{h,sub}}]$  of the accreted subhalo mass function is then assigned a dynamical time given the central halo mass bin  $[M_{\text{h,cent}}(\bar{z}), M_{\text{h,cent}}(\bar{z}) + dM_{\text{h,cent}}(\bar{z})]$  it corresponds to.

(vi) At each redshift epoch we then sum the number densities of subhalo bins  $[M_{\text{h,sub}}, M_{\text{h,sub}} + dM_{\text{h,sub}}]$  on each central halo mass

<sup>5</sup>The unevolved subhalo mass function gives the total number density of subhaloes of each mass accreted on to a given central halo over its entire growth history.

track that have not exceeded their dynamical time to create the surviving subhalo mass function.

(vii) Given the infall redshift, mass, and number densities of each subhalo mass bin  $[M_{\text{h,sub}}, M_{\text{h,sub}} + dM_{\text{h,sub}}]$  we can convolve the resulting satellite halo distribution with the SMHM relation to create the observed distribution of satellites at any epoch. In some model variants we also include additional physical processes to account for the late evolution of satellites after infall (see Paper I).

In this work we also track the number densities of subhaloes that have reached the end of their dynamical time at each epoch. At the time a subhalo reaches its dynamical time the associated satellite galaxy ‘merges’<sup>6</sup> with the central galaxy, during these mergers we inject a fraction (40 per cent) of the satellite mass to the intracluster medium (Moster et al. 2018).

### 3.2 Satellite quenching

The satellite quenching model in STEEL, presented in Paper I, has two components:

(i) A delayed-then-rapid quenching model (Wetzel et al. 2013), according to which satellite galaxies upon entering a halo continue to form stars as if they were on the star formation main sequence until their quenching time-scale,  $\tau_q$ , has elapsed. After a time  $\tau_q$  the SFR of the satellites is rapidly quenched over the fading time-scale  $\tau_f$ .<sup>7</sup>

(ii) The second component used is the halo mass-dependant cut-off (Fillingham et al. 2018), which envisions that satellite galaxies below a given stellar mass (dependant on host halo mass) are immediately quenched.

The delayed-then-rapid quenching model is improved using the latest dynamical quenching results from Cowley et al. (2019) updating the model presented in Paper I to include a redshift dependence in all quenching time-scales

$$\begin{aligned} \tau_{q,z} &= \tau_q * (1 + z_{\text{infall}})^{-3/2}, \\ \tau_{f,z} &= \tau_f * (1 + z_{\text{infall}})^{-3/2}. \end{aligned} \quad (1)$$

Additionally, we include a pre-processed fraction, inspired by the results of Wetzel et al. (2015), implementing a mass-dependent fraction such that a minimum of 30 per cent and a maximum of 60 per cent of galaxies are pre-possessed with a transitional mass range between  $10^6$  and  $10^8 M_\odot$ . The SFR of a satellite galaxy after infall is then given by

$$\text{SFR}(t, t_{\text{infall}}, M_{*,\text{infall}}) = \text{SFR}_{t_{\text{infall}}} * \begin{cases} 1, & t > t_q \\ e^{-\frac{t_q-t}{\tau_f}}, & t < t_q. \end{cases} \quad (2)$$

### 3.3 Abundance matching

In this work we populate dark matter haloes with galaxies using the abundance matching technique where galaxies are assigned to haloes by comparing the relative abundances of galaxies and haloes. For the abundance matching we use the central haloes from the HMF

<sup>6</sup>In a statistical model satellite galaxies are not strictly merging as there is no central galaxy to merge with, instead we collect statistics of merging satellites at each epoch. Using post-processing techniques we use these statistics obtain information on the average merging history of central galaxies.

<sup>7</sup>The absolute quenching time is given by  $t_q (= t_{\text{infall}} - \tau_q)$ .

described in Section 2.1, and a subhalo mass function subdivided by redshift of infall generated from STEEL. Subhaloes are assumed to follow the central SMHM relation at infall. We simplify our abundance matching by using a frozen model such that baryonic evolution after infall (stripping, star formation, etc.) is not included. The latter assumption provides a good approximation as after infall the dominant factor determining the abundances of satellite galaxies is the dynamical time and not evolutionary processes (Paper I).

To fit SMFs over multiple epochs we convolve our HMFs with a parametric SMHM relation similar to that proposed by Moster et al. (2010),

$$\begin{aligned} M_*(M_h, z) &= 2M_h N(z) \left[ \left( \frac{M_h}{M_n(z)} \right)^{-\beta(z)} + \left( \frac{M_h}{M_n(z)} \right)^{\gamma(z)} \right]^{-1} \\ N(z) &= N_{0.1} + N_z \left( \frac{z - 0.1}{z + 1} \right) \\ M_n(z) &= M_{n,0.1} + M_{n,z} \left( \frac{z - 0.1}{z + 1} \right) \\ \beta(z) &= \beta_{0.1} + \beta_z \left( \frac{z - 0.1}{z + 1} \right) \\ \gamma(z) &= \gamma_{0.1} + \gamma_z \left( \frac{z - 0.1}{z + 1} \right). \end{aligned} \quad (3)$$

In what follows we adopt both the cmodel and PYMORPH SMF described in Section 2.2 at redshift  $z = 0$  to constrain the parameters  $N, M, \beta$ , and  $\gamma$  (normalization, knee, low-mass slope, and high-mass slope). We use only the central SMF, using the Yang et al. (2012) central/satellite identification, and central HMF. The fit is performed using a Markov Chain Monte Carlo (MCMC), implemented using the PYTHON package EMCEE (Foreman-Mackey et al. 2013), over a large parameter space ( $P_{M,N,\beta,\gamma}$ ) covering all four parameters. Given a point in parameter space  $P_{M_i, N_i, \beta_i, \gamma_i}$ , the SMF is constructed using the HMF and the SMHM relation. Each bin of parent halo mass is associated with a Gaussian distribution of stellar mass with scatter 0.15 dex. This distribution is multiplied by the halo mass number density to convert to galaxy number density which are added to the relevant stellar mass bins of the SMF in construction. This operation is then repeated over all mass bins of the HMF to produce the complete central SMF. For each point  $P_{M_i, N_i, \beta_i, \gamma_i}$  in the parameter space, the SMF associated with that point is compared via a likelihood function to the observed SMF to provide the MCMC with the probability that the given point is the ‘true’ SMHM relationship.

We then fit to the Davidzon et al. (2017) data both uncorrected and corrected for the cmodel and PYMORPH fits, respectively (see Section 2.2 for details). At high redshift we use the central and subhalo mass functions initializing satellites at infall as described above.<sup>8</sup> For central haloes the method is the same as detailed above, however, as we use the total SMFs at high redshift we also include the total unevolved surviving subhalo mass function in the abundance matching. We assume that a halo before infall hosts a central galaxy; under this assumption we use the central SMHM relation to assign satellite galaxy stellar mass at the point of accretion. For the latter we must have information about the redshift of infall for subhaloes. We obtain from STEEL the unevolved

<sup>8</sup>Ideally, as for low redshift, we would use the centrals only as we are primarily concerned with the central SMHM relation. However, lacking a well-defined central SMF at high redshift, this method represents a reliable way to extend the model to higher redshifts.

**Table 1.** The abundance matching results for the cmodel and PYMORPH data. The errors are the 16th and 86th percentile from the MCMC fitting.

	$M_n$	$N$	$\beta$	$\gamma$	$M_{n,z}$	$N_z$	$\beta_z$	$\gamma_z$
cmodel	$11.91^{+0.40}_{-0.34}$	$0.029^{+0.018}_{-0.013}$	$2.09^{+1.21}_{-1.02}$	$0.64^{+0.11}_{-0.10}$	$0.52^{+0.24}_{-0.19}$	$-0.018^{+0.005}_{-0.004}$	$-1.03^{+0.049}_{-0.34}$	$0.084^{+0.20}_{-0.14}$
PYMORPH	$11.92^{+0.39}_{-0.36}$	$0.032^{+0.016}_{-0.012}$	$1.64^{+0.85}_{-0.73}$	$0.53^{+0.11}_{-0.11}$	$0.58^{+0.15}_{-0.19}$	$-0.014^{+0.007}_{-0.006}$	$-0.69^{+0.29}_{-0.36}$	$0.03^{+0.154}_{-0.147}$

surviving subhalo mass function as contributed by each redshift of infall. Each contributing part is calculated using the SMHM relation at the redshift of infall and added to the central SMF using the same method as with the centrals. The total SMF is compared, at each redshift step available, to the data via the likelihood function to give the probability that the given point is the ‘true’ evolution parameters. The abundance matching best-fitting parameters and associated errors for both the cmodel and PYMORPH are given in Table 1, and plots showing the cross-sections of the parameter space are shown in Appendix A.

In Fig. 3 we show the results of our abundance matching to the PYMORPH and cmodel central SMFs. The PYMORPH fit is steeper above the knee compared to either the cmodel or the Moster et al. (2013) model fits, as expected given the larger number density of massive galaxies found applying the Sérsic–exponential model (e.g. Shankar et al. 2014; Kravtsov, Vikhlinin & Meshcheryakov 2018). The low-mass slope for both PYMORPH and cmodel are almost identical as the galaxies in this range are not affected by the photometric choice. Differences between the fits from this work and Moster et al. (2013) are due to our selection of using only central haloes/galaxies as opposed to the total population, and the SMFs shown in the right-hand panel are lower than even cmodel are therefore missing massive galaxies.

## 4 RESULTS

### 4.1 Halo merger rates

STEEL implements a statistical dark matter merging history, thus as a very first step we check STEEL’s performance on reproducing halo merger rates as extracted from  $N$ -body dark matter-only simulations. We explore the evolution of the merger rate of haloes with a mass ratio greater than  $f = M_{\text{h,sat}}/M_{\text{h,cen}}$ . The merger rate is calculated from STEEL by integrating the ‘unevolved subhalo mass function accretion’ ( $\delta$ USHMF) above the mass ratio limit

$$\frac{dN}{dz}(M_{\text{h,cen}}) = \int_{M_{\text{h,cen}}*f}^{\infty} \delta \text{USHMF}([z, z + \delta z], M_{\text{h,cen}}, M_{\text{h,sat}}) dM_{\text{h,sat}}. \quad (4)$$

In Fig. 4 the merger rate from STEEL, shown by lines, is in good agreement with the best-fitting merger rate relations from the Millennium simulation given by Fakhouri, Ma & Boylan-Kolchin (2010), shown as shaded regions.<sup>9</sup> The slight deviation at low redshift derives from the predicted growth history of our input potential wells given by van den Bosch et al. (2014), a lower mass growth rate leads to a lower accretion rate. This deviation is due to

<sup>9</sup>It should be taken into consideration that the results from STEEL presented here and the fits from Fakhouri et al. (2010) are in different cosmologies. We show STEEL halo accretion using the Millennium cosmology used in the aforementioned work in Appendix D.

differences in the algorithms used to link haloes between simulation outputs and build merger trees used by Fakhouri et al. (2010) and van den Bosch et al. (2014)

