

How gas flows shape the stellar-halo mass relation in the EAGLE simulation

Peter D. Mitchell^{*}¹, Joop Schaye¹

¹*Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, the Netherlands*

23 March 2021

ABSTRACT

The difference in shape between the observed galaxy stellar mass function and the predicted dark matter halo mass function is generally explained primarily by feedback processes. Feedback can shape the stellar-halo mass (SHM) relation by driving gas out of galaxies, by modulating the first-time infall of gas onto galaxies (i.e., preventative feedback), and by instigating fountain flows of recycled wind material. We present a novel method to disentangle these effects for hydrodynamical simulations of galaxy formation. We build a model of linear coupled differential equations that by construction reproduces the flows of gas onto and out of galaxies and haloes in the EAGLE cosmological simulation. By varying individual terms in this model, we isolate the relative effects of star formation, ejection via outflow, first-time inflow and wind recycling on the SHM relation. We find that for halo masses $M_{200} < 10^{12} M_{\odot}$ the SHM relation is shaped primarily by a combination of ejection from galaxies and haloes, while for larger M_{200} preventative feedback is also important. The effects of recycling and the efficiency of star formation are small. We show that if, instead of M_{200} , we use the cumulative mass of dark matter that fell in for the first time, the evolution of the SHM relation nearly vanishes. This suggests that the evolution is due to the definition of halo mass rather than to an evolving physical efficiency of galaxy formation. Finally, we demonstrate that the mass in the circum-galactic medium is much more sensitive to gas flows, especially recycling, than is the case for stars and the interstellar medium.

Key words: galaxies: formation – galaxies: evolution – galaxies: haloes – galaxies: stellar content

1 INTRODUCTION

With modern multi-wavelength extra-galactic surveys it is now possible to infer the buildup of galaxy stellar masses across cosmic time (e.g., Bell et al. 2003; Drory et al. 2005; Ilbert et al. 2010; Muzzin et al. 2013). Various empirical methods have in turn been developed to link these observations to the buildup of dark matter haloes, thought to host galaxies, using predictions from the Λ cold dark matter (Λ CDM) cosmological model. These methods include halo occupation distribution modelling (e.g., Peacock & Smith 2000; Berlind & Weinberg 2002), simple abundance matching (e.g., Vale & Ostriker 2004; Conroy et al. 2006), and more sophisticated forward modelling techniques (e.g., Yang et al. 2012; Moster et al. 2018; Behroozi et al. 2019). These models convincingly demonstrate that Λ CDM is consistent with the observed abundances and clustering of galaxies.

An important product of such modelling is the inferred median relationship between galaxy stellar mass and halo mass (e.g.,

Behroozi et al. 2010; Moster et al. 2010). This is commonly expressed as the ratio of galaxy stellar mass over halo mass (i.e., proportional to the efficiency with which baryons are converted into stars) as a function of halo mass, which we hereafter refer to as the stellar-halo mass (SHM) relation. The ratio of central galaxy stellar mass to halo mass is inferred to rise roughly linearly with increasing halo mass up until the characteristic mass scale $\approx 10^{12} M_{\odot}$, above which the ratio decreases (e.g., Yang et al. 2012; Wang et al. 2013; Lu et al. 2015; Rodríguez-Puebla et al. 2017; Kravtsov et al. 2018; Moster et al. 2018; Behroozi et al. 2019). This mass dependence is shown in Fig. 1, using data for simulated galaxies from the EAGLE cosmological hydrodynamical simulation (Schaye et al. 2015).

For halo masses $< 10^{12} M_{\odot}$, the shape of the SHM relation is conventionally explained by stellar feedback; the efficiency of feedback increases with decreasing halo mass as it becomes easier energetically to overcome the gravitational potential (e.g., Larson 1974; Dekel & Silk 1986; Cole 1991; White & Frenk 1991). For halo masses $> 10^{12} M_{\odot}$, a reduction in the efficiency with which baryons are converted into stars is generally explained via

^{*} E-mail: mitchell@strw.leidenuniv.nl

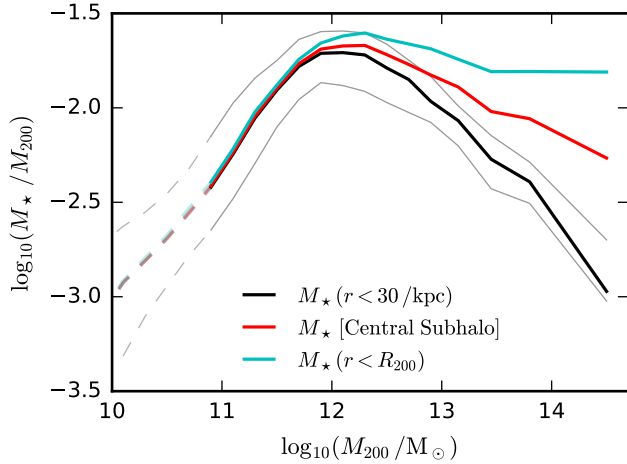


Figure 1. The SHM relation, defined as the ratio of galaxy stellar mass to halo mass, plotted as a function of halo mass for central galaxies. Data are taken from the Reference EAGLE simulation at $z = 0$ (Schaye et al. 2015). Black lines show the 16, 50, and 84th percentiles of the distribution, defining galaxy stellar mass as the mass within a spherical aperture of radius 30 kpc. The red line shows the median relation if the stellar mass is instead defined as the sum of all stellar particles within the central subhalo (this is the galaxy stellar mass definition we use in later figures). The cyan line shows the median relation if all star particles within the halo virial radius (R_{200}) are selected. Lighter dashed lines indicate the halo mass range where galaxies contain on average fewer than 100 star particles. The conversion of baryons into stars is not strongly dependent on halo mass for $M_{200} > 10^{12} M_{\odot}$, once satellite galaxies and the diffuse stellar halo are accounted for.

the longer associated radiative cooling timescales in massive, virialised systems (Rees & Ostriker 1977; White & Rees 1978), combined with the effects of efficient energy injection from an accreting central supermassive black hole (e.g., Tabor & Binney 1993; Silk & Rees 1998; Bower et al. 2006; Croton et al. 2006). Notably, as shown by the cyan line in Fig. 1, the SHM ratio actually does not decrease substantially for $M_{200} > 10^{12} M_{\odot}$ once satellite galaxies and the diffuse stellar halo are included into the SHM numerator (as they are by convention for the denominator). It is more correct therefore to state that AGN feedback inhibits star formation in massive haloes, without strongly reducing the conversion of baryons into stars compared to lower-mass haloes.

This overall picture has been fully realised and validated with various implementations of semi-analytic galaxy formation models (e.g., Bower et al. 2006; Croton et al. 2006; Somerville et al. 2008), and more recently by hydrodynamical simulations (e.g., Schaye et al. 2010; Dubois et al. 2014; Hirschmann et al. 2014; Vogelsberger et al. 2014; Schaye et al. 2015; Pillepich et al. 2018; Davé et al. 2019). Modern hydrodynamical simulations now convincingly reproduce the observationally inferred SHM relation, given uncertainties in both observations and in the energy of feedback processes that is able to mechanically power galactic outflows before being lost to radiation.

It has long been recognised that galaxies must continually accrete diffuse gas from their surrounding environments in order to explain the observed chemical abundances of stars (e.g., Larson 1972), and the relatively short gas depletion timescales of star-forming galaxies (e.g., Bauermeister et al. 2010). This picture is strongly supported by cosmological hydrodynamical simulations. Simulations also demonstrate that feedback processes could plau-

sibly affect diffuse gas accretion rates onto galaxies, both positively, by injecting metals into the circum-galactic medium, CGM, and inter-galactic medium, IGM (facilitating radiative cooling), and negatively, as galactic winds exert thermal over-pressure and kinetic ram pressure onto the surrounding gas (e.g., van de Voort et al. 2011; Faucher-Giguère et al. 2011; Nelson et al. 2015; Correa et al. 2018). In addition, gas ejected from galaxies can in principle be later re-accreted, forming a distinct galactic wind recycling contribution to galaxy growth (e.g., Oppenheimer & Davé 2008; Oppenheimer et al. 2010; Übler et al. 2014; Anglés-Alcázar et al. 2017; van de Voort 2017; Mitchell et al. 2020a).

The magnitude of these effects remains uncertain however, as does their differential impact on galaxies as a function of halo mass. It follows then that we do not have a full understanding of how the SHM relation is shaped respectively by the fluxes of accreting gas that is inflowing for the first time, accreting gas that is recycled, and gas that is outflowing from the ISM in a galactic wind. In semi-analytic galaxy formation models for example, the impact of stellar feedback on first-time inflow gas is generally neglected (see Pandya et al. 2020, for a discussion), such that the shape of the SHM relation is set by the dependence of the galaxy-scale mass loading factor (defined as the galactic outflow rate divided by the star formation rate) on halo mass for $M_{200} < 10^{12} M_{\odot}$, and by the impact of AGN suppressing cooling in more massive haloes (Mitchell et al. 2016).

For hydrodynamical simulations it is not trivial to infer the relative importance of these effects. The role of wind recycling is especially challenging to assess, given the requirement to track gas elements across the history of a simulation at sufficient time resolution to capture entrance to, and ejection from, the ISM. One approach is to vary the parameters of the feedback models implemented in simulations and interpret the resulting changes in galaxy properties (e.g., Correa et al. 2018). This will however affect both outflowing and inflowing gas at the same time, obfuscating the interpretation of the relative importance of first-time accretion, recycled accretion, and outflows at different spatial scales.

In this paper, we introduce a novel method to investigate how gaseous inflows and outflows shape the SHM relation. We make use of a complete set of measurements of galactic inflow and outflow rates (Mitchell et al. 2020a,b) from the EAGLE hydrodynamical simulation project (Schaye et al. 2015). Following Neistein et al. (2012), we then build a model that very closely follows the behaviour of EAGLE galaxies. The model is empirical in nature, but each of the terms in the model has a clear physical meaning (for example, the galactic mass loading factor). The structure of the model mimics that of conventional semi-analytic galaxy formation models, in that it tracks the integrated baryonic mass in different discrete components (i.e., CGM, ISM, stars), following the underlying hierarchical assembly of dark matter haloes and subhaloes. By varying the various terms in this model, we isolate how different parts of the network of gas flows around galaxies shape the SHM relation in the EAGLE simulation.

Since EAGLE is itself a (more complex) model, our conclusions derived from this approach will accordingly be model dependent. Nonetheless, this study still provides a physically plausible picture (in terms of feedback energetics, momentum input, etc) for the connection of the SHM relation to gaseous inflows and outflows. The model framework could be extended to consider measurements from other cosmological simulations, or constrained directly against observations using statistical inference, using simulation results as a prior (as in Mitra et al. 2015, for example).

The layout of this paper is as follows. We describe the EA-

GLE simulations, the measurements of inflow and outflow rates, and our modelling methodology in Section 2. We outline a simplified picture for the factors shaping the SHM relation in Section 3. Our main analysis of the SHM relation is presented in Section 4. A short extension of the analysis to the CGM and ISM is presented in Section 5, and we summarise our results in Section 6.

2 METHODS

2.1 Simulations

Our analysis is based on the EAGLE project (Schaye et al. 2015; Crain et al. 2015), which has been publicly released (McAlpine et al. 2016), and which includes a suite of cosmological simulations of various volumes, resolutions, and model variations. EAGLE uses a modified version of the GADGET-3 code (last described in Springel 2005) to simulate cubic, periodic regions of the Universe, solving for the equations of hydrodynamics and gravity, using smoothed particle hydrodynamics (SPH). A Λ CDM cosmological model is assumed, with parameters taken from Planck Collaboration et al. (2014). Simple “subgrid” models are included to account for the effect of physical processes that are not resolved and/or otherwise explicitly simulated. These include star formation, stellar evolution and feedback, supermassive black hole (SMBH) seeding, dynamics and growth, feedback from active galactic nuclei (AGN), and radiative cooling and heating.

Full details of each aspect of this modelling are given by Schaye et al. (2015) and references therein. A salient aspect is that both stellar and AGN feedback are modelled as thermal energy injection by a fixed temperature difference, with the temperature difference set high enough that spurious radiative cooling is mitigated given the unresolved nature of the simulated ISM (Booth & Schaye 2009; Dalla Vecchia & Schaye 2012). This in turn drives powerful galactic outflows that entrain gas in the circum-galactic medium and generally lead to high outflow rates at the halo virial radius (Mitchell et al. 2020b), and that can significantly reduce rates of first-time gas infall (Mitchell et al. 2020a; Wright et al. 2020).

The parameters of the Reference EAGLE model are calibrated such that the simulation is consistent with observed present-day star formation thresholds and kpc-scale efficiencies, such that the simulation broadly reproduces the observed galaxy stellar mass function and trends of galaxy size with stellar mass, and likewise that supermassive black hole masses are consistent with observations at a given stellar mass. We use the Reference EAGLE model for this paper, using measurements from the largest simulation with 1504^3 particles, a volume of $(100 \text{ Mpc})^3$, dark matter particle mass of $9.7 \times 10^6 M_\odot$, initial gas particle mass of $1.8 \times 10^6 M_\odot$, and maximum physical gravitational softening of 0.7 kpc.

2.2 Measurements

We make use of a comprehensive set of measurements tracking gas fluxes in EAGLE, as presented in detail by Mitchell et al. (2020a,b). In brief, we measure inflow and outflow rates at two scales. First, we measure fluxes at the halo virial radius (which we will refer to as the “halo scale”) according to which particles enter or leave the virial radius between two consecutive stored simulation outputs. We use 200 simulation outputs in total, with a spacing of $\approx 120 \text{ Myr}$ at low redshift, and with the spacing becoming finer with increasing redshift (see appendix A of Mitchell et al. 2020b).

We define the virial radius as the radius enclosing a mean overdensity that is 200 times the critical density of the Universe. We further use the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009) to define membership of particles within R_{200} to different subhaloes, distinguishing the central subhalo from satellite subhaloes on the basis of which subhalo contains the particle with the lowest value of the gravitational potential. Any particles that are within R_{200} but which are not considered bound by SUBFIND to any subhalo are redefined as belonging to the central subhalo.

We also measure fluxes at the boundary of the ISM (referred to as “galaxy scale”), according to which particles enter/leave the ISM between two consecutive outputs. The ISM is defined as particles that either pass the EAGLE star formation threshold based on density, temperature and metallicity (which captures the transition from the warm, atomic to the cold, molecular gas phase: Schaye 2004), or that otherwise have total hydrogen number density $n_H > 0.01 \text{ cm}^{-3}$ and are within 0.5 dex of the temperature floor corresponding to the equation of state imposed on the unresolved ISM (Schaye & Dalla Vecchia 2008; Mitchell et al. 2020b). In addition, we measure rates of star formation and stellar mass loss (due to stellar evolution, Wiersma et al. 2009), again defined as the change in mass between two consecutive simulation outputs.

Finally, we keep track of which gas particles were ejected from the ISM of a galaxy, and which gas particles were ejected from a subhalo. This information is then propagated through subhalo merger trees (accounting for mergers) and used to determine if gas accreted onto haloes or galaxies is recycled after having been ejected in the past. By then aggregating all these different measurements, we construct a complete description of baryonic assembly associated with dark matter haloes, expressed in terms of the total mass of the CGM (out to R_{200}), the ISM, stars, and any gas that has been ejected outwards beyond R_{200} . For simplicity, we define galaxy stellar masses as the sum of all stars within a given subhalo, without applying any spatial aperture selection.

2.3 The N12 model

Our objective is to understand the connection between the SHM relation and the underlying network of gaseous inflows, outflows, star formation, and wind recycling. One way to achieve this would involve first running multiple cosmological simulations with different choices for subgrid models and parameters, and then assessing the impact of these choices on gas flows and on stellar mass assembly. An alternative methodology that we introduce here is to first measure the average behaviour in a single cosmological simulation, and second, to then construct a model that reproduces this average behaviour, and third, to apply variations to the model in order to understand the role and relative importance of its various components.

To implement this strategy, we use the methodology introduced by Neistein et al. (2012), and we refer hereafter to our implementation of this framework as the “N12 model”. Neistein et al. (2012) demonstrated that by measuring the average inflow, star formation, and outflow rates of gas of galaxies in a cosmological simulation as a function of halo mass and redshift, it is possible (using only this averaged information plus the halo merger tree of each galaxy) to accurately reproduce the stellar mass of individual galaxies from the original simulation. They applied this framework to a simulation from the Overwhelmingly Large Simulations (OWLS, Schaye et al. 2010). Here we apply, with a number of modifications, the same methodology to the $(100 \text{ Mpc})^3$ -volume simulation from the EAGLE simulation project.

The basic structure of the N12 model resembles that of a typical semi-analytic galaxy formation model (e.g., Guo et al. 2011; Somerville et al. 2012; Lacey et al. 2016; Lagos et al. 2018). The baryonic content of each dark matter subhalo in a merger tree is split between the following components: stars (M_*), the ISM (M_{ISM}), the CGM out to R_{200} (M_{CGM}), and (newly in our implementation) a reservoir of gas that has been ejected beyond (and still resides beyond) R_{200} ($M_{\text{ej}}^{\text{halo}}$). The model tracks the mass in each of these components, and computes the various exchanges of mass between them. When two subhaloes in a merger tree “merge”¹, the masses of each baryonic component are simply summed.

