The detailed structure and the onset of galaxy formation in low-mass gaseous dark matter haloes

Alejandro Benitez-Llambay [ID]⋆ and Carlos Frenk

Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

ABSTRACT

We present a model for the formation of the first galaxies before and after the reionization of hydrogen in the early universe. In this model, galaxy formation can only take place in dark matter haloes whose mass exceeds a redshift-dependent critical value, which, before reionization, is equal (in the simplest case) to the mass at which atomic hydrogen cooling becomes effective and, after reionization, is equal to the mass above which gas cannot remain in hydrostatic equilibrium. We define the Halo Occupation Fraction (HOF) as the fraction of haloes that host a luminous galaxy as a function of halo mass. The HOF is established by the interplay between the evolution of the critical mass and the assembly history of haloes and depends on three factors: the minimum halo mass for galaxy formation before reionization, the redshift of reionization, and the intensity of the (evolving) external photoheating rate. Our fiducial model predicts a cutoff in the galaxy mass function at a present-day halo mass, $M_{200} \sim 3 \times 10^8 \, \mathrm{M_\odot}$; 100 per cent occupation at $M_{200} > 5 \times 10^9 \, \mathrm{M_\odot}$; and a population of starless gaseous haloes of present-day mass in the range $10^6 \lesssim M_{200}/\mathrm{M_\odot} \lesssim 5 \times 10^9$, in which the gas is in thermal equilibrium with the ultraviolet background radiation and in hydrostatic equilibrium in the gravitational potential of the halo. The transition between HOF = 0 and HOF = 1 reflects the stochastic nature of halo mass growth. We explore how these characteristic masses vary with model assumptions and parameter values. The results of our model are in excellent agreement with cosmological hydrodynamic simulations of galaxy formation.

Key words: galaxies: dwarf – galaxies: formation – galaxies: star formation – dark ages, reionization, first stars – dark matter .

1 INTRODUCTION

The 'Λ-cold dark matter' (ΛCDM) cosmological model makes specific and robust predictions for the growth, structure, and abundance of dark matter haloes, the sites where galaxies form. These are: (i) dark matter haloes grow hierarchically: small haloes form first and larger haloes form subsequently by mergers of smaller haloes and accretion of smooth dark matter; (ii) the internal structure of dark matter haloes is self-similar over ∼20 decades in mass (Navarro, Frenk & White 1996, 1997; Wang et al. 2020); (iii) the halo mass function rises steeply towards low masses (approximately as M^{-1}; e.g. Press & Schechter 1974; Bond et al. 1991; Jenkins et al. 2001, and references therein), implying that a large fraction of low-mass haloes must remain 'dark' at redshift $z = 0$ if ΛCDM is to be reconciled with the relatively flat faint end of the observed galaxy stellar mass function (Klypin et al. 1999; Bullock, Kravtsov & Weinberg 2000; Somerville 2002; Benson et al. 2002b).

In ΛCDM a 'cutoff' in the mass of dark matter haloes that can host galaxies is expected on general grounds. After the hydrogen emerging from the big bang recombines, it can, in principle, cool and condense into the small dark matter haloes present at that time but only in those whose potential well is deep enough to overcome the kinetic and thermal energy of the gas. Once the hydrogen is reionized at redshift, z_{re}, the hydrogen–helium plasma is heated to $T_{\mathrm{b}} \sim 2 \times 10^4$ K, preventing it from accreting on to dark matter haloes of virial temperature, $T_{200} \lesssim T_{\mathrm{b}}$ (Babul & Rees 1992; Efstathiou 1992; Quinn, Katz & Efstathiou 1996; Thoul & Weinberg 1996; Barkana & Loeb 1999), corresponding to a mass, $M_{200} \lesssim 10^{10} \, \mathrm{M_\odot}$.[1] Thus, we expect most haloes today of mass below a characteristic critical mass, M_{cr}^0, to remain dark. This conclusion agrees qualitatively with results from cosmological hydrodynamical simulations (Hoeft et al. 2006; Okamoto & Frenk 2009; Benítez-Llambay et al. 2015; Sawala et al. 2016; Fitts et al. 2017), although the exact value of the critical halo mass and its dependence on modelling details, are uncertain. A quantity often discussed in this context is the 'filtering mass' defined as the mass of those haloes that retain half of their baryons after cosmic reionization; it can be calculated using either linear theory or numerical simulations (e.g. Gnedin 2000; Benson et al. 2002a; Hoeft et al. 2006; Okamoto, Gao & Theuns 2008). The exact connection of this quantity to galaxy formation is, however, unclear.

To illustrate these ideas, let us assume that prior to reionization galaxy formation can only take place in haloes in which atomic hydrogen can cool, that is haloes of virial temperature, $T_{200} \gtrsim 7000$ K.

---

[1] We define virial quantities as those pertaining to the sphere within which the mean mass density equals 200 times the critical density of the universe and label them with the subscript 200.

⋆ E-mail: bllalejandro@gmail.com

Schematic picture of the impact of cosmic reionization on galaxy formation
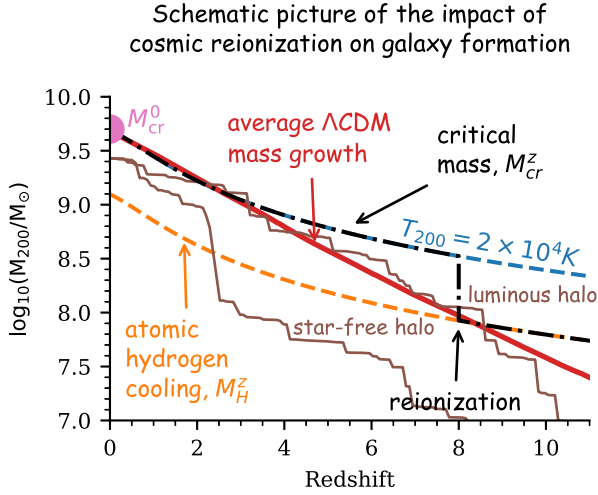

Figure 1. Redshift evolution of the critical halo mass above which atomic hydrogen cooling becomes efficient (orange dashed line), the critical mass corresponding to a fixed virial temperature, $T_{200} = T_b = 2 \times 10^4$ K (blue dashed line), and the mean mass assembly history of a ΛCDM halo of present-day mass, $M_{200} \sim 5 \times 10^9 \, M_\odot \, h^{-1}$ (red solid line). The effective critical mass for gas to collapse is shown by the dot-dashed black line. This is equal to M_H^z (equation 1) prior to the redshift of reionization, z_{re}, and to M_{cr}^z (equation 2) after z_{re}. For this particular example, the value of M_{cr}^0 constrains both, the redshift of reionization and the temperature of the intergalactic medium. The brown thin solid lines show two particular mass assembly histories of dark matter haloes of present-day mass, $M_{200} \sim 3 \times 10^9 \, M_\odot$, i.e. of mass, $M_{200} < M_{cr}^0$. One halo never exceeds M_{cr}^z and is expected to remain 'dark' at $z = 0$. The other was more massive than M_{cr}^z prior to reionization and is expected to host a luminous galaxy at $z = 0$.

The corresponding critical mass is approximately,

$$M_H^z \sim (4 \times 10^7 \, M_\odot) \left(\frac{1+z}{11} \right)^{-3/2}. \tag{1}$$

In Fig. 1 we follow the evolution of a halo of present-day mass, $M_{cr}^0 = 5 \times 10^9 \, M_\odot$, which, for illustration purposes, we take to be the critical mass above which all dark matter haloes host a luminous galaxy at $z = 0$. The red solid line shows the average mass growth (inferred from the mean mass accretion history) of a halo of that mass in ΛCDM. The halo mass required for galaxy formation to proceed is shown by the black dashed line. At $z = z_{re}$ this jumps from M_H^z to M_{cr}^z. The blue dashed line in Fig. 1 shows the evolution of M_{cr}^z, assumed to be the mass of a halo of virial temperature $T_b = 2 \times 10^4$ K.

All haloes of $M_{200} \geq M_{cr}^0$ will host a luminous galaxy at $z = 0$. Some haloes of $M_{200} \leq M_{cr}^0$ will be 'dark' but others will also host a luminous galaxy, depending on their previous history. The thin brown lines in Fig. 1 illustrate two different mass accretions histories that lead to the same halo mass at $z = 0$, $M_{200} < M_{cr}^0$. One of them never crosses the critical mass required for gas to collapse, whereas the other, although below the critical mass at $z = 0$, exceeded M_H^z before cosmic reionization. Of the two, only the later is expected to host a luminous galaxy today. This example illustrates the origin of the 'stochastic' nature of galaxy formation in dark matter haloes of mass close to the critical value (e.g. Hoeft et al. 2006; Benítez-Llambay et al. 2015; Sawala et al. 2016; Fitts et al. 2017, and references therein).

The present-day value of the critical mass, M_{cr}^0, that separates haloes that were able to form a galaxy from those that were not is a direct probe of the epoch of reionization and constrains the two parameters, z_{re} and T_b. The redshift of reionization can be estimated

by finding the time when the average mass of haloes of $M_{200} = M_{cr}^0$ first exceeded M_H^z (dashed orange line). The constraint on T_b follows from the value of M_{cr}^0 since,

$$M_{cr}^z \sim (10^{10} \, M_\odot) \left(\frac{T_b}{3.2 \times 10^4 \, K} \right)^{3/2} (1+z)^{-3/2}. \tag{2}$$

In this paper, we develop a theoretical framework to understand the onset of galaxy formation and the impact of cosmic reionization on galaxy formation and address the following questions: Is there an actual minimum halo mass at $z = 0$ below which galaxies cannot form? If so, how does this depend on the characteristic scales of cosmic reionization? Is the simple picture outlined above consistent with full hydrodynamical simulations of galaxy formation? Is our model quantitatively robust? How sensitive are the results to the underlying assumptions about galaxy formation?

