

# Galaxy formation with L-GALAXIES: modelling the environmental dependency of galaxy evolution and comparing with observations

Mohammadreza Ayromlou<sup>1</sup>,<sup>1</sup>★ Guinevere Kauffmann,<sup>1</sup> Robert M. Yates<sup>1,2</sup>, Dylan Nelson<sup>1,3</sup> and Simon D. M. White<sup>1</sup>

<sup>1</sup>Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str 1, D-85741 Garching bei München, Germany

<sup>2</sup>Department of Physics, University of Surrey, Stag Hill, Guildford GU2 7XH, UK

<sup>3</sup>Universität Heidelberg, Zentrum für Astronomie, Institut für theoretische Astrophysik, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany

Accepted 2021 April 27. Received 2021 April 27; in original form 2020 November 10

## ABSTRACT

We present a variation of the recently updated Munich semi-analytical galaxy formation model, L-GALAXIES, with a new gas stripping method. Extending earlier work, we directly measure the local environmental properties of galaxies to formulate a more accurate treatment of ram-pressure stripping for all galaxies. We fully recalibrate the modified L-GALAXIES model using a Markov Chain Monte Carlo (MCMC) method with the stellar mass function and quenched fraction of galaxies as a function of stellar mass at  $0 \leq z \leq 2$  as constraints. Due to this recalibration, global galaxy population relations, including the stellar mass function, quenched fractions versus galaxy mass, and H I mass function are all largely unchanged and remain consistent with observations. By comparing to data on galaxy properties in different environments from the SDSS and HSC surveys, we demonstrate that our modified model improves the agreement with the quenched fractions and star formation rates of galaxies as a function of environment, stellar mass, and redshift. Overall, in the vicinity of haloes with total mass  $10^{12}$  to  $10^{15} M_{\odot}$  at  $z = 0$ , our new model produces higher quenched fractions and stronger environmental dependencies, better recovering observed trends with halocentric distance up to several virial radii. By analysing the actual amount of gas stripped from galaxies in our model, we show that those in the vicinity of massive haloes lose a large fraction of their hot halo gas *before* they become satellites. We demonstrate that this affects galaxy quenching both within and beyond the halo boundary. This is likely to influence the correlations between galaxies up to tens of megaparsecs.

**Key words:** methods: analytical – methods: numerical – galaxies: evolution – galaxies: formation – large-scale structure of Universe.

## 1 INTRODUCTION

Modern theories of galaxy formation and evolution are built on top of the Lambda Cold Dark Matter ( $\Lambda$ CDM) cosmological model. One of the critical assumptions of  $\Lambda$ CDM is that cold dark matter interacts at best only very weakly with other matter except gravitationally, which has formed the basis of a number of projects that have simulated a gravity-only (dark matter only) universe (e.g. Springel et al. 2005; Boylan-Kolchin et al. 2009; Klypin, Trujillo-Gomez & Primack 2011; Riebe et al. 2011; Angulo et al. 2012; Skillman et al. 2014; Wang et al. 2020b). Above all, these simulations have shown that galaxy formation within  $\Lambda$ CDM can successfully reproduce the observed patterns of galaxy clustering within the cosmic web.

However, the Universe is observed in light coming from stars and gas, which make up galaxies. These are all made of baryonic matter, which interacts in much more complicated ways than weakly interacting gravity-only cold dark matter. Modelling these baryonic components and their interactions in our Universe is one of the most challenging topics in astrophysics. Due to the complex multiphysics and multiscale nature of the problem, analytical solutions alone

are impossible. This has motivated the development of a diverse set of simulation techniques from semi-analytical models (SAMs) (e.g. Kauffmann, White & Guiderdoni 1993; De Lucia et al. 2006; Henriques et al. 2020) to cosmological hydrodynamical simulations (e.g. Vogelsberger et al. 2014; Schaye et al. 2015; Pillepich et al. 2018) to non-periodic zoom simulations of individual regions (e.g. Grand et al. 2017; Hopkins et al. 2018; Libeskind et al. 2020).

According to the standard hierarchical structure formation theory, baryonic matter accretes into the gravitational potential wells of dark matter structures (White & Rees 1978). It then undergoes cooling and contraction, which eventually leads to the formation of stars and galaxies. Furthermore, the evolution of baryonic matter and galaxies is influenced by a set of complex physical processes (e.g. gas phase transitions, stellar evolution, galaxy mergers, and feedback processes), which motivates the development of a comprehensive theory of galaxy formation and evolution (see Mo, Van den Bosch & White 2010 for a full discussion). In addition to intrinsic physical processes, galaxy evolution is also strongly influenced by environment (Hubble & Humason 1931; Boselli & Gavazzi 2006). In other words, the properties of galaxies/subhaloes are strongly correlated with their local density (Yan, Fan & White 2013). For example, if they are within clusters (Boselli et al. 2016; Pallero et al. 2019), filaments (Sarron et al. 2019), and voids (Tavasoli, Vasei & Mohayaee

\* E-mail: ayromlou@mpa-garching.mpg.de

2013; Mosleh, Tavasoli & Tacchella 2018). Well-known examples of these phenomena are the morphology–density relation (Oemler 1974; Dressler 1980) and higher fraction of quenched galaxies within massive clusters (Kauffmann et al. 2004; Peng et al. 2010; Davies et al. 2019). Environmental processes such as ram-pressure stripping (Gunn & Gott 1972) and tidal stripping (Binney & Tremaine 1987) can strip the gas out of galaxies, affecting their gas reservoirs (Wang et al. 2020a), star formation (Donnari et al. 2020, 2021), stellar properties (Webb et al. 2020), morphology (Joshi et al. 2020), and so on.

On the theoretical side, the treatment of many of these environmental processes in hydrodynamical simulations occurs as a result of solving the equations of gravity, (magneto)hydrodynamics, and as heating and cooling together, allowing the direct modelling of both dark and baryonic matter (Hernquist & Katz 1989; Springel 2010; Crain et al. 2015). Simulations using idealized set-ups (Roediger & Brüggen 2007; Vijayaraghavan & Ricker 2017; Lee et al. 2020) as well as full cosmological simulations (Pallero et al. 2019; Yun et al. 2019) reveal the strong dependency of galaxy properties on their environment. Although large-volume hydrodynamical simulations are powerful tools for understanding the physics of galaxies, they are computationally expensive and typically take millions of CPU hours for a single simulation (Nelson et al. 2019; Pillepich et al. 2019; Dubois et al. 2020). Alternatively, one can model galaxy formation and evolution, employing modern, computationally efficient semi-analytical techniques.

A SAM implements the key physical processes involved in galaxy formation and evolution on top of halo merger trees constructed by Monte Carlo methods or extracted from dark-matter-only (DMO) simulations (Kauffmann et al. 1993, 1999; Somerville & Primack 1999; Cole et al. 2000; Springel et al. 2001; De Lucia et al. 2006). As there is no hydrodynamical interaction in the DMO simulations on which SAMs are implemented, baryonic environmental processes do not occur naturally in SAMs. Therefore, they need to be modelled explicitly using analytical approximations. Early SAMs assume that the entire hot gas reservoir of a galaxy is stripped once it infalls past the virial radius of a larger halo (e.g. Kauffmann et al. 1999; Croton et al. 2006). Instantaneous hot gas stripping for satellites within the virial radius remains a common approach in some recent SAMs as well (Lacey et al. 2016; Lagos et al. 2018). Most of modern SAMs, including the Munich model (L-GALAXIES; Henriques et al. 2020), have replaced this prescription by a gradual, time-evolving, gas stripping that allows satellite galaxies/subhaloes to retain a fraction of their gas over longer time-scales (Font et al. 2008; Tecce et al. 2010; Guo et al. 2011; Croton et al. 2016; Stevens, Croton & Mutch 2016; Cora et al. 2018).

Traditionally, SAMs model environmental processes only for satellite galaxies within the halo boundary (typically the virial radius) of their host haloes. That is partly because estimating ram-pressure beyond this scale is difficult without the direct measurement of the density and velocity of the local environment of galaxies. Accurate object-by-object comparisons between the L-GALAXIES SAM (Henriques et al. 2015, H15 hereafter) and the IllustrisTNG hydrodynamical simulations show that environmental effects influence galaxies up to much larger halocentric distances (several virial radii) in the hydrodynamical simulation compared to the SAM (Ayromlou et al. 2021). Furthermore, several observational (Hansen et al. 2009; von der Linden et al. 2010; Lu et al. 2012; Wetzel, Tinker & Conroy 2012) and theoretical (Balogh et al. 1999; Bahé et al. 2012) studies also show that environmental effects might act beyond the halo boundary. Without accurate modelling of environmental processes for all galaxies, it is not possible to robustly reproduce the properties of galaxies which reside in different environments.

In earlier work, we devised a new method to measure, locally and accurately, ram-pressure for all galaxies in the L-GALAXIES SAM using the particle data of the underlying DMO simulation (Ayromlou et al. 2019). Here, we employ this method and update L-GALAXIES by extending ram-pressure stripping to all simulated galaxies, regardless of their environment. Taking advantage of the computational efficiency of L-GALAXIES, one can quickly implement and test new ideas and theories in galaxy formation and evolution, as is done in this work. Here, we probe the role of environment in galaxy evolution by investigating the idea of galaxies becoming ram-pressure stripped *prior to* and after infall into massive and intermediate-mass haloes. To do so, we calibrate our updated L-GALAXIES model against a set of observational constraints using a full Markov Chain Monte Carlo (MCMC) approach. We then measure the actual amount of the stripped gas for every galaxy on the fly as the SAM is run. We finally contrast our model predictions with a number of other observations, specifically focusing on the role of environment, and explore the impact of our new stripping model on galaxy properties and the galaxy population as a whole.

This paper is structured as follows: In Section 2, we explain the base L-Galaxies (2020) model and the definition of the local background environment (LBE). In Section 3, we introduce a variant of the L-GALAXIES model with a novel hot gas stripping method. The observations used in this work and the mock catalogues we make to compare our model with those observations are described in Section 4.1. Section 5 is dedicated to studying the environmental dependency of galaxy evolution in our model, comparing both with previous versions and with observations. Finally, we conclude and summarize our results in Section 6.

## 2 METHODOLOGY

### 2.1 Simulations and galaxy formation model

#### 2.1.1 Simulations and subhalo identification

In this work, we use the particle and halo merger tree data of the Millennium and Millennium-II simulations (Springel et al. 2005; Boylan-Kolchin et al. 2009). Both simulations are rescaled to the Planck cosmology (Planck Collaboration 2016) from their original older cosmology, applying the method introduced by Angulo & White (2010) as updated by Angulo & Hilbert (2015). The properties of the simulations are given in Table 1. In all simulation snapshots, dark matter haloes are identified using a Friends Of Friends (FOF) algorithm (Davis et al. 1985). Each FOF halo has one central subhalo and its other subhaloes are categorized as satellites.

All subhaloes are detected using the SUBFIND algorithm (Springel et al. 2001). The algorithm sets a minimum of 20 particles for each subhalo to be included in the simulation catalogue. For every FOF halo, there is a virial radius,  $R_{200}$ , defined as the radius in which the matter density is 200 times the critical density of the Universe. The mass within  $R_{200}$  is known as the virial mass,  $M_{200}$ . Although it is common to consider  $R_{200}$  as the halo boundary, the FOF halo can extend beyond  $R_{200}$ . Therefore, satellite subhaloes of an FOF halo can exist beyond this scale as well.

#### 2.1.2 L-GALAXIES SAM of galaxy formation

L-GALAXIES is a semi-analytical galaxy formation model that uses a set of equations to model baryonic physics on top of dark matter halo merger trees. More recent versions of the model typically run on

**Table 1.** Parameters used in the Millennium (Springel et al. 2005) and Millennium-II (Boylan-Kolchin et al. 2009) simulations in their original (WMAP1, Spergel et al. 2003) and rescaled cosmology (Planck1, Planck Collaboration 2016) based on Angulo & White (2010) and Angulo & Hilbert (2015).

Simulation	$\Omega_m$	$\Omega_b$	$\Omega_\Lambda$	$H_0[\text{km s}^{-1} \text{Mpc}^{-1}]$	$n_s$	$\sigma_8$	$N_{\text{particles}}$	$m_{\text{particle}}[\text{M}_\odot]$	$l_{\text{box}}[\text{Mpc}]$	$N_{\text{snapshot}}$	$z_{\text{initial}}$
Millennium (WMAP1)	0.25	0.045	0.75	73	1	0.9	$2160^3$	$1.18 \times 10^9$	684	64	127
Millennium (Planck1)	0.315	0.049	0.685	67.3	0.96	0.826	$2160^3$	$1.43 \times 10^9$	714	64	56
Millennium-II (WMAP1)	0.25	0.045	0.75	73	1	0.9	$2160^3$	$9.42 \times 10^6$	137	68	127
Millennium-II (Planck1)	0.315	0.049	0.685	67.3	0.96	0.826	$2160^3$	$1.14 \times 10^7$	143	68	56

top of the halo merger trees from the Millennium and Millennium-II simulations (Springel et al. 2005; Croton et al. 2006; De Lucia et al. 2006; Bertone, De Lucia & Thomas 2007; Boylan-Kolchin et al. 2009; Guo et al. 2011; Yates et al. 2013; Henriques et al. 2015, 2020). In this paper, we fully calibrate a variant of the latest version of the model (Henriques et al. 2020, H20 hereafter), in which galaxy discs are resolved into a set of concentric rings and a novel gas stripping method is implemented to improve the model’s predictions for the environmental dependence of galaxy properties. Following previous versions, we also run the model on top of the Millennium and Millennium-II simulations. A full description of the H20 model is given in the supplementary material of H20.<sup>1</sup>

In L-GALAXIES, baryonic matter bound to each galaxy/subhalo is divided into seven main components: hot gas, cold gas (partitioned into H I and H<sub>2</sub>), stellar disc, bulge stars, halo stars, the supermassive black holes, and ejected material. The model initiates baryonic physics for every subhalo by seeding it with the expected fraction of diffuse hot (non-star-forming) gas at its formation time. This gas radiates energy and cools to form cold gas, containing neutral and molecular hydrogen components. Once the H<sub>2</sub> surface mass density exceeds a certain limit, stars are born. Both the cold gas and stellar components in the galactic disc are followed in 12 concentric rings and all the relevant physical processes happen within these rings.

Galaxy mergers play a key role in starbursts and the formation and growth of supermassive black holes at the centres of galaxies. Feedback processes such as stellar and black hole feedback quench star formation for low-mass and massive galaxies, respectively. The energy released by supernova feedback heats the cold gas and pushes it out of the galactic disc into the hot gas component. The remaining energy is able to drive the material of the hot gas component out of the subhalo into a reservoir of ejected material where it is no longer available for cooling. The time-scale for the reincorporation of the ejected gas is assumed to be proportional to  $1/M_{200}$  (see Fig. 1 in the supplementary material). As a result, gas returns to massive haloes quickly, while a fraction of the gas ejected out of low-mass haloes may never return.

Another important set of physical processes that can quench star formation is environmental effects such as tidal and ram-pressure stripping. Early versions of L-GALAXIES completely stripped hot gas out of satellite galaxies once they fell within  $R_{200}$ . Modern versions, since Guo et al. (2011), have gradual hot gas stripping, and therefore satellite galaxies are able to retain a fraction of their hot gas. The latest version of the model (H20) implements tidal stripping for satellite galaxies within the halo boundary,  $R_{200}$ , and limits ram-pressure stripping to satellites within  $R_{200}$  of massive clusters with  $M_{200} > M_{\text{r.p.}}$ , where  $\log_{10}(M_{\text{r.p.}}/M_\odot) = 14.7$ . This

ram-pressure stripping threshold,  $M_{\text{r.p.}}$ , was a free parameter in the H15 and H20 model calibrations, and the resulting value was found necessary in order to avoid having too many low-mass, red galaxies. We note that although this results in good agreement with the observed quenched fraction, the approach is merely a numerical fix and is not physical. In this work, we completely remove this mass threshold and extend ram-pressure stripping to all galaxies in the simulations. The implementation of the new gas stripping method is fully explained in Ayromlou et al. (2019) and also briefly in Section 3 of this paper.

In L-GALAXIES there are three types of galaxies: (i) central galaxies, also known as type 0 galaxies, which reside in the central subhalo of each FOF halo. (ii) Satellite galaxies with an identified subhalo, i.e. detected by the subhalo finder algorithm (here SUBFIND), which comprise all other subhaloes in an FOF besides the central. These are called type 1 galaxies. (iii) Satellite galaxies without an identified subhalo, i.e. their subhaloes have been tidally disrupted below the simulation resolution and are no longer detectable. These are called orphan, or type 2, galaxies. The model tracks the position and velocity of type 2 galaxies by following the most bound particle of their former subhalo. This way, L-GALAXIES continues the evolution of orphan galaxies even when their subhaloes are no longer detectable. In the rest of this paper, we use the phrases ‘subhalo’ and ‘galaxy’ interchangeably with the exception of ‘orphan galaxies’.

## 2.2 LBE measurements

In order to model ram-pressure, we must quantify the local environmental properties of galaxies. To do so, we follow the method introduced by Ayromlou et al. (2019), which measures the LBE of each galaxy directly from the particle data of the simulation. This background is defined within a spherical shell surrounding the galaxy and its subhalo. Following Ayromlou et al. (2019), we choose the background shell’s radii to exclude the galaxy and its subhalo, and keep the LBE sufficiently local. We choose the inner radius,  $r_{\text{in}} = 1.25 r_{\text{subhalo}}$  and the outer radius,  $r_{\text{out}} = 2 r_{\text{subhalo}}$ , where  $r_{\text{subhalo}}$  is the subhalo radius and is defined as the distance between the most bound and the most distant subhalo particles. In order to have proper statistics, we set a minimum of  $n_{\text{min}} = 30$  for the number of particles within the shell, which results in a statistical error smaller than  $1/\sqrt{n_{\text{min}}} \sim 20$  per cent. If there are fewer than  $n_{\text{min}}$  particles within the shell, we allow the outer radius to extend as needed. The shell density,  $\rho_{\text{shell}}$ , is the total mass within the shell divided by its volume. The mean velocity of the shell,  $\bar{v}_{\text{shell}}$ , is calculated by averaging over the velocities of all the shell particles.