The mass exchanges that take place between the baryonic components are given by the following set of ordinary differential equations:

$$\begin{bmatrix} \dot{M}_{\text{CGM}} \\ \dot{M}_{\text{ISM}} \\ \dot{M}_{\text{ej}}^{\text{halo}} \\ \dot{M}_* \end{bmatrix} = \begin{bmatrix} f_{\text{acc}}^{\text{halo}} \frac{\Omega_b}{\Omega_m - \Omega_b} \dot{M}_{\text{DM}}^{\text{1st}} - \frac{G_{\text{ret}}^{\text{gal}}}{t} M_{\text{ej}}^{\text{gal}} \\ \frac{G_{\text{ret}}^{\text{gal}}}{t} M_{\text{ej}}^{\text{gal}} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -F_{\text{CGM}}^{\text{pr}} \frac{G_{\text{acc}}^{\text{gal}}}{t} & (\eta^{\text{gal}} - \eta^{\text{halo}}) \frac{G_{\text{SF}}}{t} & \frac{G_{\text{ret}}^{\text{halo}}}{t} & 0 \\ F_{\text{CGM}}^{\text{pr}} \frac{G_{\text{acc}}^{\text{gal}}}{t} & -(1 - R + \eta^{\text{gal}}) \frac{G_{\text{SF}}}{t} & 0 & 0 \\ 0 & \eta^{\text{halo}} \frac{G_{\text{SF}}}{t} & -\frac{G_{\text{ret}}^{\text{halo}}}{t} & 0 \\ 0 & (1 - R) \frac{G_{\text{SF}}}{t} & 0 & 0 \end{bmatrix} \begin{bmatrix} M_{\text{CGM}} \\ M_{\text{ISM}} \\ M_{\text{ej}}^{\text{halo}} \\ M_* \end{bmatrix},$$

where M_{CGM} , M_{ISM} , $M_{\text{ej}}^{\text{halo}}$ and M_* are respectively the masses in the CGM, ISM, the reservoir of gas ejected from the halo, and in stars. In addition, we also track the mass of gas that has been ejected from the ISM of progenitor galaxies in a galactic wind, $M_{\text{ej}}^{\text{gal}}$. $M_{\text{ej}}^{\text{gal}}$ is comprised of gas that is part of the CGM, or that is part of the ejected gas reservoir located beyond the virial radius, and so is not mutually exclusive from M_{CGM} and $M_{\text{ej}}^{\text{halo}}$. $M_{\text{ej}}^{\text{gal}}$ is itself governed by the equation:

$$\dot{M}_{\text{ej}}^{\text{gal}} = \eta^{\text{gal}} \frac{G_{\text{SF}}}{t} M_{\text{ISM}} - \frac{G_{\text{ret}}^{\text{gal}}}{t} M_{\text{ej}}^{\text{gal}}. \quad (2)$$

$\dot{M}_{\text{DM}}^{\text{1st}}$ in Eqn. 1 is the “smooth” accretion rate of dark matter particles that are entering the halo for the first time. We define “smooth” accretion as any particles that are accreted through R_{200} while not considered bound to any other subhalo with a mass greater than $9.7 \times 10^8 M_{\odot}$, which corresponds to the mass of 100 dark matter particles at fiducial EAGLE resolution (see discussion in section 2.3 of Mitchell et al. 2020a). Ω_m is the cosmic matter density, and Ω_b is the cosmic baryonic matter density. t in both Eqns 1,2 is the age

of the Universe at a given redshift. $F_{\text{CGM}}^{\text{pr}}$ is the mass fraction of the CGM that has *not* been ejected from the ISM of a progenitor galaxy of the current subhalo. If we split $M_{\text{ej}}^{\text{gal}}$ between gas that belongs to the CGM, and gas that belongs to the ejected gas reservoir beyond R_{200} , such that $M_{\text{ej}}^{\text{gal}} = M_{\text{ej}}^{\text{gal}}(r < R_{200}) + M_{\text{ej}}^{\text{gal}}(r > R_{200})$, then $F_{\text{CGM}}^{\text{pr}}$ is defined as $F_{\text{CGM}}^{\text{pr}} \equiv \frac{M_{\text{CGM}} - M_{\text{ej}}^{\text{gal}}(r < R_{200})}{M_{\text{CGM}}}$. The remaining terms in Eqn. 1 are then coefficients that represent the efficiency of various physical processes, such as the efficiency of first-time gas accretion at the virial radius ($f_{\text{acc}}^{\text{halo}}$), the efficiency of galactic outflows (η^{gal}), and so on. The meaning of these various terms is as follows:

- **First-time gas accretion at R_{200} ($f_{\text{acc}}^{\text{halo}}$):** the source term for baryonic accretion is $\dot{M}_{\text{CGM}} = f_{\text{acc}}^{\text{halo}} \frac{\Omega_b}{\Omega_m - \Omega_b} \dot{M}_{\text{DM}}^{\text{1st}}$, which represents the smooth accretion of gas that has not been ejected from a progenitor subhalo in the past. We assume that gas accretion traces dark matter accretion at this radius since, in the absence of strong feedback effects, we expect that R_{200} approximately marks the spatial scale where a virial shock (if present) develops, and correspondingly marks the location where thermal pressure gradients are expected to significantly decouple the further inwards accretion of gas relative to that of collisionless dark matter. We choose to express this as a function of first-time smooth dark matter accretion (rather than the more conventional choice of simply the total dark matter accretion rate) because a significant fraction of accreting dark matter particles at the virial radius have already crossed this threshold in the past (e.g., Wright et al. 2020) and are being re-accreted, at which point the dynamics should be already decoupled from that of the gas. $f_{\text{acc}}^{\text{halo}}$ is then a coefficient that represents any deviation from the case where first-time gas accretion traces that of dark matter. In Mitchell et al. (2020a), we show that $f_{\text{acc}}^{\text{halo}}$ drops below unity in low-mass haloes ($M_{200} < 10^{11} M_{\odot}$), and Wright et al. (2020) demonstrate explicitly that this is because of feedback processes in the EAGLE simulations (i.e. “preventative” feedback; see also van de Voort et al. 2011). In higher-mass haloes, $f_{\text{acc}}^{\text{halo}}$ can actually exceed unity as gas that is prevented from being accreted onto low-mass progenitors at early times can catch up to cumulative dark matter accretion by being accreted onto more massive descendant haloes later on (Mitchell et al. 2020a).

- **First-time galaxy-scale accretion ($G_{\text{acc}}^{\text{gal}}$):** gas in the CGM is then assumed to infall onto the ISM at the rate $\dot{M}_{\text{ISM}} = F_{\text{CGM}}^{\text{pr}} G_{\text{acc}}^{\text{gal}} M_{\text{CGM}}/t$, where $F_{\text{CGM}}^{\text{pr}} M_{\text{CGM}}$ is the mass in the CGM that has not been present in the ISM of a progenitor galaxy of the current subhalo in the past, $G_{\text{acc}}^{\text{gal}}$ is the dimensionless efficiency of first-time gaseous infall from the CGM to the ISM, and t is the age of the Universe at a given redshift. $G_{\text{acc}}^{\text{gal}}$ is related to the characteristic timescale for the CGM to be depleted by gas accretion onto the ISM ($\tau_{\text{acc}}^{\text{gal}}$) by $G_{\text{acc}}^{\text{gal}} \equiv t/\tau_{\text{acc}}^{\text{gal}}$. We choose to express the model coefficients related to timescales ($G_{\text{acc}}^{\text{gal}}$, G_{SF} , $G_{\text{ret}}^{\text{gal}}$, $G_{\text{ret}}^{\text{halo}}$) in this way to ensure first that they are dimensionless (factoring out the zeroth order time dependence as galactic and halo-scale timescales generally scale linearly with the age of the Universe), and second to ensure that higher values of G intuitively imply higher efficiencies of gas accretion, star formation, or wind recycling.

- **Star formation (G_{SF}):** gas within the ISM of galaxies is converted into stars at the rate $\dot{M}_* = G_{\text{SF}} M_{\text{ISM}}/t$. In this case $G_{\text{SF}} = t/\tau_{\text{SF}}$, where $\tau_{\text{SF}} \equiv M_{\text{ISM}}/\dot{M}_*$ is the conventional characteristic gas depletion time for the ISM to be depleted by star formation.

¹ This refers to the moment at which a satellite subhalo is sufficiently disrupted (by tidal effects) that it can no longer be identified by SUBFIND, which may or may not correspond to the exact moment of a particular definition of a galaxy-galaxy merger. Satellite subhaloes that vanish but then reappear at a later snapshot are not merged.

• **Stellar recycling (R):** stars return mass to the ISM at a rate given by $\dot{M}_{\text{ISM}} = R G_{\text{SF}} M_{\text{ISM}}/t$, where R is the recycled fraction. R should depend on the full star formation and chemical enrichment history of the galaxy, but (for reasons of numerical convenience) we make the assumption that R can be parametrised as a function of the current star formation rate ($G_{\text{SF}} M_{\text{ISM}}/t$).

• **Galaxy- and halo-scale outflows (η^{gal} and η^{halo}):** gas is ejected from the ISM in galactic winds with a rate $\dot{M}_{\text{ISM}} = -\eta^{\text{gal}} G_{\text{SF}} M_{\text{ISM}}/t$, where η^{gal} is the galaxy-scale dimensionless mass loading factor, and $G_{\text{SF}} M_{\text{ISM}}/t$ is the star formation rate. Similarly, gas is ejected from the CGM to the ejected gas reservoir outside R_{200} with a rate $\dot{M}_{\text{CGM}} = -\eta^{\text{halo}} G_{\text{SF}} M_{\text{ISM}}/t$, where η^{halo} is the halo-scale dimensionless mass loading factor.

• **Galaxy- and halo-scale gas recycling ($G_{\text{ret}}^{\text{gal}}$ and $G_{\text{ret}}^{\text{halo}}$):** gas that has been ejected from the ISM is assumed to return at the rate $\dot{M}_{\text{ISM}} = G_{\text{ret}}^{\text{gal}} \dot{M}_{\text{ej}}^{\text{gal}}/t$, where $G_{\text{ret}}^{\text{gal}}$ is the dimensionless efficiency of galaxy-scale wind recycling. Similarly, gas that has been ejected from the CGM to outside R_{200} is assumed to return to the CGM at the rate $\dot{M}_{\text{CGM}} = G_{\text{ret}}^{\text{halo}} \dot{M}_{\text{ej}}^{\text{halo}}/t$, where $G_{\text{ret}}^{\text{halo}}$ is the dimensionless efficiency of halo-scale gas recycling.

Fiducial values for these various terms are computed for individual EAGLE galaxies as described in Section 2.2. Following Neistein et al. (2012), we then compute averages for each quantity as a function of halo mass and redshift. Within a given mass and redshift bin, we compute the weighted mean of the associated numerator and denominator of each term. Taking for example the galaxy-scale outflow efficiency, this is computed as

$$\langle \eta^{\text{gal}} \rangle (M_{200}, z) = \frac{\sum_i^N w_i \dot{M}_{\text{out},i}}{\sum_i^N w_i \dot{M}_{\star,i}}, \quad (3)$$

where \dot{M}_{out} is the outflow rate and \dot{M}_{\star} is the star formation rate, summing over the N galaxies that are present in a given bin. Each redshift bin contains ≈ 25 simulation snapshots, and so the same galaxy can appear multiple times inside the same bin. As such, and given that we are computing separate means for the numerator and denominator, we are implicitly averaging (in this case) outflow rates and star formation rates over the full redshift interval encompassed by each bin, smoothing out the phase differences between star formation and outflow events that exist within a given bin. Using mean statistics is preferable for low-mass galaxies that at our resolution have a mix of zero and non-zero star formation/inflow/outflow rates within a given halo mass bin. The weights used in the average (w_i) are set to be the inverse of the total number of galaxies identified at that snapshot, such that each snapshot contributes equally to the average.

In addition to the various mass exchanges described previously, we have also measured additional terms including the mass that is ejected/reaccreted from galaxies and haloes that does not pass the velocity cuts described in Section 2.2, and is therefore not considered as a genuine outflow, and can be thought of instead as a combination of noise as particles fluctuate across the somewhat arbitrary boundaries we define for the ISM and CGM, and low-velocity fountain flows. We also measure the gas accretion rates onto galaxies and haloes that has been ejected from non-progenitor galaxies and haloes (i.e., “galactic” or “halo” transfer, e.g., Anglés-Alcázar et al. 2017; Borrow et al. 2020; Mitchell et al. 2020a). We include both these “failed-outflow” and “transfer” terms in the N12 model, averaging them as a function of

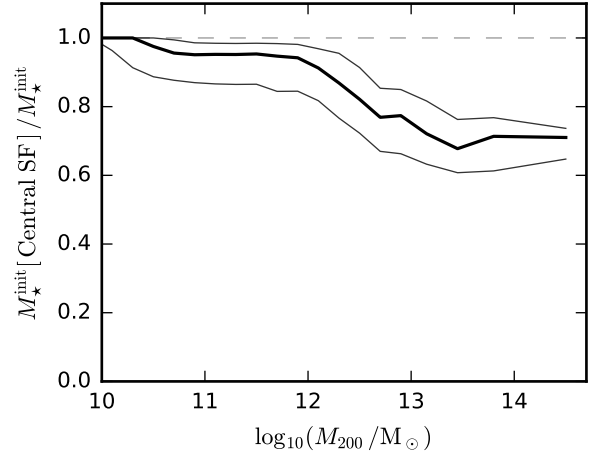


Figure 2. The fraction of star formation that occurs within progenitor subhaloes that are classed as central (rather than as satellites), plotted as a function of halo mass at $z = 0$. Note that this fraction is higher than the “in situ” fraction (not plotted), which includes only stars within the main progenitor subhalo. Fractions are computed for galaxies that are central at $z = 0$, summing over all progenitor subhaloes. Solid lines show the 16, 50, and 84th percentiles of the distribution. On average, the total mass formed in central progenitors is between 70 % and 100 %, with the fraction decreasing with increasing halo mass. This is relevant to our analysis since any variations to the N12 model will not affect the stars that are forming within subhaloes that are classified as satellites.

halo mass and redshift as with the main model terms. In practice the transfer terms are simply folded into $G_{\text{acc}}^{\text{gal}}$ and $G_{\text{acc}}^{\text{halo}}$ (and are adjusted accordingly if $G_{\text{acc}}^{\text{gal}}$, $G_{\text{acc}}^{\text{halo}}$ are themselves adjusted). The failed-outflow terms are assumed to scale linearly with \dot{M}_{ISM} at the galaxy scale, and with \dot{M}_{CGM} at the halo scale, such that $\dot{M}_{\text{ISM}} = (G_{\text{out,fail}}^{\text{gal}} - G_{\text{ret,fail}}^{\text{gal}}) \dot{M}_{\text{ISM}}/t$, and $\dot{M}_{\text{CGM}} = (G_{\text{out,fail}}^{\text{halo}} - G_{\text{ret,fail}}^{\text{halo}}) F_{\text{CGM}}^{\text{pr}} \dot{M}_{\text{CGM}}/t$, in addition to the mass exchanges described in Eqn. 1. Including these extra terms slightly improves the accuracy of our fiducial model, but they are generally subdominant compared to the main terms.

Unlike Neistein et al. (2012), we have not computed inflow/outflow rates for satellite galaxies, and as such we compute average values of each term in Eqn. 1 for central galaxies only². In our implementation of the N12 model, we evolve a given satellite galaxy simply by adjusting the mass of each tracked component to match the actual change in mass of each component of the galaxy between snapshots (we also account for mass exchanges between central and satellite subhaloes in this manner). This means that if we choose to adjust a given term in our model, the change will only affect the evolution of central galaxies at that snapshot. To gauge the importance of this choice for the final stellar mass of galaxies, Fig. 2 shows the fraction of the final stellar mass that was formed in progenitors while they were classified as centrals (as opposed to satellites), plotted for descendant galaxies that are centrals at $z = 0$. This fraction decreases from nearly 100 % in low-mass central galaxies to ≈ 70 % in galaxy groups and clusters. As such, any modifications we make to the model will have a reduced impact on stellar mass assembly for haloes with mass $\gg 10^{12} M_{\odot}$. Note that the fraction of star formation occurring in centrals is larger than

² Note that we do track the ejection/recycling of particles in EAGLE satellite galaxies, such that they are correctly added to the ejected gas reservoir of a central galaxy after a merger.

the more commonly considered “in-situ fraction” (i.e. the fraction of stars that were formed within the main progenitor of the current galaxy), since typically most of the stars that are accreted onto the central galaxy in a merger will have formed before the secondary merging galaxy was classified as a satellite.

As a final detail, the computation of the quantity $F_{\text{CGM}}^{\text{pr}}$ that appears in Eqn. 1 requires some additional steps. We introduce this complexity in order to be able to cleanly separate first-time versus recycled gas accretion onto galaxies and haloes. Specifically, we want to ensure that increasing the efficiency of galactic outflows (η^{gal}) in our model does not implicitly increase the rate of first-time gas infall from the CGM. From the Reference EAGLE simulation, we compute two terms in addition to those shown in Fig. 3, representing the probability that halo-scale outflows and gas recycling includes gas that was ejected from the ISM in the past. Full details are presented in Appendix A.