In Section 2, we introduce our model to calculate how reionization affects galaxy formation in low-mass haloes. We describe the numerical method and our simulations in Section 3. We perform a comparison between our model and high-resolution cosmological hydrodynamics simulations in Section 4. We discuss our results in Section 5 and conclude in Section 6.

2 MODEL

We will discuss later the conditions required for gas to collapse in a halo and make a galaxy. Here, we introduce a model to investigate the effects of reionization and of the presence of an external ultraviolet background (UVB) on the structure of gaseous haloes. The model applies strictly to dark matter haloes whose gaseous component has not yet collapsed to form a galaxy by z_{re}.

As discussed in the Introduction, cosmic reionization drastically reduces the gas content of dark matter haloes of virial temperature $T_{200} \lesssim T_b \sim 2 \times 10^4$ K, or equivalently, mass, $M_{200} \lesssim M_{cr}^z$ (equation 2). This is the mass scale below which the halo free-fall time, t_{ff}, becomes longer than the sound crossing time of the gas in the halo, t_{sc}.[2] At $z = 0$ these haloes have mass, $M_{200} \lesssim M_{cr}^0 \sim 10^{10} \, M_\odot$, but at $z = 10$, the critical mass is $M_c^{10} \sim 5 \times 10^8 \, M_\odot$. Any gas that settles in haloes of virial temperature $T_{200} \lesssim T_b$ is expected to be in thermal (rather than virial) equilibrium with the external ionizing radiation at a temperature $T \sim T_b$.

The distribution of gas that collapses into these low-mass haloes after reionization can be fully determined assuming that: (i) the gas reaches hydrostatic equilibrium within the (spherically symmetric) dark matter halo, and (ii) the contribution of the gas to the gravitational potential is negligible everywhere (see e.g. Ikeuchi 1986; Rees 1986). Under these assumptions, the hydrostatic equilibrium equation for the gas reads

$$\left(\frac{\tilde{T}}{\tilde{\rho}} + \frac{d\tilde{T}}{d\tilde{\rho}} \right) \frac{d\tilde{\rho}}{d\tilde{r}} = -2 \frac{\tilde{M}(\tilde{r})}{\tilde{r}^2}, \tag{3}$$

where $\tilde{\rho} = \rho/\bar{\rho}$, $\tilde{M} = M/M_{200}$, $\tilde{r} = r/r_{200}$, $\tilde{T} = T/T_{200}$, $\bar{\rho} = \rho_c \Omega_b$ are dimensionless variables; for the virial temperature, T_{200}, we adopt the definition $T_{200} = \mu m_p / 2k_B V_{200}^2$, with μ, m_p, and k_B being the mean molecular weight, the proton mass, and the Boltzmann constant, respectively; $V_{200}^2 = GM_{200}/r_{200}$ is the halo circular velocity

---

[2]The condition, $t_{ff} < t_{sc}$, yields the well-known criterion for dark matter haloes to be affected by cosmic reionization, $V_{200} < c_s$, where V_{200} is the halo circular velocity and $c_s \sim (10 \, \text{km s}^{-1})(T_b/10^4 \, \text{K})^{1/2}$ is the sound speed of the gas.
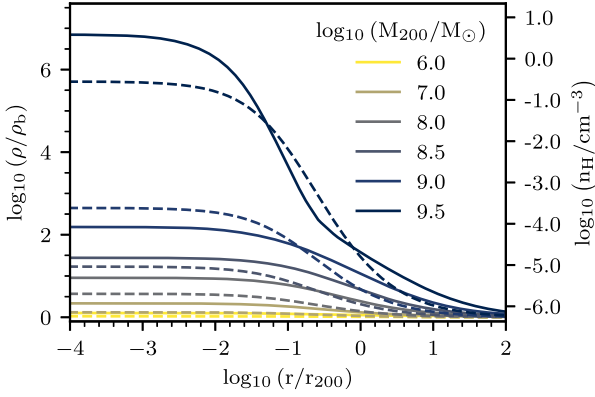
Figure 2. Gas density profile of gaseous haloes in hydrostatic equilibrium with the dark matter potential at redshift $z = 0$, for haloes of different mass, as indicated in the legend. The dashed lines show the isothermal case, for which the temperature of the gas is independent of density and equal to $T = 2 \times 10^4$. The solid lines correspond to the density–temperature relation, $T(\rho)$, consistent with the gas being in thermal equilibrium with the external Haardt & Madau (2001) UVB radiation field at $z = 0$ (see also Fig. 4).

at r_{200}. We have also made the assumption that the gas temperature is a function of density, $\tilde{T} = \tilde{T}(\tilde{\rho})$.

2.1 Isothermal case

For the simplest case in which the temperature of the gas is independent of density (i.e. for an isothermal gas), the solution of this equation yields

$$\tilde{\rho}(\tilde{r}) = \exp\left\{2\left(\frac{T_{200}}{T_b}\right)\int_{\tilde{r}}^{\infty}\frac{\tilde{M}(r')}{r'^2}\,dr'\right\}, \quad (4)$$

where we assume that $\tilde{\rho} = 1$ sufficiently far from the halo ($\tilde{r} \to \infty$), i.e. the outer pressure of the halo is equal to that of the intergalactic medium. The enclosed gas mass,

$$\tilde{M}_{gas}(\tilde{r}) = \left(\frac{3\Omega_b}{\Delta}\right)\int_0^{\tilde{r}}\tilde{\rho}(\tilde{r}')\tilde{r}'^2\,d\tilde{r}', \quad (5)$$

in which $\Delta = 200$, must be calculated numerically.

For a dark matter halo described by a Navarro–Frenk–White (NFW) density profile (Navarro et al. 1996, 1997), the gas density profile is

$$\tilde{\rho}(\tilde{r}) = \exp\left\{2\left(\frac{T_{200}}{T_b}\right)\frac{1}{f_c}\frac{\ln(1 + c\tilde{r})}{\tilde{r}}\right\}, \quad (6)$$

in which $f_c = \ln(1 + c) - c/(1 + c)$, $c = r_{200}/r_s$ is the concentration of the halo, and r_s is the characteristic radius of the dark matter profile.

The dashed lines in Fig. 2 show the gas density profile (equation 6) for different halo masses, assuming a constant concentration, $c = 10$, independently on halo mass.[3] The slope of the gas density profile becomes increasingly shallow towards the centre, and converges to a well-defined value, ρ_0, at $r = 0$. The finite maximum central gas density is a consequence of the cuspy nature of ΛCDM haloes, which gives rise to a maximum central gravitational acceleration. In addition, the gas density profile settles at the mean density of the universe ($\tilde{\rho} = 1$) further out, by construction.

---

[3]Note that the concentration of low-mass ΛCDM haloes is only weakly dependent on halo mass (see e.g. Ludlow et al. 2014; Wang et al. 2020), so this is a sensible assumption for our purposes.
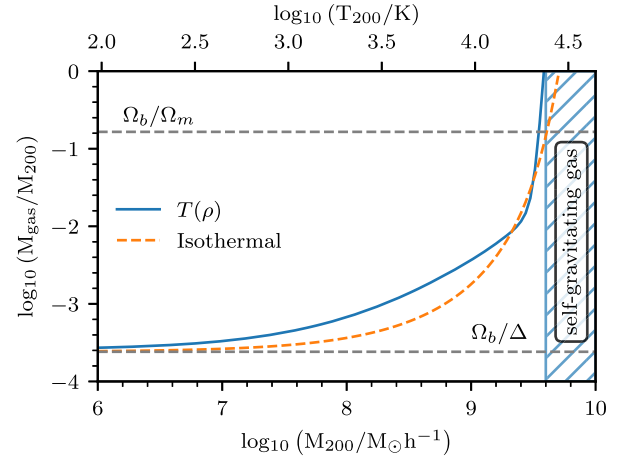

Figure 3. Gas mass of starless gaseous haloes, in units of halo mass, as a function of halo mass, for the isothermal case ($T_b = 2 \times 10^4$ K; orange dashed line) and the general model ($T(\rho)$; solid blue line). Gas in haloes more massive than, $M_{cr}^0 \sim 5 \times 10^9$ $M_\odot$ h^{-1}, at $z = 0$ is expected to become self-gravitating and undergo gravitational collapse. The gas fraction for low-mass haloes approaches the lower bound, Ω_b/Δ, which is the enclosed gas mass inside r_{200} of a halo embedded in an unperturbed background. The scale on the top indicates the virial temperature, T_{200}.

The total gas mass within the virial radius, as a function of halo mass (equation 5), is shown by the dashed orange line in Fig. 3. The gas mass grows exponentially with halo mass and asymptotically approaches the value $M_{200}\Omega_b/\Delta$ for low-mass haloes, as expected.[4] For more massive haloes, the gas mass becomes a steep function of halo mass and the amount of gas needed to balance the gravitational acceleration of the halo eventually exceeds both, the baryon fraction of the universe (horizontal dashed line) and the dark matter content of the halo, after which equation (5) is no longer valid. Gaseous haloes more massive than $M_{200} \sim 5 \times 10^9$ $M_\odot$ h^{-1} will thus become self-gravitating and continue their gravitational collapse to form a central luminous galaxy.[5]

Self-gravitating gaseous haloes develop inner densities that exceed 10^5 times the mean baryon density of the universe at $z = 0$, or $\rho/m_p = n_H \sim 1.0$ cm^{-3} (see scale on the right-hand panel in Fig. 2). This is encouraging, as this estimate is comparable to the minimum density above which gas is expected to become self-shielding, molecular, and self-gravitating, a regime in which our model is not designed to work (see e.g. Sykes et al. 2019). This number also agrees well with the neutral hydrogen column density, N_{HI}, above which damped Ly α (DLA) systems become rare (e.g. Schaye 2001; Krumholz et al. 2009; Jorgenson et al. 2014). Indeed, for a halo of mass $M_{200} = 5 \times 10^9$ $M_\odot$, assuming that all hydrogen inside the central core is neutral, we obtain a central H I column density $N_{HI} \sim 10^{22}$ cm^2.