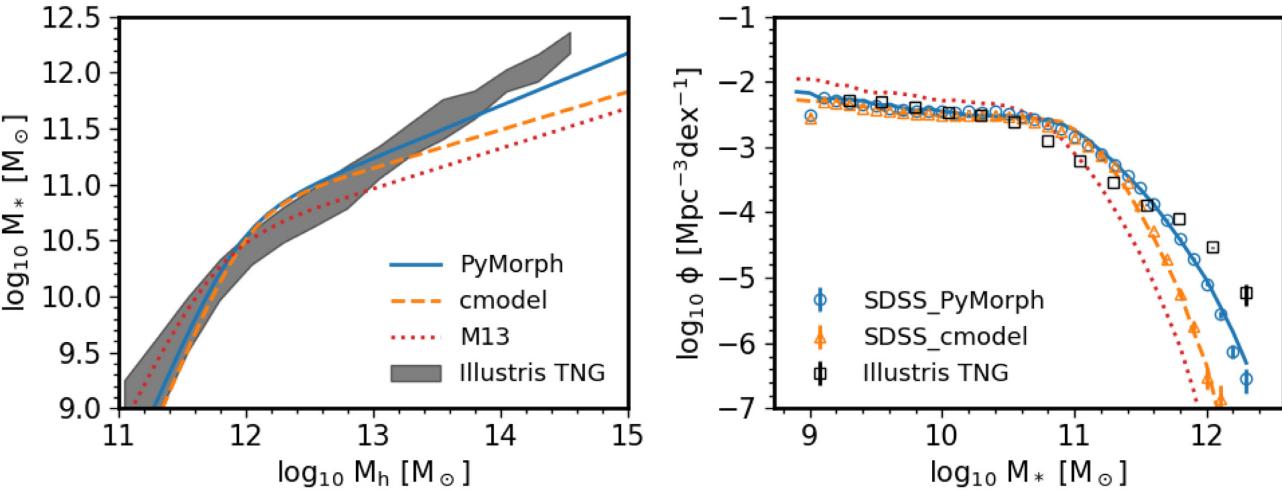
### 4.2 High-redshift clusters

We here extend the group and cluster satellite richness analysis from Paper I to high redshift. In Paper I it was found that dynamical friction and, to a second order, abundance matching, are the dominant factors in the distribution of satellite galaxies in groups and clusters above  $M_{*,\text{sat}} > 10^{10} M_\odot$ . In this section, for a more rounded view of the satellite galaxy population, we display the results for the full STEEL model which includes star formation, dynamical quenching, and stripping to evolve satellites after infall. The latter effects, despite being of lower order than dynamical friction or abundance matching, are included to be able to compare to data other than cluster richness, such as the satellite sSFR distribution.

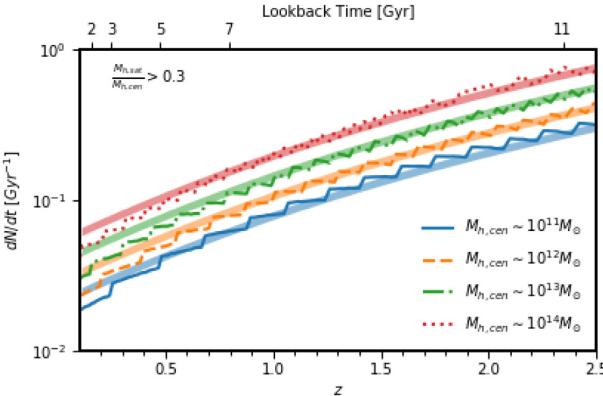
Fig. 5 shows the satellite number density per halo mass bin. For each central halo mass the cosmic number density, similarly to the number density presented in the cumulative SMFs, is calculated for satellites above a mass threshold for each halo mass bin. The predicted halo richness from STEEL, using the PYMORPH SMHM relation, is shown in this plot as solid lines. Low-redshift SDSS data are shown as a grey band, cluster data detailed in Section 2.3 are open symbols. The predictions from the Illustris TNG100 simulation (Nelson et al. 2018; Springel et al. 2018) are shown with crosses. The markers and lines in the figure are colour coded based on redshift, as indicated by the colour bar on the right.

In Fig. 5 we see that at higher redshift there are fewer massive satellites overall. This reduction is caused by several contributing effects. First, at high redshift there is an absence of high-mass haloes, which have not had time to form. Due to the lack of hosts at higher masses, the right hand side of the distribution tightens. Secondly, at high redshift the HMF is lower at any given mass, causing a shift of the satellite host halo distributions towards lower masses. Finally, as the process of formation and merging takes several gigayears to complete massive satellites are found less frequently at high redshift, and thus the number densities of the most massive satellites reduce faster than the lower mass ones.

We show STEEL is a good match to the Wang et al. (2016) cluster at redshift  $z = 2.5$ . We also achieve an adequate match to the cluster survey from Wen & Han (2018), especially in the mass range above  $M_{\text{sun}} > 10^{10.5} M_\odot$ . We also achieve similar results to the Illustris TNG100 simulation, though the TNG100 output is marginally higher at all redshifts. STEEL improves upon TNG100 in terms of the shape and breadth of this distribution. For example, at high mass and redshift STEEL resolves the turnover for satellites above  $M_{\text{sun}} > 10^{11} M_\odot$ , whereas TNG100 is too limited in volume to cover the high-mass ranges. The limitations in volume prevent TNG100 from simulating clusters such as those presented in Fig. 5 (Wang et al. 2016; Wen & Han 2018). In this respect STEEL becomes an excellent bridge between the capabilities of a high-resolution



**Figure 3.** Left: The SMHM relation at redshift  $z = 0.1$ . The PYMORPH (blue solid line) and cmodel (orange dashed line) fits from this work are both for central haloes/galaxies, the fit from Moster et al. (2013) (hereafter M13, red dotted line) is for all haloes/galaxies. The grey band is the relation from Illustris TNG100. Right: SMFs created using the central HMF and the three SMHM relations compared to PYMORPH (blue circles) and cmodel (orange triangles) central SMFs. The black squares are the SMF from Illustris TNG100.



**Figure 4.** The evolution of merger rate per Gyr at fixed halo mass. Lines are from STEEL, shaded bands are the analytic fits from Fakhouri et al. (2010). Halo masses shown are  $M_{h,\text{cen}} : 10^{11}, 10^{12}, 10^{13}, 10^{14} M_\odot h^{-1}$  as labelled.

hydrodynamical simulation and the massive cluster observations at high redshift.

#### 4.3 Connecting central mass accretion to star formation rate

We calculate from STEEL, using the PYMORPH SMHM relation, the relative contributions to the average stellar mass growth of central galaxies from satellites and star formation history (SFH), as shown in Fig. 6. For three galaxy mass bins ( $10^{11}, 10^{11.5}, 10^{12} M_\odot$ ) selected at  $z = 0.1$ , the average growth history (total, solid lines) is derived by following the host halo mass track, and the stellar mass track is implied by imposing abundance matching at all redshifts. The stellar mass history assigned by abundance matching, is naturally independent of any galaxy merger modelling assumptions from STEEL. The total accretion from satellites (accretion, dashed lines) is computed from the expected satellite accretion along halo mass tracks. For each galaxy an SFH (dotted lines) may then be calculated. The SFR is tuned such that it provides the correct SFH to account for the difference between the mass growth expected from abundance matching and the cumulative satellite stellar mass

accretion.<sup>10</sup> When calculating this difference we also take into account the stellar mass-loss rate (MLR) due to stellar recycling using the relations from Moster et al. (2018),

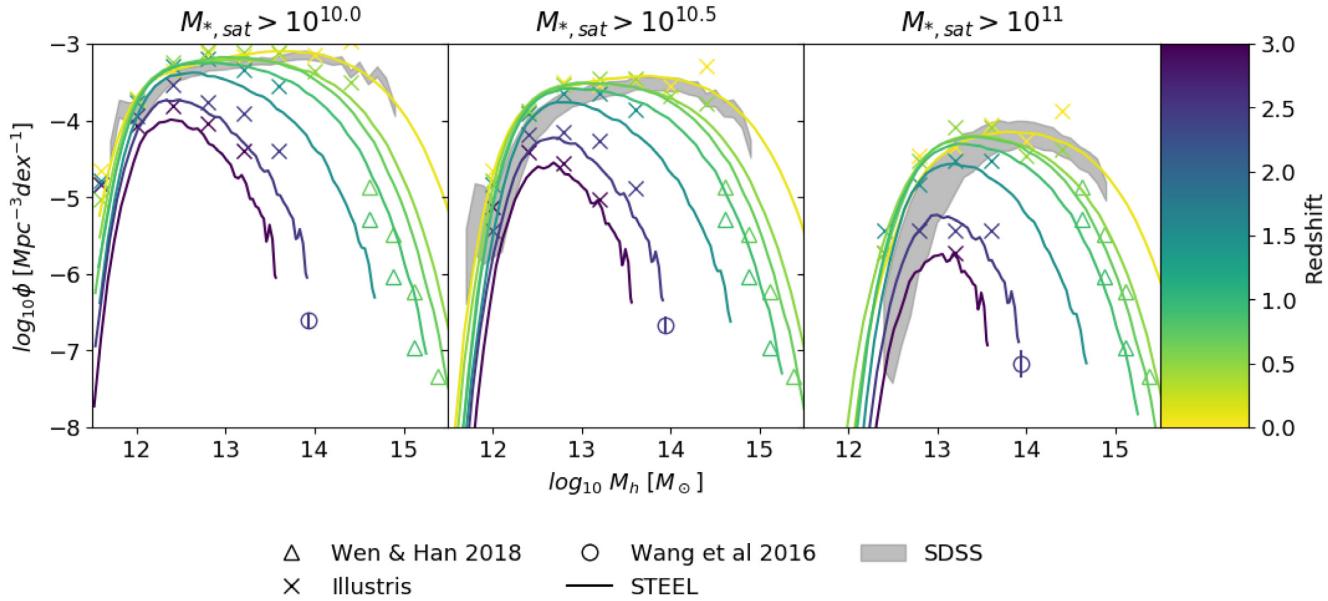
$$f(\tau_{\text{ml}}) = 0.05 \ln \left( \frac{\tau_{\text{ml}}}{1.4 \text{ Myr}} + 1 \right), \quad (5)$$

$$\text{MLR}(t) = \frac{\sum_{t'=\text{infall}}^t \text{SFH}(t')(f[t' - (t - \delta t)] - f[t' - t])}{\delta t}. \quad (6)$$