2.4 Measured coefficients

Fig. 3 shows the range of averaged coefficients that appear in our implementation of the N12 model, excluding the highest redshift bin for visual clarity. As discussed in Mitchell et al. (2020a,b), most of the coefficients shown in Fig. 3 exhibit a dependence on halo mass that will preferentially suppress the formation of stars in either (or both) low-mass haloes ($M_{200} \ll 10^{12} M_{\odot}$) and high-mass haloes ($M_{200} \gg 10^{12} M_{\odot}$).

Halo-scale first time gas accretion ($f_{\text{acc}}^{\text{halo}}$, upper-left panel) increases in efficiency with increasing halo mass, and will therefore suppress star formation (relatively) in low-mass haloes. As discussed in Mitchell et al. (2020a); Wright et al. (2020), feedback reduces first-time gas accretion rates at R_{200} for $M_{200} < 10^{11} M_{\odot}$ in EAGLE, but this delayed early accretion is compensated for by increased first time accretion rates onto more massive haloes (that are naturally the descendants of the lower-mass haloes), in excess of the simple expectation following from first time dark matter accretion rates. Galaxy-scale first time gas accretion ($G_{\text{acc}}^{\text{gal}}$, upper-middle panel) drops sharply in efficiency for $M_{200} > 10^{12} M_{\odot}$ (due to inefficient radiative cooling and the impact of AGN feedback), and will therefore suppress star formation in the most massive haloes. The efficiency of star formation per unit ISM mass (G_{SF} , upper-right panel) generally increases with increasing halo mass, implying that star formation is preferentially suppressed in low-mass haloes (we will show later however that G_{SF} does not shape the SHM relation in the Reference EAGLE simulation).

The galaxy- and halo-scale outflow mass loading factors (η^{gal} , η^{halo} , central and middle-right panels) depend negatively on halo mass for $M_{200} < 10^{12} M_{\odot}$, reflecting the relative ease with which stellar feedback can eject baryons from low-mass galaxies and haloes. For $M_{200} > 10^{12} M_{\odot}$, the mass loading factors instead increase with increasing halo mass (especially for η^{halo}), due to AGN feedback. Gas ejection via outflows should therefore suppress star formation in both low and high-mass haloes. Note that η^{halo} is generally larger than η^{gal} at low redshift (see Mitchell et al. 2020b, for a discussion). On the other hand, η^{gal} more directly affects the ISM, and so it is not obvious which of these two terms is more important for the SHM relation.

Finally, galaxy-scale wind recycling ($G_{\text{ret}}^{\text{gal}}$, lower-left panel) peaks in efficiency at $M_{200} \approx 10^{12} M_{\odot}$, and could enhance star formation at this mass scale. Halo-scale wind recycling ($G_{\text{ret}}^{\text{halo}}$, lower-middle panel) increases monotonically in efficiency with increasing halo mass, and would be expected therefore to tilt the SHM relation in towards a more positive slope.

It is also apparent that the redshift evolution of the dimensionless terms can vary in sign from one process to another: for example the dimensionless infall efficiency ($G_{\text{acc}}^{\text{gal}}$) decreases with decreasing redshift, whereas the dimensionless star formation efficiency (G_{SF}) increases with decreasing redshift. It is not obvious how these opposing trends will combine to produce the resulting redshift evolution of the SHM relation.

3 BASIC EXPECTATIONS FOR THE SHM RELATION

Before applying the full N12 model, we first use the measurements of inflow and outflow rates in EAGLE to outline a basic expectation for the relative importance of “preventative feedback”, galactic outflows, and wind recycling in shaping the mass dependence of the SHM relation. Specifically, if we ignore the flows within the circum-galactic medium, we can reduce the relevant equations to the following.

First, we can assume that first-time gas accretion onto the ISM of galaxies tracks the accretion of gas at the halo virial radius, such that the galactic first-time accretion rate is given by $\dot{M}_{\text{ISM}} = f_{\text{prev}} f_b \dot{M}_{200}$, where $f_b \equiv \frac{\Omega_b}{\Omega_m}$ is the cosmic baryon fraction. Here, f_{prev} represents the combined effects of gravitational heating and feedback in reducing the accretion rates of gas onto galaxies, referred to collectively as preventative feedback.

Second, if we make the assumption that galaxy formation is self-regulated (e.g., Finlator & Davé 2008; Schaye et al. 2010; Davé et al. 2012; Lilly et al. 2013), such that the gas accretion rate onto the ISM balances the star formation plus outflow rate (neglecting stellar mass loss for simplicity), then the resulting mass conservation equation is

$$\dot{M}_{\star} + \dot{M}_{\text{out}} = (1 + \eta^{\text{gal}}) \dot{M}_{\star} = f_{\text{prev}} f_b \dot{M}_{200} + \dot{M}_{\text{ret}}, \quad (4)$$

where \dot{M}_{out} is the galaxy-scale outflow rate, η^{gal} is the galaxy-scale mass loading factor, and \dot{M}_{ret} is the rate of recycled gas accretion. As in Eqn. 1, we can assume that galactic wind recycling can be parametrised as $\dot{M}_{\text{ret}} = G_{\text{ret}}^{\text{gal}} \dot{M}_{\text{ej}}^{\text{gal}} / t$, where $G_{\text{ret}}^{\text{gal}}$ is the characteristic gas return efficiency, t is the age of the Universe at a given redshift, and $\dot{M}_{\text{ej}}^{\text{gal}}$ is the mass in the gas reservoir that has been ejected from the ISM. This mass is in turn given by the integral $\dot{M}_{\text{ej}}^{\text{gal}} = \int (\eta^{\text{gal}} \dot{M}_{\star} - \dot{M}_{\text{ret}}) dt$. By substituting in Eqn. 4, we have $\dot{M}_{\text{ej}}^{\text{gal}} = \int \left[\frac{\eta^{\text{gal}}}{1 + \eta^{\text{gal}}} (f_{\text{prev}} f_b \dot{M}_{200} + \dot{M}_{\text{ret}}) - \dot{M}_{\text{ret}} \right] dt$. If we make the approximations that $\eta^{\text{gal}} \gg 1$ then this reduces to $\dot{M}_{\text{ej}}^{\text{gal}} = \int f_{\text{prev}} f_b \dot{M}_{200} dt$, which further reduces to $\dot{M}_{\text{ej}}^{\text{gal}} = f_{\text{prev}} f_b \dot{M}_{200} t$ if \dot{M}_{200} and f_{prev} are assumed constant as a galaxy evolves. Substituting this back into Eqn. 4 yields

$$\dot{M}_{\star} \sim \frac{1}{(1 + \eta^{\text{gal}})} \left(1 + G_{\text{ret}}^{\text{gal}} \right) f_{\text{prev}} f_b \dot{M}_{200}. \quad (5)$$

While the approximations used are somewhat crude, the resulting expression has the advantage that the differential effects of galactic outflows (η^{gal}), preventative feedback (f_{prev}), and wind recycling ($G_{\text{ret}}^{\text{gal}}$) are cleanly separated. If we further make the approximation that these three terms are constant as galaxies evolve, then it follows finally that

$$\frac{M_{\star}}{M_{200}} \sim f_{\text{prev}} f_b \frac{1}{(1 + \eta^{\text{gal}})} \left(1 + G_{\text{ret}}^{\text{gal}} \right). \quad (6)$$

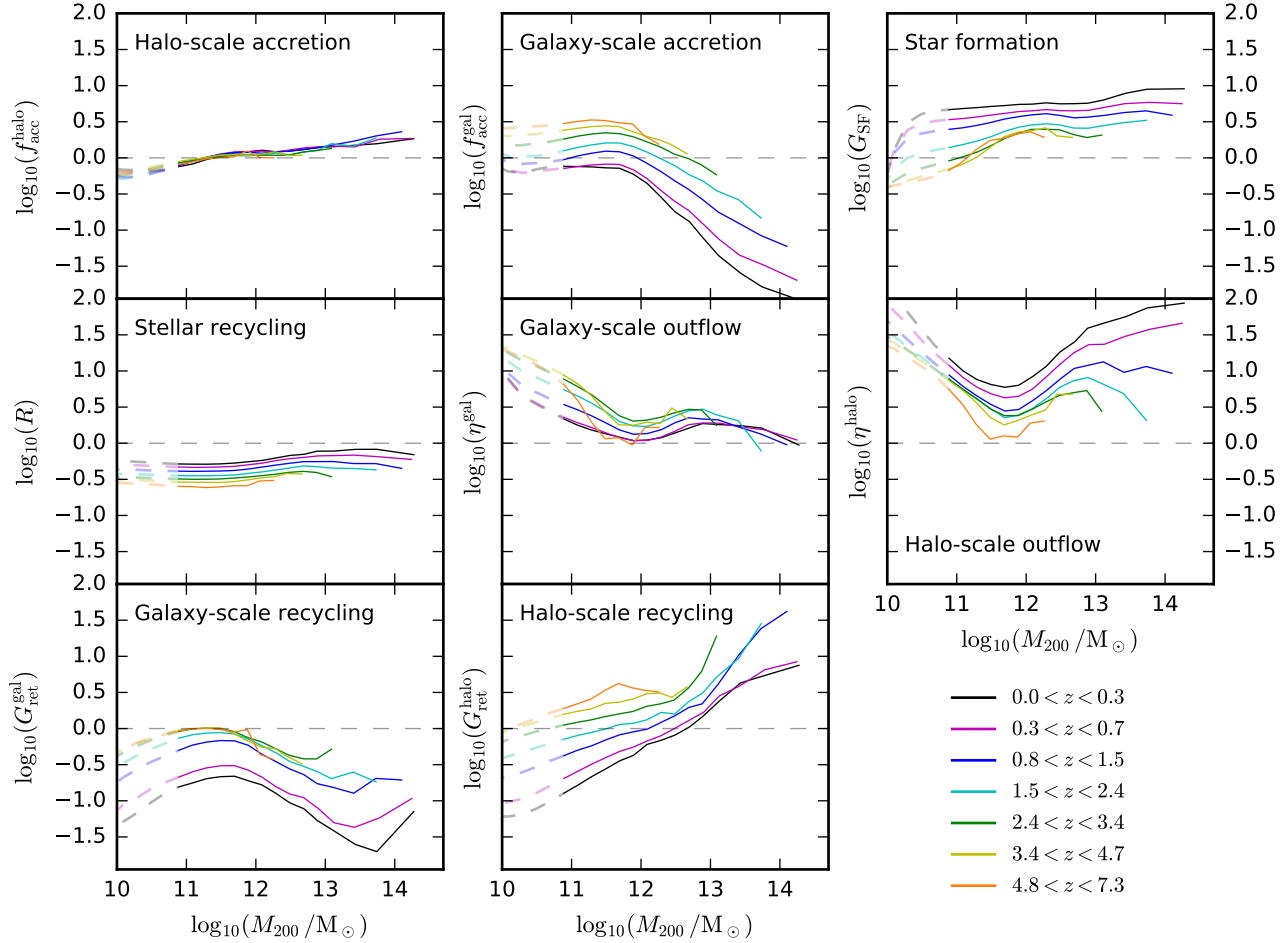


Figure 3. An overview of the various dimensionless coefficients that appear in our version of the N12 model (Eqn. 1), plotted as a function of halo mass. Each panel corresponds to a given coefficient, and each line colour corresponds to a specific redshift bin (as labelled in the bottom-right corner). A common x and y-axis range is used for all panels. Solid (dashed) lines indicate the halo mass range for which galaxies in the underlying reference EAGLE simulation contain on average more (fewer) than one hundred stellar particles.

Fig. 4 shows how the different terms in Eqn. 6 add multiplicatively to give an approximate prediction for the SHM relation. We choose to show the terms (and the EAGLE SHM) evaluated at $z \approx 1$, which is approximately the mid point for the evolution that each term undergoes at fixed halo mass (see Fig. 3). In addition, since Eqn. 4 ignores stellar recycling, we show the EAGLE SHM relation for stellar masses before any stellar mass loss (i.e., we sum the initial mass of each star particle in a galaxy).

Considering first the term that relates to wind recycling ($1 + G_{\text{ret}}^{\text{gal}} \equiv 1 + t/\tau_{\text{ret}}^{\text{gal}}$, cyan line), this term saturates when $G_{\text{ret}}^{\text{gal}} \ll 1$ ($\tau_{\text{ret}}^{\text{gal}} \gg t$), since wind recycling is of little importance if the material ejected in galactic winds takes on average longer than the age of the Universe to return. As such, while $G_{\text{ret}}^{\text{gal}}$ itself clearly follows the approximate shape of the SHM relation (see Fig. 3, lower-left panel), the long recycling timescales in EAGLE are expected to blunt the actual impact of wind recycling on the shape of the SHM relation. Recycling is more efficient (relative to the Hubble time) at high redshift, and so may play a larger role in this regime. Even for the values at $z \approx 1$ shown in Fig. 4, there is however still an appreciable curvature in the cyan line that contributes to the peak of the SHM ratio at a halo mass slightly below $10^{12} M_{\odot}$.

Considering second the preventative feedback term (f_{prev} ,

blue line in Fig. 4), it is clearly implied that this term is the most important (over ISM ejection via outflows and wind recycling) for setting the overall shape of the SHM relation. f_{prev} peaks strongly at a halo mass slightly below $10^{12} M_{\odot}$, and its effect does not saturate for low values of f_{prev} , since first-time inflows represent the source term for galactic gaseous assembly.

Thirdly, the ISM ejection outflow term ($1/(1 + \eta^{\text{gal}})$, red line in Fig. 4) has an apparent intermediate level of importance between the wind recycling and preventative feedback terms. This term peaks at $M_{200} \sim 10^{12} M_{\odot}$, at a slightly higher halo mass relative to the other two terms.

Combining these terms together (solid black line in Fig. 4), Eqn. 6 predicts an SHM that is qualitatively similar to that of EAGLE (dashed black line), but with various quantitative differences. Relative to EAGLE, Eqn. 6 predicts an SHM that peaks at slightly lower halo masses, that has a steeper slope for halo masses below the peak mass, and that is more sharply peaked. These differences stem from the fact that we are ignoring the redshift evolution that each term undergoes, from the simplifying approximations invoked to derive Eqn. 6, and from the fact that we are ignoring accretion of satellite galaxies and subhaloes (which affect the galaxy stellar mass and halo mass respectively). Ignoring satellite galaxies and

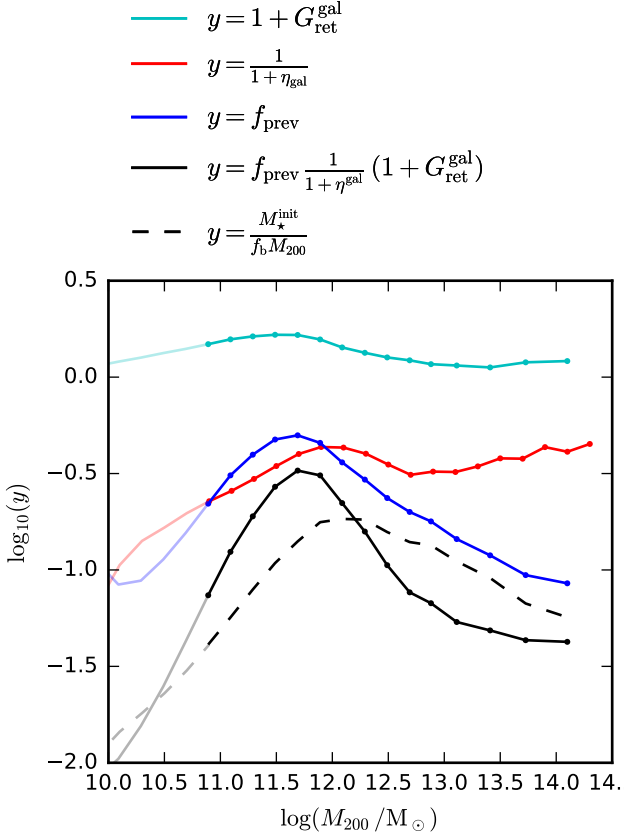


Figure 4. A simplified view of the relative importance of preventative feedback due to inefficient radiative cooling and outflows (blue), ejection of the ISM by galaxy-scale outflows (red), and galaxy-scale wind recycling (cyan) in shaping the overall SHM relation (black). Specifically, subsets of terms from Eqn. 6 are plotted as a function of halo mass, with the solid black line showing the combined expression, which represents an approximate prediction for the shape of the SHM relation. The dashed black line shows the actual SHM relation from EAGLE for reference. Eqn. 6 ignores stellar recycling, and so we show stellar masses from EAGLE that do not include stellar mass loss (M_{\star}^{init}). For each line, lighter shades indicate the halo mass range for which galaxies contain on average fewer than 100 star particles. All terms and masses are plotted given their measured values averaged over $0.8 < z < 1.5$. Taken at face value, the preventative feedback term f_{prev} (blue line) is primarily responsible for shaping the SHM relation, with ISM ejection (red line) playing a lesser but still important role, and with wind recycling (cyan) also playing an again lesser (but not negligible) role. We will revisit this basic picture using the full N12 model in Section 4.

subhaloes is expected to be a poor approximation for high-mass haloes, where accreted stars formed ex-situ are thought to form a large fraction of the total stellar mass (e.g., De Lucia & Blaizot 2007; Moster et al. 2018), which is indeed the case for EAGLE (Qu et al. 2017; Clauwens et al. 2018).