2.2 Thermal equilibrium with the UVB

More generally, the interplay between gas cooling and photoheating from the external UVB links the gas temperature to its density (see

---

[4]The enclosed gas mass within the virial radius must scale linearly with halo mass for a flat gas density profile.

[5]Note that this is particularly true for dark matter haloes with virial temperature $10^4 \lesssim T_{200} \lesssim 10^5$ K, for which the cooling time is much shorter than the Hubble time, so that the gas in the halo is effectively at $T \sim 10^4$ K.
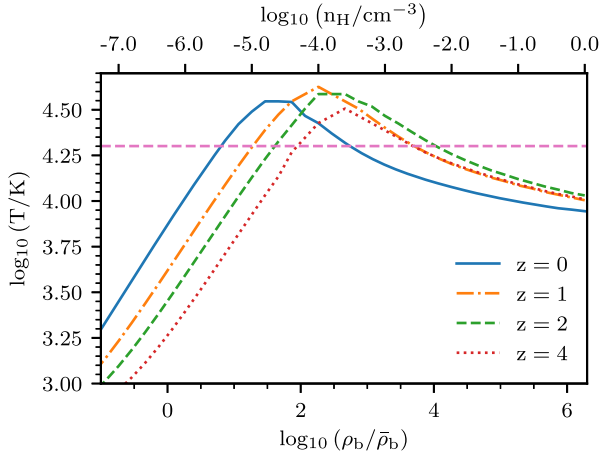
Figure 4. Temperature–density relation for the gas in starless dark matter haloes of virial temperature, $T_{200} \lesssim 10^4$ K. The lines show the $T(\rho)$ relation expected for gas in thermal equilibrium with the redshift-dependent Haardt & Madau (2001) UVB radiation field (see Section 2). The horizontal dashed line shows the temperature assumed for the isothermal case ($T_b = 2 \times 10^4$ K), in which the entire gas content of the universe is at a single constant temperature, independently of density.

e.g. Theuns et al. 1998; Barkana & Loeb 1999) and so equation (3) must be solved taking this into account.

The processes that determine the density–temperature relation, $T(\rho)$, of gas in dark matter haloes of virial temperature, $T_{200} \lesssim 10^4$ K, depend on the density of the gas. For high-density gas, $n_H \gtrsim 5 \times 10^{-4}$ cm^{-3}, the cooling and heating time-scales are both shorter than the Hubble time and $T(\rho)$ is determined by the balance between cooling and photoheating. For low-density gas, the cooling time-scale is longer than the Hubble time and $T(\rho)$ is determined exclusively by the photoheating time-scale.

Fig. 4 shows $T(\rho, z)$ calculated for gas of primordial composition, using the publicly available cooling tables compiled by Wiersma, Schaye & Smith (2009),[6] which assume an optically thin gas in ionization equilibrium illuminated by the Haardt & Madau (2001) UVB radiation field. (For completeness, we list the values of the relation in Table A1 in Appendix A). For comparison, the horizontal dashed line in the figure shows the value of $T_b = 2 \times 10^4$ K previously assumed for the isothermal case. Although for gas above the mean cosmic density the temperature is not far from T_b, there are clear differences.

The $z = 0$ density–temperature relation results in the gas density profiles shown by the solid lines in Fig. 2, and in the present-day gas mass shown by the solid blue line in Fig. 3. The results of the simple isothermal case follow quite closely those of this model for which the temperature is not constant. Note, however, that the background temperature assumed for the isothermal case, T_b, is a free parameter. Had we chosen a different value, the results would have been different. The $T(\rho)$ relation in Fig. 4 removes any freedom in the modelling. We will focus on the results for this general case, in which $T \equiv T(\rho)$, from now on.

Our model exhibits a number of interesting properties:

(i) There is a wide range in halo mass at $z = 0$ ($10^6 \lesssim M_{200}/M_\odot h^{-1} \lesssim 5 \times 10^9$) for which gas neither leaves the halo completely nor collapses further. We previously identified these systems and named them REionization Limited H I Clouds (REL-HICs) (Benítez-Llambay et al. 2017). Systems of this kind were first
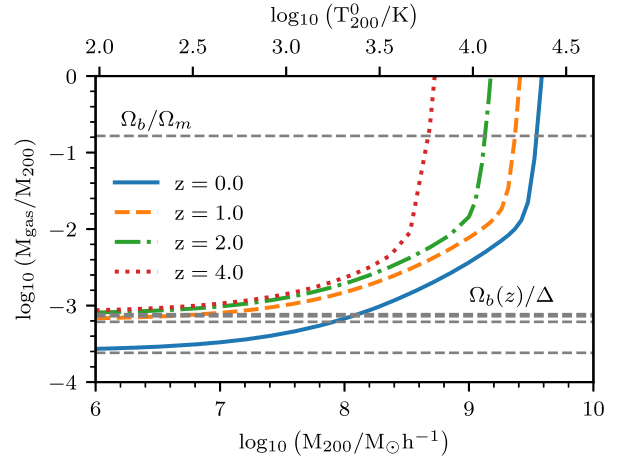
Figure 5. Gas mass, in units of the halo mass, as a function of halo mass, for starless gaseous haloes at different redshifts. These are derived from the model of Section 2, which assumes that the cosmic gas follows the redshift- and density-dependent $T(\rho, z)$ relation shown in Fig. 4. Note that the gas mass converges to different values for low-mass haloes as a result of the cosmic expansion.

explored by Rees (1986) and Ikeuchi (1986) in the context of the 'minihalo model of the Ly α forest'.

(ii) Dark matter haloes less massive than $M_{200} \sim 10^6$ $M_\odot$ h^{-1} at $z = 0$ are expected to be devoid of gas; their potential wells are not deep enough to balance the pressure of the gas.

(iii) Gas in dark matter haloes more massive than $M_{200} \sim 5 \times 10^9 M_\odot$ h^{-1} at $z = 0$ cannot provide enough pressure support to prevent further gravitational collapse; these haloes are expected to collapse, fragment, and form stars.

2.3 Redshift evolution

The redshift evolution of the model is determined by the time variation of the external pressure of the intergalactic medium, as the model assumes that the external pressure is equal to that of the mean intergalactic medium: $P(z) \propto \rho_b(z) T_b(z)$. Thus, the external pressure for the gas in the halo must be calculated taking into account the redshift evolution of the Haardt & Madau (2001) photoheating rate (see Fig. 4) and the cosmic expansion.

As an example, Fig. 5 shows the gas mass (in units of the halo mass), as a function of halo mass, at four different redshifts, $z = (0, 1, 2, 4)$. It is clear that, at fixed halo mass, the total gas mass of these low-mass haloes evolves with time. At high redshift, the universe is denser and so the external pressure on the gaseous haloes is greater. As a result, the critical mass above which gas is expected to lose pressure support shifts to lower halo masses. This invites us to define the redshift-dependent critical mass for gas to undergo gravitational collapse, M_{cr}^z, as the halo mass above which the fraction of baryons needed to ensure hydrostatic equilibrium exceeds the universal fraction, $f_b = \Omega_b/\Omega_m$, indicated in the figure as a dashed line. Clearly, the exact value of this fraction is not crucial as long as $\tilde{M}_{gas} = M_{gas}/M_{200} \gtrsim 0.1$.

The blues lines in Fig. 6 show the redshift evolution of the critical mass that results from our model, expressed either as a virial temperature (top), or as a gas mass (bottom).[7] Note that the critical
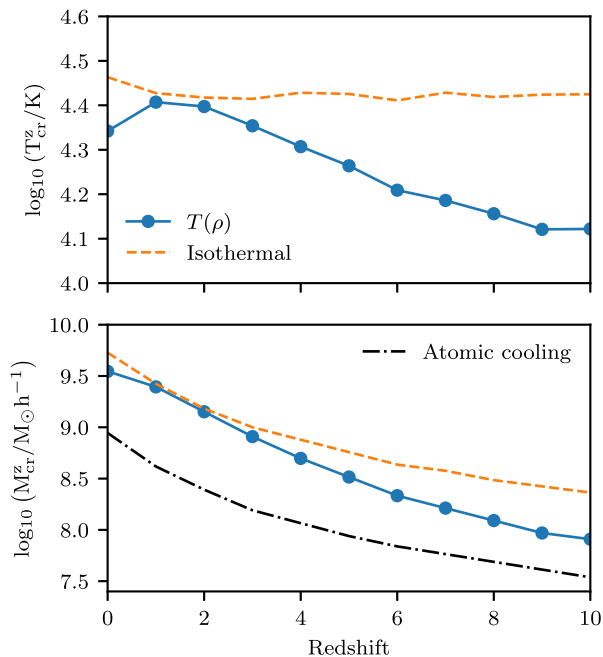
Figure 6. Critical virial temperature (top) and critical gas mass (bottom), as a function of redshift. The orange dashed line shows the result for the isothermal case, assuming $T_{\rm b} = 2 \times 10^4$ K. The solid blue line shows the result of the general model of Section 2, which has no free parameters (except for the intensity of the UVB radiation field). The black dot-dashed line in the bottom panel shows the critical halo mass above which atomic hydrogen cooling becomes efficient (equation 1).

virial temperature for gas to become self-gravitating and collapse is not a constant, contrary to common assumptions (e.g. Bullock et al. 2000; Okamoto & Frenk 2009; Bose et al. 2020; Graus et al. 2019). Moreover, the critical temperature exhibits a maximum at $z \sim 2$, when the Haardt & Madau (2001) photoheating rate peaks.