We choose the inner radius of the background shell larger than the subhalo size to remove all the detected (via SUBFIND) bound subhalo particles. However, the background shell might still contain subhalo-associated particles which are not detected as a part of the

<sup>1</sup>An updated version of this Supplementary Material, including a full description of our local environment estimation method and our stripping recipes, is appended to the online version of this paper.

**Table 2.** Summary of model updates regarding gas stripping as described in Section 3.

Ram-pressure stripping		
Model ->	H20	This work
Formula	Equation (6)	Equation (6): $R_{\text{rp}} = \left( \frac{GM_g M_{\text{hotgas}}}{4\pi R_g R_{\text{hotgas}} \rho_{\text{LBE, hotgas}} v_{\text{gal, LBE}}^2} \right)^{1/2}$
Implemented for	Satellite subhaloes within $R_{200}$ of massive clusters	All satellite and central galaxies
$\rho_{\text{LBE, hotgas}}$ (Local density)	Isothermal $\rho \propto r^{-2}$	Local density measured directly from the particle data of simulations
$v_{\text{gal, LBE}}$ (galaxy's velocity relative to its environment)	Virial velocity of the satellite's host halo	The velocity of galaxy relative to its local environment, measured from the particle data of simulations
$R_g$ (known radius to estimate gravity)	Satellite galaxies: $R_{200}$ at infall time	Satellite galaxies: half-mass radius $R_{\text{halfmass}}$
	Central galaxies: N/A	Central galaxies: $R_{200}$
$M_g$ (known mass to estimate gravity)	Satellite galaxies: $M_{200}$ at infall time	Satellite galaxies: Total mass within $R_{\text{halfmass}}$
	Central galaxies: N/A	Central galaxies: $M_{200}$
Tidal stripping		
Formula	Equation (8)	Equation (8): $R_{\text{tidal}} = \frac{M_{\text{DM}}}{M_{\text{DM, infall}}} R_{\text{hot, infall}}$
Implemented for	Satellite galaxies within $R_{200}$ of haloes	All FOF satellite galaxies
The stripping radius and the fate of the stripped gas		
Satellite within $R_{200}$	$R_{\text{strip}} = \min(R_{\text{rp}}, R_{\text{tidal}})$ The stripped gas is added to the hot gas of its host halo After stripping: $R_{\text{hotgas}} = \min(R_{\text{hotgas, old}}, R_{\text{strip}})$	Same as H20
Satellite beyond $R_{200}$	N/A	$R_{\text{strip}} = \min(R_{\text{rp}}, R_{\text{tidal}})$ The stripped gas stays outside its host halo and is allowed to fall inside $R_{200}$ if the galaxy itself falls inside $R_{200}$ of its host halo. After stripping: $R_{\text{hotgas}} = \min(R_{\text{hotgas, old}}, R_{\text{strip}})$
Central galaxy	N/A	$R_{\text{strip}} = R_{\text{rp}}$ If the galaxy falls into $R_{200}$ of a halo later, the stripped gas is also allowed to fall into that halo. After stripping: $R_{\text{hotgas}} = R_{200}$

subhalo by the subhalo finder algorithm. These particles contaminate the background shell and cause misleading values for its density and velocity, and in Ayromlou et al. (2019) we devised a Gaussian mixture method to remove this contribution. In this paper, we use the decontaminated, pure LBE properties to measure ram-pressure for every simulation galaxy.

Our final step is to convert the LBE properties taken from DMO simulations to local estimates of the properties of the gas, needed to estimate ram-pressure. For both the LBE velocity and density, we assume that gas and dark matter follow each other. The LBE velocity is therefore  $\vec{v}_{\text{LBE, gas}} = \vec{v}_{\text{LBE}}$ . To derive the gas density we take slightly different approaches for central and satellite galaxies. For central galaxies, we multiply the total LBE density,  $\rho_{\text{LBE}}$ , by the cosmic baryon fraction  $\Omega_b$ , while for satellites we multiply it by  $f_{\text{hotgas}}$ , which is the hot gas fraction of the central subhalo of their parent host FOF halo. This follows the L-GALAXIES assumption that both dark matter and hot gas within haloes have isothermal ( $\rho \propto r^{-2}$ ) density profiles. The details of the method are explained in section 3 of Ayromlou et al. (2019), while Ayromlou et al. (in preparation) extends this method to measure the local gaseous environment of galaxies in the IllustrisTNG and Eagle hydrodynamical simulations.

### 3 NEW MODEL UPDATES

#### 3.1 Hot gas stripping

Haloes accrete hot gas (and dark matter), a process known as cosmic infall. In L-GALAXIES, the amount of this gas is set to be proportional to the amount of infalling dark matter. The hot gas radius,  $R_{\text{hotgas}}$ , of every central subhalo is set equal to the virial radius ( $R_{200}$ ) of its host halo. Satellite subhaloes are assumed to have no gas infall, and their hot gas radius is determined based both on their hot gas radius at infall time and on the environmental effects they experience.

If a subhalo moves fast enough in a dense environment, it can lose gas due to stripping processes such as tidal and ram-pressure stripping. In the H20 version of L-GALAXIES, these processes only occur to satellites within the halo  $R_{200}$  (tidal stripping) and only in satellites within  $R_{200}$  of massive clusters with  $\log_{10}(M_{200}/M_{\odot}) > 14.7$  (ram-pressure stripping). We follow the method introduced by Ayromlou et al. (2019) to calculate ram-pressure stripping more self-consistently, removing these artificial boundaries in position and halo mass. We extend stripping processes so that they happen to all galaxies, regardless of whether they are a central or satellite. Below we explain our implementation of ram-pressure and tidal



stripping of hot gas in the new variant (see Ayromlou et al. 2019 for a more extensive derivation). The changes are also summarized in Table 2.

### 3.1.1 Ram-pressure stripping

If the ram-pressure force from a subhalo's environment becomes stronger than the subhalo's self-gravity on its gas, the gas beyond a given scale, called the stripping radius  $R_{\text{rp}}$ , will be stripped out of the subhalo. The stripping radius is defined as the radius where these two forces are equal. We calculate the ram-pressure using the formula derived by Gunn & Gott (1972)

$$P_{\text{rp}} = \rho_{\text{LBE, gas}} v_{\text{gal, LBE}}^2, \quad (1)$$

where  $\rho_{\text{LBE, gas}}$  and  $v_{\text{gal, LBE}}$  are the gas density of the galaxy's LBE, and the velocity of the galaxy relative to its local environment, respectively. Both are measured directly using the simulation particle data as described in Section 2.2. In contrast, L-GALAXIES models prior to this work have adopted  $\rho_{\text{LBE, gas}} \propto r^{-2}$  and  $v_{\text{gal, LBE}}$  equal to the virial velocity ( $V_{200}$ ) of the satellite's host halo, with no ram-pressure force for central galaxies nor satellites beyond  $R_{200}$ .

For the gravitational restoring force we adopt the approach suggested by McCarthy et al. (2008) as extended in Ayromlou et al. (2019). This approach was implemented in L-Galaxies by Guo et al. (2011). However, it was limited to satellite galaxies within the halo  $R_{200}$ . The gravitational restoring force per unit area is

$$F_g(r) = g_{\text{max}}(r) \rho_{\text{hotgas}}^{\text{proj}}(r), \quad (2)$$

where  $g_{\text{max}}(r)$  is the maximum gravitational acceleration of the subhalo on its hot gas and  $\rho_{\text{hotgas}}^{\text{proj}}(r)$  is the subhalo's projected hot gas density. Considering a spherically symmetric isothermal profile with  $\rho_{\text{hotgas}} \propto r^{-2}$  for the hot gas,  $g_{\text{max}}(r)$  equals

$$g_{\text{max}}(r) = \frac{GM_{\text{subhalo}}(r)}{2r^2}, \quad (3)$$

where  $M_{\text{subhalo}}(r)$  is the subhalo mass within the radius  $r$ . Taking a similar isothermal  $\rho_{\text{subhalo}} \propto r^{-2}$  profile for the subhalo, we can calculate the subhalo mass at any given scale as

$$M_{\text{subhalo}}(r) = M_g \frac{r}{R_g}, \quad (4)$$

where  $M_g$  is the mass within  $R_g$  and both are known variables which we use to estimate  $M_{\text{subhalo}}(r)$ . In earlier L-GALAXIES versions,  $R_g$  and  $M_g$  were taken to be  $R_{200}$  and  $M_{200}$  at infall time for satellite galaxies. Although these are reasonable estimates, they represent the satellite subhalo at an earlier time and ignore its evolution. Therefore, for satellite galaxies we take  $R_g$  and  $M_g$  to be the half-mass radius,  $R_{\text{halfmass}}$ , and the total mass within  $R_{\text{halfmass}}$  radius ( $M_{\text{halfmass}}$ ) at every redshift, thus representing the subhalo's current properties rather than its properties at infall. For central galaxies, we take  $R_g = R_{200}$  and  $M_g = M_{200}$ .

Assuming an isothermal ( $\rho_{\text{hotgas}} \propto r^{-2}$ ) profile for the subhalo's hot gas, with the boundary condition of having total mass of  $M_{\text{hotgas}}$  within the hot gas radius,  $R_{\text{hotgas}}$ , the 2D projected hot gas density is given as

$$\rho_{\text{hotgas}}^{\text{proj}}(r) = \frac{M_{\text{hotgas}}}{2\pi R_{\text{hotgas}} r}. \quad (5)$$

Finally, the ram-pressure radius, the scale on which the ram-pressure equals the gravitational restoring force per unit area, is calculated to

be

$$R_{\text{rp}} = \left( \frac{GM_g M_{\text{hotgas}}}{4\pi R_g R_{\text{hotgas}} \rho_{\text{LBE, hotgas}} v_{\text{gal, LBE}}^2} \right)^{1/2}. \quad (6)$$

We apply the above formula to all the subhaloes uniformly.

### 3.1.2 Tidal stripping

Unlike ram-pressure that strips gas from subhaloes, tidal stripping is able to strip both dark and baryonic matter out of subhaloes and galaxies. As the total mass change of satellite subhaloes is known from our DMO simulations, we use it to estimate the amount of the stripped gas. We assume that the fraction of gas lost is the same as the fraction of dark matter lost

$$\frac{M_{\text{hot}}(R_{\text{tidal}})}{M_{\text{hot, infall}}} = \frac{M_{\text{DM}}}{M_{\text{DM, infall}}}, \quad (7)$$

where  $M_{\text{DM, infall}}$  and  $M_{\text{hot, infall}}$  are the satellite's virial mass and hot gas mass at infall. In addition,  $M_{\text{DM}}$  and  $M_{\text{hot}}(R_{\text{tidal}})$  are the satellite's mass and hot gas mass after tidal stripping. The only unknown variable in equation (7) is  $M_{\text{hot}}(R_{\text{tidal}})$ . Assuming an isothermal profile ( $\rho \propto r^{-2}$ ), we calculate the tidal stripping radius beyond which the gas is stripped

$$R_{\text{tidal}} = \frac{M_{\text{DM}}}{M_{\text{DM, infall}}} R_{\text{hot, infall}}, \quad (8)$$

where  $R_{\text{hot, infall}}$  is the satellite's hot gas radius at infall time.

As mentioned in Section 2.1.2, orphan galaxies do not have an identified subhalo, and therefore their subhalo mass is set to zero. Equation (7) then implies that they will lose all their hot gas as well. As a result, orphan galaxies are entirely empty of hot gas due to tidal stripping. We note that the fraction of orphan galaxies depends on the resolution of the underlying simulation and the subhalo finder employed to detect subhaloes. Using a higher resolution simulation leads to a lower fraction of orphan galaxies and subsequently higher accuracy in gas-relevant processes. Furthermore, simply changing the minimum number of particles required to detect a subhalo, as well as using different subhalo finding techniques, can change the number of orphan galaxies. As mentioned in Section 3.3, we use the L-Galaxies run on the Millennium simulation for galaxies with  $\log_{10}(M_*/M_\odot) > 9$  and its run on the higher resolution Millennium-II simulation for lower mass galaxies. This decreases the errors relevant to orphan galaxies (see also Guo et al. 2011). In some studies, analytic formulae of the evolution of the subhalo mass over time are used to extend the estimation of the subhalo mass beyond the point where it is no longer detectable in the simulation, and in this case orphan galaxies retain a non-zero fraction of their hot gas (e.g. see Luo et al. 2016). However, we do not adopt any such scheme in this work.

### 3.1.3 Stripping implementation and the fate of the stripped gas

Satellite galaxies are subject to both tidal and ram-pressure stripping. Therefore, we take the stripping radius to be the smaller of the tidal stripping radius and the ram-pressure stripping radius. For central galaxies, the stripping radius is equal to the ram-pressure stripping radius, as we do not consider them subject to tidal stripping. After stripping, we assume the density profile of the remaining gas remains isothermal with  $\rho \propto r^{-2}$ . We note that ram-pressure stripping is the dominant effect for most galaxies.

In addition to hot gas stripping we also strip the ejected reservoirs of galaxies. The fraction of gas stripped from the ejected reservoir of a galaxy is assumed to be the same as the fraction of its stripped hot

gas. Stripping of material that has been expelled from the galaxy by feedback processes and is assumed to lie at least temporarily in this ejecta reservoir plays a critical role in the evolution of galaxies (see Section 5.1).

For every galaxy, if  $R_{\text{strip}} < R_{\text{hotgas}}$ , all the gas beyond  $R_{\text{strip}}$  is removed. For satellite galaxies within the halo boundary,  $R_{200}$ , the stripped gas goes directly to the hot gas component of their host halo. For central galaxies and satellites beyond the halo  $R_{200}$ , the stripped gas does not go to any halo immediately, but we keep track of it through time. If the galaxy subsequently falls into  $R_{200}$  of a halo, then the stripped gas is added to the host halo's hot gas with the condition that the host halo's baryon fraction does not exceed the cosmic value. After stripping, the new hot gas radius of a satellite will be the minimum of its former hot gas radius and  $R_{\text{strip}}$ . On the other hand, as long as a galaxy is categorized as a central galaxy, since it accretes hot gas from its environment, its hot gas radius is set to its FOF  $R_{200}$ . We note that the gas stripped out of a galaxy never returns to that galaxy.

### 3.1.4 Tracking gas stripping

We measure the amount of gas lost due to stripping in different conditions and separate it from other intrinsic physical processes. In order to evaluate the impact and predictions of our method quantitatively, we define the 'cumulative stripped fraction', or simply the 'stripped fraction' of a subhalo at a given redshift as

$$f_{\text{stripped}}(z) = M_{\text{stripped}}(z) / (f_b M_{200, \text{max}}), \quad (9)$$

where  $M_{200, \text{max}}$  is the maximum  $M_{200}$  of the main progenitor of the subhalo through time.  $M_{\text{stripped}}(z)$  is the total mass which has been stripped from the subhalo since its formation time until redshift  $z$ . This includes gas stripped from the hot gas component and the ejected reservoir including material stripped while the subhalo was a central, unless we explicitly decompose the two. This is a cumulative quantity that we track for every galaxy across time.

As an overview of our new model, Fig. 1 illustrates a massive cluster with  $\log_{10}(M_{200}/M_{\odot}) \sim 14.5$  and five galaxies that end up as its satellites by  $z = 0$ . Each galaxy is shown by one colour, and in the top panel, the transparency of that colour shows the cumulative stripped fraction of the galaxy (equation 9) at each simulation snapshot. The formation redshifts, and the redshifts at which each galaxy starts to be significantly stripped, are indicated by arrows. It is clear that stripping processes start long before galaxies fall within the virial radius ( $R_{200}$ ) of their present-day halo (dashed circle at the centre). We find that, on average, a satellite galaxy loses more than 70–80 per cent of its hot gas due to stripping prior to its infall (see Section 5.1). Our five sample galaxies have experienced their first infall into an 'FOF halo' at  $z \sim 0.15$  (#1, orange),  $z \sim 1.6$  (#2, green),  $z \sim 4.3$  (#3, blue),  $z \sim 0.2$  (#4, purple), and  $z \sim 0.8$  (#5, grey). Galaxies #2 (green) and #3 (blue) are pre-infall satellites, i.e. they are already satellites of other haloes at  $z > 0$ , prior to their infall into the central cluster. These two galaxies have lost a considerable fraction of their hot gas through pre-processing in their initial hosts.

On the other hand, the three other galaxies (#1, #4, and #5) are categorized as central galaxies until they fall directly into their  $z = 0$  host halo. These galaxies have lost their hot gas via ram-pressure stripping when moving fast through the dense environment beyond the halo boundary of their present-day host halo. The middle panels of Fig. 1 show the position of the sample galaxies in the cosmic web at two different redshifts,  $z = 2$  and  $z = 0.26$ . They illustrate the times when the green (#2) and purple (#4) galaxies (left-hand panel)

and the orange galaxy (#1) are strongly stripped. For example, galaxy #1 moves through the overdensity around the central host halo, but extending far beyond its virial radius, at which there is no abrupt change in density. All five galaxies pass through dense environments at some time in their history, losing a large fraction of their gas.<sup>2</sup> It is clear that the matter distribution of the proto-cluster is very extended, but there is no significant change in the density of the central region of the halo.

Finally, the bottom panels of Fig. 1 depict the ram-pressure (left), the cumulative stripped fraction (middle), and the gas mass (right) of all the cluster's  $z = 0$  surviving satellite subhaloes as a function of redshift. The previously highlighted galaxies are illustrated as thick curves, while the rest of the galaxies are shown with thin grey lines. The vertical lines correspond to the redshifts where the five sample galaxies start becoming significantly stripped. There is a strong correlation between high values of ram-pressure and the times when galaxies lose gas. Overall, Fig. 1 depicts the importance of modelling stripping beyond the virial radius, showing that galaxies can even lose their entire hot gas reservoir while still centrals, without ever infalling into a more massive host halo.