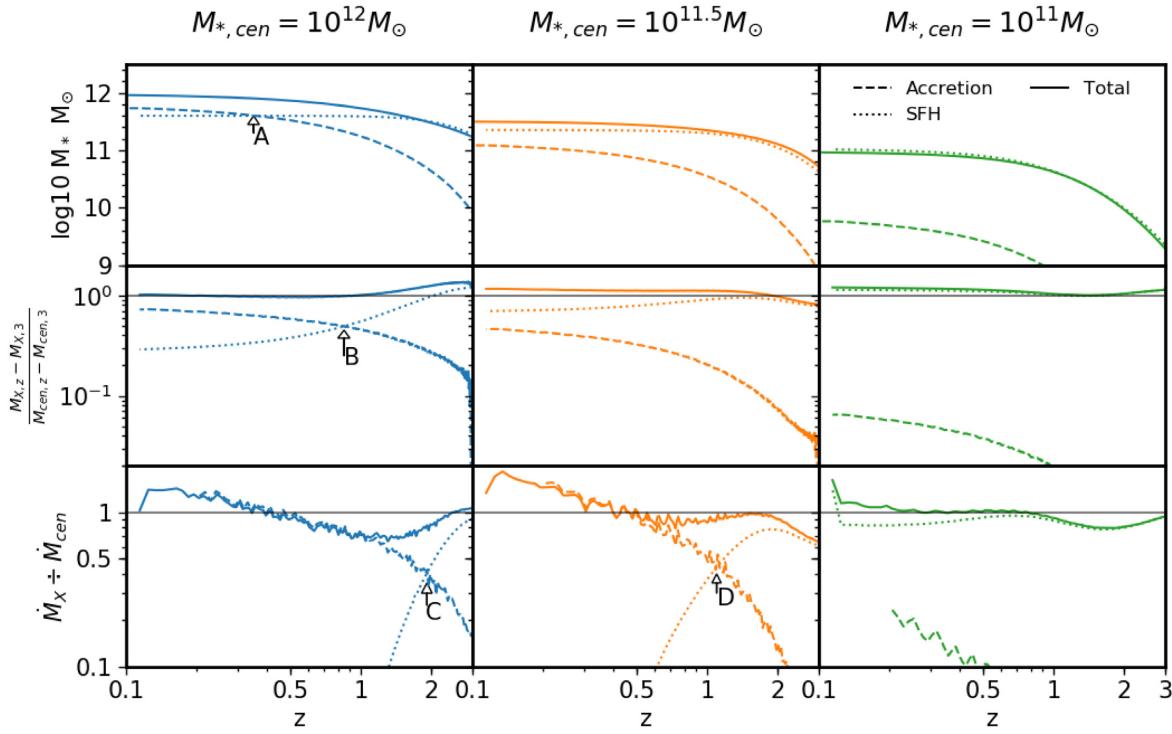
The SFR–stellar mass relation derived from this method is fitted with a double power law that evolves with redshift (for more details on the fit see Appendix B), as shown in Fig. 7. At lower redshift the normalization decreases, the peak of the distribution shifts to lower masses, and the turnover after the peak is steeper. In Fig. 7 we also show the same three galaxy population tracks from Fig. 6, discussed below, as black lines. These tracks show how the galaxy population evolves in SFR with redshift. The population tracks show a gradual increase in SFR and then a turnover before dropping sharply, as they transition to a satellite accretion-dominated regime. It is found that smaller galaxies grow for longer time-scales with increasing star formation, whilst larger galaxies start with higher SFR and transition to an accretion-dominated phase much earlier in time.

The top row of Fig. 6 shows the total mass of the galaxy and the total contributed by both accretion and the SFH. The middle row shows the fractional contributions from star formation and satellite accretion from  $z = 3$ . The bottom row shows the instantaneous mass growth from star formation and satellite accretion. There are two definitions we can use to determine the epoch after which a galaxy transitions into a merger-dominated state. First, we could define the ‘cumulative transition’ as when the galaxy has accreted more mass than it has created from star formation processes

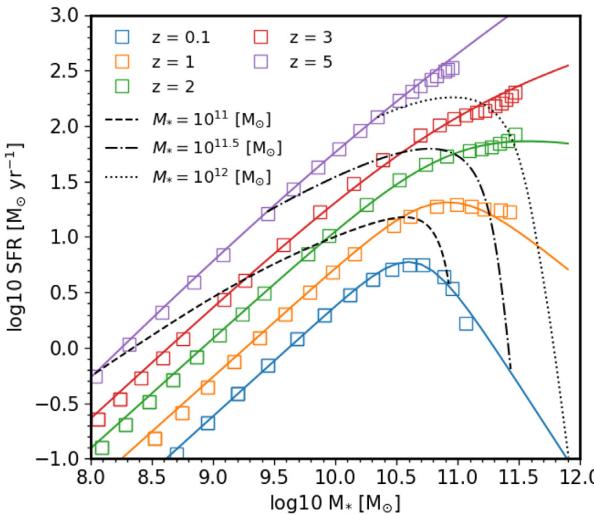
<sup>10</sup>This method directly links the SFR to the accreted mass from satellites. However, in our model satellites to grow in mass after infall (i.e. ‘non-frozen’ to follow the terminology of Paper I), we therefore recalculate the full satellite accretion on to the central galaxies updating their mass using the new SFR. Using the updated accretion the SFR is recalculated beginning an iterative process. However, this iterative process of recalculation ends after one loop as the rederived accretion is found to be nearly identical, as expected from the results of Paper I.



**Figure 5.** The number density distribution of satellites per parent halo mass predicted from STEEL, using the PYMORPH SMHM relation, at multiple redshift epochs (solid lines). The grey band is the data from SDSS at redshift  $z = 0.1$ . Also included are the high-redshift cluster data from Wang et al. (2016) (circles) and Wen & Han (2018) (triangles). We also compare to the outputs from the Illustris simulation using the TNG100 data (crosses). Each data point and line are given a colour associated with their redshift (the bar on the right provides the colour coding key).



**Figure 6.** Three ‘mass tracks’ are shown that have central galaxy masses at redshift  $z = 0.1$  of  $M_{*,\text{cen}} = 10^{12}, 10^{11.5}$ , and  $10^{11} [\text{M}_\odot]$  in blue, orange, and green, respectively. The satellite galaxy accretion is shown for evolved satellites with a dashed line, and the mass from star formation shown with a dotted line. The top panels show the total mass of the central (solid lines) and the total mass gained from accretion or star formation. The middle panels show the fraction of the total galaxy mass formed from satellite accretion or star formation since redshift  $z = 3$ . The bottom panels show the ratio of the mass accretion rate from satellite galaxies, the SFR, and the mass growth rate of the central galaxy predicted by abundance matching. The black horizontal lines in the second and third rows are at unity. The solid lines showing the sum of the other two factors should be close to or on the unity lines. The labels A and B point to where the cumulative mass from accretion overtakes the cumulative mass from star formation. The labels C and D point to where the instantaneous accretion overtakes the SFR.



**Figure 7.** The SFR–stellar mass relation derived from following central galaxy populations along halo mass histories at redshifts  $z = 0.1, 1, 2, 3, 5$ . The data extracted from the post-processing of STEEL are shown by coloured crosses and the double power-law fits are shown as lines in corresponding colours. The three black lines are the evolution of the galaxy populations selected at redshift  $z = 0.1$  with masses  $M_* = 10^{11}, 10^{11.5}, 10^{12} (\text{M}_\odot)$  presented in Fig. 6.

(Points A and B). Secondly, we define the ‘instantaneous transition’ as the epoch when the growth rate from mergers overtakes the growth rate from star formation (Points C and D). More massive galaxies transition earlier to merger dominated growth under both definitions. However, all galaxies transition earlier under the second (instantaneous) definition. The masses shown in Fig. 6 show three cases of relevant galaxy accretion tracks. The  $M_*^{z=0} = 10^{12} \text{ M}_\odot$  galaxy growth curve at low redshift is always dominated by satellite accretion. In the top and middle rows we see that more mass has been accreted than produced by star formation, and in the bottom row we see the accretion rate overtook the SFR at redshift  $z = 2$ . The  $M_*^{z=0} = 10^{11.5} \text{ M}_\odot$  galaxy growth curve has more mass created from star formation than satellite accretion. However, the galaxy population has a higher rate of accretion rate than SFR since redshift  $z = 1$ . The final population shown at  $M_*^{z=0} = 10^{11} \text{ M}_\odot$  is star formation dominated under both cumulative and instantaneous definitions. At redshift  $z = 0$  we find the transition masses for the total mass ratio and the instantaneous ratio to be at  $M_* = 10^{11.7} \text{ M}_\odot$  and  $M_* = 10^{11.1} \text{ M}_\odot$ , respectively.

We show in Fig. 8 the satellite accretion using the same template as Fig. 6. In Fig. 8 we obtain a lower limit for the accretion rate by including stripping but not star formation in the satellites thus minimizing their mass through environmental processes. We find for the high-mass galaxies, which are above the knee of the SMHM relation, even the lower limit for the accretion has an instantaneous rate greater than the growth rate of the galaxy as seen in the bottom row. This makes the cmodel SMHM relation used within our dark matter accretion model *non-physical*: steeper SMHM relations, such as the one found with the PYMORPH photometry, are favoured by hierarchical assembly. We recall, as explained in Fig. 1, that too flat SMHM relations introduce global stellar mass growth histories that are even lower than what is expected from total satellite accretion rendering the models internally inconsistent. For completeness we also tested a range of dynamical time options. In all cases, even when the merging

time is increased by a factor of two,<sup>11</sup> the accretion rate exceeds the growth rate and cmodel photometries can be excluded. Further to this we also tested variations on the mass-loss in mergers and the amount of mass lost to tidal stripping of the satellites; in the case where the tidal is doubled with respect to the reference model and the mass-loss during a merger is set to 60 per cent up from 40 per cent, the cmodel photometries remain internally inconsistent. We are confident that under all circumstances cmodel photometries can be considered internally inconsistent.

#### 4.4 Specific star formation rate distribution

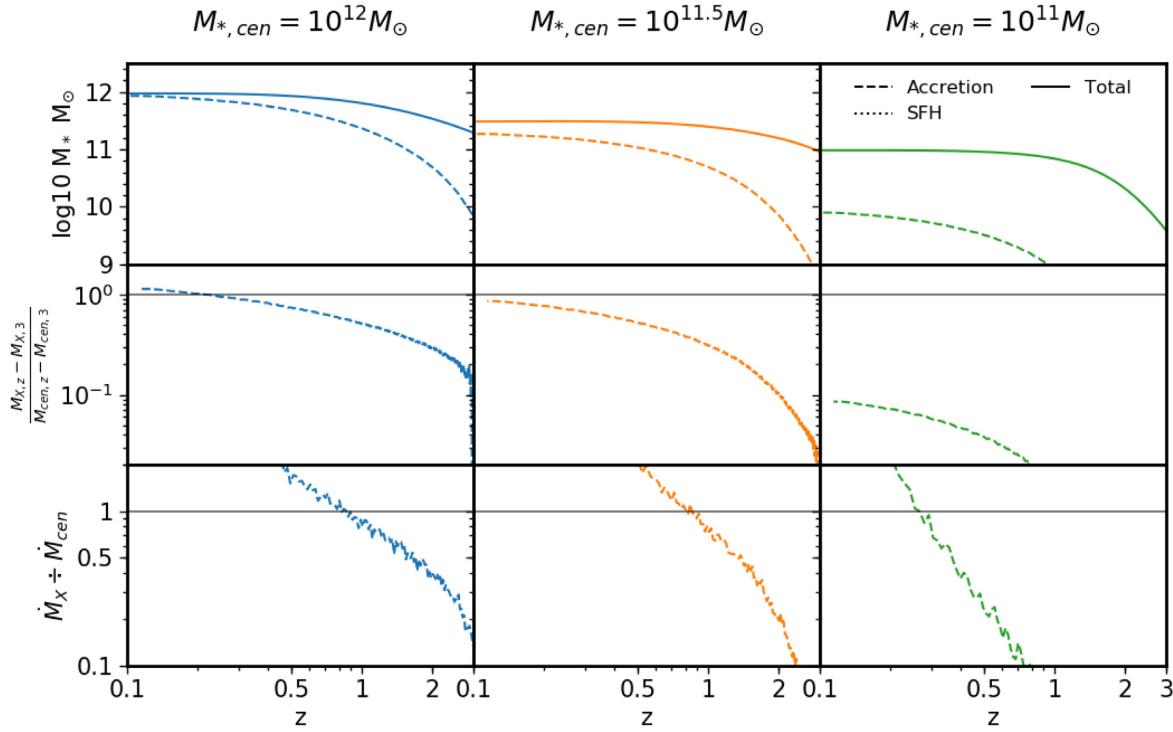
Fig. 9 shows the sSFR distribution of satellites in three mass ranges, as labelled, chosen to probe transitions found in observational data (Bernardi et al. 2011, 2014; Cappellari et al. 2013). The solid blue line and the dashed black lines show the satellite and central sSFR from STEEL, respectively, while the grey histogram shows the satellites from SDSS and the unfilled histogram shows the centrals in SDSS.