For comparison, we also show the critical mass and virial temperature for the particular isothermal case in which $T_{\rm b} = 2 \times 10^4$ K. For this particular case, the critical temperature remains constant, as expected. Interestingly, for this particular value of $T_{\rm b}$, the isothermal case resembles our model reasonably well. Note, however, that it can depart from our model by more than a factor of 3 at high redshift. The black-dashed line in the bottom panel shows the critical halo mass above which atomic hydrogen cooling becomes efficient (equation 1), and corresponds to the critical halo mass for galaxy formation in the absence of an external UVB radiation field.

We conclude this section by highlighting another central feature of our model:

(i) Dark matter haloes less massive than $\sim 5 \times 10^9$ M$_\odot$ h^{-1} at $z = 0$ that did not exceed either the atomic hydrogen cooling limit prior to cosmic reionization or the critical mass imposed by the external UVB afterwards will remain 'dark' at $z = 0$; otherwise they will be luminous (see also Fig. 1). Galaxy formation below the (present-day) mass scale $M_{\rm cr}^0 \sim 5 \times 10^9$ M$_\odot$ h^{-1} is thus largely stochastic and dependent on the assembly history of individual ΛCDM haloes.

We now turn our attention to a comparison between the predictions of our models and the results of high-resolution hydrodynamical cosmological simulations. We introduce the simulation in Section 3.1, provide details of the identification of substructures in Section 3.2, and present the detailed comparison in Section 3.3.

3 TECHNIQUES

3.1 Simulation details

We consider a random realization of a periodic cosmological cubic volume of side 20 Mpc, with initial conditions obtained using the publicly available code MUSIC (Hahn & Abel 2011). The simulation volume is filled with 1024^3 dark matter particles and the same number of gas particles, so that the dark matter and gas particle masses are: $m_{\rm drk} = 1.98 \times 10^5$ M$_\odot$ h^{-1} and $m_{\rm gas} = 3.68 \times 10^4$ M$_\odot$ h^{-1}, respectively. The Plummer-equivalent gravitational softening, ϵ, adopted in our simulation never exceeds 1 per cent of the mean interparticle separation. This gives $\epsilon \sim 195$ pc for both the gas and the dark matter particles. We assume cosmological parameters consistent with early Planck results (Planck Collaboration XVI 2014).[8]

We performed the simulation with the version of the code P-Gadget3 (last described in Springel 2005) previously used for the EAGLE project (Crain et al. 2015; Schaye et al. 2015). We refer the reader to the original papers for further details about the physics modules included in the simulation. We describe here only those relevant for our analysis.

The simulation includes radiative cooling and a uniform, isotropic, redshift-dependent UVB spectrum (Haardt & Madau 2001). The simulation follows star formation, stellar evolution, supernova feedback, and chemical enrichment. Gas cooling is calculated using the Wiersma et al. (2009) cooling tables, which include cooling due to several metal species. For our purposes, however, cooling by primordial hydrogen and helium are the only relevant processes.

Reionization is modelled by turning on the isotropic Haardt & Madau (2001) UVB at the redshift of reionization, assumed to be $z_{\rm re} = 11.5$.[9] In order to ensure that gas is quickly heated to $\sim 10^4$ K at high redshift, an energy of 2 eV per proton mass is injected instantaneously at $z = z_{\rm re}$.

Star formation proceeds stochastically according to the Kennicut–Schmidt law in gas particles whose density is above a threshold, $n_{\rm H, 0}$. In the EAGLE simulations, the density threshold for star formation is a function of metallicity. We do not make this assumption here and instead assume a constant value, $n_{\rm H,0} = 1.0$ cm^{-3}, motivated by the discussion of Section 2 (see also Fig. 2). The results of our simulation should not be strongly dependent on the assumed value of $n_{\rm H,0}$ provided this value is comparable to or larger than $n_{\rm H,0} \sim 1$ cm^{-3}, above which gas becomes self-gravitating in low-mass haloes (e.g. Benítez-Llambay et al. 2019). Once the gas in the simulation becomes self-gravitating, its maximum density is limited either by resolution or by the adopted threshold.

The EAGLE model imposes a temperature floor, $T(\rho) = T_0(n_{\rm H}/n_{\rm H, 0})^{\gamma - 1}$ for high-density gas, $n_{\rm H} > n_{\rm H, 0}$, where $\gamma = 4/3$ is the ratio of specific heats, and $T_0 = 8000$ K. We adopt this temperature floor to prevent the development of extremely high densities, which might impact the run time of the simulation. This density-dependent temperature floor affects only star-forming gas, so it should not affect the comparison between the starless haloes of our model and those from the simulation.

3.2 Identification of haloes in the simulation and mass assembly histories

Dark matter haloes are identified using the publicly available code `HBT+` (Han et al. 2018). In short, `HBT+` uses an input catalogue of friends-of-friends haloes, constructed using the standard linking length, $b = 0.2$ (Davis et al. 1985), and determines which particles are gravitationally bound to each halo, discarding the rest. The unbinding procedure starts at the earliest snapshot ($z = 20$ in our case) and the list of gravitationally bound particles in each halo is then passed to the next snapshot to identify the descendant halo. When multiple progenitors end up in the same descendant, a main progenitor is determined according to mass, and the remaining particles are subject to the unbinding procedure to create satellite subhaloes. Finally, the bound particles that do not belong to any satellite are assigned to the main progenitor. These steps are repeated between consecutive snapshots until the final snapshot of the simulation is reached.

Applying this algorithm to all the simulation outputs results in a catalogue of central and satellite galaxies at redshift $z = 0$ along with their time evolution. We shall refer to the evolution of the mass of the main progenitor as the Mass Assembly History. High completeness of the $z = 0$ halo and subhalo catalogues can be achieved by sampling the evolution from redshift $z = 20$ to $z = 0$ with 60 snapshots (Han et al. 2012). Our simulation has more than twice this number.

The `HBT+` catalogue provides an extensive set of halo properties, such as position, velocity, virial mass, gas mass, stellar mass, etc. It also provides a list of particles (gas, dark matter, and stars) gravitationally bound to each central and satellite halo. We shall refer to these particles as the bound particles.

3.3 Comparison between our model and our ΛCDM hydrodynamics simulation

The properties of RELHICs (REionization Limited H I Clouds; Benítez-Llambay et al. 2017) – starless haloes whose gas is in thermal equilibrium with the external UVB radiation field and in hydrostatic equilibrium within the halo – provide an important test of our model. The model (Section 2) predicts the existence of these RELHICs in the mass range $10^6 \lesssim M_{200}/M_\odot\,h^{-1} \lesssim 5 \times 10^9$ at $z = 0$. In order to test the model we select all field dark matter haloes in our simulation of mass, $M_{200} \geq 3 \times 10^7\,M_\odot\,h^{-1}$, that contain at least one bound gas particle and no stars. These haloes are resolved with more than 150 dark matter particles in our simulation.

Fig. 7 shows the bound gas mass within the virial radius of the selected haloes, as a function of halo mass at redshifts $z = 0$, 1, and 2. The running medians and the 10–90th percentiles of the distributions are shown by the orange circles and shaded regions, respectively.

The overall agreement between the gas mass of the simulated RELHICs and the predictions of our model is remarkable, even in systems that contain fewer than 10 bound gas particles. Both, the zero-point and the slope of the relation agree. This is not entirely surprising when we consider that the thermodynamic properties of the gas inside starless systems are largely established by external processes. One of these is the presence of an external gravitational potential sourced by (well-resolved) dark matter haloes; the other is the interplay between photoheating and cooling. The agreement between our model and the simulation demonstrates that the gas particles inside the RELHICs trace the local gas density fluctuations of the intergalactic medium induced by the presence of a dark matter halo.

This result extends those of Benítez–Llambay et al. (2017), where we probed this regime using a limited sample of haloes

taken from a zoom-in simulation of the Local Group at redshift $z = 0$ only. Our simulation validates the assumptions made in Section 2 in a cosmological context, for a large sample of haloes as a function of time. Interestingly, the number of gaseous haloes decreases dramatically as the halo mass approaches the critical mass, where the model predicts a sharp upturn in gas mass (see Fig. 5). This is because the fraction of dark matter haloes that host a luminous galaxy increases steadily with halo mass above $M_{200} > 3 \times 10^8\,M_\odot\,h^{-1}$ and becomes unity above the critical mass. As a consequence, the number of starless gaseous haloes drops significantly near the critical mass. This validates our interpretation that the gas in RELHICs is expected to lose pressure support and collapse once the halo mass $M_{200} \gtrsim M_{cr}^z$. We will return to this point in Section 4.

A further test of the model is provided by the detailed radial structure of the RELHICs, which we characterize by the radii, r_f, encompassing a fraction, f, of the total gas mass. For the haloes in the simulation, r_f is obtained from the gas particles bound to the systems while for the model, r_f is calculated numerically. In order to measure r_f reliably in the simulation, we restrict our analysis to systems that contain more than 100 bound gas particles at the redshift of interest.

The top panel of Fig. 8 shows r_f, for $f = (0.9, 0.5, 0.2)$, as a function of halo virial mass at the same three redshifts as before. The orange circles and shaded regions show the running median and 10–90th percentiles for the simulated haloes. The blue lines show the results of our model. The agreement between the simulated density profiles and the predictions of the model of Section 2 is very good. This is a non-trivial result that further validates the model and demonstrates that the gaseous structure of RELHICs is well understood. The bottom panel of the same figure shows the concentration of RELHICs, defined as the ratio, r_{90}/r_{20}, as a function of halo mass. For comparison, the horizontal dashed line shows the expected concentration if the gas density profile were constant, i.e. $r_{90} = (9/2)^{1/3}r_{20}$.