### 3.2 Gas infall into haloes

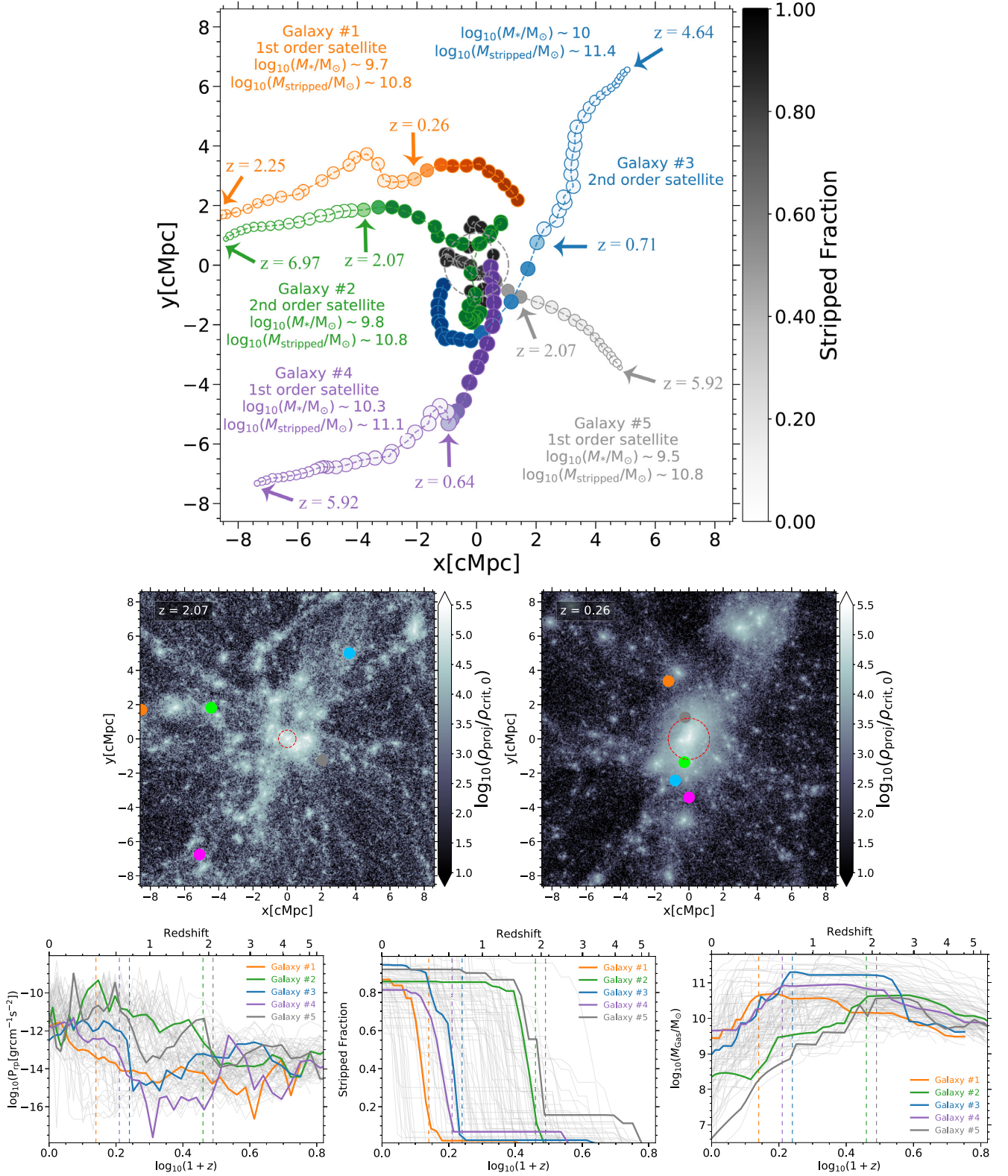
The L-GALAXIES infall recipe in H20 is that the gas accretion rate on to haloes is equal to the accretion rate of total matter times the cosmic baryon fraction. It is possible that the baryon fraction exceeds the cosmic value (see fig. 8 of Ayromlou et al. 2021). There are two reasons for this discrepancy, as described by Yates, Thomas & Henriques (2017). The first and more dominant effect is due to the implementation of environmental effects such as tidal and ram-pressure stripping. Satellite galaxies in the infall regions lose dark matter due to tidal effects in the DMO simulation on which L-GALAXIES is run. The lack of environmental stripping effects for gas beyond  $R_{200}$  can give these satellites a baryon fractions greater than the cosmic value when they fall into  $R_{200}$  of their host halo. This artificially increases the baryon fraction of their host, given that its infall has already topped-up the halo's baryon budget to the cosmic baryon fraction. In this work, we fully resolves this issue by having gas stripping for satellites beyond the halo  $R_{200}$  (see Section 3.1.3).

The second reason is that total halo mass can decrease with time because of changes in morphology or halo concentration or simply through numerical fluctuations (see De Lucia et al. 2004 for subhalo mass fluctuations). At the same time, in L-GALAXIES the halo baryonic mass within  $R_{200}$  is kept unchanged by construction, causing an increase in the baryon fraction. Following the prescription of Yates et al. (2017), we correct the input halo merger trees to prevent the  $M_{200}$  from decreasing with time. This accounts for any artificial decrease in  $M_{200}$  measured when determining  $R_{200}$  based on the assumption of spherical symmetry for haloes.

### 3.3 Model calibration

Like all galaxy formation models and simulations, L-GALAXIES has a number of parameters (e.g. the star formation efficiency) that need to be fit. Therefore, we calibrate our new model against a set of observational constraints using the MCMC approach developed by Henriques et al. (2009) and used by previous L-GALAXIES versions (e.g. Guo et al. 2013; H20). To speed up the model calibration, the

<sup>2</sup>A video showing the stripping process for our sample galaxies is given in the supplementary material of this paper.



**Figure 1.** Illustration of gas stripping in our model. Top panel: Five sample galaxies in the vicinity of a massive cluster with  $\log_{10}(M_{200}/M_\odot) \sim 14.5$ . Each colour corresponds to a galaxy, and different circles of the same colour demonstrate the evolution of that galaxy at different redshifts. The size of each circle represents, qualitatively, the galaxy's stellar mass. Shown by the colourbar, the transparency of colours denotes the fraction of the gas stripped (equation 9). Here, first-order satellites are satellites of the central cluster at  $z = 0$  which have not been satellites of other haloes previously, while second-order satellites were previously satellites of other haloes at higher redshifts. The stellar mass and total stripped mass of each galaxy are quoted at  $z = 0$ . Middle panel: Matter distribution at two redshifts of the simulation, depicting the positions of the five galaxies in the cosmic web. All coordinate units are comoving. Bottom panel: ram-pressure (left), stripped fraction (middle), and total gas mass (right) versus redshift. Each curve corresponds to a galaxy that is currently a satellite of the central cluster. The dashed vertical lines correspond to the redshift where our five chosen galaxies start getting stripped significantly.



MCMC approach runs the model on a representative set of halo merger trees (see Henriques et al. 2013) instead of on the full simulations, a scheme we also adopt. In addition, we extend the previous MCMC method to also incorporate our new treatment of environmental effects.

We use six independent observational constraints: the stellar mass function and the fraction of quenched galaxies, each at  $z = 0, 1, 2$ . Unlike H20, we do not use the H I gas mass function at  $z = 0$  as an observational constraint. However, our final model output for the H I mass function is in relatively good agreement with observations and H20 (bottom left panel of Fig. 7). Choosing proper observational constraints and weighting them to converge to an acceptable fit can be a tricky task. For instance, weighting all observational constraint equally would result in a rather bad fit for the stellar mass function at  $z = 0$ , especially for galaxies with  $10 < \log_{10}(M_*/M_\odot) < 11$ . Ultimately, exploring different weightings for different data sets, we find it best to give the highest weight to the observational constraints at  $z = 0$ . Furthermore, at  $z = 0$  itself, we give the stellar mass function a higher weight than the quenched fraction. To properly fit the  $z = 0$  stellar mass function for  $M^*$  galaxies, we give an additional weight to the stellar mass function at  $10 < \log_{10}(M_*/M_\odot) < 11$ . At  $z > 0$ , observational constraints are weighted equally.

We run the MCMC for several tens of thousands of steps, i.e. we execute our model with different free parameters tens of thousands of times. During the calibration, we use L-GALAXIES run on the Millennium simulation for galaxies with  $\log_{10}(M_*/M_\odot) > 9$  and L-GALAXIES run on Millennium-II for lower stellar masses. This stellar mass transition value is chosen following H20 and also by monitoring the approximate stellar mass where the two runs converge for a few smaller runs of the model (see also Guo et al. 2011; H20). We note that our model has one fewer free parameter than H20, as we no longer limit ram-pressure stripping to satellites of massive clusters. Table 3 compares our best-fitting parameters to previous models. The relevant equation for each free parameter is also given in that table. The description of the equations can be found in the supplementary material of this paper. In addition, histograms of the PDFs of accepted MCMC steps for all free parameters are shown in a corner plot in Fig. 2. 2D contour plots show the marginalized (and normalized) 2D posterior distributions for all the possible free parameter pairs of our model. Furthermore, 1D histograms depict the individual constraints on each parameter. The vertical dashed lines show the best fit from our model (green) and from H20 (purple), and the value quoted indicates the best-fitting value (in linear units).

Our best fits to the observational constraints are shown in Fig. 3 in comparison with observations and the H20 model. The top panels of Fig. 3 show that our fits (orange lines) to the stellar mass functions at different redshifts are slightly better than H20, in particular for low-mass galaxies. In addition, the bottom panels show that we have a higher fraction of massive quenched galaxies at higher redshifts, in better agreement with observational data. This is primarily due to the higher AGN feedback efficiency parameter (see Table 3). Overall, our model fits the targeted observational constraints successfully.

We emphasize that we *do not* include any environment-dependent quantities in the observational constraints used in our model calibration. As a result, we retain all environmental dependencies and correlations related to galaxy evolution as *predictions*, rather than the result of fits to observations. Our observational constraints, i.e. the stellar mass function and quenched fraction, are global values and do not distinguish between field galaxies and those in dense environments. In the rest of this work, we study our model's predictions regarding the properties of galaxies in different environments.

## 4 OBSERVATIONAL DATA AND MOCK CATALOGUES

### 4.1 Observational data

#### 4.1.1 The fraction of quenched galaxies at $z = 0$

For consistency and to ensure the most robust comparisons between our model and empirical data, we carry out a new analysis of observations which constrain galaxy quenched fractions as a function of environment. Our galaxy sample at  $z = 0$  is taken from the Sloan Digital Sky Survey Data Release 7 (SDSS-DR7, Abazajian et al. 2009), with stellar masses and star formation rates calculated using the methodologies described in Kauffmann et al. (2003), Brinchmann et al. (2004), and Salim et al. (2007). We cross-match the SDSS-DR7 catalogue with the group catalogue from Yang et al. (2005, 2007) to identify the most massive galaxy of each group, the so-called central galaxy. We note that there are a lot of complexities associated with group-finding in observational data (e.g. see Bravo et al. 2020; Tinker 2020). We estimate the halo  $M_{200}$  and  $R_{200}$  using our galaxy formation model by fitting the central galaxy stellar to halo mass relation as described in Section 4.2.2 and Appendix A.

For each halo, we find galaxies with line-of-sight velocity separation from the central galaxy in the halo  $|v_{\text{gal,LOS}} - v_{\text{halo,LOS}}| \leq \pm 10V_{200,\text{halo}}$  which lie within a 2D projected halocentric distance of  $10R_{200}$ . Here,  $V_{200,\text{halo}}$  is the halo virial velocity and is defined as  $V_{200,\text{halo}} = \sqrt{GM_{200,\text{halo}}/R_{200,\text{halo}}}$ . We note that for most of our analysis we only take galaxies with  $|v_{\text{gal,LOS}} - v_{\text{halo,LOS}}| \leq \pm 2V_{200,\text{halo}}$  (e.g. see Figs 9 and 10). In this work, we limit our SDSS galaxy sample to galaxies with  $\log_{10}(M_*/M_\odot) \geq 9.5$  and our observational halo sample to haloes with  $\log_{10}(M_{200}/M_\odot) \geq 12$ . The redshift interval of our host haloes is taken as  $0.01 \leq z \leq 0.04$ . Finally, we adopt the definition that galaxies with  $\log_{10}(\text{SSFR}/\text{yr}^{-1}) < -11$  are quenched.

A similar analysis has been carried out by Wetzel et al. (2012, 2014), where they use the SDSS data and a modified Yang et al. (2007) halo catalogue to study the properties of galaxies in different environments. In contrast, our new analysis is based on a different halo catalogue, halo mass derivation method, stellar and halo mass bins, and so on. Furthermore, the analysis in this paper is designed so that the results can be compared directly and self-consistent with our SAM model predictions.

#### 4.1.2 The fraction of quenched galaxies at $z > 0$

At  $z > 0$ , we compare our results with the fraction of quenched galaxies inferred from the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP) survey data of optical broadband photometry from Pintos-Castro et al. (2019). Their analysis is based on the HSC-SSP project (Aihara et al. 2018). In their study, galaxy clusters are taken from the SpARCS (Muzzin et al. 2009; Wilson et al. 2009) fields, using a modified cluster red sequence algorithm by Gladders & Yee (2000, 2005) and Muzzin et al. (2008) that employs photometric redshifts. Furthermore, their galaxy sample is complete for  $\log_{10}(M_*/M_\odot) > [8.6, 8.6, 9.5, 9.7]$  at  $z \sim [0.4, 0.6, 0.8, 1]$ , respectively (see Pintos-Castro et al. 2019 for more details).

### 4.2 Mock catalogues

In order to make a fair comparison with observations, we create mock catalogues both from our own model and from that of H20. The details are explained in this subsection.



**Table 3.** Free parameters used in the MCMC model calibration in this work and previous models (Guo et al. 2013; H15; H20). Those values labelled as ‘F.P.’ are fixed parameters, which are not taken from the MCMC best fit. A full model description is given in the Supplementary Material appended to the online version of this paper.

Model parameter	Equation in L-GALAXIES	G13	H15	H20	This work	Units
$\alpha_{\text{SF}}$ (Star formation efficiency)	$\Sigma_{\text{SFR}} = \alpha_{\text{SF}} \Sigma_{\text{H}_2} / t_{\text{dyn}}$	0.011	0.025	0.06	0.073	
$\alpha_{\text{SF, burst}}$ (Star formation burst efficiency)	$M_{*, \text{burst}} = \alpha_{\text{SF, burst}} \left( \frac{M_1}{M_2} \right)^{\beta_{\text{SF, burst}}} M_{\text{cold}}$	0.56	0.60	0.50	0.116	
$\beta_{\text{SF, burst}}$ (Star formation slope)	same as above	0.70	1.9	0.38	0.674	
$k_{\text{AGN}}$ (Radio feedback efficiency)	$\dot{M}_{\text{BH}} = k_{\text{AGN}} \left( \frac{M_{\text{hot}}}{10^{11} M_{\odot}} \right) \left( \frac{M_{\text{BH}}}{10^8 M_{\odot}} \right)$	N/A.	$5.3 \times 10^{-3}$	$2.5 \times 10^{-3}$	$8.5 \times 10^{-3}$	$M_{\odot} \text{ yr}^{-1}$
$f_{\text{BH}}$ (Black hole growth efficiency)	$\Delta M_{\text{BH, Q}} = \frac{f_{\text{BH}} (M_{\text{sat}} / M_{\text{cen}}) M_{\text{cold}}}{1 + (V_{\text{BH}} / V_{200c})^2}$	0.03	0.041	0.066	0.011	
$V_{\text{BH}}$ (Quasar growth scale)	same as above	280	750	700	1068	$\text{km s}^{-1}$
$\epsilon_{\text{reheat}}$ (Mass-loading efficiency)	$\epsilon_{\text{disc}} = \epsilon_{\text{reheat}} \left[ 0.5 + \left( \frac{V_{\text{max}}}{V_{\text{reheat}}} \right)^{-\beta_{\text{reheat}}} \right]$	4.0	2.6	5.6	9.7	
$V_{\text{reheat}}$ (Mass-loading scale)	same as above	80	480	110	119	$\text{km s}^{-1}$
$\beta_{\text{reheat}}$ (Mass-loading slope)	same as above	3.2	0.72	2.9	2.9	
$\eta_{\text{eject}}$ (Supernova ejection efficiency)	$\epsilon_{\text{halo}} = \eta_{\text{eject}} \left[ 0.5 + \left( \frac{V_{\text{max}}}{V_{\text{eject}}} \right)^{-\beta_{\text{eject}}} \right]$	0.18	0.62	5.5	9.56	
$V_{\text{eject}}$ (Supernova ejection scale)	same as above	90	100	220	172	$\text{km s}^{-1}$
$\beta_{\text{eject}}$ (Supernova ejection slope)	same as above	3.2	0.80	2.0	1.88	
$\gamma_{\text{reinc}}$ (Ejecta reincorporation)	$t_{\text{reinc}} = \gamma_{\text{reinc}} \frac{10^{10} M_{\odot}}{M_{200c}}$	N/A.	$3.0 \times 10^{10}$	$1.2 \times 10^{10}$	$6.6 \times 10^9$	1/yr
$M_{\text{TP}}$ (Ram-pressure threshold)	N/A	0.0	$1.2 \times 10^4$	$5.1 \times 10^4$	N/A	$10^{10} M_{\odot}$
$R_{\text{merger}}$ (Major-merger threshold)	N/A	0.30	0.1 (F.P.)	0.1 (F.P.)	0.1 (F.P.)	
$\alpha_{\text{friction}}$ (Dynamical friction)	$t_{\text{friction}} = \alpha_{\text{friction}} \frac{V_{200c}^2 r_{\text{sat}}^2}{G M_{\text{sat}} \ln(1 + M_{200c} / M_{\text{sat}})}$	2.0	2.5	1.8	0.312	
$f_{\text{z, hot, TypeII}}$	N/A	N/A.	N/A.	0.3 (F.P.)	0.3 (F.P.)	
$f_{\text{z, hot, TypeIa}}$	N/A	N/A.	N/A.	0.3 (F.P.)	0.3 (F.P.)	
$v_{\text{inflow}}$	$v_{\text{inflow}} = r/t_v$	N/A.	N/A.	1000 (F.P.)	1000 (F.P.)	$\text{km s}^{-1} \text{ Mpc}^{-1}$

#### 4.2.1 Projection in 2D

In order to compare with the SDSS data at  $z = 0$ , to locate galaxies in the vicinity of each halo in our model, we first transform the galaxy distribution into redshift/velocity space. This is done using their positions and peculiar velocities along the  $z$ -axis of the simulation domain, accounting for expansion due to the Hubble flow. For each simulated halo we locate nearby galaxies with  $|v_{\text{gal, LOS}} - v_{\text{halo, LOS}}| \leq \pm 2V_{200, \text{halo}}$  within  $10R_{200}$  projected halocentric distance.

In both our mock catalogues and the SDSS data, we take the most massive galaxy of each halo as its central galaxy. However, the algorithms employed to find the most massive galaxy of each group are not the same. We note that matching these algorithms may improve the comparison between our mock catalogues and the SDSS observations.