Despite these simplifications, Eqn. 6 does clearly capture the basic behaviour. That said, the “preventative feedback” term in Eqn. 6 effectively combines most of the gas flows that affect the circum-galactic medium into a single term. In the framework of the equilibrium model of Davé et al. (2012), f_{prev} conceptually represents the effects of feedback processes in slowing the first-time accretion rates of gas onto galaxies. As we shall show in the following section, once f_{prev} is split into different (time and mass-dependent) parts, including a reduction of gas inflow rates at the virial radius

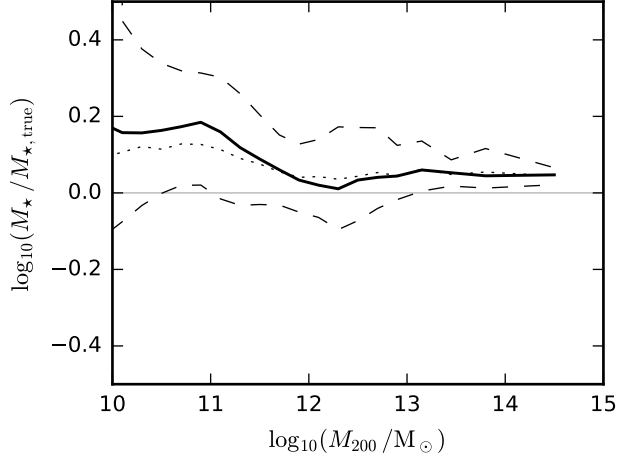


Figure 5. The residual in stellar mass between the N12 model and EAGLE at $z = 0$, plotted for central galaxies as a function of halo mass. Solid lines show the median, dashed lines show the 16 and 84th percentiles of the distribution, and the dotted line shows the mean residual. Overall, the N12 model reproduces EAGLE stellar masses to within a few tens of percent; the largest deviation is seen in very low-mass haloes, where the N12 model over-predicts the stellar mass by $\approx 40\%$. We regard the level of agreement to be sufficient for our analysis.

($f_{\text{acc}}^{\text{halo}}$ in Eqn. 1), first-time infall from the CGM to the ISM ($G_{\text{acc}}^{\text{gal}}$), halo-scale outflows (η^{halo}), and halo-scale wind recycling ($G_{\text{ret}}^{\text{halo}}$), it becomes apparent that (of these terms) it is the ejection of gas beyond the halo virial radius combined with inefficient gas infall in very massive haloes that is primarily responsible for shaping the SHM relation in EAGLE.

4 RESULTS

Fig. 5 demonstrates the accuracy with which the fiducial implementation of the N12 model is able to reproduce the SHM relation from the original reference EAGLE simulation. The two agree to within 10% on average for $M_{200} > 10^{12} M_{\odot}$. Agreement is worse at lower halo masses, with the model biased slightly high by $\approx 40\%$ for $M_{200} < 10^{11} M_{\odot}$. Overall, the median SHM relation is captured sufficiently accurately for our purposes.

4.1 What sets the normalisation of the SHM relation?

As a first application of the N12 model, we can simply multiply a single term that appears in Eqn. 1 by a constant factor, and thereby gauge the relevant importance of that term to the normalisation of the SHM relation. We consider both positive and negative factors of two, four and eight³. The resulting SHM relations for central galaxies are presented in Fig. 6. For reference, grey ticks show the expected change in the SHM relation if changes in galaxy stellar

³ Since the changes considered can be quite extreme, we choose here to disable the mass evolution of satellite galaxies (i.e., their star formation, inflow, outflow rates are all set to zero), which otherwise are evolved to exactly follow changes of the original satellites from EAGLE (see Section 2.3). Without this change, star formation within satellites can in some cases become the dominant mode of galaxy stellar mass growth (even for central galaxies due to mergers), obfuscating our interpretation of the effect of changing a model term.

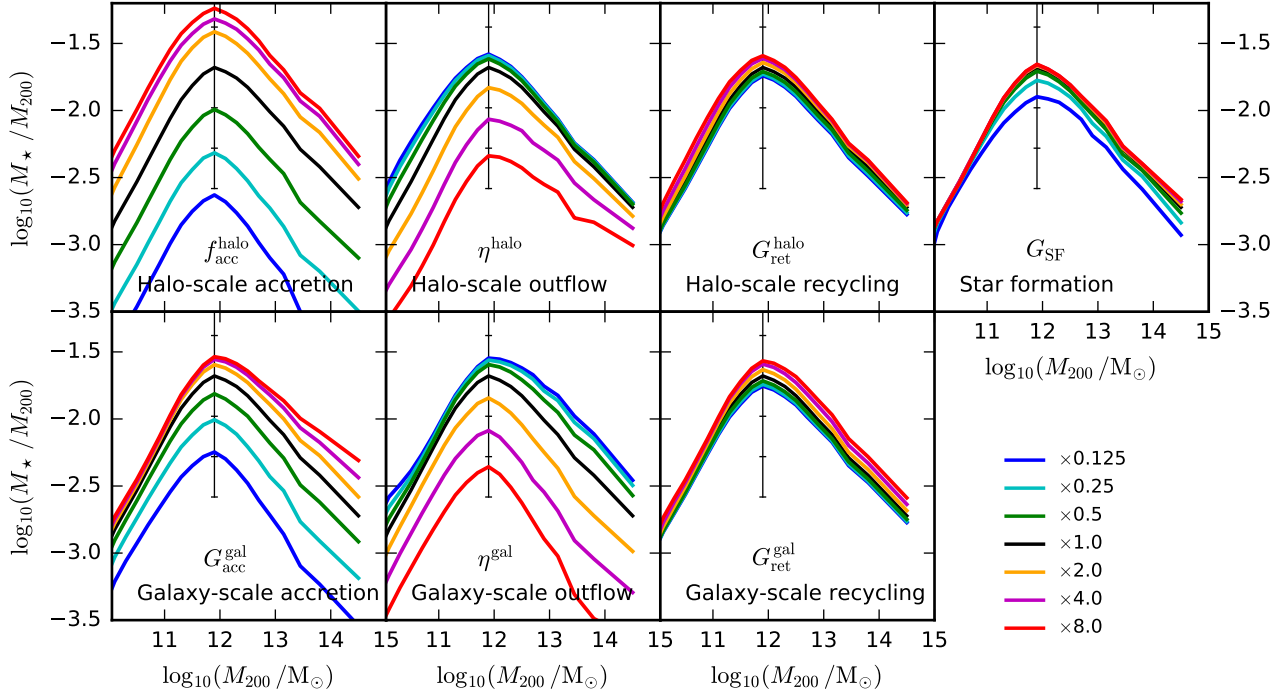


Figure 6. The impact of multiplying individual terms in our implementation of the N12 model by a constant (see Eqn. 1). The ratio of galaxy stellar mass to halo mass is plotted as a function of halo mass for central galaxies at $z = 0$. The black line shows the fiducial model, and coloured lines show the model with a single term multiplied by the factor indicated in the legend. Each panel corresponds to a different term being modified. The horizontal notches connected to vertical grey lines indicate how the SHM relation would change if there were a linear response to the multiplying constant. There are clear differences in the relative importance of the different terms in the model. For example, changing the outflow efficiency terms has a much larger effect than changing the star formation and recycling terms.

mass would scale linearly with changes to the model terms. This is approximately the case for the first-time gas accretion term ($f_{\text{acc}}^{\text{halo}}$, upper-left panel) since this acts as the source term (via the term $f_{\text{acc}} \frac{\Omega_b}{\Omega_m - \Omega_b} \dot{M}_{\text{DM}}^{\text{1st}}$ in the expression for \dot{M}_{CGM} , see Eqn. 1). For positive modifications to $f_{\text{acc}}^{\text{halo}}$, we choose to saturate $f_{\text{acc}}^{\text{halo}}$ at the maximum recorded value for a given redshift bin, in order to prevent a scenario where haloes contain significantly more than their share of the cosmic baryon fraction.

Modifying any other term in the model produces a clearly sub-linear response in the SHM relation. In addition, it is evidently easier to reduce the efficiency of the conversion of baryons into stars than it is to increase it; positive responses of the SHM relation to a given fractional change in the value of a term in Eqn. 1 are generally smaller (fractionally) in magnitude than the response to the corresponding fractional change that produces a negative response. At $M_{200} \sim 10^{12} M_{\odot}$, this can be partly explained as a saturation in the maximum possible conversion of baryons into stars; increasing the SHM normalization by a factor ten would result in over-shooting the cosmic baryon fraction ($f_b = 0.16$) at this mass scale. This is clearly not the entire explanation however, since (for example) no model variation produces a value of $\log_{10}(M_{\star}/M_{200}) > -3$ for $M_{200} \sim 10^{10} M_{\odot}$, which is well short of the cosmic baryon fraction.

4.1.1 Outflows

We consider first the two outflow terms (η^{gal} , η^{halo} , second column panels), which are the dimensionless mass loading factors de-

fined at the galaxy scale (gas leaving the ISM) and at the halo scale (gas moving beyond R_{200}). Increasing η^{gal} or η^{halo} produces the strongest response in the SHM relation, when compared to other model terms. The response is still sub-linear however: increasing η^{gal} by a factor two reduces galaxy stellar masses by $\approx 30\%$, a factor of 1.5. For the galaxy-scale outflow, the basic expectation is that $\dot{M}_{\star} \propto \frac{1}{1 + \eta^{\text{gal}}}$ (i.e. gas in the ISM either goes into stars or into an outflow). This only elicits a linear response if $\eta^{\text{gal}} \gg 1$, which is not the case over the entire evolution of a given galaxy (see Fig. 3). In addition, while increasing η^{gal} in isolation does reduce the star formation rate, the reduced star formation in turn decreases the outflow rate at the virial radius, which in turn will result in enhanced rates of first-time gas accretion onto the ISM.

This can be further generalised to consider gas accretion onto the ISM. In the regime where all of the gas in the CGM will be accreted onto the ISM within a Hubble time, or will be ejected, we can make the approximation that $\dot{M}_{\text{acc}}^{\text{halo}} \approx \dot{M}_{\text{acc}}^{\text{gal}} + \dot{M}_{\text{CGM,out}}$, where $\dot{M}_{\text{acc}}^{\text{halo}}$ is the gas accretion rate onto the halo, $\dot{M}_{\text{acc}}^{\text{gal}}$ is the gas accretion rate from the CGM onto the ISM, and $\dot{M}_{\text{CGM,out}}$ is the CGM outflow rate. If we apply a similar approximation to the ISM ($\dot{M}_{\text{acc}}^{\text{gal}} \approx (1 + \eta^{\text{gal}})\dot{M}_{\star}$), it follows that $\dot{M}_{\text{acc}}^{\text{halo}} \approx \dot{M}_{\star}(1 + \eta^{\text{gal}} + \eta^{\text{halo}})$, since halo-scale outflows are assumed to be proportional to the star formation rate in our model. If these approximations hold, we see that changing η^{halo} will only yield a linear response in \dot{M}_{\star} if $\eta^{\text{halo}} \gg 1 + \eta^{\text{gal}}$ (and similarly for η^{gal}). In other words, it does not matter for the SHM relation what value η^{halo} takes if the vast majority of the gas accreted onto galaxies is ejected in a galaxy-scale outflow, such that the star formation rate (and therefore the

halo outflow rate) is very small compared to the accretion rate onto the ISM. In EAGLE, $\eta^{\text{halo}} \gg 1 + \eta^{\text{gal}}$ is generally satisfied at low redshift, but not for $z > 1$ (Fig. 3).

Furthermore, reducing the values of either outflow term will essentially produce no response if the mass loading factors are $\ll 1$, because in that limit gas consumption rather than outflows limits the fuel for star formation. This is indeed seen in Fig. 6, where the SHM relation saturates to a constant at a given halo mass as η^{halo} and η^{gal} are reduced significantly below their fiducial values.

4.1.2 Star formation

For the model terms that are related to timescales ($G_{\text{acc}}^{\text{gal}}, G_{\text{SF}}, G_{\text{ret}}^{\text{gal}}, G_{\text{ret}}^{\text{halo}}$), their relative importance to galaxy stellar mass growth will generally depend on whether they represent a significant bottleneck relative to the Hubble time. For example, recall that $G_{\text{SF}} \equiv t \dot{M}_*/M_{\text{ISM}} \equiv t/\tau_{\text{SF}}$, where t is the age of the Universe and τ_{SF} is the conventional ISM depletion timescale. The exact value of G_{SF} is unimportant for galaxy star formation rates if $G_{\text{SF}} \gg 1$ (i.e. if the ISM depletion time is much shorter than the Hubble time), which is generally the case for high-mass galaxies (Fig. 3). Furthermore, the effective ISM depletion timescale is shortened by $1/(1 + \eta^{\text{gal}})$ due to galactic outflows, especially for low-mass galaxies where $\eta^{\text{gal}} \gg 1$. As such, Fig. 6 shows that the only case for which the value of G_{SF} has any impact on the SHM relation is when G_{SF} is reduced to $\approx 10\%$ of its fiducial value, at which point a given hydrogen atom accreted onto the halo would on average have to spend a significant fraction of a Hubble time in the ISM before being locked into a forming star.

4.1.3 Galaxy-scale accretion

On similar grounds, we expect that the exact value of the galaxy-scale accretion efficiency term $G_{\text{acc}}^{\text{gal}}$ would be unimportant if it is $\gg 1$. As a reminder, this term represents the efficiency with which circum-galactic gas that has not yet been part of the galaxy's ISM is able to accrete onto the ISM, defined relative to the Hubble time. In low-mass haloes ($M_{200} < 10^{12} M_{\odot}$) where the virial temperature is not high enough to push beyond the peak of the radiative cooling curve, the basic expectation is that gravitational timescales of order the halo dynamical time limit gas infall from the CGM to the ISM. Since the dynamical time scales as $1/\sqrt{\rho}$ and halo densities exceed the cosmic mean by $\sim 10^2$, this would give $G_{\text{infall}} \sim 10$. As discussed in Mitchell et al. (2020a), feedback processes in EAGLE inhibit gas infall, pushing $G_{\text{acc}}^{\text{gal}}$ to lower values (Fig. 3; see also, e.g., Nelson et al. 2015; Pandya et al. 2020), which can even be less than unity for low-mass galaxies at $z < 1$ (i.e. the average infall time is longer than the Hubble time at low redshift). For high-mass haloes, long characteristic radiative timescales plus AGN feedback push $G_{\text{acc}}^{\text{gal}}$ to very low values (one of the requirements for creating passive central galaxies).

As such, we expect that adjusting $G_{\text{acc}}^{\text{gal}}$ in Fig. 6 by a given factor will yield a larger response than adjusting G_{SF} , which is indeed the case. The largest response is seen for the most massive haloes, where the fiducial values of $G_{\text{acc}}^{\text{gal}}$ are lowest. For $M_{200} \sim 10^{12} M_{\odot}$, the effect of increasing $G_{\text{acc}}^{\text{gal}}$ quickly saturates as typical values of $G_{\text{acc}}^{\text{gal}}$ are generally much larger than unity. Decreasing $G_{\text{acc}}^{\text{gal}}$ produces a much larger fractional response, as expected.

4.1.4 Wind recycling

Fig. 6 shows that modifying either of the gas recycling efficiencies ($G_{\text{ret}}^{\text{gal}}, G_{\text{ret}}^{\text{halo}}$) elicits only a weak response in the SHM relation. For recycling timescales that are very long compared to the Hubble time ($G_{\text{ret}} \ll 1$), recycling becomes negligible for stellar mass growth, and the exact values of the recycling terms will be unimportant (see Eqn. 6); the same is true if the outflow ejection terms (η) are small. Fig. 3 shows that the galaxy-scale recycling term ($G_{\text{ret}}^{\text{gal}}$) is much less than unity at low redshift, but is marginally efficient (≈ 1) for $z \geq 1$, and indeed in Mitchell et al. (2020a) we show explicitly that galaxy-scale recycling in EAGLE plays a secondary but still important role for stellar mass assembly.

Decreasing the efficiency of gas recycling in our model is therefore expected to quickly saturate as recycling becomes negligible. Conversely, increasing $G_{\text{ret}}^{\text{gal}}$ will saturate if this results in $G_{\text{ret}}^{\text{gal}} \gg G_{\text{SF}}$, as in this case the star formation efficiency G_{SF} becomes the bottleneck, with recycled gas spending most of its time in the ISM rather than in the CGM. From Fig. 3, this will start to be the case for $z > 1.5$, but not at lower redshifts. In addition, while increasing $G_{\text{ret}}^{\text{gal}}$ does by definition allow more of the gas ejected from the ISM to participate in later star formation, this also depletes the CGM of galactic wind material. This in turn means that more of the CGM that was not ejected in a galactic wind will be ejected outside of the virial radius (i.e. preventative feedback for first-time infalling gas becomes more effective).