We conclude that the detailed structure of RELHICs, i.e. post-reionization starless gaseous haloes in ΛCDM, can be understood from three simple considerations:

(i) The dark matter profiles of dark matter haloes are well described by the NFW form.
(ii) Gas in RELHICs is in thermal equilibrium with the external UVB radiation field.
(iii) The external pressure of RELHICs is equal to that of the intergalactic medium.

3.4 Synthetic mass assembly histories

The limited resolution of our simulation precludes following the formation of dark matter haloes below a minimum mass of a few $10^6\,M_\odot\,h^{-1}$. This imposes a maximum (mass-dependent) redshift above which our simulation is no longer valid. In order to overcome this limitation, we construct a set of synthetic assembly histories using the Extended Press–Schechter formalism (EPS; Bond et al. 1991; Bower 1991) and the method proposed by Lacey & Cole (1993). Specifically, for 50 $z = 0$ mass bins equally spaced logarithmically in the range $10^6 \leq M_{200}/M_\odot\,h^{-1} \leq 10^{12}$, we calculate 500 ΛCDM assembly histories per mass bin, from redshift $z = 36$ to $z = 0$, in 406 equally spaced redshift bins. We shall refer to these synthetic assembly histories as 'EPS assembly histories'.

As we demonstrate in Appendix C, our set of EPS assembly histories agree well with those measured in the simulation down to the resolution limit of the simulation ($M_{200} \sim 10^6\,M_\odot\,h^{-1}$).
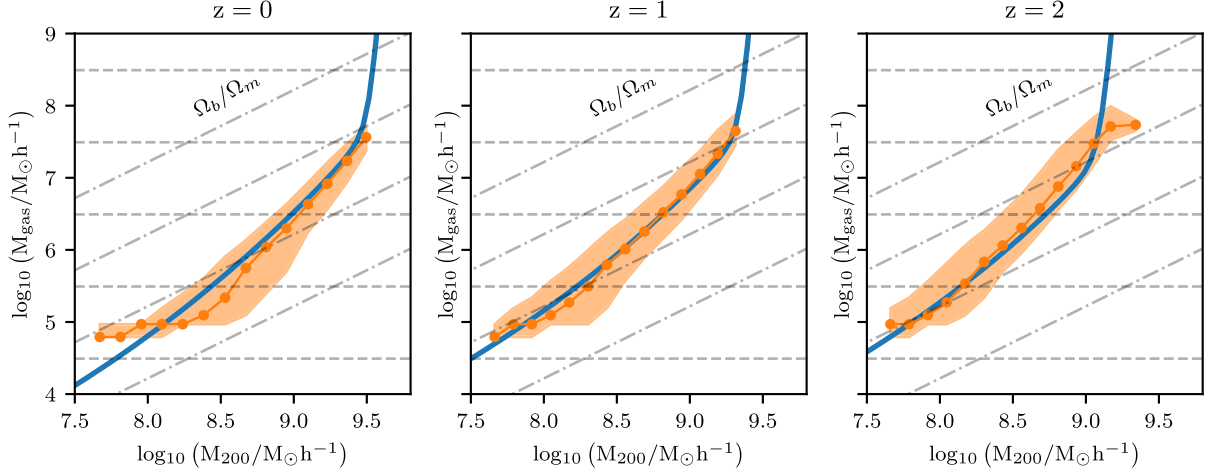
Figure 7. Gas mass within the virial radius of simulated starless gaseous haloes (RELHICs) as a function of halo mass at redshifts $z = 0$, 1, and 2. The blue solid lines show the gas mass derived from the model of Section 2 (see Fig. 5), which assumes gas in thermal equilibrium with the Haardt & Madau (2001) UVB radiation field and in hydrostatic equilibrium within its dark matter halo. The running median and the 10–90th percentiles of the simulation data are shown by orange circles and the shaded regions, respectively. Note that our model has no free parameters other than the intensity and shape of the UVB background. The agreement with the simulation results, and the correct redshift-dependence, are remarkable. The oblique dot-dashed grey lines correspond to multiples of the universal baryon fraction, $\alpha \times \Omega_b/\Omega_m$, with $\alpha = (1, 0.1, 0.01, 0.001)$. The horizontal dashed lines indicate, from top to bottom, the gas mass that corresponds to 10^4, 10^3, 10^2, 10, and 1 gas particles.
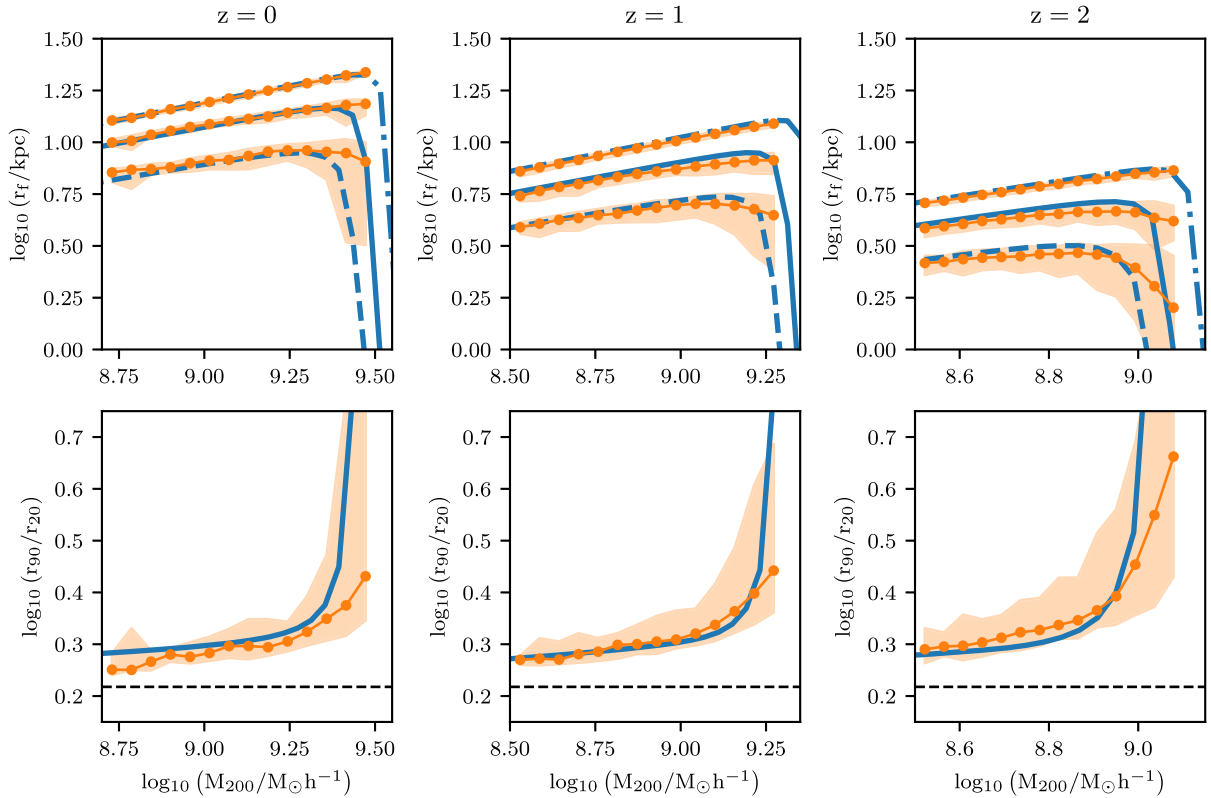

Figure 8. Top panels: the radius, r_f, encompassing different fractions, f, of the total gas mass of starless gaseous haloes (RELHICs) that contain more than 100 gas particles, as a function of halo mass and redshift. From top to bottom the different blue curves show the result of our analytical model of Section 2 for $f = (0.9, 0.5, 0.2)$. The orange circles and shaded regions show the running median and 10–90th percentiles of the distributions, respectively. Different panels correspond to three different redshifts as indicated. Bottom panels: concentration, r_{90}/r_{20}, of the gaseous profiles, also as a function of halo mass and redshift. For comparison, the horizontal dashed lines in the bottom panels show the trivial relation expected for a constant gas density profile (i.e. $r_{90}/r_{20} = (9/2)^{1/3}$). The agreement between the simulated RELHICs and the analytical model below the critical mass is remarkable.
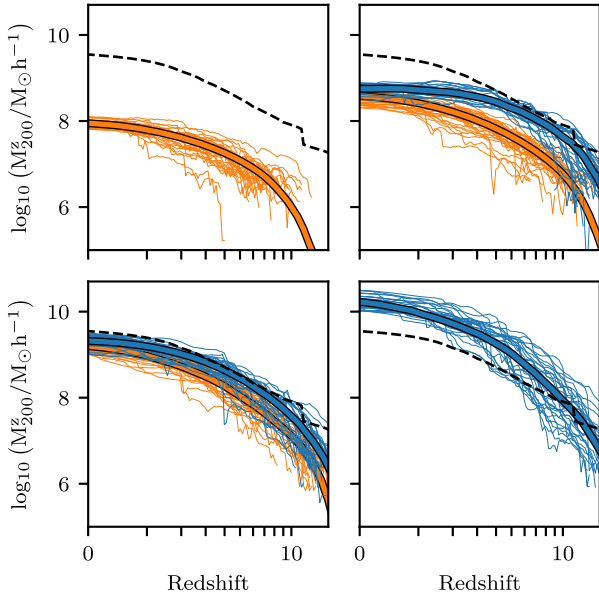
Figure 9. Mass assembly histories of haloes in our cosmological hydrody-namics simulation in four bins of present-day halo mass. The thin orange lines correspond to a random sample of field dark matter haloes that do not host a luminous galaxy at $z = 0$. The thin blue lines correspond to haloes that do contain a galaxy, i.e. more than one stellar particle, at their centre. The thick colour lines show the median of each set. The black dashed line indicates the critical mass for galaxy formation derived from our model (Section 2). The limit imposed by our model and the outcome of the simulation agree remarkably well, suggesting that the onset of galaxy formation can be understood in terms of our simple model.