STEEL accurately captures the key trends in the distributions, such as bimodality, which is seen in both the central and satellite populations. The central population below  $M_* = 10^{10.5} (\text{M}_\odot)$  is mostly star forming whereas the satellites show signs of quenching. In the intermediate-mass range a fraction of the centrals become quenched and the satellites show a strong quenching effect. In the highest mass range all galaxies show strong quenching features with little star formation. Whilst still not an exact match to the SDSS distribution, we find that including a redshift dependence in the dynamical quenching provides a better fit than the model used in Paper I. The central sSFR is calculated using the SFR presented in Section 4.3, which use the PYMORPH SMHM relation. Each central mass is assigned an SFR with a scatter of 0.2 dex. To account for the fraction of galaxies that are quenched via mergers at each stellar mass we modify the assigned SFRs by setting a fraction of galaxies equal to the elliptical fraction from Section 5.3 to have an sSFR of  $10^{-12} (\text{yr}^{-1})$  with a scatter of 0.2 dex and in turn increase the SFR of the remaining galaxies to maintain the same average SFR for the population. This approach tests if mergers alone can account for the bimodality found in the central sSFR, the high-mass centrals  $> 10^{11.3} (\text{M}_\odot)$ , but produces an inadequate fit to the SDSS centrals at masses lower than  $10^{10.5} (\text{M}_\odot)$ . The discrepancies in the location of the star-forming population are likely caused by the imperfect fit to observed SFR as seen in Fig. 10 and the deficit of quenched galaxies in the lower mass cuts is likely due to causes of quenching that are not merger related (e.g. AGN feedback).

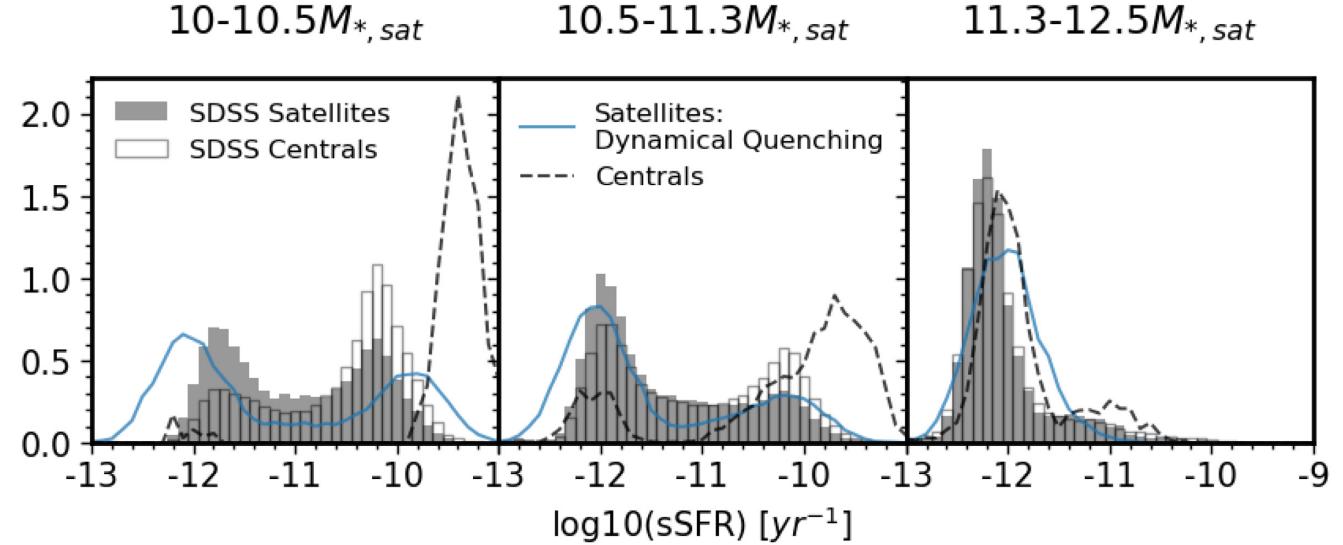
## 5 DISCUSSION

Using STEEL we have presented a consistent picture of group and cluster richness across several orders of magnitude in mass, and satellite accretion histories over 11 Gyr of the Universe’s history. It is essential for a model that aims to predict the hierarchical growth of galaxies, that both the central and satellite SMFs are well reproduced at all redshifts. STEEL uses state-of-the-art statistical accretion histories and powerful abundance matching techniques to ensure we have the essential consistency with observed galaxy number densities by design.

<sup>11</sup>Dynamical time factors higher than 2 are shown to not reproduce the satellite distributions in Paper I



**Figure 8.** Same format as Fig. 6 but for the cmodel photometry. It is clear that this model is internally non-physical as the accretion via satellites (dashed lines) rapidly overshoots the total growth in stellar mass (solid lines) implied by the underlying growth host halo growth, as evident in the middle and bottom rows.



**Figure 9.** We show the sSFR of satellites and centrals compared to SDSS in three mass bins selected to mirror proposed breaks in the galaxy main sequence. The SDSS data for satellites and centrals are filled and unfilled histograms, respectively. The STEEL result for the satellites is the solid blue line and the post-processed central result is the dashed black line.

### 5.1 High-redshift clusters

Galaxy groups and clusters represent an excellent laboratory to test theories on galaxy evolution. The rich cluster environments are observable up to high redshift and contain some of the most massive galaxies. Exploring the richness of the environments around massive galaxies provides an excellent constraint to hierarchical assembly predicted by  $\Lambda$ CDM cosmology at the most extreme

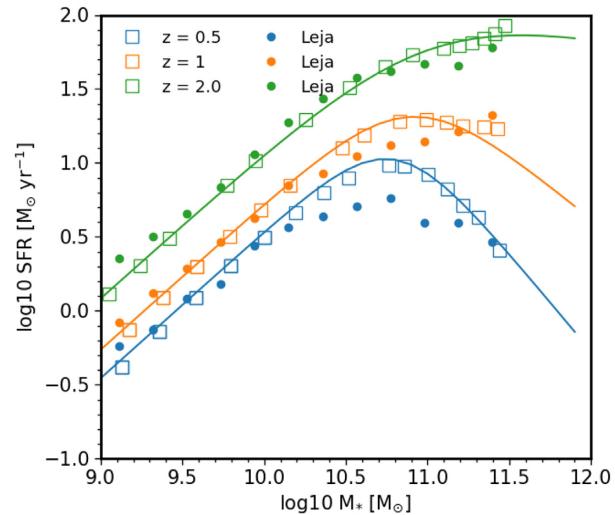
masses (Shankar et al. 2015). In this work we show a singular cluster reported in Wang et al. (2016). Other models (e.g. Henriques et al. 2015) have been unable to reconcile this cluster within a  $\Lambda$ CDM framework. We found in Fig. 5 that STEEL is able to predict the existence of such massive objects. However we concur with Wang et al. (2016) that these objects are rare and their absence in traditional simulations could be simply attributed to poor statistics and not necessarily to the implied physical model. With large-scale

surveys such as EUCLID coming online, a well-tuned statistical model could more easily place robust constraints on high-redshift cluster formation.

## 5.2 Central assembly

In Section 4.3 we found one of the major factors in regulating the *in situ* and *ex situ* accretion pathways to be the *shape* of the SMHM relation. A shallower low-mass slope causes larger amounts of satellite accretion as smaller haloes, with much higher number density, are initialized with larger satellite galaxies. Similarly to Shankar et al. (2006) and Moster et al. (2018), we find the high-mass slope to undergo only a small amount of evolution with increasing redshift, this implies the growth of central galaxies is directly linked to the steepness of the high-mass slope and the growth of the host halo. The flatter the high-mass slope of the SMHM relation, the less growth is expected in stellar mass following the assembly of the host dark matter halo. In turn, a weak evolution in the stellar mass content of the central galaxy can be in tension with what is expected from satellite accretion, especially for the most massive galaxies. We discussed that the slope of the high-mass end of the SMF and implied slope of the SMHM relation strongly depend on the choice of light profile, background subtraction, and mass-to-light ratios. However, not all resulting SMFs provide physically self-consistent results in a  $\Lambda$ CDM Universe. Steeper SMHM relations, such as those predicted by PYMORPH-based SMFs (Bernardi et al. 2013), produce more consistent central and satellite accretion stellar mass growths. In addition to models with different SMHM slopes, we also tested models with the dynamical time varied by  $\pm 20$  per cent, within the range of possible dynamical times predicted in Paper I constrained by satellite richness. This relatively modest alteration has a minor effect on the satellite accretion rate and mass contribution to the central. In this work we find the transitional stellar mass, above which dry mergers progressively become the major contributor to galaxy growth, to be  $M_* = 10^{11.1}$ , see Fig. 6. The latter is consistent with previous findings (e.g. Bernardi et al. 2011; Cappellari 2013; Shankar et al. 2013).

By following the statistical dark matter accretion histories we were able to use the central mass tracks and abundance matching to obtain a growth history for central galaxies. Subtracting from the latter at each time-step the cumulative stellar mass from satellite accretion, we created a ‘SFR’ interpreted as the remaining mass required to build the central mass. Our methodology is similar to the continuity approach based on Leja et al. (2015) used in Paper I, but with the key difference that here we follow halo growth instead of galaxy number density. The resulting SFR for galaxies is notably different to that of Tomczak et al. (2014), used in Paper I. At all redshifts the turnover is notably different, with SFR for masses above the turnover decreasing sharply at low redshift. For masses below the turnover, at  $z < 1$  the SFR is lower by 0.3 dex, and at  $z > 1$  the SFR is higher by 0.1–0.2 dex. Recent work, where the SFHs are properly accounted for when measuring SFRs, has suggested that the previous determinations of SFRs using UV+IR are 0.1–1 dex too high (Leja et al. 2019) and cannot be reconciled with the growth of the SMF (Leja et al. 2015; Lapi et al. 2017). Our SFR is consistent with the results of Leja et al. (2019), as reported in Fig. 10. The excellent match to Leja et al.’s independent estimates further supports the idea that a more robust method to derive more reliable SFRs is to follow galaxy assembly along host halo growth histories (see e.g. Moster et al. 2018).