For our comparison with the HSC data (Pintos-Castro et al. 2019) at  $z > 0$ , we adopt the projection made by Pintos-Castro et al. (2019) within a redshift slice of

$$\Delta z = z_{\text{halo}} \pm 0.05 \times (1 + z_{\text{halo}}). \quad (10)$$

The physical distance corresponding to  $\Delta z$  could exceed our simulation box size for some redshifts. Therefore, for simplicity we project the entire simulation box along the  $z$ -axis of the simulation

volume ( $l_{\text{box}} \sim 700 \text{ Mpc}$  for the Millennium simulation in the Planck cosmology).

#### 4.2.2 Deriving halo mass and radius

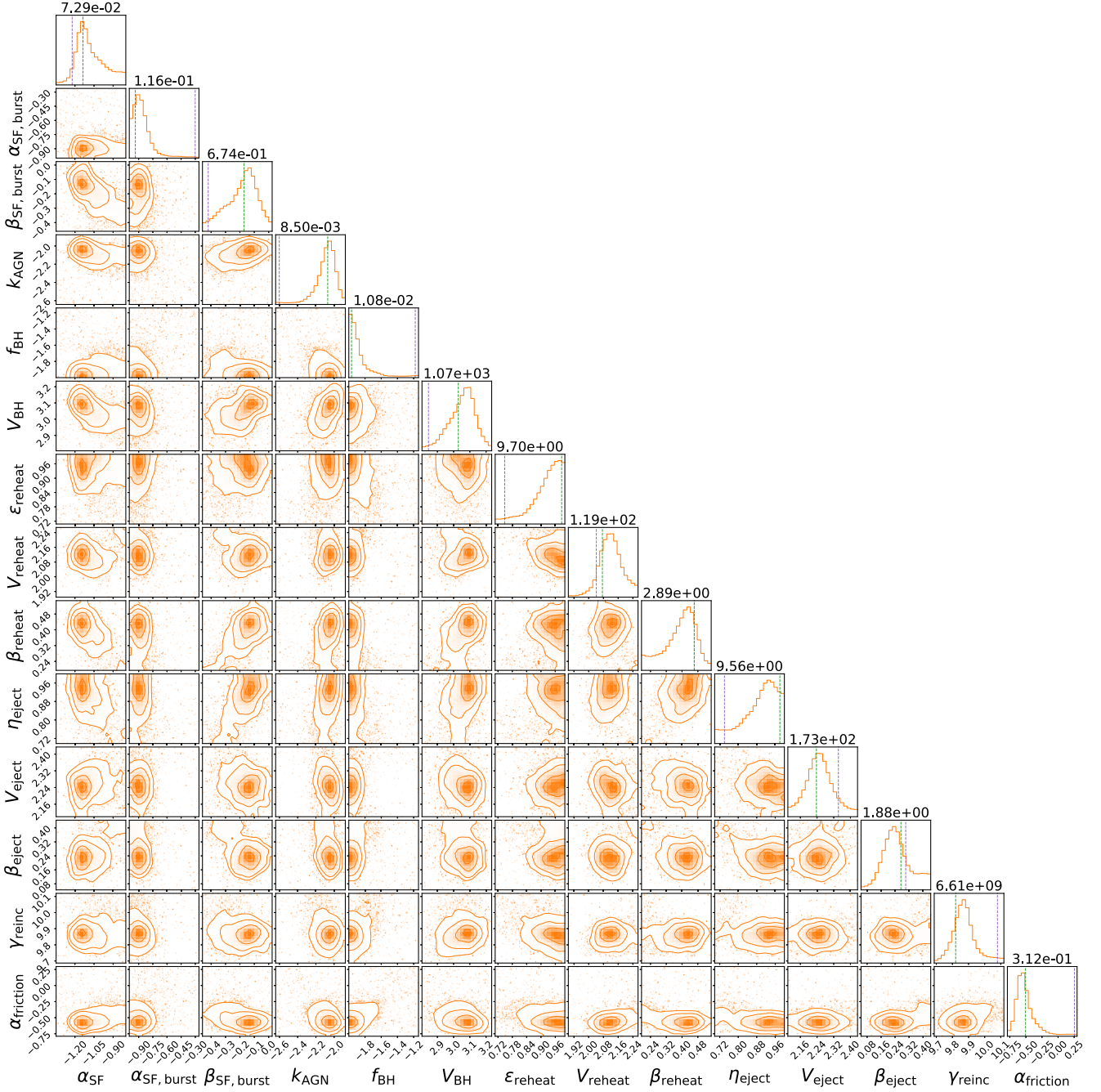
For each central galaxy in both models and observations we calculate a halo mass,  $M_{200}$ , from its stellar mass with the fitting formula below (motivated by Yang, Mo & van den Bosch 2003; Moster et al. 2010):

$$\log_{10}(M_{200}/M_{\odot}) = \alpha_1 \log_{10}(M_{*}/M_{\odot}) + \beta_1, \quad (11)$$

where  $\alpha_1$  and  $\beta_1$  are free parameters which we derive using our final, calibrated model output (see Figs A1 and A2 in Appendix A):  $\alpha_1 = 1.65$ ,  $\beta_1 = -5.16$  for  $\log_{10}(M_{*}/M_{\odot}) \geq 10.5$ , and  $\alpha_1 = 0.80$ ,  $\beta_1 = 3.70$  for  $\log_{10}(M_{*}/M_{\odot}) < 10.5$ . The above equation is valid for galaxies with  $\log_{10}(M_{*}/M_{\odot}) > 9.5$ . The stellar mass at which  $\alpha_1$  and  $\beta_1$  change is  $\log_{10}(M_{*}/M_{\odot}) = 10.5$  which roughly corresponds to haloes with  $\log_{10}(M_{200}/M_{\odot}) \sim 12.2$ . For each halo,  $R_{200}$  is calculated from  $M_{200}$  through

$$R_{200} = \left( \frac{3M_{200}}{4\pi \times 200\rho_{\text{crit}}} \right)^{1/3}. \quad (12)$$

We note that independent values of halo mass and radius are given in the Yang et al. (2005, 2007) catalogue as well. These were calculated



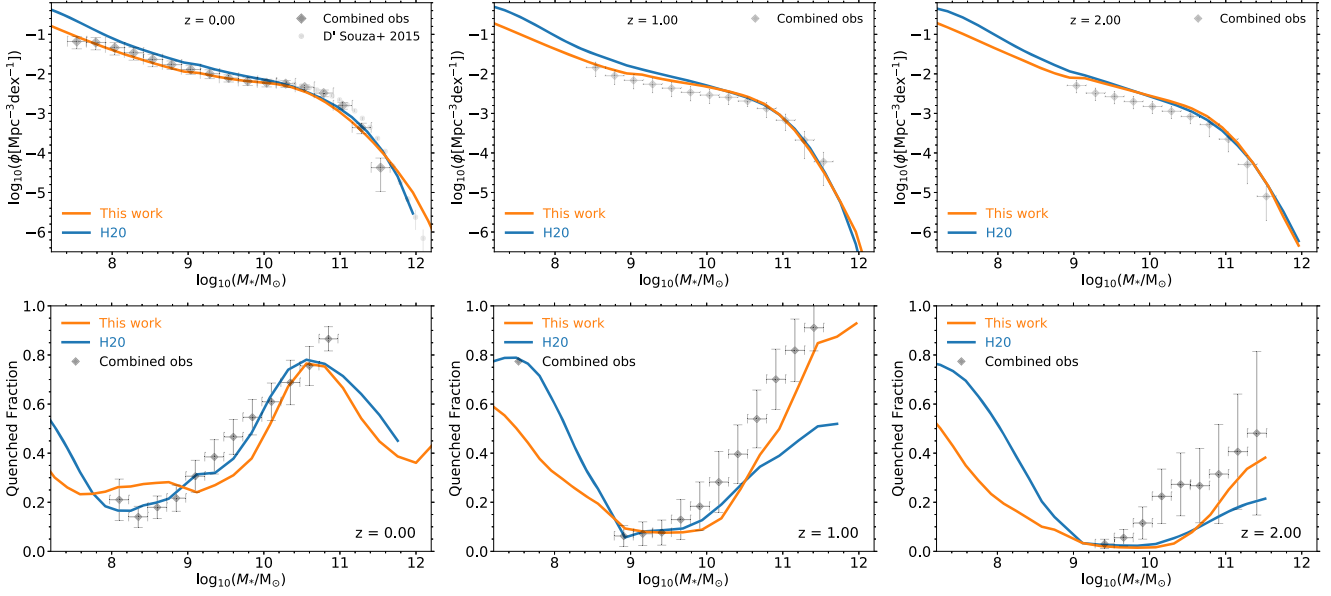
**Figure 2.** Histograms of the weighted accepted MCMC steps (after burn-in) for all the free parameters used in this model. The green dashed lines correspond to our best fit for each parameter and the purple dashed lines are the best fit in H20. The x-axes and y-axes in all panels are reported in logarithmic units. The value above each 1D histograms corresponds to the best fit (in linear units) of that parameter in this work. This figure is generated using  $\sim 23\,000$  MCMC steps.

from abundance matching using the observed stellar mass function and an analytical halo mass function. In contrast, here we use the median stellar mass to halo mass ratio at each stellar mass bin in our model, which does not invoke the same assumptions as abundance matching. In practice, the differences between their values and ours are quite small.

For our direct comparisons with the SDSS observations, wherever the halo mass,  $M_{200}$ , and the halo radius,  $R_{200}$ , are needed, we derive  $M_{200}$  using equation (11). The corresponding  $R_{200}$  is then derived using equation (12). We do this for the data *as well as* for the models.

We note that such a derivation of  $M_{200}$  and  $R_{200}$  leads to considerable scatter in halo mass and radius compared with the actual values measured directly from the simulation. This is explained in more detail in Appendix A (see Figs A1 and A2).

For our comparison with the HSC data (Pintos-Castro et al. 2019) at  $z > 0$ , we use the direct simulation halo masses rather than making a mock catalogue. This is because there are many complexities in the technique that work employs to derive halo masses, which are beyond the level we wish to emulate in this paper. We note that this makes our high-redshift comparisons less reliable (see Section 5).



**Figure 3.** Model results for the quantities used as observational constraints in the MCMC in comparison with observations and with the H20 model. The top panels show the stellar mass functions and the bottom panels show the fraction of quenched galaxies. The combined observations used in the MCMC calibration are taken from H15 (see their appendix 2 for more details) and include Baldry, Glazebrook & Driver (2008), Baldry et al. (2012), Li & White (2009) for SMF at  $z = 0$ , Marchesini et al. (2009, 2010), Ilbert et al. (2010, 2013), Domínguez Sánchez et al. (2011), Muzzin et al. (2013), Tomczak et al. (2014) for the SMF at  $z > 0$ , Bell et al. (2003), Baldry et al. (2004) for the quenched fraction at  $z = 0$ , and Muzzin et al. (2013), Ilbert et al. (2013), Tomczak et al. (2014) for the quenched fraction at  $z > 0$ .

#### 4.2.3 Homogenization of the halo mass distributions between models and observations

In general, the halo mass distribution within a given halo mass bin (e.g.  $14 < \log_{10}(M_{200}/M_{\odot}) < 15$ ) can differ between the observations and the models. This bias could affect our results and conclusions. Therefore, we resample the distribution of simulated haloes to match the observed distributions.

Assume that we want to compare the quantity ‘ $Q$ ’ (e.g. the fraction of quenched galaxies) between a model and an observation, as a function of  $M_{200}$ . We divide each halo mass bin (e.g.  $14 < \log_{10}(M_{200}/M_{\odot}) < 15$ ) in the model into smaller sub-bins and calculate the quantity  $Q$  within those sub-bins. The final value of the quantity  $Q$  within the larger bin is

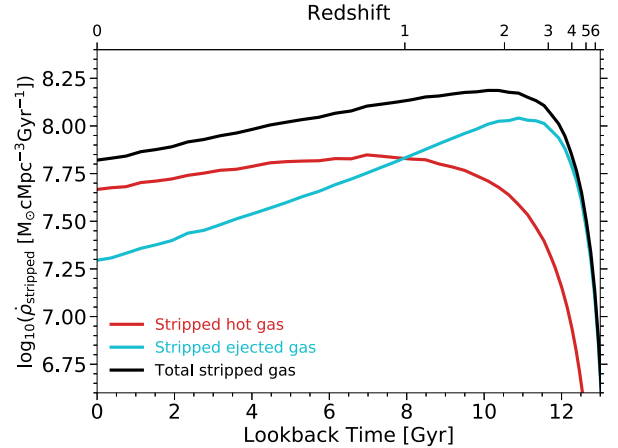
$$Q = \sum_{i=1}^n Q_i W_i, \quad (13)$$

where  $Q_i$  is the value (mean, median, etc.) of  $Q$  within the  $i$ th sub-bin and ‘ $n$ ’ is the number of sub-bins. In addition,  $W_i$  is the weight of the  $i$ th sub-bin, defined as the ratio of the fraction of haloes in the  $i$ th sub-bin of the observational sample to the fraction of haloes in the  $i$ th sub-bin of the model. This homogenization is only done for our comparison with the SDSS data at  $z = 0$ .

## 5 RESULTS

### 5.1 Gas stripping through time

As a first, global exploration of the role of gas stripping, we define the stripping rate density,  $\dot{\rho}_{\text{stripped}}(t)$ , as the stripped mass in the entire



**Figure 4.** Time evolution of the stripping rate density as defined in equation (14). The red and cyan lines correspond to the stripping rate density of the hot gas (red) and the ejected reservoir (cyan), respectively. Summing the two, the black line is the total gas mass stripping rate density.

simulation per unit time (Gyr), normalized to the simulation volume:

$$\dot{\rho}_{\text{stripped}}(t) = \frac{1}{V_{\text{simulation}}} \sum_{i=1}^{N_{\text{gal}}} \dot{m}_{\text{stripped},i}(t), \quad (14)$$

where  $\dot{m}_{\text{stripped},i}(t)$  is the gas stripping rate of the  $i$ th galaxy per unit time, and the sum is over all galaxies in the simulation ( $N_{\text{gal}}$ ). In addition,  $V_{\text{simulation}}$  is the comoving simulation volume.

Fig. 4 shows  $\dot{\rho}_{\text{stripped}}(t)$  as a function of lookback time (lower x-axis) as well as redshift (upper x-axis) in our model’s run on the Millennium simulation. Considering the redshift evolution of  $\dot{\rho}_{\text{stripped}}$ , most of the stripped gas at  $z > 1.2$  is due to the stripping of

the ejected material (cyan line) rather than the hot halo gas (red line), demonstrating a strong correlation between feedback processes and gas stripping at high redshifts. At  $z < 1.2$ , stripping of the hot gas is the dominant stripping process.

In L-GALAXIES, the outflows contributing to the ejected reservoir are caused by supernova feedback and are present mostly for low-mass galaxies. We note that AGN feedback in our model prevents the hot gas from cooling but does not eject it outside the halo. Therefore, it does not contribute to the material in the ejected reservoir. Some hydrodynamical simulations (e.g. IllustrisTNG) have strong outflows caused by black hole feedback in massive galaxies. As discussed in Ayromlou et al. (2021), the presence of ejective black hole feedback results in a significant enhancement in ram-pressure stripping for massive galaxies. We will consider the possibility of ejective black hole feedback and study its influence on gas stripping in L-GALAXIES in future work.

The total stripping rate density (black line) does not decrease monotonically with redshift, but it has a maximum at  $z \sim 2$ . The same is true for the ejected stripped density rate (blue line) with a maximum at  $z \sim 2$ , implying a close relationship to the peak of cosmic star formation rate density which also occurs at  $z \sim 2$  (see Fig. 13). Due to the maximal total star formation rate at  $z \sim 2$  there is an abundance of powerful stellar feedback activity in L-GALAXIES, and the largest fraction of ejected material at this cosmic epoch. This results in the suppression of the hot gas density in galaxies and, as a result, the suppression of the gravitational force between the galaxies and their hot halo gas (equation 2). Therefore, the hot and ejected gas of galaxies are more easily stripped.

Fig. 5 shows the median cumulative stripped fraction (equation 9) up to  $z = 0$ , versus halocentric distance, for galaxies in the vicinity of clusters (right row), groups (middle row), and lower mass haloes (left row). We stack all galaxies in the vicinity of haloes as a function of distance, and in each panel the curves show median values at each distance bin. The solid and dotted lines correspond to central and satellite galaxies, respectively.

The top panel of Fig. 5 shows the total stripped fraction of galaxies by  $z = 0$ . We also divide the stripped fraction into two categories based on whether they were central or satellite galaxies when stripping occurred. In particular, the stripped fractions in the middle and bottom rows are calculated only when a galaxy is either a satellite, or a central, respectively. Therefore, the top panel is the sum over the middle and bottom panels. In the rest of this subsection, we use Fig. 5 as a reference to describe the model predictions for gas stripping in different types of galaxies.

### 5.1.1 Gas stripped out of satellite galaxies

The middle panel of Fig. 5 shows that  $z = 0$  satellite galaxies (dotted lines) lose a fraction of their gas after infall, i.e. when they are no longer categorized as the central galaxy of their own haloes. There is a trend with distance for galaxies near FOF haloes: galaxies closer to the halo centre have lost more gas. That is a direct influence of environmental processes, as galaxies closer to the halo centre reside in denser environments and are more strongly stripped. There is a minimum amount of stripping at  $R/R_{200} \sim 1-3$ . This is where the most distant satellites of the FOF haloes reside, i.e. the characteristic scale at which galaxies change type from central to satellite according to the halo finder algorithm. Among all satellites, the most distant are the least influenced by environmental effects. Beyond this scale, satellite galaxies belong to other FOF haloes and are mostly closer to the centre of their respective hosts. As a result, they are more

strongly stripped. The resulting stripped fraction increases with distance beyond this characteristic scale, until reaching a constant, mass-dependent, global value.

Galaxies that are centrals at  $z = 0$  could have been satellites at some point in their history. These galaxies are mainly fly-by or splashback systems, i.e. former satellites of other central haloes. Fly-by galaxies enter a halo, traverse it, and then leave without becoming bound to the halo, while splashback galaxies remain bound to that halo and can reach apocentre before possibly returning. The solid lines in the middle row of Fig. 5 represent the gas stripped out of these two kinds of galaxies. They are the central galaxy of their own haloes at  $z = 0$ , and the stripped gas shown with solid lines is calculated when they were satellites of other haloes at  $z > 0$ . There is a characteristic scale at  $R/R_{200} \sim 1.5-2$  where these galaxies show any signs of having lost gas, which corresponds to the splashback radius of the nearby central halo at  $z = 0$ .