4 RESULTS

4.1 Mass assembly histories and the onset of galaxy formation

Prior to reionization, we will assume that galaxies can only form in haloes in which atomic hydrogen cooling becomes efficient, as discussed in the Introduction. (We will test the impact of a different assumption in Section 4.3.) Once the universe is reionized, we can follow the fate of the gas in haloes using our model of Section 2. Gas is not able to collapse into every dark matter halo so not all haloes will host a visible galaxy. We define the Halo Occupation Fraction (HOF) as the fraction of dark matter haloes that host a luminous galaxy, as a function of redshift and halo mass. Similar definitions have been adopted by other authors to quantify the ability of dark matter haloes to host stars prior to reionization (e.g. Xu et al. 2016).

The onset of galaxy formation is closely related to the assembly history of ΛCDM haloes and to the redshift-dependent critical mass for gas to lose pressure support and collapse. As we have seen, after reionization the critical mass is set by the presence of the UVB radiation field; before reionization further assumptions are required.

Fig. 9 shows the assembly history of a random selection of haloes identified in our ΛCDM cosmological hydrodynamics simulation, split into four bins according to their mass at $z = 0$. A variety of different formation histories can give rise to the same final mass. The black dot-dashed line shows the critical mass for galaxy formation derived in Section 2, assuming a constant concentration, $c = 10$, for all haloes, a redshift of reionization, $z_{\rm re} = 11.5$ (i.e. consistent with that in our simulation) and a critical mass for galaxy formation prior to cosmic reionization described by equation (1). Haloes that do not host a luminous galaxy at $z = 0$ are shown by orange lines; those

that have a stellar component (of at least one star particle) at $z = 0$ are shown by blue lines. The medians of each set are shown by the thick lines of the corresponding colour.

The top left-hand panel shows that the mass of all dark matter haloes with present-day mass, $M_{200} \sim 10^8\,{\rm M}_\odot\,h^{-1}$, always remain below the critical mass required for galaxies to form. As expected, none of these haloes hosts a galaxy in our simulation. More massive haloes may or may not cross the critical mass for galaxy formation; only those whose mass exceeded, or was comparable to the critical mass some time in the past will host a luminous galaxy at $z = 0$. Remarkably, we find in the simulation that indeed the only dark matter haloes that contain a galaxy at their centre at $z = 0$ are those whose mass was above or close to the critical mass in the past (see e.g. top right-hand panel). Moreover, we find that galaxy formation actually begins soon after the haloes cross this critical mass (see Appendix B).

As the present-day halo mass increases, so does the fraction of haloes that cross the critical mass. If the mass today is large enough so that every dark matter halo once exceeded the critical mass, then the fraction of dark matter haloes hosting a luminous galaxy today becomes unity. This trend was already seen in a number of studies of the formation of dwarf galaxies in cosmological hydrodynamical simulations (e.g. Hoeft et al. 2006; Okamoto et al. 2008; Okamoto & Frenk 2009; Benítez-Llambay et al. 2015; Sawala et al. 2016; Fitts et al. 2017, and references therein), but the exact value of the critical mass for galaxy formation is somewhat uncertain, although it is not far from the mass corresponding to a virial temperature, $T_{200} \sim 10^4$ K. This uncertainty is removed by our model which can be used to understand quantitatively the HOF.

The key to understanding the HOF is the mass accretion history. To demonstrate this, we select all field dark matter haloes in our simulation identified by HBT+ and compare their growing mass, determined by their individual assembly histories, to the critical mass, $M_{\rm cr}^z$. If the mass of a simulated halo exceeds the critical mass at some redshift, we label this halo as luminous, that is, we assume that gas will be able to collapse and form a galaxy at its centre. Otherwise we label the halo as dark. We then calculate the HOF as a function of halo mass, based on this classification only and compare it to the fraction of haloes that actually host a galaxy in the simulation.

Fig. 10 compares the HOF measured in three ways: directly in our cosmological simulation at $z = 0$, from our model applied to the assembly histories inferred in the simulation and from our model applied to synthetic assembly histories derived from EPS theory. The agreement amongst all three is excellent. This demonstrates that the HOF is a result of the existence of a critical halo mass for galaxies to form and the stochastic nature of the growth of ΛCDM haloes. We thus conclude the following (where the numbers are valid for the Haardt & Madau 2001 UVB assumed in our simulation):

(i) At $z = 0$ field dark matter haloes less massive than $3 \times 10^8\,{\rm M}_\odot\,h^{-1}$ do not host a luminous galaxy. The gravitational force in these haloes is not strong enough to overcome the pressure of the gas imparted by the UVB.

(ii) All dark matter haloes more massive than $M_{200} > 5 \times 10^9\,{\rm M}_\odot\,h^{-1}$ host a galaxy at $z = 0$. Gas pressure cannot prevent gravitational collapse. Once dark matter haloes exceed this mass, they will inevitably form a galaxy at their centre.

(iii) The HOF in the mass range $3 \times 10^8 \lesssim M_{200}/{\rm M}_\odot\,h^{-1} \lesssim 5 \times 10^9$ is determined by the scatter in the assembly histories of ΛCDM haloes.

Two main assumptions lie behind these conclusions: (i) galaxy formation is largely determined by atomic hydrogen cooling prior
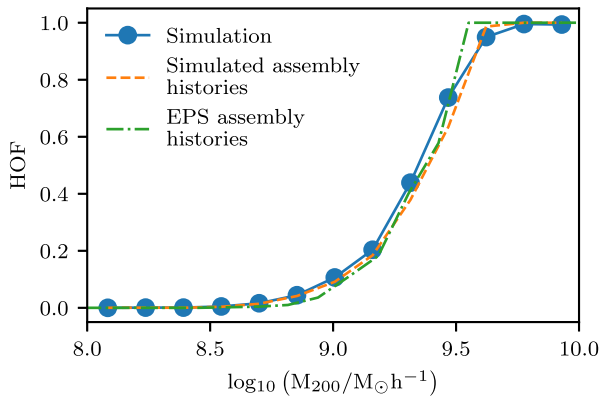
Figure 10. Comparison of the $z = 0$ halo occupation fraction (HOF) measured directly in our cosmological simulation (blue circles), derived from the actual halo assembly histories measured in the simulation (orange dashed line) and derived assuming EPS assembly histories (green dot-dashed line). The agreement amongst all the curves demonstrates that the onset of galaxy formation in our cosmological hydrodynamical simulation is correctly described by the simple model of Section 2 coupled to the assembly histories of ΛCDM haloes.

to cosmic reionization; (ii) cosmic reionization occurred at $z_{\rm re} = 11.5$. We now discuss how sensitive our conclusions are to variations of these assumptions. First we consider the impact of varying the redshift of reionization and then of relaxing our criterion for galaxy formation in low-mass haloes prior to reionization. To this end, we will use our set of EPS assembly histories.

4.2 The effect of $z_{\rm re}$ on the $z = 0$ halo occupation fraction

To assess the effect of the redshift of reionization, $z_{\rm re}$, on the HOF at $z = 0$, we use our set of EPS assembly histories and vary the value of $z_{\rm re}$. This has a direct impact on the value of the critical mass for galaxy formation. For $z > z_{\rm re}$, we assume that galaxy formation occurs only in haloes in which atomic hydrogen cooling is efficient. We model the effect of reionization by assuming that the cosmic gas is suddenly placed on the density–temperature relation shown in Fig. 4, for the appropriate redshift.

Fig. 11 shows the result of this calculation. The HOF is clearly sensitive to the assumed value of $z_{\rm re}$ unless reionization occurred either early or very late. For example, the bottom panel of Fig. 11 shows that for $z_{\rm re} > 10$ (thick line in the top panel), the halo mass at which only half the dark matter haloes host a luminous galaxy is $M_{200,50} \sim 10^{9.4}\,{\rm M}_\odot\,h^{-1}$. On the other hand, if the universe is reionized at $z = 4$, the value of $M_{200,\,50}$ is a factor of 4 lower. In the limiting case of a universe that experiences very late reionization (or no reionization), the assumption that galaxies form through atomic hydrogen cooling yields $M_{200,50} \sim 10^{8.6}\,{\rm M}_\odot\,h^{-1}$. For comparison, the left red dashed line in the top panel shows the HOF measured in a zoom-in simulation of the Local Group carried out by Sawala et al. (2016) that does not include reionization. Reassuringly, our model reproduces their simulation results very well. We also plot the result of our own simulation, in which cosmic reionization occurs at $z_{\rm re} = 11.5$, i.e. early enough that the measured HOF is independent of the exact value of $z_{\rm re}$.