**Figure 10.** We show the SFR–stellar mass relationship from Fig. 7 at redshifts  $z = 0.5, 1, 2$  (blue, orange, and green, respectively, STEEL data are crosses and fits are solid lines). In this plot we compare with the observed SFR from Leja et al. (2019) shown as filled circles with corresponding colours denoting corresponding redshift.

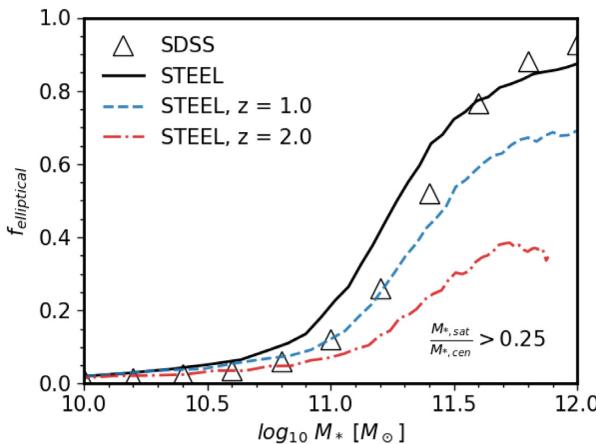
## 5.3 Central morphologies

Mergers are thought to be one of the drivers for morphological transformation, size growth, and other galaxy changes (Bournaud, Jog & Combes 2007; Hopkins et al. 2009, 2010a; Shankar et al. 2011; Fontanot et al. 2015). Broadly inspired by the results of hydrodynamic simulations, a number of analytically based cosmological models have generally assumed that major mergers, in particular, with a mass ratio of at least  $M_{\text{sat}}/M_{\text{cen}} > 0.25$ , are effective in destroying discs and in forming ellipticals (Baugh 2006; Malbon et al. 2007; Bower et al. 2010). Given the very promising results of STEEL in predicting satellite number densities in different environments and epochs, we here take a step further and explore whether STEEL’s cumulative number of major mergers is able to account for the local fraction of elliptical galaxies. For each central mass track we evolve the fraction of galaxies that have had a merger with stellar mass ratio greater than 0.25 since redshift  $z = 3$ . Fig. 11 shows the probability/fraction of central galaxies that have experienced a merger above the mass threshold of 0.3 at redshifts  $z = 0.1, 0.65, 1.75$ , while the black triangles show the T-Type-selected elliptical fraction from the SDSS catalogue. We find that applying this simple recipe to the merging number densities from STEEL creates a good match to the elliptical fraction in the local universe.

Despite the noticeably good agreement between model predictions and data in Fig. 11, we stress that different input SMHM relations can, as proven in this work, substantially affect the accretion rate which in turn will modify the number of galaxies experiencing major mergers. It follows that any cosmological galaxy evolution model that uses mergers as a physical driver for galaxy transformation should first simultaneously and self-consistently closely reproduce SMFs, the SMHM relation, and satellite distributions at high redshift.

## 6 CONCLUSIONS

In this second paper on our STatistical sEmi-Empirical model, STEEL, we proved that STEEL can successfully reproduce galaxy satellite richness also at high redshifts. Its innovative design,



**Figure 11.** We show at three redshift steps the predicted fraction of ellipticals as a function of stellar mass. The lines are the predictions from STEEL and the triangles are the T-Type selected elliptical fraction from SDSS at redshift  $z = 0.1$ .

unconstrained by volume or mass resolution, allows STEEL to predict the number densities of even the rarest objects in the Universe at the highest redshifts, a fundamental test for dark matter and galaxy evolution theories though currently inaccessible by cosmological hydrodynamic simulations.

Given the success of STEEL in reproducing satellite richness at different cosmic epochs and environments, we can in turn predict more reliable galaxy merger rates, using central growth rates implied by the central mass track of our statistical dark matter accretion histories and abundance matching. We found that SMHM relations with shallow high-mass slopes create central growth histories that are physically inconsistent with the expected satellite merger rate. We found that steeper SMHM relations at the high-mass end, as induced by the latest determinations of the SMFs based on Sersic-exponential photometry, are favoured against shallower SMHM relations, based on outdated determinations of the SMF. The total stellar mass growth of a galaxy is mostly due to satellite mergers and/or star formation. A flatter SMHM relation, however, naturally implies, for a given increase in host halo mass, a much weaker growth in the stellar mass of the central galaxy than in the case of a steep SMHM relation. The accretion via satellites could then be substantial enough to overshoot the moderate growth in the central galaxy rendering the model internally physically inconsistent.

By safely assuming the difference in central growth rate ( $\dot{M}$ ) and satellite accretion rate is attributable to the star formation in the central galaxy, we predict SFHs and an SFR–stellar mass relations. The latter approach is qualitatively similar to a continuity equation (e.g. Leja et al. 2019, and Paper I), but more accurate as it is developed along the accretion tracks of host haloes so better follows galaxy populations. We find our resulting SFHs to be in excellent agreement with the latest cutting-edge observational measurements by Leja et al. (2019), based on multiparameter Bayesian analysis.

Finally, following traditional models of galaxy evolution, we use our improved galaxy merger rate to predict the fraction of central galaxies as a function of mass that have been transformed into ellipticals via major mergers, (where the stellar mass ratio of the central to satellite is greater than 1/3). We find this fraction to be in excellent agreement with centrals selected as ellipticals via T-Type in SDSS. We use this elliptical fraction and our derived SFR to create a distribution of sSFRs. We find this basic and common assumption to form ellipticals in analytic cosmological models to

be sufficient to also reproduce the bimodality in SFRs of massive galaxies above  $M_* > 10^{11.3} \text{ M}_\odot$ . Below this stellar mass threshold, we find a too high fraction of star-forming galaxies, which implies additional quenching mechanisms, beside major mergers, must be included in the models.

Our results are of the utmost importance to predict robust and self-consistent SMHM relations and to generate reliable mock catalogues for the next generation of extragalactic surveys such as Euclid and LSST.

## ACKNOWLEDGEMENTS

We acknowledge extensive use of the Python libraries ASTROPY, MATPLOTLIB, NUMPY, PANDAS, and SCIPY. PJG acknowledges support from the STFC for funding this PhD. FS acknowledges partial support from a Leverhume Trust Research Fellowship and from the European Union’s Horizon 2020 programme under the AHEAD project (grant agreement #654215).

## REFERENCES

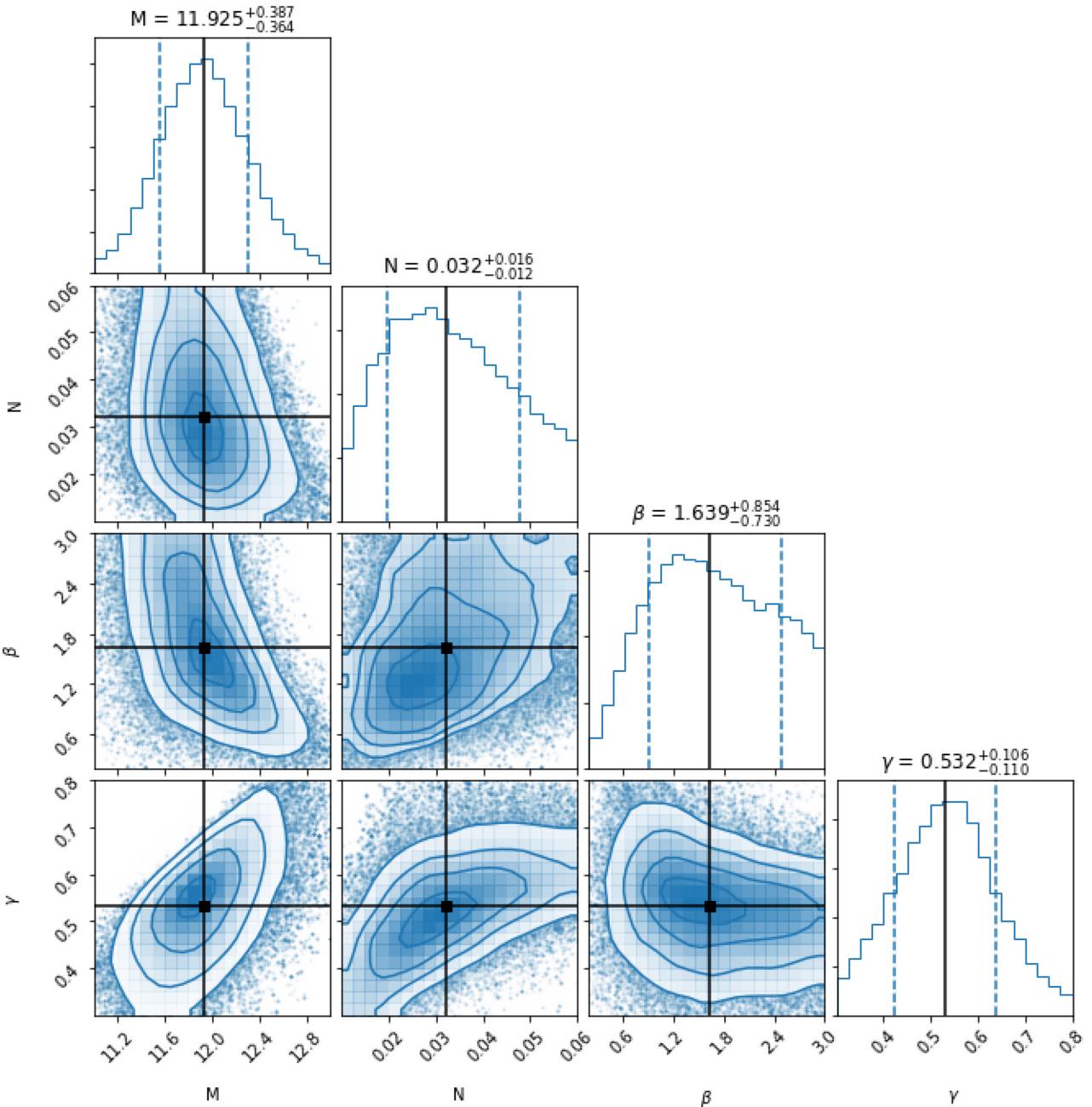
- Abazajian K. N. et al., 2009, *ApJS*, 182, 543
- Abolfathi B. et al., 2017, *ApJS*, 235, 42
- Asquith R. et al., 2018, *MNRAS*, 480, 1197
- Baugh C. M., 2006, *Rep. Prog. Phys.*, 69, 3101
- Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, 488, 3143
- Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, *ApJS*, 149, 289
- Bernardi M., Roche N., Shankar F., Sheth R. K., 2011, *MNRAS*, 412, L6
- Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, *MNRAS*, 436, 697
- Bernardi M., Meert A., Vikram V., Huertas-Company M., Mei S., Shankar F., Sheth R. K., 2014, *MNRAS*, 443, 874
- Bernardi M., Meert A., Sheth R. K., Huertas-Company M., Maraston C., Shankar F., Vikram V., 2016, *MNRAS*, 455, 4122
- Bernardi M., Fischer J.-L., Sheth R. K., Meert A., Huertas-Company M., Shankar F., Vikram V., 2017, *MNRAS*, 468, 2569
- Bournaud F., Jog J., Combes F., 2007, *A&A*, 476, 1179
- Bower R. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, *MNRAS*, 407, 2017
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Buchan S., Shankar F., 2016, *MNRAS*, 462, 2001
- Cappellari M., 2013, *ApJ*, 778, L2
- Cappellari M. et al., 2013, *MNRAS*, 432, 1862
- Cattaneo A., Mamon G. A., Warnick K., Knebe A., 2011, *A&A*, 533, A5
- Chabrier G., 2003, *PASP*, 115, 763
- Cowley W. I. et al., 2019, *ApJ*, 874, 114
- Davidzon I. et al., 2017, *A&A*, 605, A70
- de Vaucouleurs G., 1948, *Ann. Astrophys.*, 11, 247
- Dekel A. et al., 2009, *Nature*, 457, 451
- Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, *MNRAS*, 456, 2486
- De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, *MNRAS*, 366, 499
- Diemer B., 2017, Technical report, COLOSSUS: A python toolkit for cosmology, large-scale structure, and dark matter haloes. Harvard-Smithsonian Center for Astrophysics, <http://arxiv.org/abs/1712.04512>
- Fakhouri O., Ma C.-P., 2010, *MNRAS*, 401, 2245
- Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, *MNRAS*, 2278, 2267
- Fillingham S. P., Cooper M. C., Boylan-Kolchin M., Bullock J. S., Garrison-Kimmel S., Wheeler C., 2018, *MNRAS*, 477, 4491
- Fontanot F., Macciò A. V., Hirschmann M., De Lucia G., Kannan R., Somerville R. S., Wilman D., 2015, *MNRAS*, 451, 2968
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Technical report, emcee: The MCMC Hammer. <http://dan.iel.fm/emceeundertheMITLicense>