### 5.1.2 Gas stripped out of central galaxies

While gas stripping out of satellite galaxies is expected, central galaxies are less often imagined to experience significant stripping. As illustrated in the bottom panel of Fig. 5, our model shows that galaxies can in fact lose a non-negligible fraction of their hot gas due to stripping while they are centrals. The stripped fraction decreases with halocentric distance for galaxies that are centrals at  $z = 0$  (solid lines), and can reach 80 per cent in the median for low-mass central galaxies in the vicinity of clusters (bottom right panel, solid lines). Interestingly, more than half of the current massive central galaxies near clusters (bottom right panel, solid purple lines) have lost more than half of their gas due to stripping beyond  $\sim 1.5R_{200}$ .

Our model predicts that  $z = 0$  satellite galaxies (dotted lines in Fig. 5) lose a large fraction of their hot gas prior to infall, when they are the central galaxy of their own halo. This is illustrated in the bottom panel of Fig. 5 (dotted lines). There is a local maximum at  $R/R_{200} \sim 1-2$ , corresponding to where these galaxies change type from central to satellite. At this scale, the overall stripped fraction is very similar for central versus satellite galaxies, reflecting the uniformity of our gas stripping approach in treating different types of galaxies. For satellite galaxies with  $R/R_{200} \lesssim 1-2$ , the fraction of the gas they have lost as a central decreases towards the centre of the halo: galaxies close to the halo centre have earlier infall times, and have spent a correspondingly longer time as satellites than more distant galaxies.

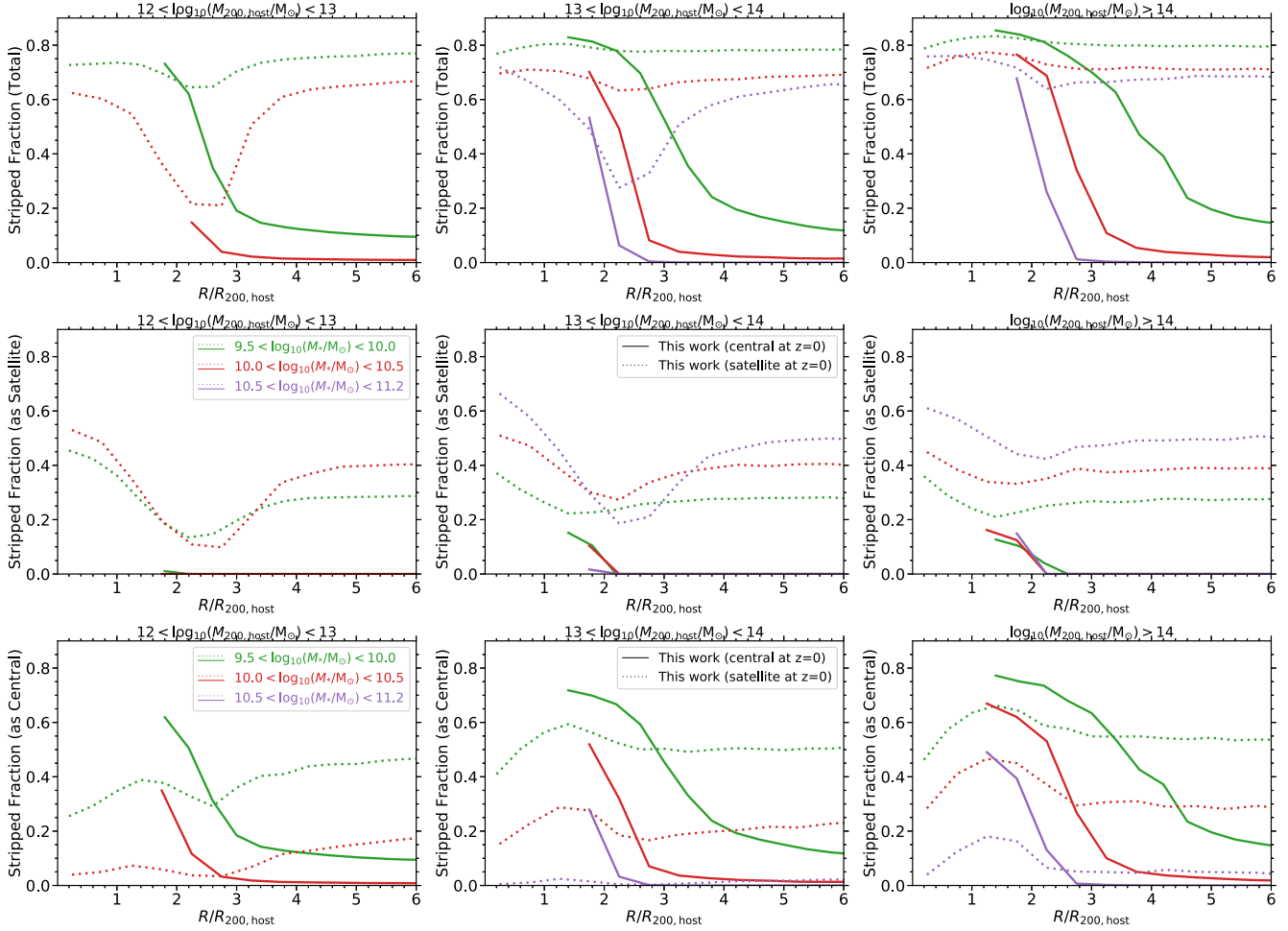
There is an overall trend with stellar mass. The top panel of Fig. 5 shows that the total stripped fraction decreases with stellar mass (green, to red, to purple). The reason is the weaker gravitational binding energy of low-mass systems. The trend is reversed for the stripped fraction *as satellites* (middle panels), because low-mass galaxies lose a larger percentage of their gas prior to infall when they were still the central galaxy of their own halo (bottom panel). Therefore, there is not much residual gas available to be stripped once they later become satellites.

Comparing the three panels in the bottom row of Fig. 5, there is a trend with halo mass. Central galaxies in the vicinity of more massive haloes are more stripped up to halocentric distances as large as  $4-5 R_{200}$ .

## 5.2 Gas content of galaxies and subhaloes

Gas content is the quantity most affected by our new stripping method. In this subsection, we analyse the predictions of our model





**Figure 5.** The median cumulative stripped fraction of gas (equation 9) in galaxies as a function of halocentric distance at  $z = 0$ . Different columns show results for different central host halo masses. In all the panels, different colours correspond to different stellar mass ranges (indicated in the legends) and different line styles correspond to central galaxies (solid) versus satellites (dotted). The top row shows the total stripped fraction, which is divided into two categories as shown in the middle and bottom rows: the cumulative gas stripped from galaxies when they were satellites (middle row) versus when they were centrals (bottom row). Galaxies experience significant stripping, both when they are satellites, and when they are centrals (see text for details).

for hot (non-star-forming) and cold (star-forming) gas, comparing to observations and the **H20** model.

### 5.2.1 Hot (non-star-forming) gas

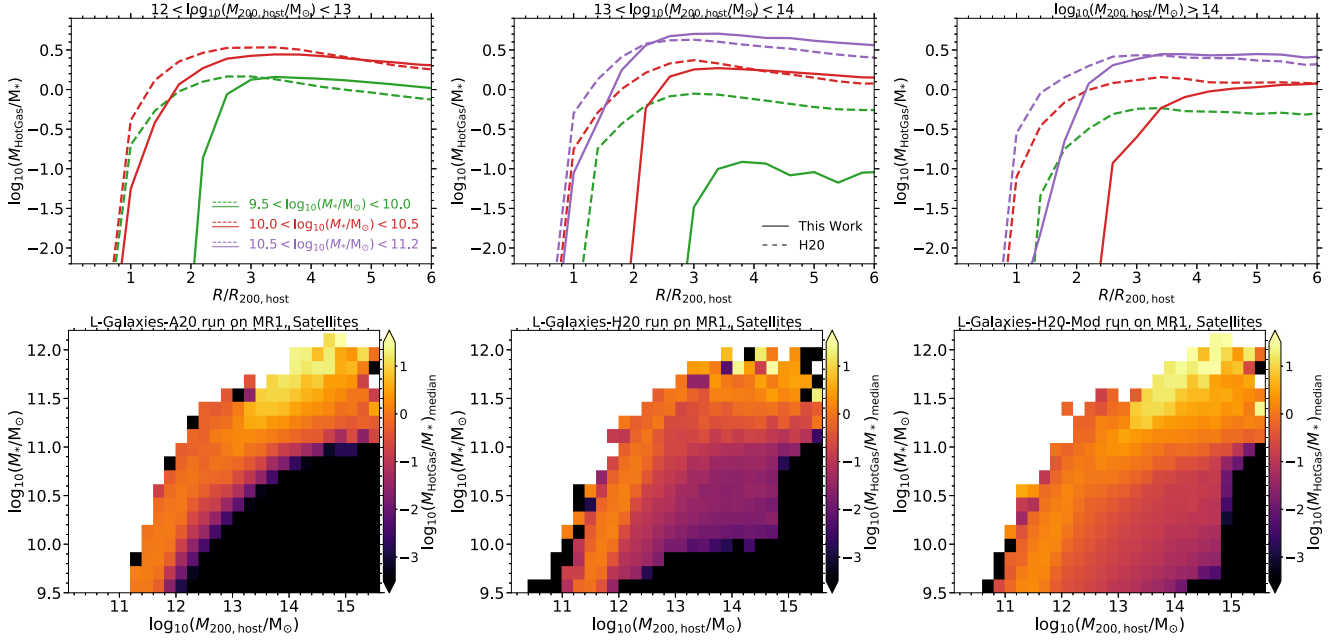
In the top panel of Fig. 6 we show the hot gas to stellar mass ratio,  $M_{\text{hotgas}}/M_{\star}$ , as a function of halocentric distance for galaxies in the vicinity of clusters (top right), groups (top middle), and lower mass haloes (top left). For both models (different line styles) and at all stellar mass ranges (different colours), the  $M_{\text{hotgas}}/M_{\star}$  ratio decreases with decreasing halocentric distance from a constant global value. The radius at which the ratio starts to decrease in our model is larger than in **H20**, showing the impact of more spatially extended gas stripping in the new model.

The hot gas to stellar mass ratio decreases with the halo mass, reflecting the fact that gas stripping processes are more efficient in the dense regions surrounding the most massive haloes. Interestingly, the difference between our model and **H20** is more significant near massive haloes (right-hand panel), showing the importance of modelling ram-pressure self-consistently in all environments.

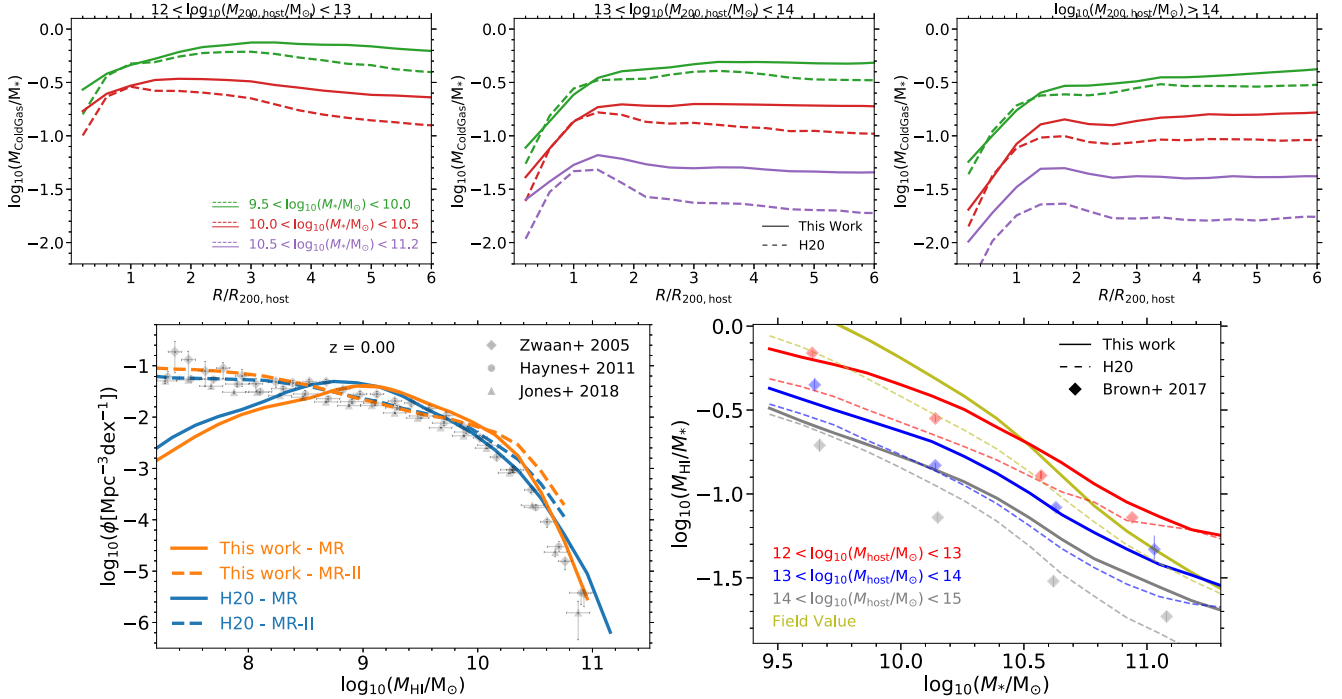
Moreover, the hot gas to stellar mass ratio increases with galaxy stellar mass. This is also a direct impact of stripping as discussed in Section 5.1 (see Fig. 5).

The dependence of this ratio on the stellar mass of satellites, as well as their host halo mass, is illustrated in the bottom panels of Fig. 6. The 2D histogram is coloured by the median value of  $\log_{10}(M_{\text{hotgas}}/M_{\star})$ . For reference, here we include a modified **H20** model (**H20-Mod** hereafter) that includes all **H20** physics but with free parameters taken from our model (Table 3) too. In this way, we check how much our new gas stripping method changes the model predictions, regardless of model calibration. In both **H20** and **H20-Mod** (middle and right-hand panels), there is an abrupt transition at  $\log_{10}(M_{200}/M_{\odot}) \sim 14.7$ . This results from **H20** only applying ram-pressure stripping to satellites within  $R_{200}$  of haloes with  $\log_{10}(M_{200}/M_{\odot}) \gtrsim 14.7$  (see Section 2.1.2). On the other hand, such a sharp transition is not present in our model, and low-mass satellite galaxies are more gas-poor than in **H20** and **H20-Mod**, due to both pre-infall and post-infall stripping.

The differences between our model (bottom left panel) and **H20-Mod** (bottom right panel) show the direct impact of our new gas



**Figure 6.** Top panel: Hot gas to stellar mass ratio as a function of halocentric distance at  $z = 0$ , as a function of host mass (three columns), stellar mass (line colours), and model (line styles). Bottom panel: 2D histograms of the hot gas to stellar mass ratio of satellite galaxies as a function of their host halo mass (x-axis) and stellar mass (y-axis), contrasting this work (left-hand panel), **H20** (middle panel), and **H20-Mod** (right-hand panel).



**Figure 7.** Top panels: Cold gas to stellar mass ratio as a function of halocentric distance at  $z = 0$ . We show the dependence on host halo mass (three columns), stellar mass (colours), and contrast our updated model with **H20** (linestyles). Bottom left panel: H I mass function from this work in comparison with **H20** and with observations from Zwaan et al. (2005), Haynes et al. (2011), and Jones et al. (2018). Bottom right panel: Average H I to stellar mass ratio of satellite galaxies versus stellar mass in our model, **H20**, and observations from Brown et al. (2017).

stripping method. On the other hand, the differences between **H20** and **H20-Mod** show the influence of using different sets of free parameters. For instance, for massive galaxies, the hot gas to stellar

mass ratio is higher in **H20-Mod** (bottom left) and our model (bottom right) than in **H20** (bottom middle). This is due to a higher AGN feedback efficiency in our model parameters (see Table 3), which

prevents hot gas from cooling and results in a higher fraction of hot gas. We note that the H20-Mod model is not calibrated and cannot be compared with observational data.

### 5.2.2 Cold star-forming gas in galaxies

The cold gas component is also affected by environmental processes. However, as we do not implement cold gas stripping directly, we do not expect our model to differ significantly from H20 in this regard. The top panel of Fig. 7 shows the median cold gas to stellar mass ratio as a function of halocentric distance. As before, galaxies in the vicinity of haloes are stacked and the lines represent the median values at each radial distance bin.

In both models,  $M_{\text{coldgas}}/M_*$  increases with distance and reaches a constant global value at  $R/R_{200} \sim 1-2$ . The trend with distance roughly follows the trend of  $\log_{10}(M_{\text{hotgas}}/M_*)$  (Fig. 6, top panel). However, the influence of environment on the cold gas extends out to smaller radii than is seen for the hot gas. Only when galaxies run out of hot gas does cold gas component cease to grow. Subsequently, as star formation continues, the cold gas mass decreases further until either it is exhausted, or the galaxy merges.

As illustrated in Fig. 7, the field value of the cold gas to stellar mass ratio is higher in our model than in H20, in contrast to what is expected from our extended gas stripping method. This discrepancy is due to the recalibration of the model. To address this, we have compared our results with H20-Mod (a modified H20 model that includes all H20 physics but with free parameters taken from our model) in our additional analysis. As expected, galaxies in our model have a significantly lower  $M_{\text{coldgas}}/M_*$  than H20-Mod.

We note that the cold gas illustrated in the top panels of Fig. 7 is the total cold gas mass, i.e. the sum of H I and H<sub>2</sub> gas masses. Our additional analyses on these two gas types show that the average value of the H<sub>2</sub> to stellar mass ratio of galaxies in the vicinity of massive haloes in our model is lower than in H20. Due to our H<sub>2</sub>-based star formation prescription (see Section 2.1.2 in this work and also section S9 in the supplementary material), this results in lower star formation and higher quenched fraction in our model, which we show in Section 5.3.