We conclude that if galaxy formation is largely determined by atomic hydrogen cooling prior to cosmic reionization, then:

(i) the $z = 0$ HOF is insensitive to the redshift of reionization provided $z_{\rm re} > 10$ and $z_{\rm re} < 2$.
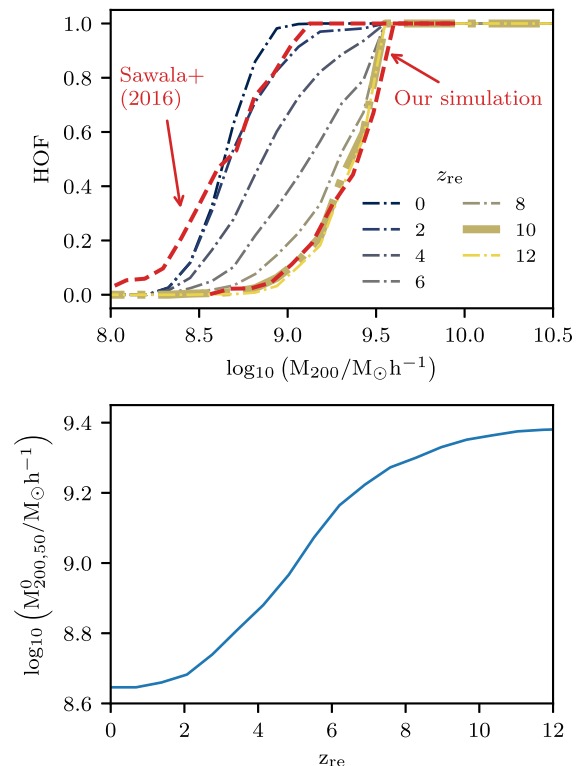

Figure 11. Top panel: the effect of varying the redshift of reionization on the $z = 0$ halo occupation fraction. If reionization occurred early, $z_{\rm re} \gtrsim 10$, the $z = 0$ HOF is largely insensitive to the exact value of $z_{\rm re}$. On the other hand, if $z_{\rm re} \lesssim 10$, then $z_{\rm re}$ has a dramatic effect on the $z = 0$ HOF. In a universe that never experienced reionization and in which galaxies formed through atomic hydrogen cooling, all haloes above $M_{200} \gtrsim 10^{8.9}\,{\rm M}_\odot\,h^{-1}$ would host a luminous galaxy. As a comparison, we show the result of a zoom-in simulation of the Local Group by Sawala et al. (2016) that does not include reionization (left red dashed line). The agreement between that simulation and the predictions using our framework is reassuring. We also plot the result of our simulation, in which reionization occurs at $z_{\rm re} = 11.5$, a redshift at which the HOF is expected to become largely independent of the exact value of $z_{\rm re}$. Bottom panel: the present-day halo mass at which the halo occupation fraction, HOF = 0.5, as a function of $z_{\rm re}$. If galaxies form though atomic hydrogen cooling prior to reionization, then the HOF becomes independent of the value of $z_{\rm re}$ for both, $z_{\rm re} > 10$ and $z_{\rm re} < 2$.

(ii) the $z = 0$ HOF depends strongly on the redshift of reionization if $2 < z_{\rm re} < 10$.

The model of Section 2 thus provides a simple framework to account for both the qualitative and quantitative effect of $z_{\rm re}$ on the HOF. We now turn our attention to the impact of varying the critical mass for galaxy formation.

4.3 Pre-reionization galaxy formation and the halo occupation fraction

We now investigate the sensitivity of the HOF to assumptions about how galaxy formation proceeds prior to reionization. For the purposes of this discussion, we fix the redshift of reionization to $z_{\rm re} = 11.5$.

Fig. 12 shows the $z = 0$ HOF obtained from our model, using our set of EPS assembly histories for different assumptions. Specifically, we consider the limiting case where all dark matter haloes reach $z_{\rm re}$ without hosting a luminous galaxy (green solid curve), as well as the cases when galaxy formation before $z_{\rm re}$ occurs in haloes of mass
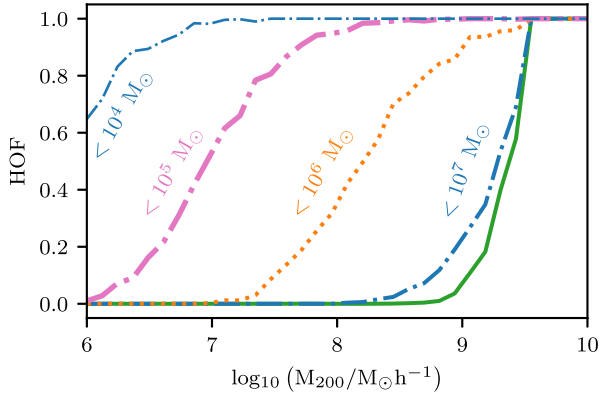
Figure 12. Halo Occupation Fraction (HOF) at the present day as a function of halo mass predicted by our model (Section 2), assuming a redshift of reionization, $z_{re} = 11.5$. The green solid line shows the limiting case of assuming that all dark matter haloes reach z_{re} without hosting a luminous galaxy; the other curves assume that all haloes that exceeded the quoted mass prior to reionization host a galaxy by z_{re}. The $z = 0$ HOF depends strongly on assumptions about galaxy formation prior to reionization, but only if galaxy formation proceeds in haloes well below the atomic hydrogen cooling limit at z_{re}. For $z_{re} \sim 11.5$, this corresponds to $M_{200} \sim 10^7 \, h^{-1} \, \mathrm{M_\odot}$.

$M_{200} > 10^7 \, \mathrm{M_\odot} \, h^{-1}$ (blue dot-dashed line), $10^6 \, \mathrm{M_\odot} \, h^{-1}$ (orange dotted line), $10^5 \, \mathrm{M_\odot} \, h^{-1}$ (pink dot-dashed line), and $10^4 \, \mathrm{M_\odot} \, h^{-1}$ (blue dot-dashed line). Not surprisingly, these different assumptions about pre-reionization galaxy formation have a dramatic impact on the HOF.

The critical mass for atomic hydrogen cooling to become efficient at $z = 11.5$ is $M_H^{11.5} \sim 3 \times 10^7 \, \mathrm{M_\odot} h^{-1}$. For this value of z_{re}, the present-day HOF varies little if we assume that all dark matter haloes reach z_{re} without a luminous galaxy in them or only those with $M_{200} \sim M_H^{z_{re}}$. However, if we allow less massive haloes to be populated with luminous galaxies prior to reionization – such as haloes of mass $M_{200} = 10^6 \, \mathrm{M_\odot} \, h^{-1}$ at z_{re} – then the HOF becomes very sensitive to our assumptions.

One caveat of the exercise carried out in Fig. 12 is that the halo occupation fraction prior to reionization may not be a simple step function. Although the formation of the first galaxies depends primarily on halo mass, it may also depend on secondary parameters possibly related to the local environment (e.g. Couchman & Rees 1986; Yoshida et al. 2003, 2006, and references therein). Recent numerical simulations by Wise et al. (2014) show, however, that although the halo occupation fraction is not a step function prior to reionization, the transition between luminous and starless haloes occurs over a very narrow range of halo mass at $z = 15$ (roughly a factor of 3 in halo mass below the atomic hydrogen cooling limit).

We will next show that if the redshift of reionization is lower than $z_{re} = 10$, the HOF depends quite strongly on the assumptions about galaxy formation before reionization whereas if the redshift is higher, the converse is true.

4.4 Combined impact of z_{re} and pre-reionization galaxy formation

We now explore how the $z = 0$ HOF depends on variations of both the assumed redshift of reionization, z_{re}, and the pre-reionization critical mass for galaxy formation, $M_c^{z_{re}}$. For each value of z_{re}, we compute the HOF by comparing the individual EPS assembly histories to the critical mass derived from our model (Section 2), assuming that haloes more massive than $M_c^{z_{re}}$ at z_{re} are already luminous at z_{re}. We
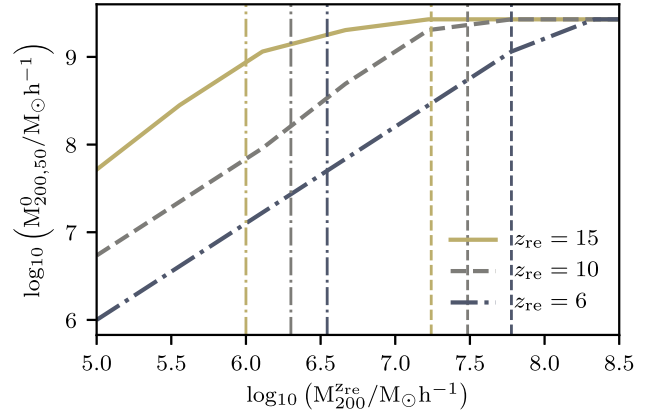
Figure 13. Halo mass at $z = 0$ at which 50 per cent of haloes host a galaxy, as a function of the halo mass at z_{re} above which all haloes are assumed to have become luminous prior to z_{re}. Different colours correspond to different values of the redshift of reionization, z_{re}, as indicated in the legend. As an example, if all haloes more massive than $10^6 \, \mathrm{M_\odot}$ already host a luminous galaxy by $z_{re} = 10$, then 50 per cent of the haloes with halo mass, $M_{200} \sim 6 \times 10^7 \, \mathrm{M_\odot}$ at $z = 0$ will also host a luminous galaxy. The vertical dashed lines show the critical atomic cooling mass at z_{re} (i.e. lines of constant temperature, $T_{200} = 8000 \, \mathrm{K}$). Vertical dot-dashed lines correspond to a constant virial temperature, $T_{200} = 1000 \, \mathrm{K}$.

express the outcome of this calculation as the halo mass at $z = 0$ at which only half of the dark matter haloes are luminous, $M_{200,50}^0$, as a function of both, z_{re} and $M_c^{z_{re}}$. This is shown in Fig. 13 for three different assumed values of z_{re}.

Fig. 13 shows that both z_{re} and $M_c^{z_{re}}$ affect the characteristic mass scale at which 50 per cent of the haloes are luminous at $z = 0$. As discussed in the previous section, if reionization occurred early, the characteristic halo mass at $z = 0$ at which 50 per cent of the haloes remain dark is roughly independent of assumptions about early galaxy formation, as long as galaxies form in haloes more massive than $10^6 \, \mathrm{M_\odot}$ prior to $z_{re} = 15$. However, if reionization occurs much later, then the critical halo mass for galaxy formation before reionization has a much greater impact on the $z = 0$ HOF.