- 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
- Groenewald D. N., Skelton R. E., Gilbank D. G., Loubser S. I., 2017, *MNRAS*, 467, 4101
- Grylls P. J., Shankar F., Zanisi L., Bernardi M., 2019, *MNRAS*, 483, 2506 (Paper I)
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, 451, 2663
- Hopkins P. F. et al., 2009, *MNRAS*, 397, 802
- Hopkins P. F., Younger J. D., Hayward C. C., Narayanan D., Hernquist L., 2010a, *MNRAS*, 402, 1693
- Hopkins P. F. et al., 2010b, *ApJ*, 715, 202
- Jiang F., van den Bosch F. C., 2016, *MNRAS*, 458, 2848
- Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, *MNRAS*, 457, 4340
- Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, *Astron. Lett.*, 44, 8
- Lapi A. et al., 2011, *ApJ*, 742, 24
- Lapi A., Mancuso C., Bressan A., Danese L., 2017, *ApJ*, 847, 13
- Leja J., van Dokkum P. G., Franx M., Whitaker K. E., 2015, *ApJ*, 798, 115
- Leja J. et al., 2019, *ApJ*, 877, 140
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119
- Malbon R. K., Baugh C. M., Frenk C. S., Lacey C. G., 2007, *MNRAS*, 382, 1394
- Matharu J. et al., 2019, *MNRAS*, 484, 595
- McAlpine S. et al., 2015, *MNRAS*, 450, 1937
- Meert A., Vikram V., Bernardi M., 2015, *MNRAS*, 446, 3943
- Menci N., Gatti M., Fiore F., Lamastra A., 2014, *A&A*, 569, A37
- Mendel J. T., Simard L., Palmer M., Ellison S. L., Patton D. R., 2014, *ApJS*, 210, 3
- Moster B. P., Somerville R. S., Maulbetsch C., Van Den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903
- 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
- Nelson D. et al., 2018, *MNRAS*, 475, 624
- Parkinson H., Cole S., Helly J., 2008, *MNRAS*, 383, 557
- Planck Collaboration XIII, 2015, *A&A*, 594, A13
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Schawinski K. et al., 2014, *MNRAS*, 440, 889
- Shankar F., 2013, *Class. Quantum Gravity*, 30, 244001
- 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., 2011, *MNRAS*, 428, 109
- Shankar F., Marulli F., Bernardi M., Mei S., Meert A., Vikram V., 2013, *MNRAS*, 128, 109
- Shankar F. et al., 2014, *MNRAS*, 439, 3189
- Shankar F. et al., 2015, *ApJ*, 802, 73
- Springel V. et al., 2018, *MNRAS*, 475, 676
- Tomczak A. R. et al., 2014, *ApJ*, 783, 85
- van Dokkum P. G. et al., 2010, *ApJ*, 709, 1018
- van den Bosch F. C., Jiang F., Hearn A., Campbell D., Watson D., Padmanabhan N., 2014, *MNRAS*, 445, 1713
- Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
- Wang T. et al., 2016, *ApJ*, 828, 56
- Wen Z. L., Han J. L., 2018, *MNRAS*, 481, 4158
- Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, *MNRAS*, 432, 336
- Wetzel A. R., Deason A. J., Garrison-Kimmel S., 2015, *ApJ*, 807, 49
- Wright E. L. et al., 2010, *AJ*, 140, 1868
- Yang X., Mo H. J., Van Den Bosch F. C., Zhang Y., Han J., 2012, *ApJ*, 752, 41
- Zavala J., Avila-Reese V., Firmani C., Boylan-Kolchin M., 2012, *MNRAS*, 427, 1503

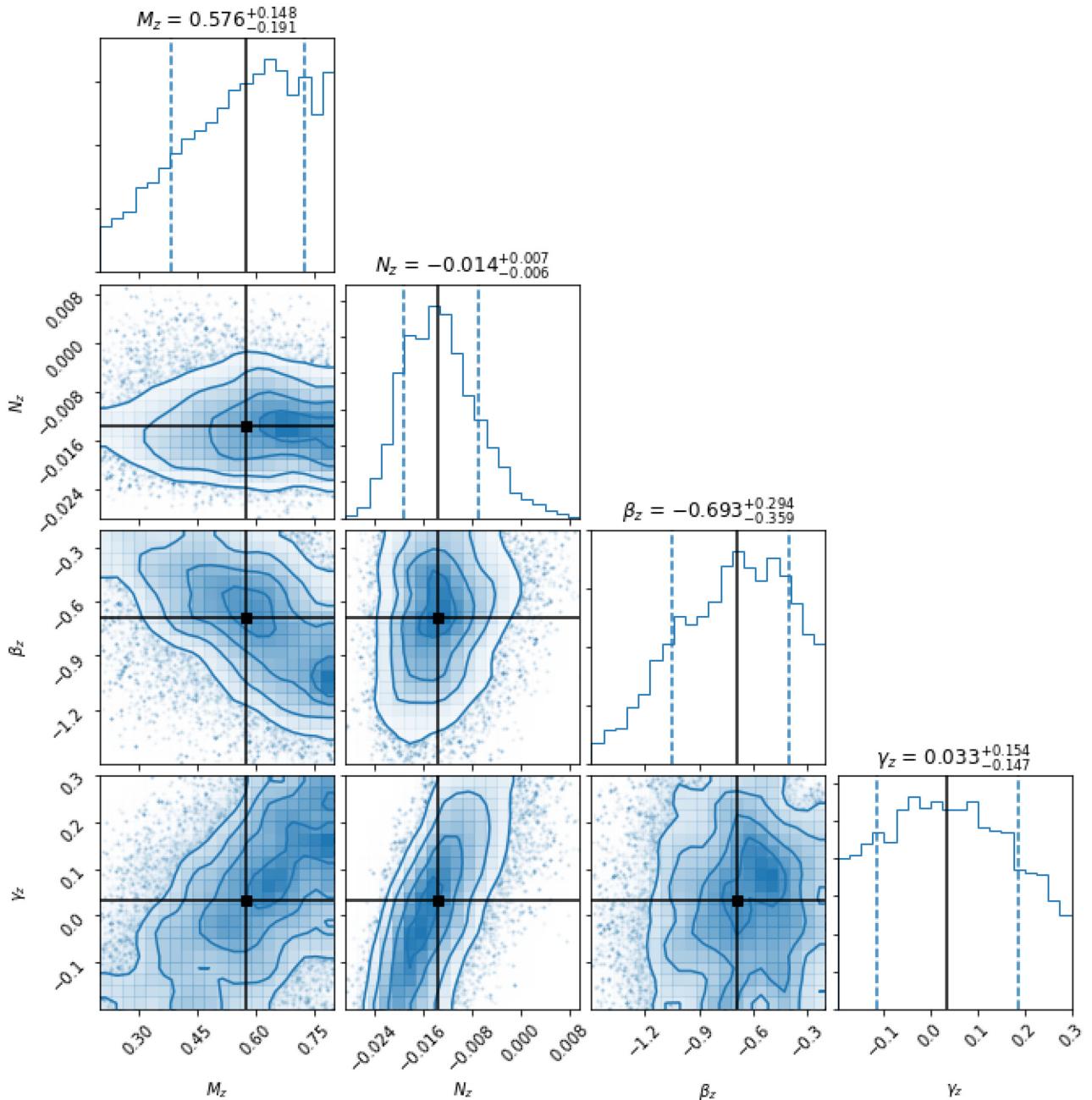
## APPENDIX A: ABUNDANCE MATCHING MCMC

Fig. A1 shows the redshift  $z = 0.1$  output from the MCMC abundance matching fits. It becomes immediately obvious that the low-mass slope ( $\beta$ ) is poorly constrained however the impact on the SMF is limited within the margin of error. The position of the knee ( $M$ ) is well constrained against both the normalization ( $N$ ) and the high-mass slope ( $\gamma$ ). The shape of the constraint between the normalization and gamma emanates from the need to produce high-mass galaxies, if the normalization is decreased the slope must increase to ensure enough haloes produce massive galaxies.

Fig. A2 shows the redshift  $z > 0.1$  output from the MCMC abundance matching fits. All parameters have low evolution and the SMHM relation evolves only weakly with redshift. For  $M$ ,  $\beta$ , and  $\gamma$  where the distributions are wide or close to the prior we have tested wider priors and insignificant change is found.



**Figure A1.** We show the MCMC parameter space for the redshift  $z = 0.1$  fit. The position of the knee ( $M$ ), the normalization ( $N$ ), the low-mass slope ( $\beta$ ), and the high-mass slope ( $\gamma$ ) are shown from left to right. Columns are titled with the best-fitting values and 16th/84th percentile errors. The black lines show the best-fitting value with a black square at intersections, the 16th/84th percentiles are shown with blue dashed lines on the histograms.



**Figure A2.** We show the MCMC parameter space for the high redshift  $z > 0.1$  fit. The evolution of: the position of the knee ( $M_z$ ), the normalization ( $N_z$ ), the low-mass slope ( $\beta_z$ ), and the high-mass slope ( $\gamma_z$ ) are shown from left to right. Columns are titled with the best-fitting values and 16th/84th percentile errors. The black lines show the best-fitting value with a black square at intersections, the 16th/84th percentiles are shown with blue dashed lines on the histograms.