Fig. 3 shows that halo-scale recycling ($G_{\text{ret}}^{\text{halo}}$) becomes very inefficient in low-mass haloes ($M_{200} < 10^{12} M_{\odot}$) at low redshift. For high-mass haloes, gas accretion onto the ISM ($G_{\text{acc}}^{\text{gal}}, G_{\text{ret}}^{\text{gal}}$) instead becomes very inefficient, rendering the high efficiency of halo-scale recycling in this regime somewhat unimportant for stellar mass growth. In addition, increasing $G_{\text{ret}}^{\text{halo}}$ to arbitrarily large values will quickly saturate in effect since the galaxy-scale gas accretion timescales ($G_{\text{acc}}^{\text{gal}}, G_{\text{ret}}^{\text{gal}}$) are always at least comparable to the Hubble time for all halo masses, at which point the time spent outside R_{200} for a given ejected hydrogen atom is not an important bottleneck for stellar mass growth, compared to the time spent in the CGM afterwards. We have checked that if we artificially boost the efficiencies of gas accretion onto the ISM ($G_{\text{acc}}^{\text{gal}}, G_{\text{ret}}^{\text{gal}}$) to very high values (such that the time spent outside R_{200} is always the bottleneck), the SHM relation does respond strongly to the chosen value of $G_{\text{ret}}^{\text{halo}}$. Under physically sensible conditions, we therefore find that the exact efficiency of halo-scale recycling is not important for galaxy stellar mass growth (at least within the scenario for gas flows around galaxies presented by the EAGLE simulation).

4.1.5 Summary

Putting this all together, what is important (and what is not important) for setting the overall normalisation of the SHM relation, for a given amount of first-time cosmological gas accretion at the virial radius? In terms of which model terms produce the strongest response, the mass ejected from the ISM per unit star formation is important for all galaxies, as is the energy ejected from the ISM that goes into powering outflows at larger spatial scales (especially for $M_{200} < 10^{12} M_{\odot}$). The efficiency with which gas is accreted for the first time from the CGM onto the ISM is almost as important (especially for $M_{200} > 10^{12} M_{\odot}$). The efficiency of gas recycling is less important in comparison (but not negligible), and finally the chosen value of the star formation efficiency is inconsequential, unless it drops to very low values that correspond to gas consumption timescales similar to or greater than the age of the Universe.

On the other hand, we can also ask what is important for suppressing the maximum allowed value of M_*/M_{200} (set by $f_{\text{acc}}^{\text{halo}}$) down to the value seen for the fiducial model at $M_{200} \approx 10^{12} M_{\odot}$. Framed in this way, the two wind recycling terms are actually of comparable importance to the outflow terms and the galaxy-scale gas accretion term. This is because none of the terms (except for $f_{\text{acc}}^{\text{halo}}$) are capable of producing a strong positive response in terms of M_*/M_{200} . For outflows, this is mostly because varying one term (i.e., galaxy or halo-scale) in isolation is compensated for by the other term. We will show in the next subsection that varying both outflow terms at the same time can elicit a much larger positive response at lower halo masses ($M_{200} \ll 10^{12} M_{\odot}$).

4.2 What sets the shape of the SHM relation?

Fig. 7 shows the impact of removing the halo mass dependence of terms in the N12 model by fixing them to their values corresponding to the halo mass $M_{200} = 10^{12} M_{\odot}$. In doing so, we can assess the respective role of different terms in setting the shape of the SHM relation.

4.2.1 First-time gas accretion

The upper-left panel of Fig. 7 focuses on first-time inflowing gas, which is set in the N12 model by $f_{\text{acc}}^{\text{halo}}$ (first-time gas accretion at the virial radius) and $G_{\text{acc}}^{\text{gal}}$ (first-time infall from the CGM to the ISM). From Fig. 3, we can see that $f_{\text{acc}}^{\text{halo}}$ exhibits a shallow and positive dependence on halo mass, since feedback processes moderately inhibit first-time gas accretion onto low-mass haloes in EAGLE, and this slightly increases the first-time infall of gas onto more massive descendant haloes as a result (Wright et al. 2020; Mitchell et al. 2020a). As such, fixing $f_{\text{acc}}^{\text{halo}}$ to its value at $M_{200} = 10^{12} M_{\odot}$ (dashed blue line in Fig. 7) does not have a dramatic effect on the SHM relation, but does slightly increase the stellar mass in haloes with $M_{200} < 10^{12} M_{\odot}$. Interestingly, fixing $f_{\text{acc}}^{\text{halo}}$ has negligible impact on more massive haloes. This can be explained by the low galaxy accretion efficiencies ($G_{\text{acc}}^{\text{gal}}$, $G_{\text{ret}}^{\text{gal}}$) measured for these haloes, such that star formation that occurred in low-mass progenitor galaxies (that later merge) dominates the final stellar mass in the central galaxies of very massive haloes⁴. In other words, most of the stars in dark matter haloes with $M_{200} \gg 10^{12} M_{\odot}$ are formed from gas that was accreted for the first time onto less massive progenitor haloes with $M_{200} \sim 10^{12} M_{\odot}$ at earlier times.

As a mirror image to fixing $f_{\text{acc}}^{\text{halo}}$, fixing $G_{\text{acc}}^{\text{gal}}$ (dotted blue line in the top-left panel of Fig. 7) does not impact low-mass haloes, but has the strongest impact of any of the model terms on the stellar mass in haloes with $M_{200} > 10^{12} M_{\odot}$. Fig. 3 shows that for $M_{200} < 10^{12} M_{\odot}$, $G_{\text{acc}}^{\text{gal}}$ exhibits barely any dependence on halo mass (which can be explained if gas infall in low-mass haloes is, as expected, connected to gravitational timescales, which are scale-free and therefore independent of halo mass; Mitchell et al. 2020a), and as such does not affect the shape of the SHM relation in this regime. For $M_{200} > 10^{12} M_{\odot}$, $G_{\text{acc}}^{\text{gal}}$ decreases sharply with increasing halo mass as longer radiative cooling times and AGN feedback take effect, and as such fixing $G_{\text{acc}}^{\text{gal}}$ to its value for $M_{200} = 10^{12} M_{\odot}$ significantly increases galaxy stellar masses in high-mass haloes. Fixing both $G_{\text{acc}}^{\text{gal}}$ and $f_{\text{acc}}^{\text{halo}}$ simultaneously

(solid blue line) has little additional impact, since the halo mass dependence of the two terms affect the SHM relation in different regimes.

4.2.2 Outflows

Removing the halo mass dependence of the mass loading factors for galaxy or halo-scale outflows (η^{gal} and η^{halo} , respectively the dotted and dashed red lines in the upper-right panel of Fig. 7) impacts the shape of the SHM relation for $M_{200} < 10^{12} M_{\odot}$, which is expected since both are negative functions of halo mass in this range (see Fig. 3). The halo mass dependence of the mass loading factors is steeper for halo-scale outflows (η^{halo}), and so accordingly fixing this to its value at $M_{200} = 10^{12} M_{\odot}$ has a larger effect. Interestingly, fixing both η^{gal} and η^{halo} simultaneously (solid red line) has a much larger effect than fixing one of the two, for $M_{200} < 10^{12} M_{\odot}$. As discussed in Section 4.1, decreasing one of η^{gal} or η^{halo} in isolation actually has the effect of increasing the importance of the other term (for example, decreasing η^{gal} in isolation increases \dot{M}_* , which in turn increases the halo-scale outflow rate, which in turn will decrease the rate of gas infall onto the ISM). For $M_{200} > 10^{12} M_{\odot}$, fixing either or both of η^{gal} and η^{halo} has little effect. η^{gal} depends only weakly on halo mass in this range; η^{halo} has a stronger mass dependence but, as discussed earlier, in this halo mass range gas infall from the CGM onto the ISM is extremely inefficient (and in any case halo-scale recycling is very efficient in this regime), rendering halo-scale outflows irrelevant for stellar mass growth.

4.2.3 Recycling

Removing the halo mass dependence of the efficiencies of galaxy or halo-scale gas recycling ($G_{\text{ret}}^{\text{gal}}$ and $G_{\text{ret}}^{\text{halo}}$, respectively the dotted and dashed green lines in the bottom-left panel of Fig. 7) impacts the SHM relation for $M_{200} < 10^{12} M_{\odot}$, but not at higher halo masses. In the low-halo mass regime, both the galaxy- and halo-scale recycling efficiencies increase with halo mass (see Fig. 3), such that fixing them to their values at $M_{200} = 10^{12} M_{\odot}$ increases galaxy stellar masses, softening the slope of the SHM relation.

4.2.4 Summary & discussion

Putting all these components together, the lower-right panel of Fig. 7 shows respectively how the mass dependencies of outflows, (first-time) accretion, and recycling shape the SHM relation in EAGLE. The yellow line shows the effect of fixing all of these terms simultaneously to their values at $M_{200} = 10^{12} M_{\odot}$, such that there is effectively no explicit halo-mass dependence in the N12 model. Accordingly, the SHM relation becomes nearly flat, at least for $M_{200} < 10^{13} M_{\odot}$. The SHM relation still drops at higher halo masses, in small part because our model variations do not affect star formation within satellite galaxies (see Fig. 2), and in larger part because massive haloes are less relaxed and therefore contain a larger mass fraction within satellite subhaloes (and because these satellites are better resolved in more massive haloes).

This is usually of secondary importance for galaxy stellar masses, since low-mass satellite galaxies ($M_* \ll 10^{10} M_{\odot}$) are extremely inefficient at forming stars, meaning that satellites only contribute significantly to the total stellar mass within R_{200} for haloes that are massive enough to in turn host very massive satellite galaxies, as seen in Fig. 1. For the model considered here, if all

⁴ In addition, star formation that occurs in satellites is unaffected by our modifications to terms in the N12 model, which is a $\approx 30\%$ effect in massive haloes, see Fig. 2.

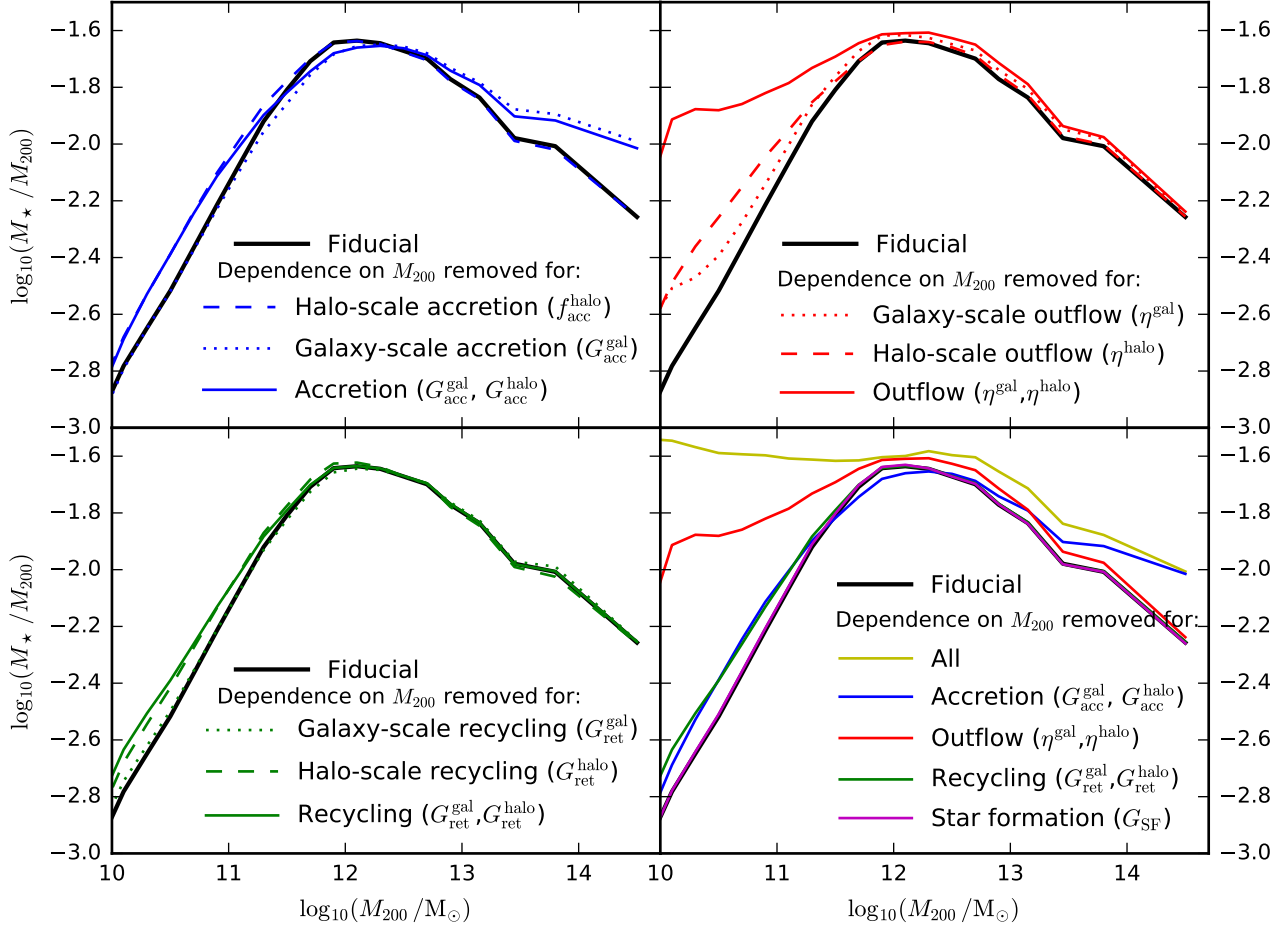


Figure 7. The impact of removing the halo mass dependence of one (or more) term(s) in the N12 model (see Eqn. 1), by fixing its value(s) to the one(s) corresponding to $M_{200} = 10^{12} M_{\odot}$ for each redshift bin. The median ratio of galaxy stellar mass to halo mass is plotted as a function of halo mass for central galaxies at $z = 0$. Black lines show the fiducial model and coloured lines show the model with the term indicated in each panel held fixed. The “All” case (yellow line, bottom-right panel) indicates that all of the considered modifications are applied (i.e. $f_{\text{acc}}^{\text{halo}}$, $G_{\text{acc}}^{\text{gal}}$, G_{SF} , $\eta_{\text{gal}}^{\text{gal}}$, $\eta_{\text{halo}}^{\text{halo}}$, $G_{\text{ret}}^{\text{gal}}$, $G_{\text{ret}}^{\text{halo}}$ are all fixed at their respective values at $M_{200} = 10^{12} M_{\odot}$). These modifications indicate how different processes in EAGLE shape the dependence of galaxy stellar mass on halo mass. Roughly speaking, the mass dependence of outflows plays the largest role for $M_{200} < 10^{12} M_{\odot}$, but halo-scale gas accretion (i.e. preventative feedback at R_{200}) and halo-scale gas recycling also contribute. The mass dependence of galaxy-scale gas accretion (i.e., the efficiency of first-time CGM infall onto the ISM) plays the largest role in shaping the SHM relation for $M_{200} > 10^{12} M_{\odot}$, but outflows also contribute.

of the N12 model terms are fixed, then low-mass satellites are dramatically more efficient in forming stars, such that the higher mass fraction in sub-structures found in group and cluster-mass haloes is instead the main reason for why there is relatively less stellar mass in the central subhalo. This point is further explained in Appendix B.

Relative to this baseline scenario (and relative to our fiducial implementation of the N12 model), the mass dependencies of the outflow terms (red) are clearly the most important for shaping the low-mass slope of the SHM relation, but still only account for about one half of the slope in this regime, with the first-time accretion (blue) and recycling (green) terms clearly all playing a role. The high-mass slope of the SHM relation is set mostly by the mass dependence of the first-time accretion terms (specifically gas accretion onto the ISM, $G_{\text{acc}}^{\text{gal}}$), with the outflow terms playing a secondary role. Newly in this panel, we also show the effect of fixing the star formation efficiency term (G_{SF} , purple), which has no discernible effect.

Finally, if we compare to the simplified picture set out in

Eqn. 6 and Fig. 4 of Section 3 (note that a flat line in Fig. 4 means the corresponding process does not affect the shape of the SHM relation, whereas the opposite is true for a flat line in Fig. 7), we see that the full N12 model has more or less borne out our basic expectations. Naturally, the full model is much more successful at reproducing the EAGLE SHM relation because it accounts for the full redshift evolution of different processes at a given halo mass, and also accounts for the effects of galaxy and subhalo mergers. Compared to Fig. 4, the full N12 model shows that galaxy-scale recycling is indeed a secondary (but not negligible) effect for the low-mass slope of the SHM relation. At higher masses, Fig. 4 implies that galaxy-scale recycling should also be mildly important, which is not the case in the full N12 model. This is (in part) because in deriving Eqn. 6 we made the approximation that $\eta_{\text{gal}}^{\text{gal}} \gg 1$, which is not the case for high-mass galaxies (see Fig. 3). Relative to Eqn. 6, the full N12 model separates out the “preventative feedback” term (f_{prev}) into halo-scale accretion ($f_{\text{acc}}^{\text{halo}}$), outflows ($\eta_{\text{halo}}^{\text{halo}}$) and recycling ($G_{\text{ret}}^{\text{halo}}$), and also galaxy-scale accretion ($G_{\text{acc}}^{\text{gal}}$). With these terms fully separated, we can now appreciate that it is the halo mass

dependence of the halo-scale outflows that provides the most important contribution to f_{prev} for the low-mass SHM slope, and that the mass dependence of the first-time galaxy-scale infall efficiency term ($G_{\text{acc}}^{\text{gal}}$) is most important for the high-mass SHM slope.