The bottom left panel of Fig. 7 shows the H I mass function for galaxies in our model (orange lines), H20 (blue lines), and observations (grey points, Zwaan et al. 2005; Haynes et al. 2011; Jones et al. 2018). Two different runs are shown for each model: L-GALAXIES run on the Millennium and Millennium-II simulations. Due to the better resolution of the Millennium-II simulation we can extend the model results down to lower H I masses. On the other hand, the Millennium simulation provides better statistics for larger  $M_{\text{H I}}$  values. Both our model and H20 are in relatively good agreement with observations. However, we note that H20 uses the H I mass function as an observational constraint for their model calibration, while we do not.

Finally, we study the average H I gas to stellar mass ratio of satellite galaxies as a function of stellar mass in our model, H20, and observations from the ALFALFA Survey (Brown et al. 2017) in the bottom right panel of Fig. 7. The general observed trend with stellar mass is well produced in both our model and H20. In addition, similar to observations,  $M_{\text{H I}}/M_*$  decreases with halo mass (different colours) in the models as well. Our model's predictions are in good agreement with observations for satellites within groups and lower mass haloes. For cluster satellites, however, the model overpredicts the H I to stellar mass ratio. Adding cold gas stripping to the model

may result in lower  $M_{\text{H I}}/M_*$  in cluster satellites and therefore better model predictions.

## 5.3 Star formation and galaxy quenching

### 5.3.1 Direct model predictions

We move on to study the star formation activity, and quenching, of galaxies. The top panel of Fig. 8 shows the fraction of quenched galaxies,  $f_q$ , as a function of halocentric distance at  $z = 0$ . For each halo mass bin (panel) all galaxies in the vicinity of haloes are stacked. The results are shown for different stellar mass bins (different colours) for our model (solid lines) and for H20 (dashed lines). In the vicinity of clusters (top right),  $f_q$  decreases with distance in both models and flattens at a distance that is always larger than the halo boundary,  $R_{200}$ . There is a clear trend with halo mass: galaxies in the vicinity of more massive haloes are more quenched. This follows directly from what we showed in Section 5.2 (Figs 6 and 7) regarding the gas content of galaxies.

The top panel of Fig. 8 often reveals a minimum in the quenched fraction at  $R/R_{200} \sim 1-2$ . This is where the most distant satellites of FOF haloes reside, i.e. where they are the least influenced by their environment. At halocentric radii smaller than this scale, satellite galaxies are strongly influenced by their host halo. Beyond this scale, satellites belong to other FOF haloes and could be close to their own centrals and therefore more strongly influenced by environmental processes. The contribution of those satellites at  $R/R_{200} > 2$  in the total quenched fraction (i.e. centrals and satellites together) is sufficient to cause a quenched fraction larger than the value at  $R/R_{200} \sim 1-2$ , causing the minimum. We note that this directly follows the minimum of the cumulative stripped fraction as shown by the dotted lines in the top panel of Fig. 5.

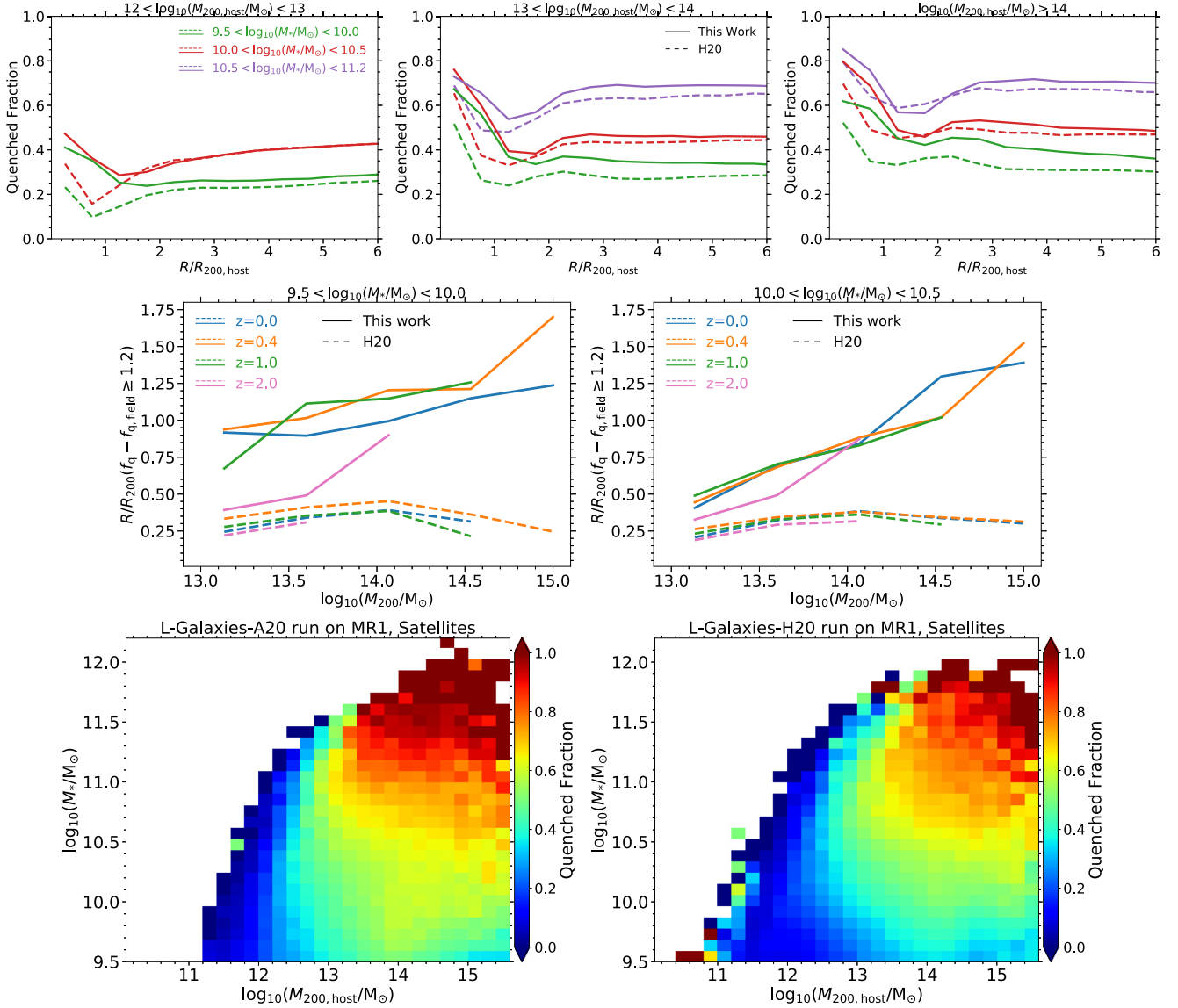
Comparing with H20, gas stripping in our model affects  $f_q$  to larger halocentric distances. In order to quantify this, we define a characteristic radius called  $R_{c,20 \text{ per cent}}$  at which the quenched fraction is 20 per cent above its field value,

$$f_q(R_{c,20 \text{ per cent}}) - f_{q,\text{field}} = 0.2, \quad (15)$$

where  $f_{q,\text{field}}$  is the fraction of quenched field galaxies, calculated for galaxies with halocentric distance in the range of  $5 < R/R_{200} < 10$ . The middle row of Fig. 8 shows  $R_{c,20 \text{ per cent}}/R_{200}$  (normalized to the host halo  $R_{200}$ ) for our model (solid lines) and H20 (dashed lines) as a function of halo mass. The results are shown for two different stellar mass ranges (middle left and middle right panels) at different redshifts (different colours).

Our model quenches galaxies in the vicinity of massive haloes to much larger halocentric distances than H20 at all redshifts. At  $z > 2$ , there is no distance within which galaxies become 20 per cent more quenched than field galaxies. This indicates that the influence of environmental processes on galaxy quenching is non-negligibly visible only since  $z \sim 2$ . At  $z \lesssim 1$ ,  $R_{c,20 \text{ per cent}}/R_{200}$  in our model is on average four times larger than in H20. We also observe a trend with halo mass: galaxies in the vicinity of more massive haloes are quenched up to larger fractions of their  $R_{200}$ . The same trend is not present in H20 or is very weak. Moreover, there is no significant correlation with redshift, except for  $z \gtrsim 2$  where  $R_{c,20 \text{ per cent}}/R_{200}$  is either small (pink lines, middle row of Fig. 8) or is not present. Comparing the two panels in the middle row of Fig. 8,  $R_{c,20 \text{ per cent}}/R_{200}$  decreases with stellar mass, i.e. low-mass galaxies are strongly influenced by their environment out to larger distances.

Environment can also influence intrinsic physical processes such as stellar and black hole feedback. As an example, consider galaxies



**Figure 8.** Top panel: Fraction of quenched galaxies as a function of halocentric distance at  $z = 0$ . Middle panel: Halocentric radius at which the quenched fraction is 20 percent above the field value versus the halo mass both for this work and for **H20** at different redshifts. Bottom panel: 2D histograms of the quenched fraction of satellite galaxies as a function of their host halo masses (x-axis) and their stellar masses (y-axis) in this work (left-hand panels) and in **H20** (right-hand panels) at  $z = 0$ . At  $z = 0$ , galaxies with  $\log_{10}(\text{SSFR}/\text{yr}^{-1}) < -11$  are considered as quenched and at  $z > 0$  galaxies with  $\text{SSFR}/\text{yr}^{-1} < (1 + z)/(2 \times 10^{10})$  are considered as quenched (following the quenching criteria in **H20**).

with  $10.5 < \log_{10}(M_{\text{star}}/M_{\odot}) < 11.2$  that reside in the vicinity of clusters in our model (solid purple line, top right panel). When the gas is removed, there are two main consequences: (i) their star formation will decrease due to hot gas stripping and a lack of cold gas replenishment, and (ii) the efficiency of their black hole feedback will decrease because it is a function of hot halo gas (see Table 3 and the supplementary material). The decrease in black hole feedback allows the hot gas to cool faster, which increases star formation. Such non-trivial couplings between environmental effects and feedback processes will be explored in more detail in future work.

Finally, the bottom panel of Fig. 8 shows 2D histograms of the quenched fraction of satellite galaxies as a function of stellar mass and host halo mass. In both our model (bottom left) and **H20** (bottom right), the quenched fraction increases with stellar mass. This shows the influence of intrinsic physical processes such as stellar and black hole feedback. In addition, there is a strong trend with halo mass,

showing the influence of environment. In general, satellites of more massive hosts are more quenched. Comparing the two models, our model has more quenched satellites than **H20** at all stellar and host halo masses. Significantly, satellites with  $\log_{10}(M_{\text{star}}/M_{\odot}) < 10$  residing in groups with  $13 < \log_{10}(M_{200}/M_{\odot}) < 14$  are more quenched in our model than in **H20**. This is due to gas stripping both within groups and prior to infall.

### 5.3.2 Comparing with observations at $z = 0$

We compare our results for the fraction of quenched galaxies in different environments with SDSS observations. For both the models and observations, the halo mass ( $M_{200}$ ) and radius ( $R_{200}$ ) are estimated from their stellar masses as described in Section 4.2.2. The simulations are transformed into redshift/velocity space based



on the method explained in Section 4.2.1. For every dark matter halo, we project the outputs of both our model and H20 along the z-axis of the simulation volume in velocity space with the thickness of the projected slice taken as  $|v_{\text{gal,LOS}} - v_{\text{halo,LOS}}| \leq \pm 2V_{200,\text{halo}}$ . For observations, the velocity separations are calculated along the line of sight using the galaxy redshifts, and the 2D projected distances are calculated from the sky coordinates.

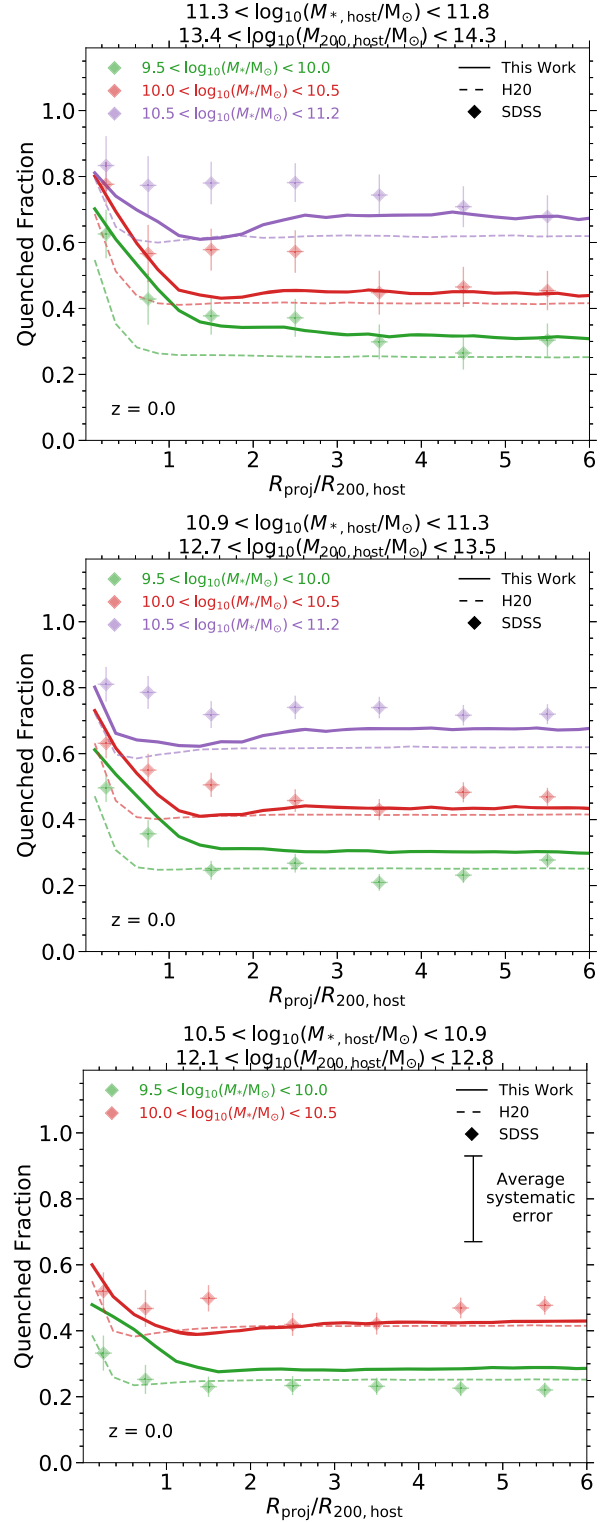
Fig. 9 shows the fraction of quenched galaxies as a function of projected halocentric distance in our model (solid lines), in H20 (dashed lines), and in the SDSS observations (points with errorbars). Each panel corresponds to a host stellar mass (or equivalently halo mass) bin indicated at the top of each panel. The results are shown for three different stellar mass ranges, specified by different colours. We first note that both models reproduce the field quenched fractions as a function of stellar mass. In all the halo mass ranges, a clear trend with distance is present both in the models and in the observations: the quenched fraction decreases with halocentric distance and reaches a constant, field value at some radius usually larger than the halo boundary,  $R_{200}$ .

For almost all stellar and host mass regimes our model shows better agreement with observations than H20. In the vicinity of massive haloes (top panel), the environmental dependence of quenched fraction extends to larger radius in our model and in SDSS compared to H20. Looking specifically at low-mass galaxies in the vicinity of clusters (top panel, green lines and points), our model and SDSS show that up to 60–70 percent of galaxies near the halo centre are quenched, while this value is 10–20 percent lower in H20. While our model is in better agreement for intermediate-mass (red lines and points) and massive (purple lines and points) galaxies, it underpredicts the fraction of quenched massive and intermediate-mass galaxies at  $1 < R/R_{200} < 3$ .

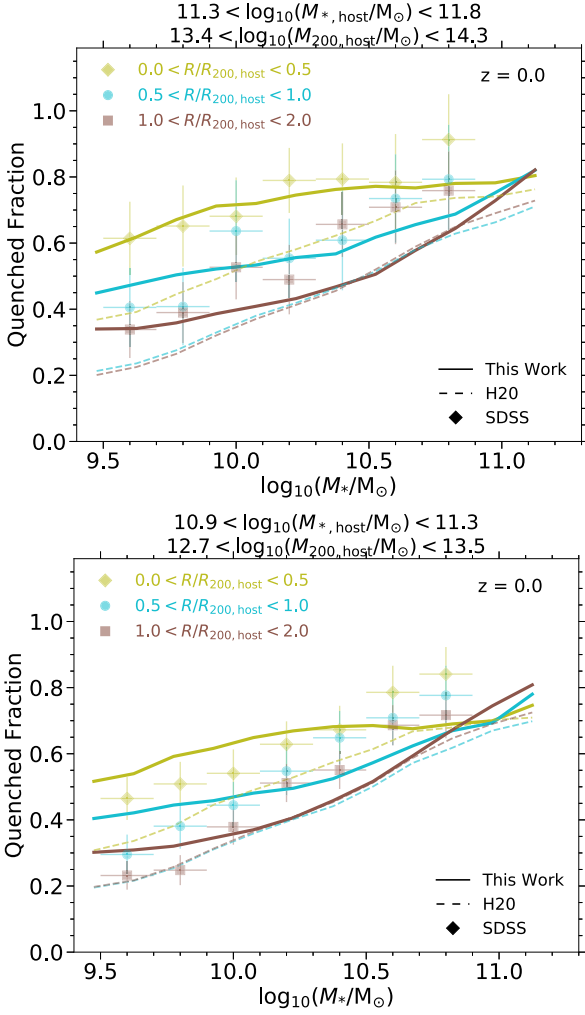
We suggest two reasons for this discrepancy which can motivate future model developments. The first is a lack of cold gas stripping in our model, which can be accommodated within our LBE framework, while the second is a lack of ejective AGN feedback. It is possible that strong ejective feedback can push gas outside the halo boundary. Infalling galaxies passing through this ejected gas experience substantial enhancement in ram-pressure and can therefore lose a larger fraction of their halo gas, as we see in the IllustrisTNG simulations (Ayromlou et al. in preparation). Furthermore, in massive galaxies with  $\log_{10}(M_*/M_\odot) > 10.5$ , AGN feedback makes extended subhalo gas less bound which results in an enhancement of ram-pressure stripping and thus galaxy quenching (Ayromlou et al. 2021).

Considering galaxies surrounding groups (middle panel) and lower mass haloes (bottom panel; Fig. 9), our model is again in reasonably good agreement with observations, while H20 exhibits weaker effects within  $R_{200}$  and the dependence of the quenched fraction of distance flattens at smaller distances comparing to our model and SDSS. The persistence of the variation in quenched fraction out to large halocentric distances is even more marked in the SDSS data, a point we will come back to later.