## APPENDIX B: FITTING THE STAR FORMATION RATES

The fit to the SFR we derive in Section 4.3 is given by the following, equation B1,

$$\text{SFR}(M_*, z) = 2N(z) \left[ \left( \frac{M_*}{M_n(z)} \right)^{-\alpha(z)} + \left( \frac{M_*}{M_n(z)} \right)^{\beta(z)} \right]^{-1}$$

$$\log_{10} N(z) = 10.65 + 0.33z - 0.08z^2$$

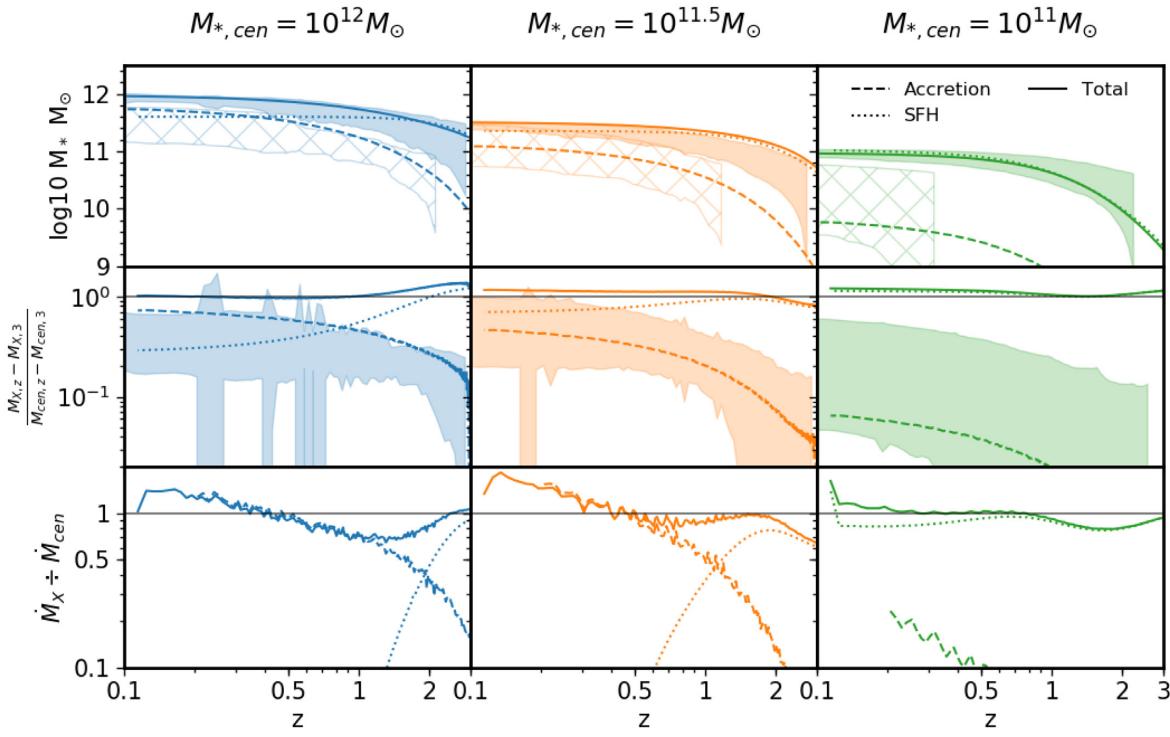
$$\log_{10} M_n(z) = 0.69 + 0.71 * z - 0.088z^2$$

$$\alpha(z) = 1.0 - 0.022z + 0.009z^2$$

$$\beta(z) = 1.8 - 1.0 * z - 0.1z^2. \quad (\text{B1})$$

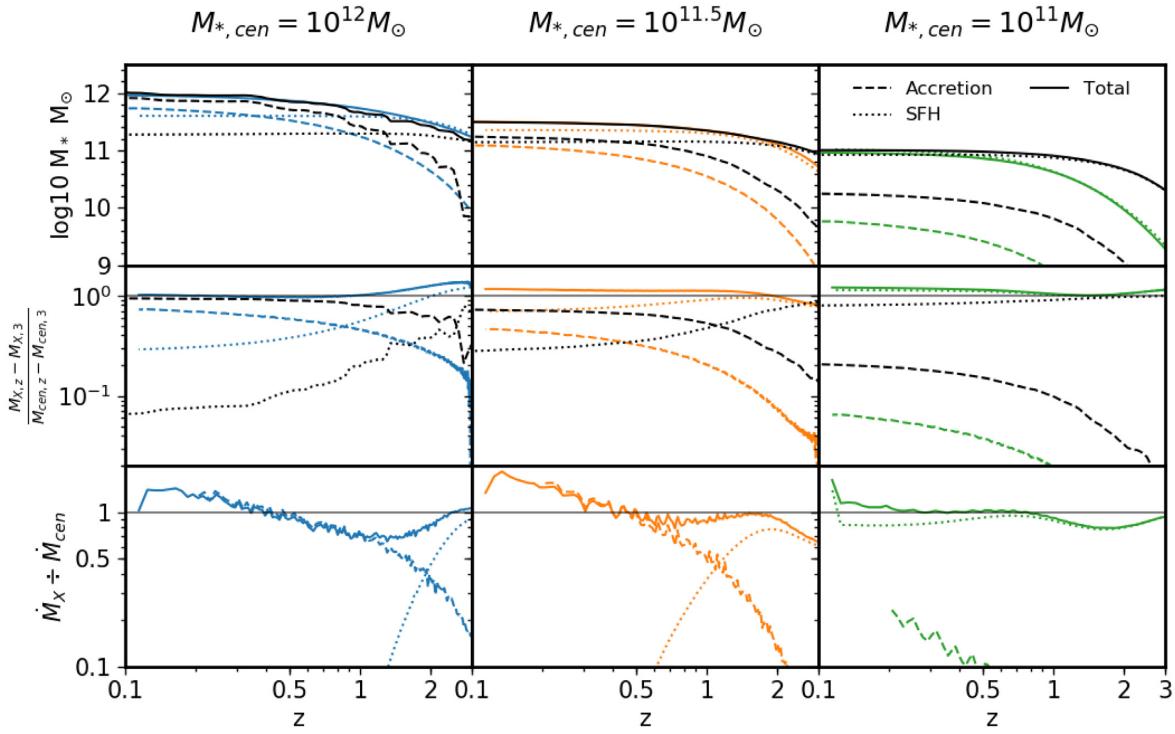
## APPENDIX C: COMPARISON OF IN SITU/EX SITU GROWTH TO OTHER MODELS

In Fig. C1 we show the *in situ* versus *ex situ* growth with the same model as shown in Fig. 6, we add to this plot data extracted from the Illustris TNG100 simulation. In Fig. 3 we see Illustris has a shallower slow mass slope and a steeper high mass slope such that more stellar mass is mapped into haloes of all sizes. We see the change in both of these slopes reflected in the accretion histories, first, for the lower mass galaxies (see  $\log_{10} M_{*,\text{cen}} = 10^{11}$ ) closer to the SMHM knee we find enhanced accretion due to the larger masses from more minor mergers. Secondly the high-mass slope is a direct result of the accretion, to support the same merger assembly with the higher mass galaxies in the satellite haloes above the knee where galaxy growth is dominated by the accretion the galaxy growth with halo size must be enhanced.



**Figure C1.** As in Fig. 6 three ‘mass tracks’ are shown that have central galaxy masses at redshift  $z = 0.1$  of  $M_{*,\text{cen}} = 10^{12}, 10^{11.5}, \text{ and } 10^{11} (M_\odot)$  in blue, orange, and green, respectively. The satellite galaxy accretion is shown for evolved satellites with a dashed line and the mass from star formation shown with a dotted line. The top panel shows the total mass of the central (solid line) and the total mass gained from accretion or star formation. The middle panel shows the fraction of the total galaxy mass formed from satellite accretion or star formation since redshift  $z = 3$ . The bottom panel shows the ratio of the mass accretion rate from satellite galaxies, the SFR, and the mass growth rate of the central galaxy predicted by abundance matching. In the top panel the shaded regions are galaxies selected from the Illustris simulation; the hashed region is then the satellite accretion from Illustris, in the middle panel the shaded region is the ratio of satellite accretion from Illustris. The grey lines in the second and third panel are at unity, the solid lines showing the sum of the other two factors should therefore be close to or on these lines.

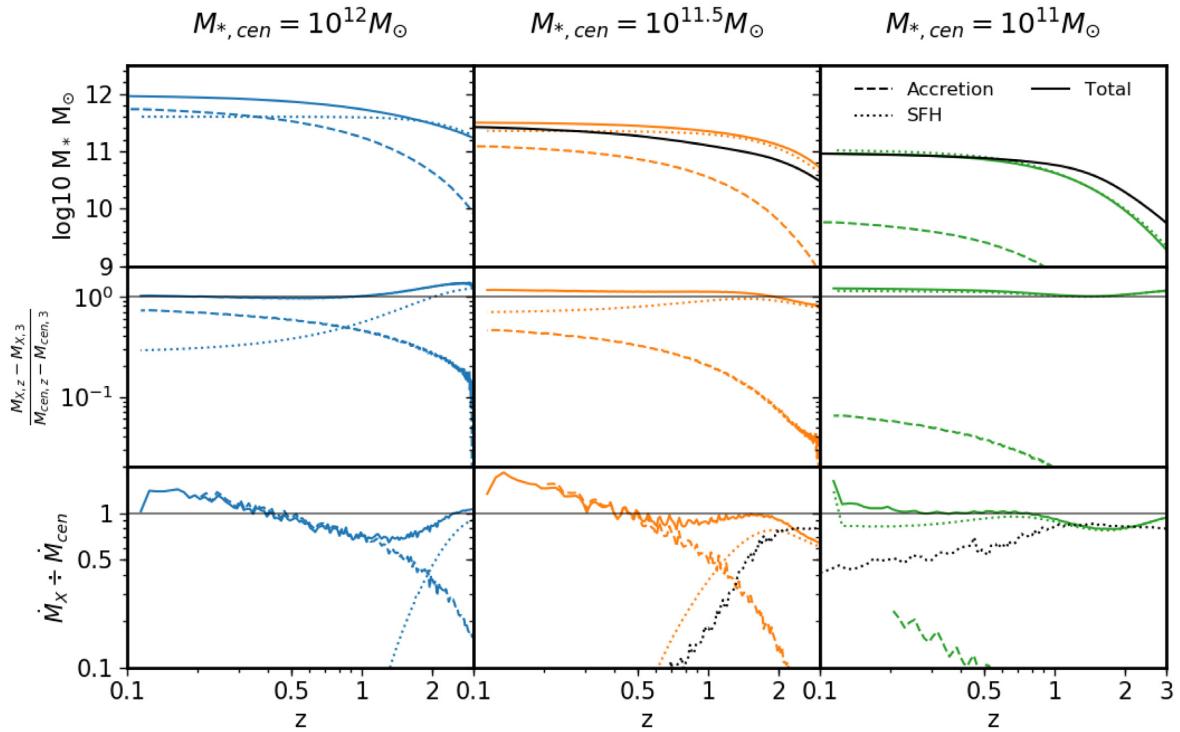
In Fig. C2 we show the *in situ* versus *ex situ* growth with the same model as shown in Fig. 6, we add to this plot data from the EMERGE model from Moster et al. (2018) shown as black lines. The solid lines show the total galaxy mass followed back selecting populations by mass at redshift  $z = 0.1$ . The dotted and dashed lines show the amount of galaxy mass formed *in situ* and *ex situ*, respectively. In all cases EMERGE predicts satellite accretion becomes the dominant mass growth pathway at higher redshifts than STEEL. In the third column we see that EMERGE and STEEL also disagree about the mass growth history of  $\log 10 M_{*,cen} = 11$  galaxies, however, both models agree that the dominant mass growth path of galaxies at this mass are *in situ* processes.