4.3 What sets the redshift evolution of the SHM relation?

To explore how (and which) physical processes in EAGLE shape the evolution of the SHM relation, we fix a single term in the N12 model to its value at $z = 1$ (but keeping the halo mass dependence), thereby removing any redshift dependence for that term. Note that for the terms that are related to timescales ($G_{\text{acc}}^{\text{gal}}$, G_{SF} , $G_{\text{ret}}^{\text{gal}}$, $G_{\text{ret}}^{\text{halo}}$), the associated timescale is therefore still assumed to scale with the Hubble time. For example, the star formation efficiency within the ISM ($G_{\text{SF}} \equiv t M_{\star} / M_{\text{ISM}} \equiv t / \tau_{\text{SF}}$) is fixed such that the ISM depletion time τ_{SF} scales with the Hubble time, and so is shorter at higher redshift.

Fig. 8 shows the results of this exercise, comparing EAGLE (lower-middle-left panel) to different variations of the N12 model. The fiducial model (being approximate in nature) slightly overpredicts the evolution at fixed halo mass compared to EAGLE, particularly for $M_{200} \sim 10^{11} M_{\odot}$, but is nonetheless a reasonable representation of the underlying simulation. In the base EAGLE simulation, stellar masses increase at fixed halo mass with decreasing redshift, at least for $z < 3$ and for $M_{200} \gtrsim 10^{11} M_{\odot}$, with the clearest redshift evolution seen for $M_{200} \sim 10^{12} M_{\odot}$. This could imply that low-redshift haloes are more efficient at forming stars at fixed halo mass, but we will show presently that this interpretation is flawed.

Stepping through the various model terms in turn, we see that fixing either G_{SF} (star formation efficiency within the ISM) or η^{gal} (galaxy-scale outflows) reduces the level of SHM evolution. Conversely, fixing either $G_{\text{acc}}^{\text{gal}}$ (first-time infall from the CGM to ISM) or η^{halo} (halo-scale outflows) increases the level of SHM evolution. This can be visually appreciated by comparing to the grey error bars, which show the range in evolution for different halo masses in the fiducial model. The interpretation of these trends is straightforward, as visual inspection of Fig. 3 shows that these four model terms evolve at fixed halo mass in a manner that is consistent with the trends seen in Fig. 8. Halo-scale outflows become more efficient at lower redshift, but galaxy-scale outflows are more efficient at higher redshift. Star formation in the ISM is more efficient (relative to the Hubble time) at lower redshift, whereas first-time gas accretion onto galaxies is more efficient at higher redshift. We note from Fig. 3 that the efficiencies of galaxy and halo-scale gas recycling also vary comparably with redshift (and in opposite directions relative to each other), but we find that fixing these terms has negligible effect on the evolution of the SHM relation (probably because of their relatively weak impact, see Fig. 6), and so we omit them from Fig. 8 to save space.

Interestingly, these various effects conspire to nearly cancel out for the model variant where we simultaneously fix all of the model terms to their values at $z = 1$ (“All”, lower-far-right panel), with this variant exhibiting slightly more redshift evolution than the fiducial model. This is at first glance puzzling, since if the efficiency of all terms are constant with redshift, there is no reason for there to be any evolution in the SHM relation. The explanation for this paradox is shown in the left panel of Fig. 9, which shows the evolution of the EAGLE SHM relation, but replaces M_{200} with $M_{\text{DM}}^{\text{1st}}$, the cumulative mass of dark matter accreted (for the first time) onto progenitors of the central subhalo. As a reminder, the time derivative of this quantity acts as the source term for the model; i.e., we

assume that gas being accreted for the first time at the virial radius traces the rate with which dark matter is being accreted for the first time.

Fig. 9 shows that when compared to the conventional SHM definition using M_{200} (right panel), the SHM relation defined instead with $M_{\text{DM}}^{\text{1st}}$ (left panel) exhibits almost no redshift evolution at fixed halo mass, aside from for low-mass haloes that are poorly resolved. M_{200} differs from $M_{\text{DM}}^{\text{1st}}$ in several important respects. First, M_{200} includes the mass within satellite subhaloes, whereas we show $M_{\text{DM}}^{\text{1st}}$ for the central subhalo only (note that, following convention, M_{\star} is also shown for the central subhalo only). At fixed host halo mass, the fraction of total mass in satellite subhaloes increases with increasing redshift in EAGLE. In addition, M_{200} also includes the total mass that has been stripped from satellite subhaloes ($M_{\text{DM}}^{\text{1st}}$ does not). As discussed in Appendix B, we generally expect galaxy stellar masses to better trace $M_{\text{DM}}^{\text{1st}}$ than the instantaneous mass of a subhalo after stripping, since dark matter is much more readily stripped from a satellite subhalo compared to stars.

A second factor is that significant amounts of accreted dark matter will move back out beyond R_{200} , defining the so-called “splashback” radius at larger spatial scales (e.g., Diemer 2017). Some of this ejected dark matter will then be re-accreted at later times, actually comprising the majority of the total dark matter accretion at late times (e.g., Wright et al. 2020). $M_{\text{DM}}^{\text{1st}}$ is by definition insensitive to this process, but M_{200} is affected. As a third and final consideration, M_{200} includes the contribution of baryons, and the baryon fraction evolves in EAGLE at fixed halo mass.

Putting these effects together, it is possible for the SHM relation to evolve even if the galaxy formation efficiency were to remain constant with time, due to the evolution of M_{200} for a given halo. Stars do form continuously in the EAGLE simulation, but the evolution of the SHM relation is mostly set by the evolution of M_{200} , and not by an evolving efficiency with which haloes convert gas into stars. While difficult to test observationally, the same may well hold for the SHM relation in the real Universe, when using the conventional halo mass definition of M_{200} (or similar definitions). We note that Behroozi et al. (2019) use observations to infer similar evolution of the SHM relation to EAGLE, in that they find the normalization increases as a function of cosmic time for halo masses $\sim 10^{12} M_{\odot}$ until $z = 1$, and is approximately constant afterwards for $0 < z < 1$. We have verified that the same is true in EAGLE if we use their halo mass definition (which is taken from Bryan & Norman 1998).

5 WHAT SETS THE RELATIONSHIP BETWEEN THE MASS OF THE ISM, CGM, AND HALO MASS?

While our primary focus has been the SHM relation, we now briefly extend our analysis to consider how galactic star formation, inflows, and outflows affect the relationship between halo mass and the masses of the ISM and CGM. Fig. 10 and Fig. 11 show respectively how the CGM and ISM masses at $z = 0$ respond to multiplying individual terms in our model by a constant. These figures are analogous to Fig. 6, which shows the corresponding information for galaxy stellar masses.

Starting with the CGM in Fig. 10, it is apparent that the CGM is in general more sensitive to the chosen model parameters, when compared to the mass in stars (Fig. 6). This is true for all of the model terms except for the first time halo-scale gas accretion term ($f_{\text{acc}}^{\text{halo}}$, upper-far-left panel), since this term is directly proportional to the source term in Eqn. 1. In some cases, the mass in the CGM

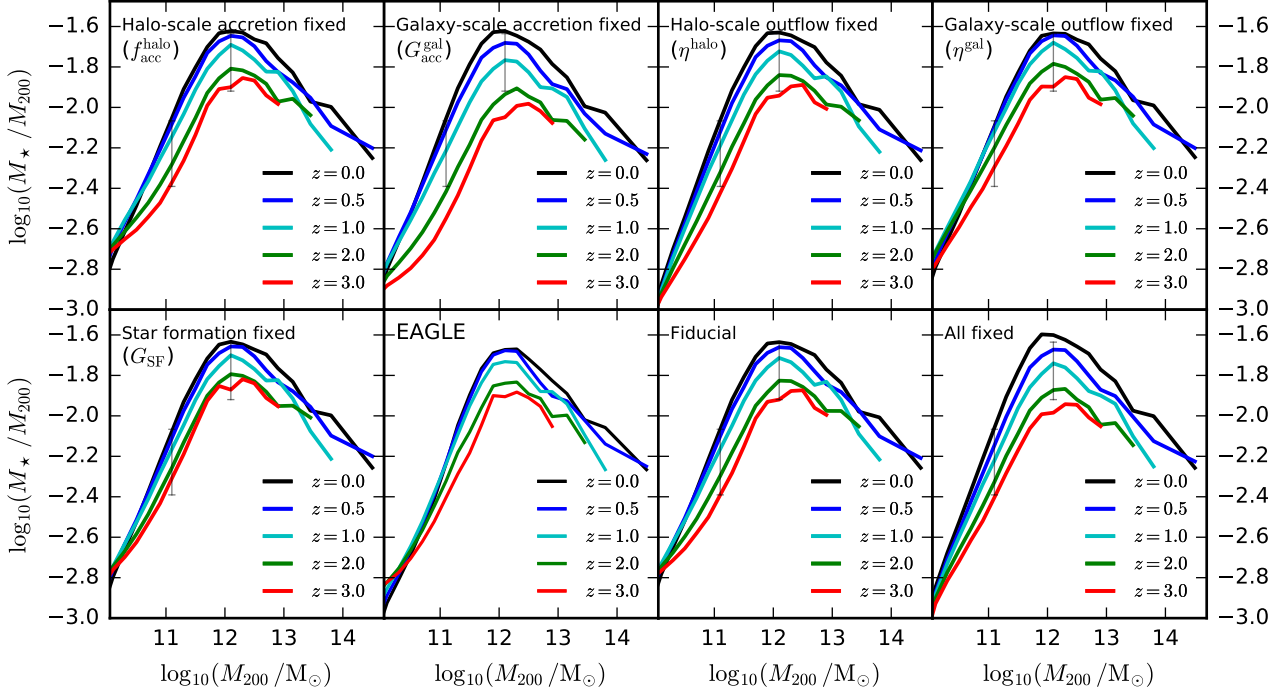


Figure 8. The impact of removing the redshift dependence of one (or more) term(s) in the N12 model by fixing its value(s) to the one(s) corresponding to $z = 1$. Note that if we hold the terms related to timescales ($G_{\text{acc}}^{\text{gal}}$, G_{SF} , $G_{\text{ret}}^{\text{gal}}$, $G_{\text{ret}}^{\text{halo}}$) constant at fixed halo mass, then the associated timescales still scale with the Hubble time. The median ratio of galaxy stellar mass to halo mass is plotted as a function of halo mass for central galaxies. Different line colours correspond to different redshifts. The lower-middle-left panel shows the evolution of the SHM relation in EAGLE, and the lower-middle-right panel shows the corresponding evolution in the N12 model. The lower-far-right (“All”) panel shows the model when each of the parameters $f_{\text{acc}}^{\text{halo}}$, $G_{\text{acc}}^{\text{gal}}$, G_{SF} , η^{gal} , η^{halo} , $G_{\text{ret}}^{\text{gal}}$, $G_{\text{ret}}^{\text{halo}}$ are held fixed to their values at $z = 1$. Other panels show the model when only one of these terms is fixed to its value at $z = 1$. To save space, we do not show panels for the recycling terms as they are indistinguishable from the fiducial model. As a reference point, the grey error bars show the evolution of the SHM relation for the fiducial N12 model at $M_{200} = 10^{11}$ and $10^{12} M_{\odot}$, over the range $0 < z < 3$. Most of the evolution of the SHM relation in EAGLE seems to be set by the evolution of the star formation efficiency (G_{SF}), galaxy-scale gas accretion ($G_{\text{acc}}^{\text{gal}}$, i.e., the efficiency of first-time infall from the CGM to the ISM), galaxy-scale outflows (η^{gal}), and halo-scale outflows (η^{halo}).

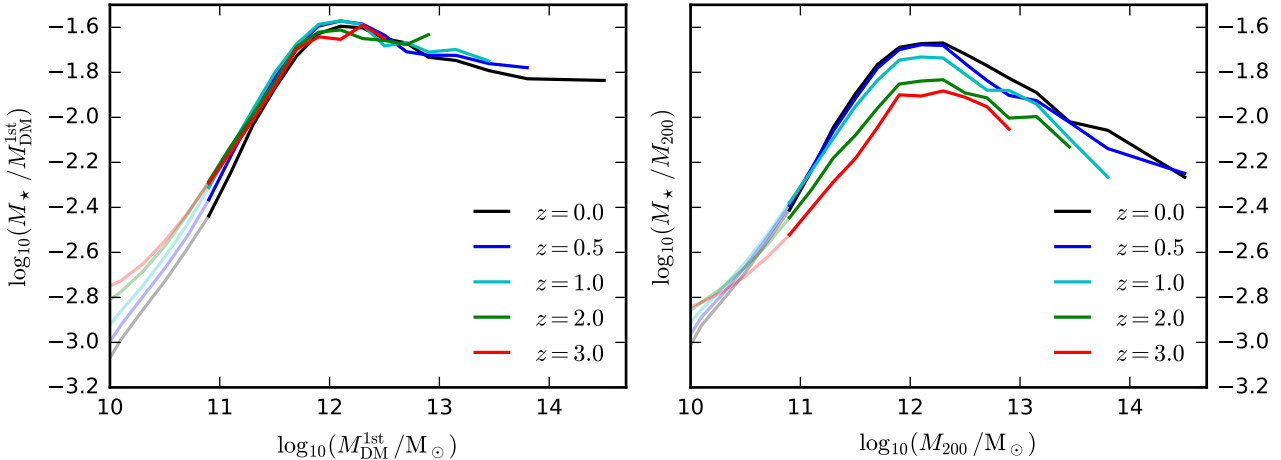


Figure 9. *Left:* the evolution of the EAGLE SHM relation for central galaxies, but replacing M_{200} with $M_{\text{DM}}^{\text{1st}}$, the cumulative mass of dark matter accreted for the first time onto progenitors of the central subhalo. *Right:* the same evolution of the SHM relation but using M_{200} as the halo mass variable, to act as a visual reference. For each line, lighter shades indicate the halo mass range for which galaxies contain on average fewer than 100 star particles. Compared to the conventional SHM definition shown in the right panel, the ratio of $M_{\star}/M_{\text{DM}}^{\text{1st}}$ exhibits almost no redshift evolution at fixed halo mass for $M_{\text{DM}}^{\text{1st}} > 10^{11} M_{\odot}$. In general, it is expected that M_{\star} should trace $M_{\text{DM}}^{\text{1st}}$ more closely than M_{200} (see main text). It follows therefore that there is overall very little evolution in the efficiency of galaxy formation with cosmic time in EAGLE, given the amount of expected baryonic accretion onto a halo of a given mass.

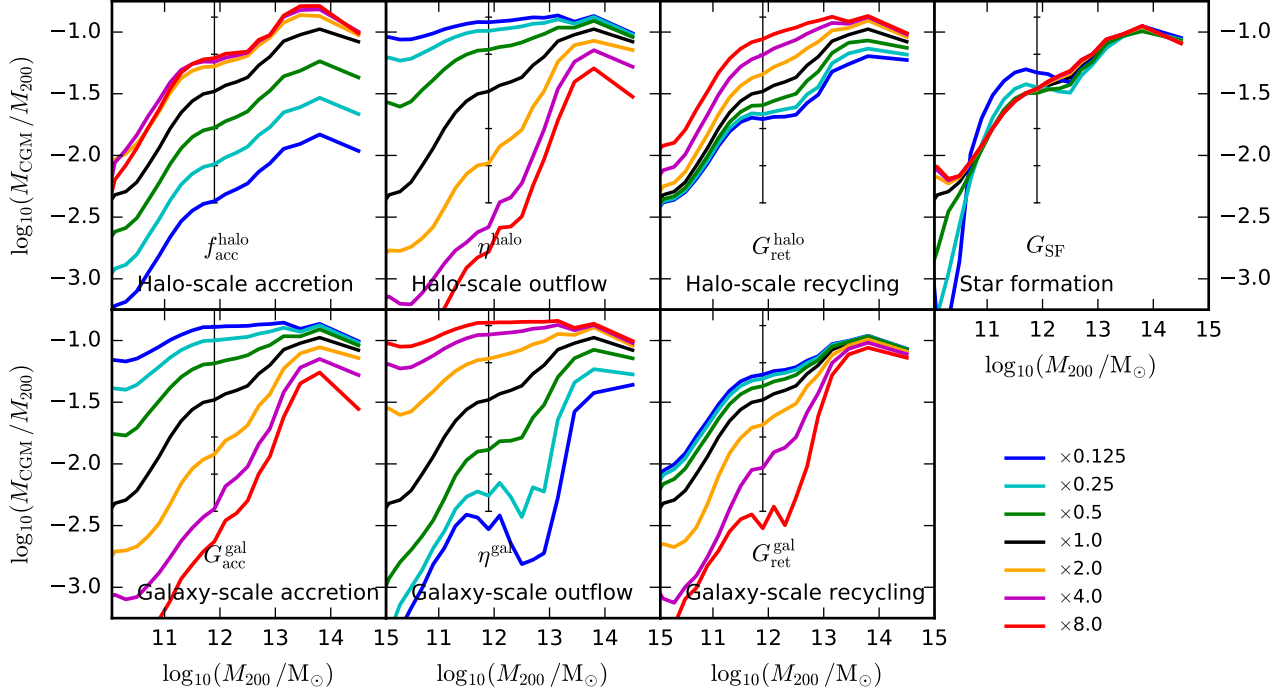


Figure 10. The impact on the mass of the CGM of multiplying individual terms in our model (see Eqn. 1) by a constant. Masses are plotted for central galaxies at $z = 0$. The figure format follows Fig. 6, replacing galaxy stellar mass with CGM mass on the y-axis. Compared to Fig. 6, it is apparent that the CGM is more sensitive to a wider range of the model terms (e.g. galactic-scale recycling).