In Fig. 10, we show the fraction of quenched galaxies, now as a function of stellar mass, for three different halocentric distance bins (different colours). Different panels correspond to different host stellar (or halo) masses. In general, the quenched fraction almost always increases monotonically with stellar mass for both the models and the observations. In the vicinity of massive haloes (top panel), our model is in relatively good agreement with observations for all the distance bins, while H20 is off by up to 20 per cent. The difference between our model and H20 decreases with stellar mass and is the largest for low-mass galaxies.



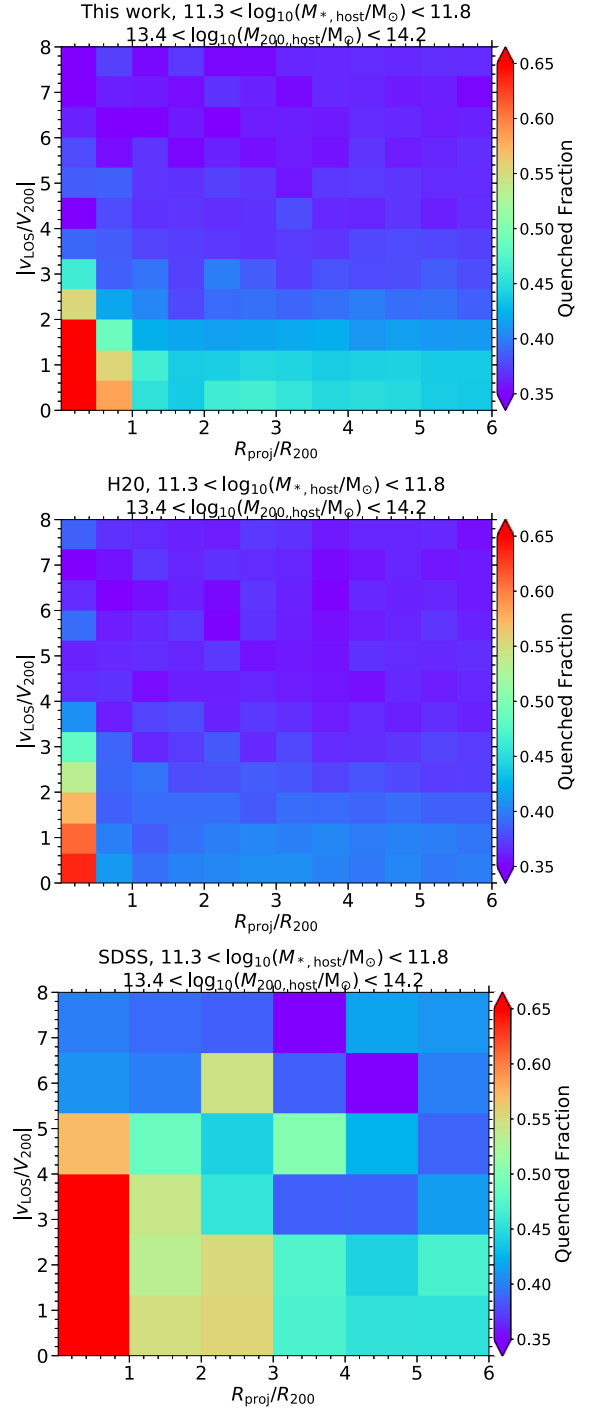
**Figure 9.** Fraction of quenched galaxies as a function of projected distance to the centre of haloes at  $z = 0$ . For both models and observations, galaxies with  $\log_{10}(\text{sSFR}/\text{yr}^{-1}) < -11$  are considered as quenched. The halo mass in both models and observations is calculated using the stellar masses of the central galaxies, based on equation (11). The error bars are binomial 95 percent confidence intervals based on a Gaussian approximation. The average systematic observational errors of the SDSS data are shown in the figure for  $1\sigma$  error of sSFR. The overall trends seen in SDSS are well reproduced by our new model.



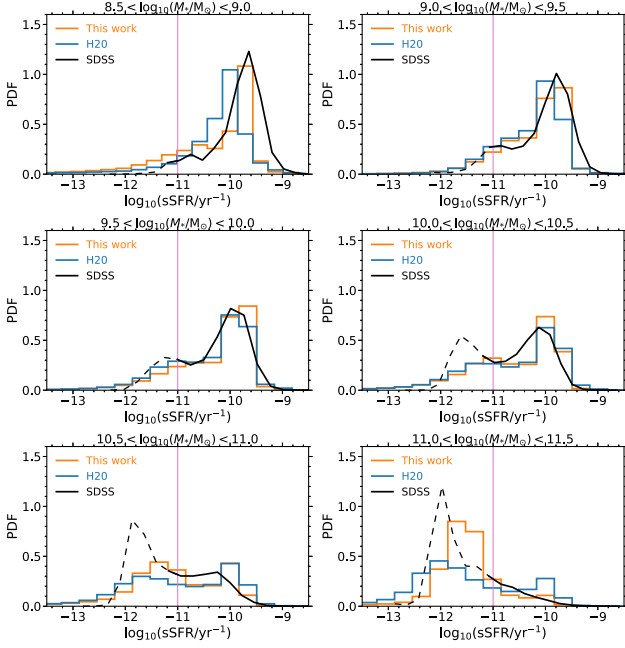
**Figure 10.** Fraction of quenched galaxies as a function of galaxy stellar mass at  $z = 0$ , comparing the two models to our analysis of SDSS data. For both models and observations, galaxies with  $\log_{10}\text{SSFR/yr} < -11$  are considered as quenched. The halo mass in both models and observations are calculated using the stellar masses of the central galaxies, based on equation (11). The error bars are binomial 95 per cent confidence intervals based on a Gaussian approximation.

Similar results are found for galaxies in the vicinity of intermediate-mass haloes (bottom panel of Fig. 10), although the difference between our model and H20 is smaller. Overall, our model predictions are in better agreement with observations than H20, and low-mass galaxies in cluster environments are the most influenced by our new gas stripping method.

Fig. 11 shows the fraction of quenched galaxies as a function of projected halocentric distance (x-axis) and line-of-sight velocity (y-axis), i.e. projected phase space (PPS; see also Oman et al. 2021). We contrast our model (top panel), H20 (middle panel), and SDSS (bottom panel). In all three cases the quenched fraction decreases both with halocentric distance and with the magnitude of the line-of-sight velocity. As  $R_{\text{proj}}/R_{200}$  and  $|v_{\text{LOS}}/v_{200}|$  increase, more field galaxies are included in each bin and the quenched fraction eventually approaches the field value. The trend with distance is stronger in our model than in H20, but is still weaker than in SDSS. In other words, real haloes seem to influence nearby galaxies out to somewhat larger distances than in either model.



**Figure 11.** The fraction of quenched galaxies in PPS, as a function of projected halocentric distance (x-axis) and line-of-sight velocity (y-axis). We compare our new model (top panel), H20 (middle panel), and SDSS (bottom panel), defining galaxies with  $\log_{10}(\text{SSFR/yr}^{-1}) < -11$  as quenched. All galaxies with  $\log_{10}(M_*/M_\odot) > 9.5$  are included. Note that the statistics are quite different: the theory panels contain  $\sim 18000$  haloes while only  $\sim 100$  haloes are available in the data. The new model better reproduces the extended quenched fractions of galaxies out to larger distances and higher line-of-sight velocities.

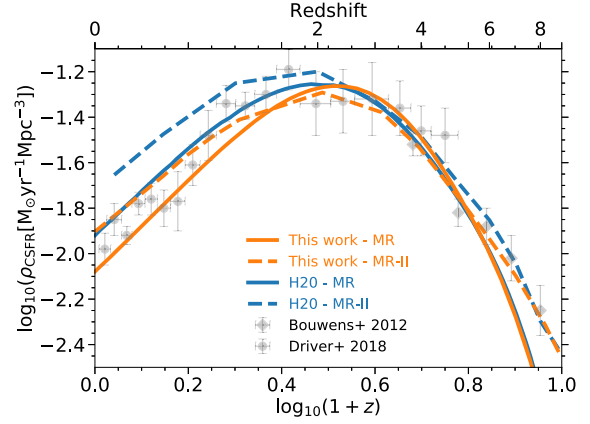


**Figure 12.** Distributions of sSFR for galaxies in this work, **H20**, and observations. Each panel corresponds to a particular stellar mass range. The SDSS data are based on Brinchmann et al. (2004) with the corrections of Salim et al. (2007). Galaxies with  $\log_{10}(\text{sSFR}/\text{yr}^{-1}) < -11$  are considered as quenched. Observational data are shown with dashed lines in the regions where there is high uncertainty in the observed specific star formation rates. Despite extensive modification to the stripping properties of galaxies, the sSFR distributions are largely unchanged.

As a final comparison with SDSS data, Fig. 12 shows the distributions of specific star formation rate in this work, in **H20**, and in SDSS observations at  $z \sim 0$ . Different panels correspond to different stellar mass ranges. SDSS data are shown with black lines, which become dashed where there is high uncertainty in observed star formation rates. The magenta vertical lines at  $\log_{10}(\text{sSFR}/\text{yr}^{-1}) = -11$  demarcate the value below which galaxies are considered as quenched. Although we have changed the properties of galaxies near massive hosts significantly, the overall specific star formation rates are rather similar. Both our model and **H20** are in relatively good agreement with observation. The noticeable difference is for massive galaxies with  $11 < \log_{10}(M_*/M_\odot) < 11.5$ , where our model shows lower specific star formation rates than **H20**. This is caused by the higher black hole feedback efficiency in our new model.

### 5.3.3 Beyond $z = 0$ : the cosmic star formation rate density

In Fig. 13, we compare the cosmic star formation rate density,  $\rho_{\text{CSFR}}$ , as a function of redshift in the two models to observations taken from Bouwens et al. (2012) and Driver et al. (2018). Overall, both models agree quite well with data. The two models are similar at redshifts greater than the peak of cosmic star formation rate density, i.e.  $z \gtrsim 2$ . On the other hand, at lower redshifts,  $\rho_{\text{CSFR}}$  in our model is smaller than in **H20**. The maximum difference between our model and **H20** occurs at  $z = 0$  where  $\rho_{\text{CSFR}}$  is 0.15 dex smaller in our model. This is mainly due to our extended gas stripping implementation but also due to our higher AGN feedback efficiency. The former mostly affects low-mass galaxies, while the latter influences star formation in more massive objects.



**Figure 13.** Cosmic star formation rate density as a function of redshift for our model, for **H20**, and for observations from Bouwens et al. (2012) and Driver et al. (2018).

### 5.3.4 Quenched fractions versus observations at $z > 0$

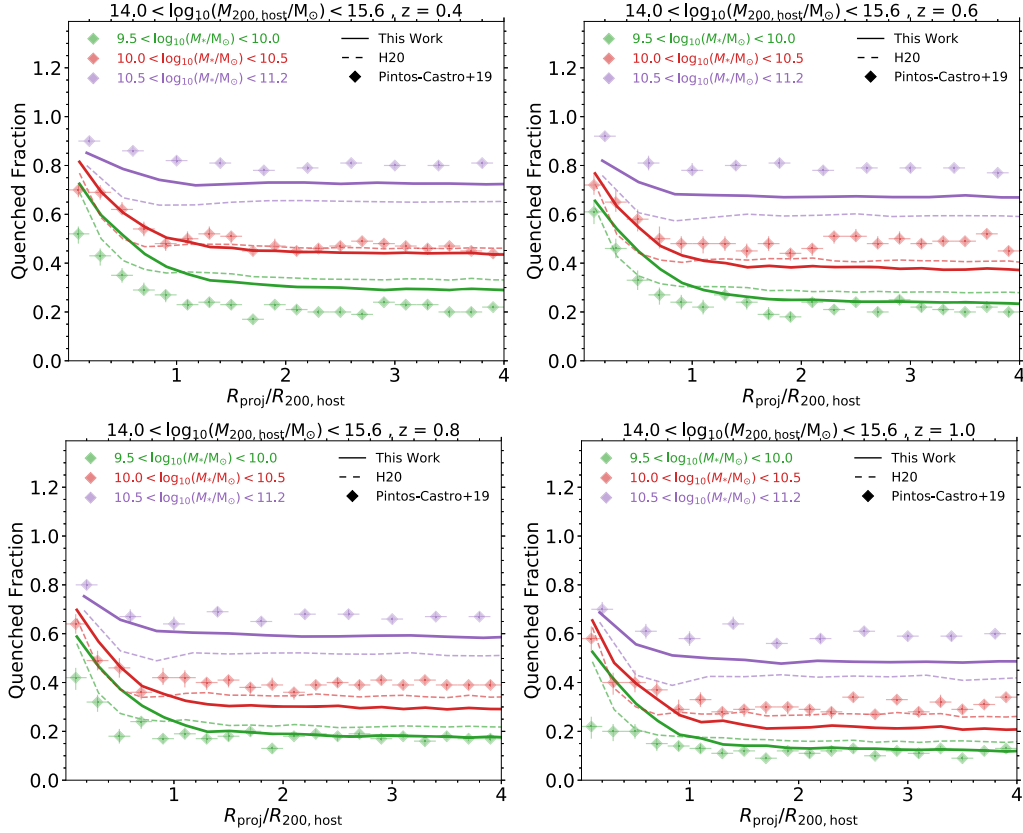
Fig. 14 shows the fraction of quenched galaxies versus projected halocentric distance in our model, in **H20**, and as observed in the HSC survey (Pintos-Castro et al. 2019) at four different redshifts. For this comparison, we take the direct output of our simulation for the halo mass and radius (Section 4.2.2) and project the whole simulation depth along the  $z$ -axis of the simulation volume as a rough approximation of the survey characteristics (Section 4.2).

At all redshifts the quenched fraction decreases with distance both for models and for observations. The agreement with observations is fair for both models, although the quenched fraction of low-mass galaxies in our model (solid green lines) is higher than in **H20** (dashed green lines) and than observed (green points) within the virial radius of clusters. At  $z = 1$  (bottom right panel), both our model and **H20** show a rather strong trend with distance for low-mass galaxies, whereas such a trend is not observed. This could be due to the presence of low-mass quenched galaxies at  $z = 1$  which fall outside the observed samples. The overall trend with stellar mass is similar for our work, for **H20** and for the observational data. At all halocentric distances and redshifts, the quenched fractions increase with stellar mass.

In both models, the quenched fraction near the cluster centre almost always decreases with redshift, i.e. lower redshift galaxies have higher quenched fractions. On the other hand, comparing observations at  $z = 0.4$  and  $z = 0.6$  (top left and top right panels), the quenched fraction is higher at higher redshift, the opposite of the trend predicted by the models, possibly due to different methods or definitions of halo mass. It would be best to compare high-redshift data with spectroscopic redshifts to confirm whether these discrepancies are real or are caused by systematic errors of some kind. Observational constraints beyond the local Universe will undoubtedly pose a challenge to theoretical models, and more sophisticated comparisons and future model improvements will further increase our understanding of the role of environment in galaxy evolution.

## 6 SUMMARY AND DISCUSSION

In this work, we study the impact of environment on the formation and evolution of galaxies. We present a variant of the Munich SAM of galaxy formation, L-GALAXIES, with a novel gas-stripping method.



**Figure 14.** Fraction of quenched galaxies as a function of projected distance to centre of clusters at different redshifts. In the models, galaxies with  $\text{SSFR} < (1 + z)/(2 \times 10^{10})$  are considered as quenched. Here, we compare to observational data at  $0.4 \leq z \leq 1.0$  from Pintos-Castro et al. (2019). Note that in these data a colour cut is used and galaxies with  $U - J > 0.88(J - V) + 0.59$  are considered as quenched. In comparison to H20, our updated model shows an environmental effect on the quenched fraction which extends to larger halocentric distance.

Following Ayromlou et al. (2019), we measure the properties of the local environment of every galaxy and subhalo directly from the particle data of the underlying  $N$ -body simulations. This enables us to devise a more accurate treatment of environmental processes, particularly ram-pressure stripping. We recalibrate the parameters of the new model using an MCMC technique and a set of observational constraints, namely the stellar mass function and quenched fraction at  $z = 0, 1, 2$ . Due to this recalibration, global properties of galaxies such as the stellar mass function, quenched fractions versus galaxy mass, and H I mass function are all largely unchanged from H20 and remain consistent with observations. Analysing the results of our new model and the standard model on which it is based (H20), our main results are as follows:

- (i) Measuring the total amount of stripped gas in our model, we find that galaxies in the vicinity of haloes with  $M_{200}/M_{\odot} > 10^{12}$  lose a large fraction (median  $\sim 80$  per cent in low-mass galaxies with  $9.5 < \log_{10}(M_{*}/M_{\odot}) < 10$ ) of their gas due to ram-pressure stripping while they were in fact central galaxies (Fig. 5).
- (ii) At high redshifts,  $z \gtrsim 1$ , most stripping is due to gas removal of previously ‘ejected’ material. At  $z \lesssim 1$ , in contrast, stripping of the hot halo gas is the dominant contributor (Fig. 4).
- (iii) The ratio of hot gas to stellar mass,  $M_{\text{hotgas}}/M_{*}$ , decreases with decreasing halocentric distance towards the centres of haloes, due to stripping. In our model, more than half of all galaxies in the vicinity of clusters and groups, up to several virial radii, are almost devoid of hot gas. The dependence of  $M_{\text{hotgas}}/M_{*}$  on environment extends to much

larger halocentric distances in our model than in H20, with low-mass galaxies being more strongly affected by their environment (Fig. 6).

- (iv) Near clusters and groups in our model, the quenched fraction decreases with halocentric distance, and it flattens to the field value only at  $R/R_{200} \lesssim 2-3$ , far beyond the halo virial radius (Fig. 8).

- (v) The characteristic halocentric distance at which the fraction of quenched galaxies is 20 per cent larger than its field value ( $R_{c,20 \text{ per cent}}$ ) is, on average, four times larger in our model than in H20. In our model,  $R_{c,20 \text{ per cent}}/R_{200}$  increases monotonically with halo mass, whereas such a trend is not present in H20 (Fig. 8).