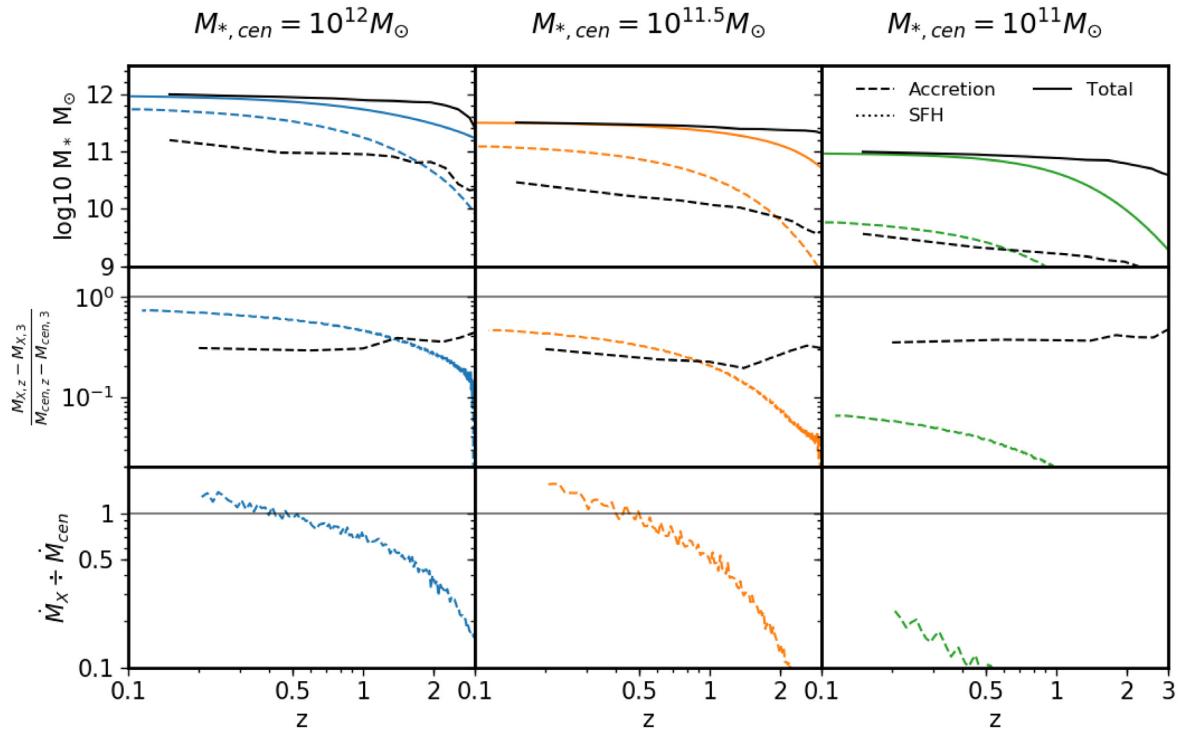
**Figure C2.** As in Fig. 6 three ‘mass tracks’ are shown that have central galaxy masses at redshift  $z = 0.1$  of  $M_{*,cen} = 10^{12}$ ,  $10^{11.5}$ , and  $10^{11}$  ( $M_\odot$ ) in blue, orange, and green, respectively. The satellite galaxy accretion is shown for evolved satellites with a dashed line and the mass from star formation shown with a dotted line. The top panel shows the total mass of the central (solid line) and the total mass gained from accretion or star formation. The middle panel shows the fraction of the total galaxy mass formed from satellite accretion or star formation since redshift  $z = 3$ . The bottom panel shows the ratio of the mass accretion rate from satellite galaxies, the SFR, and the mass growth rate of the central galaxy predicted by abundance matching. In the top and middle rows we add black lines to show the *in situ* and *ex situ* growth from EMERGE Moster et al. (2018). The grey lines in the second and third panel are at unity, the solid lines showing the sum of the other two factors should therefore be close to or on these lines.

In Fig. C3 we show for the  $\log 10 M_{*,cen} = 11.5$ , and 11 galaxies the central growth and SFR ratio from Behroozi et al. (2019). The central growth is close to that found from STEEL and the SFR transition for  $\log 10 M_{*,cen} = 11.5$  is an excellent match.

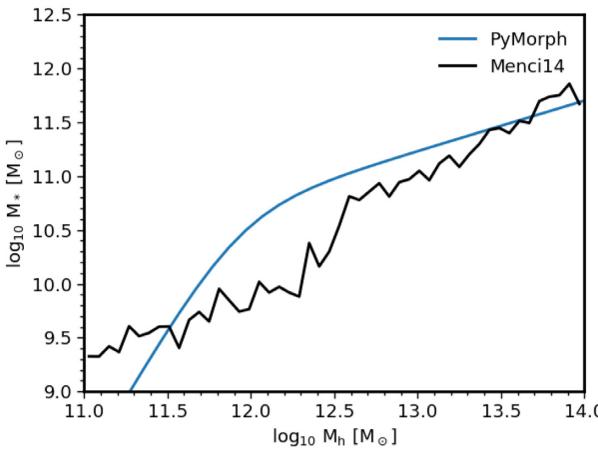
In Fig. C4 we show a comparison with the semi-analytic model described in Menci et al. (2014). At all masses the stellar growth is substantially different to STEEL and the other models shown in this appendix. Furthermore the semi-analytic model shows little change in the accreted mass ratio over cosmic time, again this is inconsistent with the findings from STEEL and the other models presented in this section. We attribute most of the differences seen here to the substantial difference in the SMHM relationship shown in Fig. C5.



**Figure C3.** As in Fig. 6 three ‘mass tracks’ are shown that have central galaxy masses at redshift  $z = 0.1$  of  $M_{*,cen} = 10^{12}$ ,  $10^{11.5}$ , and  $10^{11}$  ( $M_\odot$ ) in blue, orange, and green, respectively. The satellite galaxy accretion is shown for evolved satellites with a dashed line and the mass from star formation shown with a dotted line. The top panel shows the total mass of the central (solid line) and the total mass gained from accretion or star formation. The middle panel shows the fraction of the total galaxy mass formed from satellite accretion or star formation since redshift  $z = 3$ . The bottom panel shows the ratio of the mass accretion rate from satellite galaxies, the SFR, and the mass growth rate of the central galaxy predicted by abundance matching. In the top and bottom rows, for the  $\log 10 M_{*,cen} = 11.5$ , and  $11$ , we add black lines to show the central galaxy growth and the SFR ratio from Behroozi et al. (2019). The grey lines in the second and third panel are at unity, the solid lines showing the sum of the other two factors should therefore be close to or on these lines.



**Figure C4.** As in Fig. 6 three ‘mass tracks’ are shown that have central galaxy masses at redshift  $z = 0.1$  of  $M_{*,cen} = 10^{12}$ ,  $10^{11.5}$ , and  $10^{11}$  ( $M_\odot$ ) in blue, orange, and green, respectively. The satellite galaxy accretion is shown for evolved satellites with a dashed line and the mass from star formation shown with a dotted line. The top panel shows the total mass of the central (solid line) and the total mass gained from accretion or star formation. The middle panel shows the fraction of the total galaxy mass formed from satellite accretion or star formation since redshift  $z = 3$ . The bottom panel shows the ratio of the mass accretion rate from satellite galaxies, the SFR, and the mass growth rate of the central galaxy predicted by abundance matching. In the top and middle rows we add black lines to show the *ex situ* growth and central growth from the semi-analytic model described in Menci et al. (2014). The grey lines in the second and third panel are at unity, the solid lines showing the sum of the other two factors should therefore be close to or on these lines.

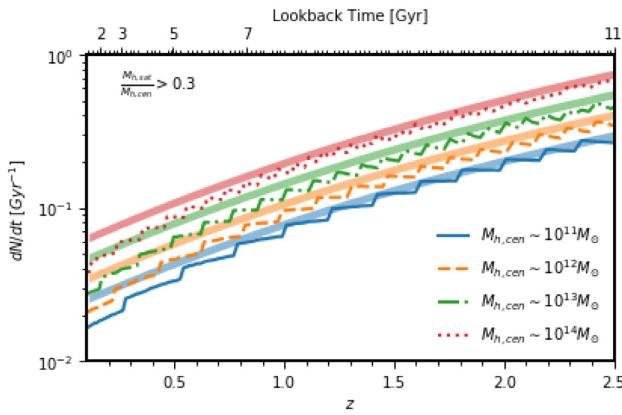


**Figure C5.** We show the SMHM relationship for PYMORPH (blue line, used in this work) and for the semi-analytic model from Menci et al. (2014) (black line).

#### APPENDIX D: HALO MERGER RATE IN THE MILLENNIUM COSMOLOGY.

The statistical dark matter accretion history used in STEEL is cosmologically flexible.<sup>12</sup> The current prescriptions for the halo growth histories and dark matter substructure (van den Bosch et al. 2014; Jiang & van den Bosch 2016) can be used for  $\Lambda$ CDM models that are within a factor two of current constraints. Using COLLOSUS (Diemer 2017) we set the cosmology used in STEEL to that of the Millennium simulation ( $\Omega_m = 0.25$ ,  $\Omega_b = 0.045$ ,  $\Omega_\Lambda = 0.75$ ,  $h = 0.73$ ). The statistical dark matter accretion history is then recalculated. Fig. D1 shows the halo merger rate from STEEL from this alternative accretion history. The halo merger rate tracks from Fakhouri & Ma (2010) shown in Figs 4 and D1 are calculated using the Millennium simulation cosmology. The deviation between the merger rate of STEEL and that of Fakhouri & Ma (2010) still remains despite the change of cosmology, as discussed in Section 3.1 this is an effect of the algorithms used to build halo merger trees.

<sup>12</sup>The eventual goal of STEEL is to be cosmologically independent such that any cosmology can be used.



**Figure D1.** The evolution of merger rate per Gyr at fixed halo mass. Lines are from STEEL with cosmology altered to that of the analytic fits from Fakhouri et al. (2010) (shaded bands). Halo masses shown are  $M_{h,\text{cen}} : 10^{11}, 10^{12}, 10^{13}, 10^{14} \text{ M}_\odot h^{-1}$  as labelled.

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