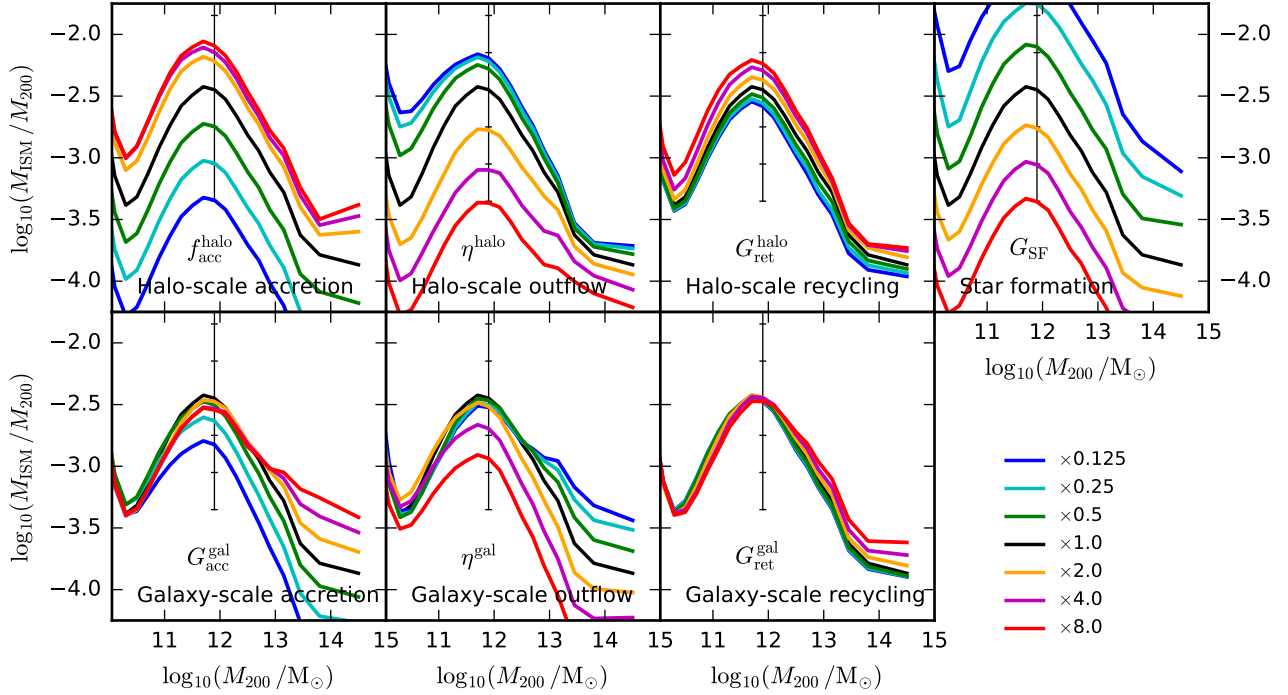


Figure 11. The same as Fig. 10, but in this case showing the ISM mass (again at $z = 0$). The ISM (unlike stars and the CGM) is very sensitive to the star formation efficiency term (G_{SF} , upper-far-right panel). The mass in the ISM is also slightly more sensitive to halo-scale outflows and recycling than the mass in stars is, but is slightly less sensitive to galaxy-scale outflows and recycling. This is because the ISM mass traces recent gas accretion and outflow activity, whereas galaxy stellar masses reflect the full time-integrated activity.

even responds super-linearly to changes in model parameters: this is the case for halo-scale outflows (η^{halo} , upper-middle-left panel), for first time galaxy-scale accretion ($G_{\text{acc}}^{\text{gal}}$, lower-far-left panel), and for galaxy-scale outflows (η^{gal} , lower-middle panel).

It is generally accepted that the CGM should be a sensitive tracer of gas flows and feedback processes, since many of the metals produced in stars are believed to be ejected from galaxies (e.g., Peebles et al. 2014), and because galaxy properties (ISM and stars) are often degenerate with respect to competing scenarios for galactic inflows, outflows and recycling (e.g., Mitchell et al. 2014, 2020b; Pandya et al. 2020). Fig. 10 bears out this expectation. Comparing Fig. 10 with Fig. 6, it is particularly noteworthy that the mass in the CGM is much more sensitive to both galaxy- and halo-scale recycling than the mass in stars is. While halo-scale recycling boosts the CGM mass at all halo masses, galaxy-scale recycling suppresses the CGM for $M_{200} < 10^{13} M_{\odot}$ and has little effect for higher halo masses.

Considering instead the mass of the ISM in Fig. 11, the most obvious difference is that the ISM mass depends negatively and approximately linearly with the value of the star formation efficiency term ($G_{\text{SF}} \equiv \dot{M}_{\star} t / M_{\text{ISM}}$, upper-far-right panel). In contrast, the mass in stars and in the CGM depends only very weakly on this parameter. This is a well known result that galaxy formation is self-regulated by feedback, in the sense that star formation rates and outflow rates (and therefore the masses of stars and the CGM) respond such that they roughly balance the galaxy-scale accretion rate (e.g., Finlator & Davé 2008; Schaye et al. 2010; Davé et al. 2012; Lilly et al. 2013; Sharma & Theuns 2020). This renders G_{SF} inconsequential for star formation rates and outflow rates unless $G_{\text{SF}} < 1$ (at which point gas consumption in the ISM is an important bottleneck). This self-regulation is achieved as a result of the ISM mass increasing or decreasing in response to the balance between accretion and the efficiencies of outflows and star formation. As such, a shorter gas consumption time scale requires less ISM to produce the star formation rate that is required for outflows plus star formation to balance inflows, meaning that the ISM mass is very sensitive to the star formation efficiency (Schaye et al. 2010; Haas et al. 2013).

Comparing the ISM mass shown in Fig. 11 to the stellar mass shown in Fig. 6, it is also notable that the ISM at $z = 0$ is more sensitive to halo-scale outflows and recycling (η^{halo} , $G_{\text{ret}}^{\text{halo}}$), whereas the stellar mass is comparatively more sensitive to galaxy-scale outflows and recycling (η^{gal} , $G_{\text{ret}}^{\text{gal}}$). Gas consumption times in the ISM are generally shorter than the age of the Universe in EAGLE (i.e., $G_{\text{SF}} > 1$, see Fig. 3), and the effective consumption time is even shorter once galaxy-scale outflows are accounted for. This means that the ISM at a given redshift is only sensitive to recent accretion, star formation, and outflow activity. In contrast, galaxy stellar masses reflect the integrated activity over the entire history of a galaxy. From Fig. 3, we see that halo-scale outflows are more efficient at low redshift in EAGLE, whereas the efficiency of galaxy-scale outflows increases with increasing redshift. This means in turn that the $z = 0$ ISM mass is more sensitive to η^{halo} , whereas the $z = 0$ stellar mass is (comparatively) more sensitive to η^{gal} , as many of the stars in galaxies at $z = 0$ were formed for $z > 1$.

6 SUMMARY

The median SHM relation, $\frac{M_{\star}}{M_{200}}(M_{200}, z)$, is a fundamental diagnostic of the efficiency of galaxy formation within the context of the working cosmological model. We have introduced a novel method

to explain how the SHM relation predicted by hydrodynamical simulations is shaped, disentangling the relative contributions of gas consumption by star formation, first-time inflows (i.e. accretion of gas that has not been ejected in the past), gas ejection due to outflows, and wind recycling. We used measurements of gas flows in the EAGLE simulations from Mitchell et al. (2020a,b), averaging as a function of halo mass and redshift. These measurements then provide coefficients in a set of coupled linear differential equations for the evolution of the mass fractions in stars, ISM, CGM, and gas that is ejected from the halo, with the first time gas accretion at the virial radius acting as the source term. By integrating these equations along the merger tree of each individual galaxy in the EAGLE simulation, we constructed a model that then computes the mass fractions for each individual EAGLE galaxy. By inspection of the model coefficients, and by modifying the coefficients by hand, we built an understanding of what shapes the SHM relation and its evolution with redshift.

When modifying individual model terms by a constant multiplicative factor, we find that the normalization of the SHM relation is most sensitive to the efficiencies of first-time gas accretion and ejection by outflows, and is less sensitive to the efficiency of wind recycling, and of gas consumption by star formation (Fig. 6). Cosmological first-time gas accretion at the halo virial radius sets the main boundary condition (the maximum possible value of $\dot{M}_{\star} / M_{200}$). For a fixed first-time halo accretion rate, \dot{M}_{\star} is suppressed by galaxy-scale outflows (for all masses), halo-scale outflows (mostly for $M_{200} < 10^{12} M_{\odot}$), and by the finite efficiency of galaxy-scale gas accretion (including the effect of preventative feedback, mostly for $M_{200} > 10^{12} M_{\odot}$). For a fixed first-time halo accretion rate, \dot{M}_{\star} is increased by gas recycling, although the effect is smaller than that of outflows. The efficiency of gas consumption in the ISM via star formation would have to be reduced dramatically for it to become a bottleneck, and therefore has little effect on the $z = 0$ SHM relation in EAGLE.

Our first main result is that the shape of the SHM relation is driven primarily by the ejection of gas via outflows for $M_{200} < 10^{12} M_{\odot}$, and by the (in)efficiency of first-time gas inflow from the CGM to the ISM in more massive haloes (Fig. 7), although the drop in $\dot{M}_{\star} / M_{200}$ for $M_{200} > 10^{12} M_{\odot}$ is actually driven mostly by the conventional choice to exclude satellite galaxies and the diffuse stellar halo from the stellar mass used in SHM relation (Fig. 1). For $M_{200} < 10^{12} M_{\odot}$, the steep $\dot{M}_{\star} / M_{200} \propto M_{200}$ slope of the SHM relation is shaped by the ejection of gas via outflows at both the galaxy (i.e., gas ejected from the ISM) and halo (i.e., gas ejected beyond the virial radius) scales. Interestingly, including the halo mass dependence of either the galaxy or halo-scale gas ejection terms in isolation is sufficient to explain most of the total effect of ejection on the SHM shape. Including the halo mass dependencies of both galaxy- and halo-scale outflow rates together has only a weak additional effect, since more efficient galaxy-scale gas ejection reduces the gas ejection rate at the virial radius (due to reduced star formation rates), and vice versa. In addition, for $M_{200} < 10^{12} M_{\odot}$ secondary roles are played by preventative feedback (i.e., a reduction in the first-time inflow of gas) at the scale of the halo virial radius, and by halo-scale recycling of ejected gas.

Our second main result is that the redshift evolution of the SHM relation (which drops with increasing redshift at fixed halo mass for $z > 1$ in EAGLE) is driven mostly by the evolution of halo properties, rather than any change in the combined net efficiency of inflows, star formation, outflows, and recycling. Specifically, the baryon fraction within R_{200} (i.e., backsplash), the mass fraction in satellite subhaloes (which contribute to M_{200} but not to

M_* , by convention), the ejection and re-accretion of dark matter at R_{200} , and satellite subhalo mass loss (baryons plus dark matter) all evolve with redshift at fixed halo mass, driving in turn the redshift evolution of the SHM relation (Fig. 9). In contrast, when, instead of M_{200} , we use the cumulative mass of dark matter that is accreted for the first time, then the SHM relation does not evolve.

Finally, we also briefly examined how star formation and gas flows affect the relationship between halo mass, and the masses of both the ISM and the CGM. Of the three baryonic components inside haloes (stars, ISM, CGM), the mass of the CGM is generally the most sensitive to the effects of inflows and outflows (Fig. 10). Notably, gas recycling (at both the halo and galaxy scales) has a much larger impact on the CGM than on the stars (Fig. 6) or on the ISM (Fig. 11). The ISM is the only reservoir that is sensitive to the ISM gas consumption efficiency via star formation (Fig. 11), reflecting the well-known result that galaxies self-regulate by increasing or decreasing star formation and outflow rates to balance changes in the galaxy-scale gas accretion rate. Compared to the mass in stars, the ISM mass at $z = 0$ is more sensitive to halo-scale gas flows (Fig. 11), whereas the stellar mass is comparatively more sensitive to galaxy-scale gas flows (Fig. 6). This is because the ISM traces only recent accretion, star formation and outflow activity, combined with the fact that galaxy-scale outflows are more efficient at high redshift in EAGLE, whereas halo-scale outflows are more efficient at low redshift (Fig. 3).

In conclusion, we have estimated the impact of different types of gas flows on the SHM relation and on the mass fractions of the ISM and CGM in the EAGLE hydrodynamical simulation. By varying the values of individual terms of a model that by design reproduces the gas flows measured for EAGLE, we investigated the effects of varying the efficiency of star formation in the ISM and the rates of, respectively, the first time gas accretion, outflows and recycling, where we distinguished gas flows on galaxy and halo scales. While cosmological first-time accretion onto the halo sets the overall amplitude of the SHM relation, its shape is modified mainly by outflows, with halo-scale outflows being relatively more important at lower masses. The outflows do not only have a direct impact on the stellar mass, they also affect it indirectly by suppressing gas accretion onto galaxies, especially for more massive objects. Recycling, either on galaxy or halo scales, has only a minor effect on the SHM relation in EAGLE, though it does have a large impact on the CGM. We also found that the evolution of the SHM relation nearly vanishes when we plot the ratio of the stellar mass to the cumulative mass of dark matter that fell in for the first time, as opposed to plotting the ratio with the total halo mass. This suggests that the evolution of the SHM relation is set by the definitions of the galaxy and halo, rather than by evolution in the true efficiency of galaxy formation.

It would be of considerable interest to apply our methodology to another state-of-the-art hydrodynamical simulation that provides a good match to the SHM relation inferred from observations. This would reveal how generic our conclusions are.

ACKNOWLEDGEMENTS

We thank Eyal Neistein for sharing the source code used for the original implementation of his modelling technique, and for directly inspiring the methodology we use in this paper. We would also like to thank John Helly for producing and sharing the halo merger trees that form the backbone of our analysis.

This work used the DiRAC@Durham facility managed by

the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

This work was supported by Vici grant 639.043.409 from the Netherlands Organisation for Scientific Research (NWO).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author. Raw particle data and merger trees for the EAGLE simulations have been publicly released (McAlpine et al. 2016).