We undertake a new analysis of SDSS galaxy data (Abazajian et al. 2009) combined with the Yang et al. (2005, 2007) halo catalogues, inferring the quenched fraction versus halocentric distance out to  $R_{\text{proj}} = 10R_{200}$  with a methodology consistent between simulations and data. Together with observational results from the HSC survey, we compare against our model predictions. Our principal results are:

- (i) The  $z = 0$  observed trend of the quenched fraction of galaxies ( $f_q$ ) versus halocentric distance (SDSS) is well reproduced in our model up to several halo virial radii, a noticeable improvement over H20. Nevertheless, the observed environmental dependency of galaxies extends to slightly larger distances (Fig. 9).
- (ii) In our model, as well as in SDSS observations, the quenched fraction near haloes increases with halo mass (Fig. 9). The strength of environmental quenching in the vicinity of haloes also decreases with increasing galaxy stellar mass, which is consistent with the observations (Fig. 10).



(iii) At higher redshifts,  $0.4 \leq z \leq 1$ , our model is in relatively good agreement with observations from the HSC survey, while quantitative differences remain, particularly within the  $R_{200}$  of clusters (Fig. 14). We note that we have not attempted to model the effect of errors on the photometric redshifts that are used in the data.

(iv) Our predicted H I gas mass function is in good agreement both with the previous model and with  $z = 0$  data. For satellite galaxies, the H I gas mass to stellar mass ratio,  $M_{\text{H I}}/M_*$ , decreases with halo mass in our model, which reasonably agrees with the ALFALFA observations at  $z = 0$ , although our model overpredicts  $M_{\text{H I}}/M_*$  in cluster satellites (Fig. 7).

The remaining tensions with data motivate two possible future model improvements: (i) stripping of the cold, star-forming gas discs in galaxies, and (ii) handling the ejection and redistribution of gas in the (sub)halo due to baryonic feedback processes. Our method for incorporating the LBE of galaxies can be naturally extended to handle cold gas stripping at smaller scales. At the same time, we have shown that the impact of environment on galaxy properties extends to much larger scales than the often assumed halo virial radius (also see Ayromlou et al. 2019), and that related effects are also present in cosmological hydrodynamical simulations (Ayromlou et al. 2021). Model improvements, incorporating insights from the use of our local environmental measurements in hydrodynamical simulations including IllustrisTNG and EAGLE (Ayromlou et al. in preparation) will be ideally suited to reveal the links between the physics of galaxy evolution and large-scale correlations (e.g. two-point correlation function of galaxies of different colour). Complex questions such as the physics behind the galactic conformity, and the observed large-scale correlation between the star formation of neighbouring galaxies (Weinmann et al. 2006; Kauffmann et al. 2013) will be interesting avenues of further investigation using our new model.

## ACKNOWLEDGEMENTS

We would like to thank the anonymous referee for his/her useful comments that helped us improve the paper. MA is grateful to Bruno Henriques for very helpful discussions about the MCMC model calibration. MA also thanks Irene Pintos-Castro and Toby Brown for kindly sharing their observational data and for fruitful discussions. MA finally thanks Abhijeet Anand, Alireza Vafaei Sadr, Hasti Khoramizhad, Wolfgang Enzi, Richard D'Souza, and John Helly for useful discussions and assistance. The analysis herein was carried out on the compute cluster of the Max Planck Institute for Astrophysics (MPA) at the Max Planck Computing and Data Facility (MPCDF).

## DATA AVAILABILITY

We have made the full output of our updated L-GALAXIES model for all the Millennium and Millennium-II snapshots publicly available at [lgalaxiespublicrelease.github.io](https://lgalaxiespublicrelease.github.io), where the H20 model is already publicly available.

## REFERENCES

Abazajian K. N. et al., 2009, *ApJS*, 182, 543  
 Aihara H. et al., 2018, *PASJ*, 70, S4  
 Angulo R. E., Hilbert S., 2015, *MNRAS*, 448, 364  
 Angulo R. E., White S. D., 2010, *MNRAS*, 405, 143  
 Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, *MNRAS*, 426, 2046

Ayromlou M., Nelson D., Yates R. M., Kauffmann G., White S. D. M., 2019, *MNRAS*, 487, 4313  
 Ayromlou M., Nelson D., Yates R. M., Kauffmann G., Renneby M., White S. D. M., 2021, *MNRAS*, 502, 1051  
 Bahé Y. M., McCarthy I. G., Crain R. A., Theuns T., 2012, *MNRAS*, 424, 1179  
 Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, *ApJ*, 600, 681  
 Baldry I., Glazebrook K., Driver S., 2008, *MNRAS*, 388, 945  
 Baldry I. K. et al., 2012, *MNRAS*, 421, 621  
 Balogh M. L., Morris S. L., Yee H., Carlberg R., Ellingson E., 1999, *ApJ*, 527, 54  
 Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, *ApJS*, 149, 289  
 Bertone S., De Lucia G., Thomas P. A., 2007, *MNRAS*, 379, 1143  
 Binney J., Tremaine S., 1987, *Galactic Dynamics*. Princeton Univ. Press, Princeton  
 Boselli A., Gavazzi G., 2006, *PASP*, 118, 517  
 Boselli A. et al., 2016, *A&A*, 596, A11  
 Bouwens R. J. et al., 2012, *ApJ*, 754, 83  
 Boylan-Kolchin M., Springel V., White S. D., Jenkins A., Lemson G., 2009, *MNRAS*, 398, 1150  
 Bravo M., Lagos C. d. P., Robotham A. S. G., Bellstedt S., Obreschkow D., 2020, *MNRAS*, 497, 3026  
 Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151  
 Brown T. et al., 2017, *MNRAS*, 466, 1275  
 Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168  
 Cora S. A. et al., 2018, *MNRAS*, 479, 2  
 Crain R. A. et al., 2015, *MNRAS*, 450, 1937  
 Croton D. J. et al., 2006, *MNRAS*, 365, 11  
 Croton D. J. et al., 2016, *ApJS*, 222, 22  
 Davies L. J. M. et al., 2019, *MNRAS*, 483, 5444  
 Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371  
 De Lucia G., Kauffmann G., Springel V., White S. D. M., Lanzoni B., Stoehr F., Tormen G., Yoshida N., 2004, *MNRAS*, 348, 333  
 De Lucia G., Springel V., White S. D., Croton D., Kauffmann G., 2006, *MNRAS*, 366, 499  
 Domínguez Sánchez H. et al., 2011, *MNRAS*, 417, 900  
 Donnari M., Pillepich A., Nelson D., Marinacci F., Vogelsberger M., Hernquist L., 2020, preprint ([arXiv:2008.00004](https://arxiv.org/abs/2008.00004))  
 Donnari M. et al., 2021, *MNRAS*, 500, 4004  
 Dressler A., 1980, *ApJ*, 236, 351  
 Driver S. P. et al., 2018, *MNRAS*, 475, 2891  
 Dubois Y. et al., 2020, preprint ([arXiv:2009.10578](https://arxiv.org/abs/2009.10578))  
 Font A. S. et al., 2008, *MNRAS*, 389, 1619  
 Gladders M. D., Yee H. K. C., 2000, *AJ*, 120, 2148  
 Gladders M. D., Yee H. K. C., 2005, *ApJS*, 157, 1  
 Grand R. J. J. et al., 2017, *MNRAS*, 467, 179  
 Gunn J. E., Gott J., Richard I., 1972, *ApJ*, 176, 1  
 Guo Q. et al., 2011, *MNRAS*, 413, 101  
 Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, *MNRAS*, 428, 1351  
 Hansen S. M., Sheldon E. S., Wechsler R. H., Koester B. P., 2009, *ApJ*, 699, 1333  
 Haynes M. P. et al., 2011, *AJ*, 142, 170  
 Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, *MNRAS*, 396, 535  
 Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, *MNRAS*, 431, 3373  
 Henriques B. M., White S. D., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, 451, 2663 (H15)  
 Henriques B. M. B., Yates R. M., Fu J., Guo Q., Kauffmann G., Srisawat C., Thomas P. A., White S. D. M., 2020, *MNRAS*, 491, 5795 (H20)  
 Hernquist L., Katz N., 1989, *ApJS*, 70, 419  
 Hopkins P. F. et al., 2018, *MNRAS*, 480, 800  
 Hubble E., Humason M. L., 1931, *ApJ*, 74, 43  
 Ilbert O. et al., 2010, *ApJ*, 709, 644  
 Ilbert O. et al., 2013, *A&A*, 556, A55

Jones M. G., Haynes M. P., Giovanelli R., Moorman C., 2018, *MNRAS*, 477, 2

Joshi G. D., Pillepich A., Nelson D., Marinacci F., Springel V., Rodriguez-Gomez V., Vogelsberger M., Hernquist L., 2020, *MNRAS*, 496, 2673

Kauffmann G., White S. D., Guiderdoni B., 1993, *MNRAS*, 264, 201

Kauffmann G., Colberg J. M., Diaferio A., White S. D., 1999, *MNRAS*, 303, 188

Kauffmann G. et al., 2003, *MNRAS*, 341, 33

Kauffmann G., White S. D., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, *MNRAS*, 353, 713

Kauffmann G., Li C., Zhang W., Weinmann S., 2013, *MNRAS*, 430, 1447

Klypin A. A., Trujillo-Gomez S., Primack J., 2011, *ApJ*, 740, 102

Lacey C. G. et al., 2016, *MNRAS*, 462, 3854

Lagos C. d. P., Tobar R. J., Robotham A. S., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018, *MNRAS*, 481, 3573

Lee J., Kimm T., Katz H., Rosdahl J., Devriendt J., Slyz A., 2020, *ApJ*, 905, 31

Li C., White S. D., 2009, *MNRAS*, 398, 2177

Libeskind N. I. et al., 2020, *MNRAS*, 498, 2968

Lu T., Gilbank D. G., McGee S. L., Balogh M. L., Gallagher S., 2012, *MNRAS*, 420, 126

Luo Y., Kang X., Kauffmann G., Fu J., 2016, *MNRAS*, 458, 366

Marchesini D., Van Dokkum P. G., Schreiber N. M. F., Franx M., Labbé I., Wuyts S., 2009, *ApJ*, 701, 1765

Marchesini D. et al., 2010, *ApJ*, 725, 1277

McCarthy I. G., Frenk C. S., Font A. S., Lacey C. G., Bower R. G., Mitchell N. L., Balogh M. L., Theuns T., 2008, *MNRAS*, 383, 593

Mo H., Van den Bosch F., White S., 2010, *Galaxy Formation and Evolution*. Cambridge Univ. Press, Cambridge

Mosleh M., Tavasoli S., Tacchella S., 2018, *ApJ*, 861, 101

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903

Muzzin A., Wilson G., Lacy M., Yee H. K. C., Stanford S. A., 2008, *ApJ*, 686, 966

Muzzin A. et al., 2009, *ApJ*, 698, 1934

Muzzin A. et al., 2013, *ApJ*, 777, 18

Nelson D. et al., 2019, *MNRAS*, 490, 3234

Oemler A., 1974, PhD thesis, California Institute of Technology

Oman K. A., Bahé Y. M., Healy J., Hess K. M., Hudson M. J., Verheijen M. A. W., 2021, *MNRAS*, 501, 5073

Pallero D., Gómez F. A., Padilla N. D., Torres-Flores S., Demarco R., Cerulo P., Olave-Rojas D., 2019, *MNRAS*, 488, 847

Peng Y.-j. et al., 2010, *ApJ*, 721, 193

Pillepich A. et al., 2018, *MNRAS*, 475, 648

Pillepich A. et al., 2019, *MNRAS*, 490, 3196

Pintos-Castro I., Yee H. K. C., Muzzin A., Old L., Wilson G., 2019, *ApJ*, 876, 40

Planck Collaboration, 2016, *A&A*, 594, A13

Riebe K. et al., 2011, preprint (arXiv:1109.0003)

Roediger E., Brüggemann M., 2007, *MNRAS*, 380, 1399

Salim S. et al., 2007, *ApJS*, 173, 267

Sarron F., Adami C., Durret F., Laigle C., 2019, *A&A*, 632, A49

Schaye J. et al., 2015, *MNRAS*, 446, 521

Skillman S. W., Warren M. S., Turk M. J., Wechsler R. H., Holz D. E., Sutter P. M., 2014, preprint (arXiv:1407.2600)

Somerville R. S., Primack J. R., 1999, *MNRAS*, 310, 1087

Spergel D. N. et al., 2003, *ApJS*, 148, 175

Springel V., 2010, *MNRAS*, 401, 791

Springel V., White S. D., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726

Springel V. et al., 2005, *Nature*, 435, 629

Stevens A. R., Croton D. J., Mutch S. J., 2016, *MNRAS*, 461, 859

Tavasoli S., Vasei K., Mohayaee R., 2013, *A&A*, 553, A15

Tecce T. E., Cora S. A., Tissera P. B., Abadi M. G., Lagos C. d. P., 2010, *MNRAS*, 408, 2008

Tinker J. L., 2020, preprint (arXiv:2007.12200)

Tomczak A. R. et al., 2014, *ApJ*, 783, 85

Vijayaraghavan R., Ricker P. M., 2017, *ApJ*, 841, 38

Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518

von der Linden A., Wild V., Kauffmann G., White S. D., Weinmann S., 2010, *MNRAS*, 404, 1231

Wang J., Xu W., Lee B., Du M., Overzier R., Shao L., 2020a, *ApJ*, 903, 103

Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020b, *Nature*, 585, 39

Webb K. et al., 2020, *MNRAS*, 498, 5317

Weinmann S. M., Van Den Bosch F. C., Yang X., Mo H., 2006, *MNRAS*, 366, 2

Wetzel A. R., Tinker J. L., Conroy C., 2012, *MNRAS*, 424, 232

Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2014, *MNRAS*, 439, 2687

White S. D. M., Rees M. J., 1978, *MNRAS*, 183, 341

Wilson G. et al., 2009, *ApJ*, 698, 1943

Yan H., Fan Z., White S. D. M., 2013, *MNRAS*, 430, 3432

Yang X., Mo H. J., van den Bosch F. C., 2003, *MNRAS*, 339, 1057

Yang X., Mo H. J., van den Bosch F. C., Jing Y. P., 2005, *MNRAS*, 356, 1293

Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, *ApJ*, 671, 153

Yates R. M., Henriques B., Thomas P. A., Kauffmann G., Johansson J., White S. D., 2013, *MNRAS*, 435, 3500

Yates R. M., Thomas P. A., Henriques B. M. B., 2017, *MNRAS*, 464, 3169

Yun K. et al., 2019, *MNRAS*, 483, 1042

Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L., 2005, *MNRAS*, 359, L30

## SUPPORTING INFORMATION

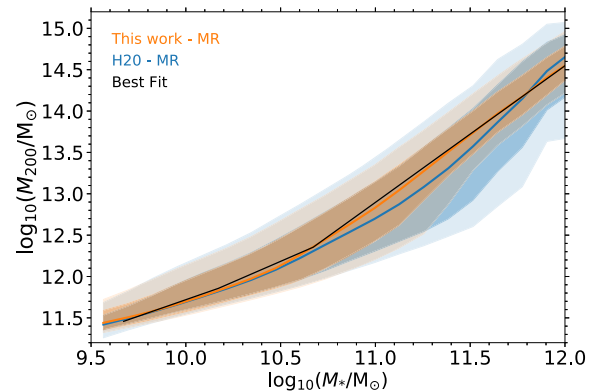
Supplementary data are available at *MNRAS* online.

### SuppMat.zip

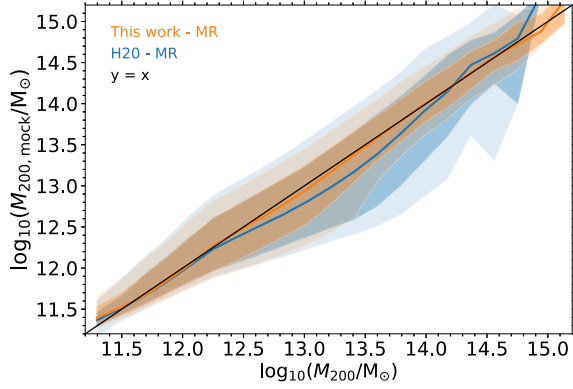
Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: DERIVING THE HALO MASS FROM STELLAR MASS

The total halo mass,  $M_{200}$ , is not observable and is usually estimated using other well-defined observables. As explained in Section 4.2, we derive the halo mass from the stellar mass of each halo's central



**Figure A1.** Halo mass as a function of stellar mass in this work (orange) and in H20 (blue). The solid black line shows our best fit using L-Galaxies (this work) to the halo mass, as given in equation (11). The orange and blue solid lines correspond to the median values, and the shaded regions illustrate  $1\sigma$  and  $2\sigma$  of the distribution. Our best fit (black line) is very close to the median value from the simulation (orange line), but slightly differs from the median line of the other model (H20).



**Figure A2.** The halo mass calculated using the stellar mass (equation 11) as a function of the halo mass measured directly from the Millennium simulation. The solid lines correspond to the median values and the shaded regions illustrate  $1\sigma$  and  $2\sigma$  of the distribution. The  $1\sigma$  scatter around  $y = x$  line is  $\sim 0.2$ – $0.3$  dex, and the  $2\sigma$  scatter is up to  $0.5$ – $0.6$  dex.

galaxy to make mock catalogues comparable with observations. Comparing these mock halo masses with their actual values from the simulation, we here explore the resulting scatter.

Fig. A1 shows the direct simulation halo mass from the Millennium simulation as a function of the stellar mass of the central

galaxy in our model (orange) and H20 (blue). The solid lines denote the median values and the light and dark shaded regions correspond to the  $1\sigma$  and  $2\sigma$  scatter of the distribution. The black line shows our best fit, which is in good agreement with our model (orange line) based on which the fitting is performed.

Additionally, Fig. A2 shows the fit halo mass from equation (11) (y-axis) as a function of the direct simulation halo mass (x-axis) in our model and H20. The two halo masses (fit and model) are in rather close agreement with the  $1\sigma$  scatter of  $0.2$ – $0.3$  dex. The H20 is a bit off near  $M_{200}/M_{\odot} \sim 10^{13}$ , which reflects different stellar mass to halo mass relationships in our model and H20. In this work, whenever we needed to convert stellar mass to halo mass, we used the fit based on our model (see Section 4.2.2).

Although this scatter is likely smaller than the typical error on stellar mass estimates from observations, its impact on the virial radius,  $R_{200}$ , could be substantial. For instance, if the mock halo mass is  $0.3$  dex larger/smaller than the true halo mass, this leads to the mock virial radius being  $\sim 20$ – $25$  per cent larger/smaller than the true virial radius. As a result, in our comparison with observations, where we report results as a function of halocentric distance (Figs 9 and 14), the trends with distance could contain  $20$ – $25$  per cent error.

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