REFERENCES

- Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Quataert E., Murray N., 2017, *MNRAS*, 470, 4698
- Bauermeister A., Blitz L., Ma C.-P., 2010, *ApJ*, 717, 323
- Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, 488, 3143
- Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, 717, 379
- Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, *ApJS*, 149, 289
- Berlind A. A., Weinberg D. H., 2002, *ApJ*, 575, 587
- Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53
- Borrow J., Anglés-Alcázar D., Davé R., 2020, *MNRAS*, 491, 6102
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *MNRAS*, 370, 645
- Bryan G. L., Norman M. L., 1998, *ApJ*, 495, 80
- Clauwens B., Schaye J., Franx M., Bower R. G., 2018, *MNRAS*, 478, 3994
- Cole S., 1991, *ApJ*, 367, 45
- Conroy C., Wechsler R. H., Kravtsov A. V., 2006, *ApJ*, 647, 201
- Correa C. A., Schaye J., van de Voort F., Duffy A. R., Wyithe J. S. B., 2018, *MNRAS*, 478, 255
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937
- Croton D. J. et al., 2006, *MNRAS*, 365, 11
- Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 426, 140
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *MNRAS*
- Davé R., Finlator K., Oppenheimer B. D., 2012, *MNRAS*, 421, 98
- De Lucia G., Blaizot J., 2007, *MNRAS*, 375, 2
- Dekel A., Silk J., 1986, *ApJ*, 303, 39
- Diemer B., 2017, *ApJS*, 231, 5
- Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497
- Drory N., Salvato M., Gabasch A., Bender R., Hopp U., Feulner G., Pannella M., 2005, *ApJ*, 619, L131
- Dubois Y. et al., 2014, *MNRAS*, 444, 1453
- Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, *MNRAS*, 417, 2982
- Finlator K., Davé R., 2008, *MNRAS*, 385, 2181
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Haas M. R., Schaye J., Booth C. M., Dalla Vecchia C., Springel V., Theuns T., Wiersma R. P. C., 2013, *MNRAS*, 435, 2955

Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304

Ilbert O. et al., 2010, ApJ, 709, 644

Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astronomy Letters, 44, 8

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

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

Larson R. B., 1972, Nature Physical Science, 236, 7

Larson R. B., 1974, MNRAS, 169, 229

Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119

Lu Z., Mo H. J., Lu Y., Katz N., Weinberg M. D., van den Bosch F. C., Yang X., 2015, MNRAS, 450, 1604

McAlpine S. et al., 2016, Astronomy and Computing, 15, 72

Mitchell P. D., Lacey C. G., Baugh C. M., Cole S., 2016, MNRAS, 456, 1459

Mitchell P. D., Lacey C. G., Cole S., Baugh C. M., 2014, MNRAS, 444, 2637

Mitchell P. D., Schaye J., Bower R. G., 2020a, MNRAS, 497, 4495

Mitchell P. D., Schaye J., Bower R. G., Crain R. A., 2020b, MNRAS, 494, 3971

Mitra S., Davé R., Finlator K., 2015, MNRAS, 452, 1184

Moster B. P., Naab T., White S. D. M., 2018, MNRAS, 477, 1822

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

Muzzin A. et al., 2013, ApJS, 206, 8

Neistein E., Khochfar S., Dalla Vecchia C., Schaye J., 2012, MNRAS, 421, 3579

Nelson D., Genel S., Vogelsberger M., Springel V., Sijacki D., Torrey P., Hernquist L., 2015, MNRAS, 448, 59

Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577

Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, MNRAS, 406, 2325

Pandya V. et al., 2020, arXiv e-prints, arXiv:2006.16317

Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144

Peeples M. S., Werk J. K., Tumlinson J., Oppenheimer B. D., Prochaska J. X., Katz N., Weinberg D. H., 2014, ApJ, 786, 54

Pillepich A. et al., 2018, MNRAS, 473, 4077

Planck Collaboration et al., 2014, A&A, 571, A16

Qu Y. et al., 2017, MNRAS, 464, 1659

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Rodríguez-Puebla A., Primack J. R., Avila-Reese V., Faber S. M., 2017, MNRAS, 470, 651

Schaye J., 2004, ApJ, 609, 667

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

Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210

Schaye J. et al., 2010, MNRAS, 402, 1536

Sharma M., Theuns T., 2020, MNRAS, 492, 2418

Silk J., Rees M. J., 1998, A&A, 331, L1

Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, MNRAS, 423, 1992

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481

Springel V., 2005, MNRAS, 364, 1105

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

Tabor G., Binney J., 1993, MNRAS, 263, 323

Übler H., Naab T., Oser L., Aumer M., Sales L. V., White S. D. M., 2014, MNRAS, 443, 2092

Vale A., Ostriker J. P., 2004, MNRAS, 353, 189

van de Voort F., 2017, Astrophysics and Space Science Library, Vol. 430, The Effect of Galactic Feedback on Gas Accretion and Wind Recycling, Fox A., Davé R., eds., p. 301

van de Voort F., Schaye J., Booth C. M., Haas M. R., Dalla Vecchia C., 2011, MNRAS, 414, 2458

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

Wang L. et al., 2013, MNRAS, 431, 648

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

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

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009, MNRAS, 399, 574

Wright R. J., Lagos C. d. P., Power C., Mitchell P. D., 2020, MNRAS

Yang X., Mo H. J., van den Bosch F. C., Zhang Y., Han J., 2012, ApJ, 752, 41

APPENDIX A:

As briefly mentioned at the end of Section 2.3, the computation of the quantity $F_{\text{CGM}}^{\text{pr}}$ that appears in Eqn 1 requires some additional steps. From the Reference EAGLE simulation, we compute two terms in addition to those shown in Fig. 3, representing the probability that halo-scale outflows and gas recycling includes gas that was ejected from the ISM in the past.

The two terms are

$$B_{\text{out}} \equiv \frac{F_{\text{CGM,out}}^{\text{gal}}}{F_{\text{CGM}}^{\text{gal}}}, \quad (\text{A1})$$

and

$$B_{\text{ret}} \equiv \frac{F_{\text{CGM,ret}}^{\text{gal}}}{F_{\text{ej,halo}}^{\text{gal}}}, \quad (\text{A2})$$

where $F_{\text{CGM}}^{\text{gal}}$ is the mass fraction of the CGM that was previously ejected from the ISM of a progenitor galaxy, and $F_{\text{CGM,out}}^{\text{gal}}$ is the fraction of the halo-scale mass outflow rate that was previously ejected from the ISM of a progenitor galaxy. Similarly, $F_{\text{ej,halo}}^{\text{gal}}$ is the mass fraction of the ejected gas reservoir (outside R_{200}) that was also ejected from the ISM of a progenitor galaxy, and $F_{\text{CGM,ret}}^{\text{gal}}$ is the fraction of the halo-scale recycled mass inflow rate that was previously ejected from the ISM of a progenitor galaxy. With these two terms, we then track separately the fraction of mass in the CGM (M_{CGM}) and in the ejected gas reservoir outside R_{200} ($M_{\text{ej}}^{\text{halo}}$) that was part of the ISM in the past. Neither term strongly affects our results, and we maintain both at their fiducial values throughout our analysis.

As with the other coefficients introduced in Eqn. 1, we compute B_{out} and B_{ret} from an EAGLE simulation, averaging as a function of halo mass and redshift. As the definition of these parameters was only introduced relatively late in the development of this project, we measure B_{out} and B_{ret} from a smaller version of the main reference EAGLE simulation, using instead a $(25 \text{ Mpc})^3$ volume. We extrapolate to higher halo masses simply by holding the values of B_{out} and B_{ret} fixed to the end points of the measured range.

These dependencies are shown in Fig. A1. Circum-galactic gas that was part of the ISM in the past is generally more likely to be ejected out of the halo than circum-galactic gas that was not formerly part of the ISM. Conversely, ejected gas residing outside

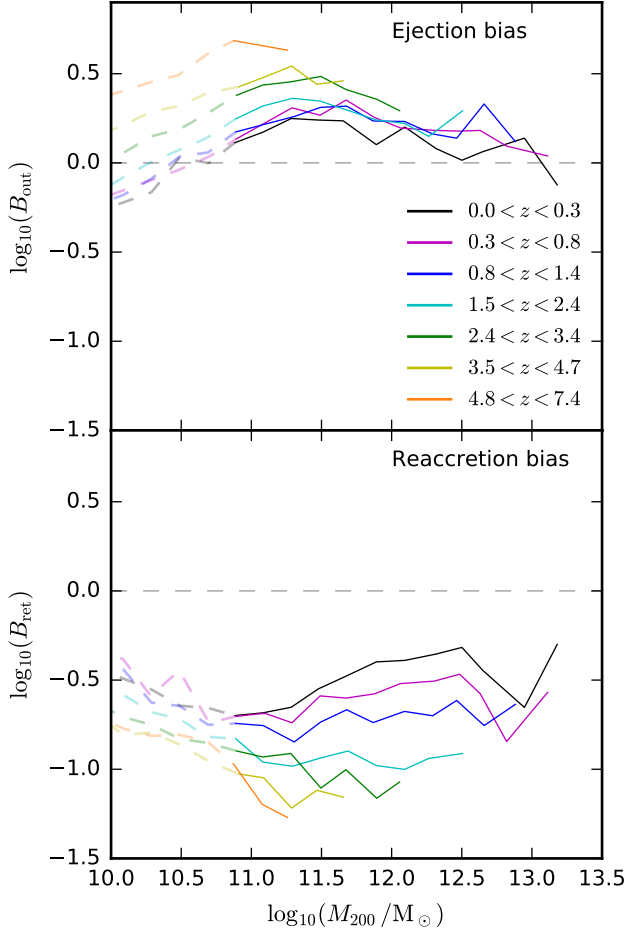


Figure A1. *Top:* the dependence of the bias parameter B_{out} on halo mass and redshift (defined in Eqn A1). B_{out} expresses whether any gas that is ejected outside of haloes is biased with respect to whether that gas was previously part of the ISM. $B_{\text{out}} > 1$ implies that former-ISM circumgalactic gas is more likely to be ejected from the halo than gas that has not been processed through the ISM. *Bottom:* the corresponding halo mass and redshift dependence of the second bias parameter B_{ret} (see Eqn A2). B_{ret} expresses whether any gas that is re-accreted onto haloes is biased with respect to whether that gas was previously part of the ISM.

R_{200} is more likely to be re-accreted if it was not part of the ISM in the past. B_{out} and B_{ret} do not have a large effect on our results for galaxy stellar masses, and we hold them at their fiducial values throughout our analysis.

APPENDIX B:

In the lower-right panel of Fig. 7 of Section 4.2, we showed that if all explicit halo mass dependence is removed from the N12 model (by fixing each model term to its value at $M_{200} = 10^{12} M_{\odot}$), the SHM relation becomes nearly flat (i.e. $M_{\star}/M_{200} \approx \text{constant}$), but still drops significantly for $M_{200} > 10^{13} M_{\odot}$. The reason for this drop at high halo masses is demonstrated in Fig. B1, which shows the fraction of mass in central subhaloes (as opposed to satellites) in the left-side panels, and the differential distribution of satellite subhalo masses in the right-side panels.

We show two distinct definitions of subhalo mass: the instantaneous dark matter subhalo mass (M_{DM} , top panels), and the cu-

mulative mass of dark matter accreted onto subhaloes ($M_{\text{DM}}^{\text{1st}}$). The latter quantity is computed by summing over all first-time dark matter accretion onto all progenitors of the final subhalo, and is directly connected to the source term in our implementation of the N12 model (see Eqn. 1). The main distinction between M_{DM} and $M_{\text{DM}}^{\text{1st}}$ is that only M_{DM} is affected by processes that remove dark matter from a subhalo, which for satellites is dominated by tidal stripping processes as the satellite orbits within the tidal field exerted by the host. On average, stars are significantly more bound than dark matter particles to a subhalo, and are accordingly much less effected by tidal stripping. As such, we would expect that satellite stellar masses would be more closely related to $M_{\text{DM}}^{\text{1st}}$ than to M_{DM} .

The upper-left panel of Fig. B1 shows that the central subhalo always dominates the instantaneous mass of dark matter haloes within EAGLE. The mass fraction within the central subhalo does decrease slightly with increasing halo mass, dropping from $\approx 100\%$ for $M_{200} = 10^{11} M_{\odot}$ to $\approx 80\%$ for $M_{200} = 10^{14.5} M_{\odot}$. This is expected, partly because higher-mass haloes are less relaxed because they collapsed more recently, which implies more substructure (leading also to a weak, negative halo concentration-mass relation), and partly since EAGLE resolves a larger number of substructures in high-mass haloes relative to low-mass haloes.

For $M_{\text{DM}}^{\text{1st}}$ (lower-left panel), satellite subhaloes make a significantly larger contribution to the total mass fraction, and actually comprise the majority of the mass for $M_{200} > 10^{13.5} M_{\odot}$. The strong dependence of the $M_{\text{DM}}^{\text{1st}}$ mass fraction on halo mass is again expected to be partly a resolution effect, which is stronger in this case. If we compare a simulation at the standard EAGLE resolution (solid lines, with numerical particle mass of $m_{\text{DM}} = 9.7 \times 10^7 M_{\odot}$) to an eight-times higher mass resolution simulation (dash-dotted lines) we see indeed that the fraction of $M_{\text{DM}}^{\text{1st}}$ is lower at higher resolution. The right panels of Fig. B1 show the differential distribution of satellite masses for the two simulations, demonstrating that the higher-resolution simulation resolves more of the low-mass substructure, as expected.

We emphasise that changes in satellite contributions to $M_{\text{DM}}^{\text{1st}}$ with changing resolution are not ordinarily a large concern for galaxy stellar masses. This is because low-mass satellites (with $M_{\text{DM}}^{\text{1st}} \ll 10^{12} M_{\odot}$) are far less efficient at forming stars than satellites with $M_{\text{DM}}^{\text{1st}} \sim 10^{12} M_{\odot}$, which are resolved in simulations like EAGLE. For the unrealistic model variation shown in Fig. 7 (yellow line), low-mass satellites are just as efficient at forming stars as high-mass satellites. In this case therefore, not resolving a population of low-mass satellites will be an issue, contributing to the decrease in the SHM relation for $M_{200} > 10^{13} M_{\odot}$, since these haloes do resolve a larger fraction of the (now-relevant) low-mass substructures.

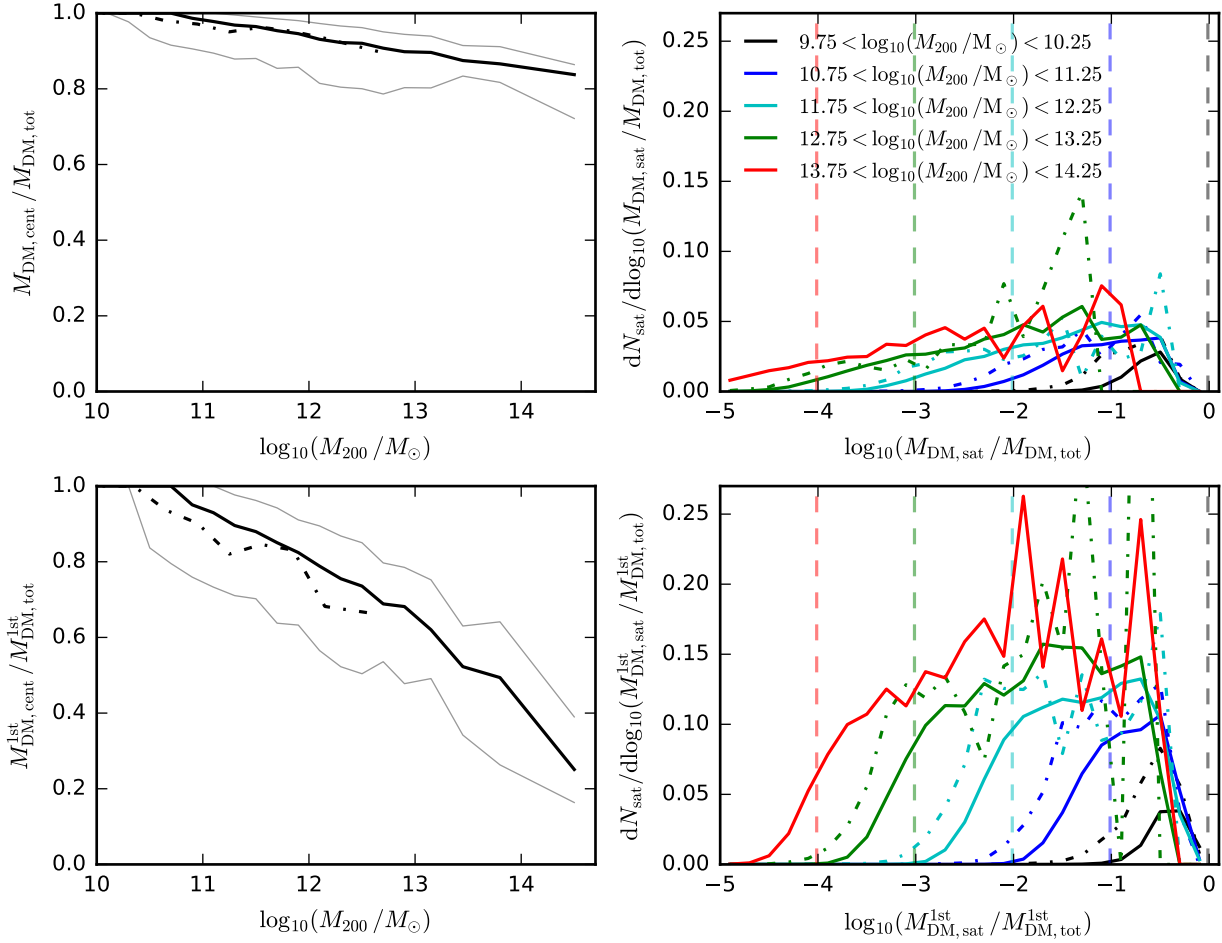


Figure B1. *Left panels:* the fraction of mass in central subhaloes at $z = 0$, compared to the total halo mass including satellite subhaloes, plotted as a function of halo mass. The top panel shows the mass fraction for the instantaneous dark matter mass as identified by SUBFIND (M_{DM}). The bottom panel shows the cumulative mass of dark matter particles that have been accreted for the first time onto a given subhalo and its progenitors ($M_{\text{DM}}^{\text{1st}}$). Solid lines show the 16, 50, and 84th percentiles for the Reference EAGLE simulation. Dash-dotted lines show the median relation for a smaller-volume simulation with 8 times higher particle resolution (Recal). For all halo masses, the instantaneous dark matter mass is always dominated by the central subhalo; satellite subhaloes generally lose most of their mass to tidal stripping. Conversely, satellite subhaloes make up the majority of $M_{\text{DM}}^{\text{1st}}$ for $M_{200} > 10^{13.5} M_{\odot}$, since $M_{\text{DM}}^{\text{1st}}$ is not affected by stripping. $M_{\text{DM}}^{\text{1st}}$ is expected to be more closely connected to the stellar mass of satellite galaxies than M_{DM} , as stars are more tightly bound on average than dark matter particles. *Right panels:* the differential distribution of satellite subhalo masses for M_{DM} (top) and $M_{\text{DM}}^{\text{1st}}$ (bottom). Different line colours indicate different bins of mass of the corresponding host halo. Dashed vertical lines indicate (for the Reference EAGLE simulation) the scale for which the satellite subhalo mass is less than the mass of 100 dark matter particles. Below this scale, satellite subhaloes are not well resolved, and so the simulation becomes incomplete. Comparing the Reference simulation (solid lines) to Recal (dash-dotted lines), it is evident that the simulation is not converged in terms of the total mass in satellites, especially for $M_{\text{DM}}^{\text{1st}}$.