Detailed cosmological simulations of the formation of the first galaxies – which include cooling processes, such as molecular hydrogen cooling, ignored by our simulation – seem to support the idea that the process of galaxy formation is largely triggered only in haloes in which atomic hydrogen cooling becomes efficient, i.e. those more massive than $2 \times 10^7 \, \mathrm{M_\odot} \, h^{-1}$ at $z \sim 15$ (see e.g. Xu, Wise & Norman 2013; Xu et al. 2016, and references therein). For such high redshift of reionization, our results are already robust even if we allow haloes 10 times smaller than this to host a galaxy by that time.

If the redshift of reionization is much lower, $z = 6$ say, then the critical halo mass at which 50 per cent of the dark matter haloes host a galaxy by today becomes independent of the exact assumptions about early galaxy formation, provided galaxy formation occurs in haloes more massive than $\sim 10^8 \, \mathrm{M_\odot} \, h^{-1}$ before reionization. This mass is not far from the atomic hydrogen cooling limit at z_{re} (vertical green dashed line), but allowing low-mass haloes with virial temperature $T_{200} \sim 1000 \, \mathrm{K}$ (vertical dot-dashed lines) to host a galaxy at z_{re}, has a large impact on the $z = 0$ HOF.

We conclude that the lower the redshift of reionization, the more dependent the HOF is on assumptions about pre-reionization galaxy formation. Neglecting cooling processes that are important below a virial temperature of $10^4 \, \mathrm{K}$, as we have done, might affect the expected $z = 0$ HOF if reionization occurred as late as $z \leq 6$.

5 DISCUSSION

The agreement between the model of Section 2 and our high-resolution cosmological hydrodynamic simulation suggests that the onset of galaxy formation in low-mass haloes, as well as the detailed structure of RELHICs, can be understood in detail in the context of the ΛCDM cosmology. The key inputs are: (i) the redshift at which hydrogen was reionized, (ii) the redshift-dependent temperature of the intergalactic medium after reionization (which is largely set by the external UVB radiation field), and (iii) the minimum halo mass in which galaxies can form prior to reionization. Some of the predictions of our model are relatively insensitive to the details of these inputs.

If reionization occurred in the redshift range, $6 < z_{\rm re} < 10$, as indicated by recent Planck results (Planck Collaboration VI 2020), then the minimum halo mass today that can host a luminous galaxy cannot be much smaller than $M_{200} \sim 10^8 \, {\rm M}_\odot \, h^{-1}$, unless galaxy formation was able to take place in haloes of virial temperature $T_{200} < 2 \times 10^3 \, {\rm K}$ prior to reionization (or $M_{200} \lesssim 3 \times 10^6 \, {\rm M}_\odot \, h^{-1}$ at $z_{\rm re} = 10$). Recent hydrodynamical simulations indicate that this is unlikely (Xu et al. 2016). Thus, the $M_{200} \sim 10^8 \, {\rm M}_\odot \, h^{-1}$ mass scale should be a general outcome of simulations of galaxy formation, regardless of the details of the physical modelling.

Shen et al. (2014) studied the formation of dwarf galaxies in seven haloes in the mass range $3 \times 10^8 \lesssim M_{200} \, {\rm M}_\odot \, h^{-1} \lesssim 2.4 \times 10^{10}$, in high-resolution zoom simulations carried out with the GASOLINE code. They assumed the Haardt & Madau (2012) UVB external radiation field and a redshift of reionization, $z_{\rm re} = 6$. They found that three of their haloes, of mass less than $M_{200} \lesssim 4.5 \times 10^8 \, {\rm M}_\odot \, h^{-1}$, remained starless at $z = 0$. They also found that the gas density in these 'dark' haloes had never exceeded $n_{\rm H} \sim 10^{-2} \, {\rm cm}^{-3}$, a value well below the imposed density threshold for star formation in their simulations ($n_{\rm H} = 100 \, {\rm cm}^{-3}$). These results are in excellent agreement with ours and are readily understood in terms of our model of Section 2. In this mass range, our model predicts a maximum central gas density almost two orders of magnitude lower than that reached in haloes that are close to the critical mass (see bottom panel of Fig. 3).

The same minimum halo mass for galaxy formation also emerges from the high-resolution zoom cosmological simulations of dwarf galaxies carried out by Munshi et al. (2019), also with the GASO-LINE code, but with different assumptions for star formation. They found that both simulations that include, and simulations that ignore molecular hydrogen cooling make galaxies only in field haloes more massive than $M_{200} \gtrsim 3 \times 10^8 \, {\rm M}_\odot \, h^{-1}$ at $z = 0$, as predicted by our model. The authors find that neglecting molecular hydrogen cooling produces twice as many luminous galaxies as when this cooling process is included. This counterintuitive result, which is at odds with the expectations of our model, may simply reflect the particular treatment of star formation in the latter simulations, which is closely tied to the fraction of molecular hydrogen. There is observational (e.g. Michałowski et al. 2015) and theoretical (e.g. Hu et al. 2016) evidence suggesting that the presence of molecules is a sufficient, but not a necessary condition for star formation, which is mostly regulated by atomic hydrogen. For example, the high-resolution cosmological hydrodynamical simulations of Xu et al. (2013) that include radiative transfer, non-equilibrium chemistry and cooling by molecular hydrogen are consistent with the onset of galaxy formation being regulated mainly by atomic, rather than molecular hydrogen cooling prior to reionization.

Thus, our results, together with those from other authors, provide strong evidence for the existence of a minimum halo mass below which galaxy formation cannot proceed. The descendents of these barren haloes have mass, $M_{200} \lesssim 10^8 \, {\rm M}_\odot \, h^{-1}$, at $z = 0$. Finding a luminous galaxy in a dark matter halo close to this mass scale should be rare, but not impossible. For example, our simulation indicates that there should be $\sim 3.4 \times 10^{-2} \, {\rm Mpc}^{-3}$ luminous galaxies inhabiting haloes of mass $M_{200} \lesssim 10^9 \, {\rm M}_\odot \, h^{-1}$. Claims challenging the ΛCDM model on the grounds that there are nearby field dwarfs surrounded by low-mass haloes of mass $M_{200} \lesssim 10^{10} \, {\rm M}_\odot$ (Ferrero et al. 2012) should therefore be treated with caution.

Finally, our results also indicate that field dark matter haloes above the mass scale, $M_{200} \gtrsim 5 \times 10^9 \, {\rm M}_\odot \, h^{-1}$, should always host a luminous galaxy at $z = 0$. As far as we are aware, no simulation of galaxy formation has ever found a dark matter halo with mass above this mass scale that remains devoid of stars at $z = 0$ (e.g. Benítez-Llambay et al. 2015; Sawala et al. 2016; Fitts et al. 2017; Munshi et al. 2019, and references therein).

6 SUMMARY AND CONCLUSIONS

We have developed a simple theoretical framework, in the context of the ΛCDM cosmological model, to follow the formation of the first galaxies before and after the reionization of hydrogen at early times. In particular, we have introduced a simple analytical model to calculate the properties of the gas component of haloes after reionization. Within this framework we have addressed two specific problems: the mass of the smallest haloes that can sustain galaxy formation, and the properties of RELHICs – starless gaseous haloes whose gas was unable to collapse by the time the universe reionized. We introduced the concept of the HOF, which we have defined as the fraction of dark matter haloes that host a luminous galaxy today as a function of halo mass.

In our model galaxy formation can only take place in haloes whose mass exceeds a redshift-dependent critical value, $M_{\rm cr}^z$. Before reionization, $M_{\rm cr}^z$ is the mass at which atomic hydrogen cooling becomes effective, $\sim 4.8 \times 10^7 \, {\rm M}_\odot$ at $z = 7$, for example. After reionization, $M_{\rm cr}^z$ is the mass above which gas, which was previously kept in equilibrium in a halo by the pressure of the UVB radiation, can no longer remain in hydrostatic equilibrium and therefore collapses to make a galaxy. The HOF is then the result of the interplay between the evolution of this critical mass and the assembly history of haloes. It is completely determined by three factors: the minimum halo mass for galaxy formation before reionization; the redshift of reionization, $z_{\rm re}$; and the intensity of the (evolving) external photoheating rate.

The properties of haloes whose gaseous component has not yet collapsed to make a luminous galaxy by $z_{\rm re}$ are determined by the effects of reionization. After reionization, the temperature of the gas in the high-density intergalactic medium is set by the balance between radiative cooling and photoheating and, in low density gas, by photoheating alone. We developed a simple analytical model to study how the properties of the ionized intergalactic medium affects the properties of gaseous haloes as a function of time.

We validated our model by comparing with results of a cosmological hydrodynamics simulation that models the main astrophysical processes involved in galaxy formation. The main conclusions arising from our model may be summarized as follows.

(i) If, prior to reionization, galaxies form predominantly from gas that condenses through atomic hydrogen cooling, then the HOF at $z = 0$ is almost insensitive to $z_{\rm re}$ provided $z_{\rm re} > 10$. We tested two limiting cases, $z_{\rm re} > 10$ and $z_{\rm re} < <10$, with our high-resolution

cosmological hydrodynamical simulation and found that the HOF agrees remarkably well with the predictions of our model (see Fig. 11).

(ii) If the redshift of reionization is $z_{re} > >10$, then the $z = 0$ HOF becomes insensitive to the details of how gas cools within haloes before reionization. For example, if $z_{re} = 15$, and galaxies were able to form before this redshift in haloes more massive than $10^6 \, M_\odot \, h^{-1}$ (in which molecular hydrogen cooling could be relevant) there would be very little change in the $z = 0$ HOF.

(iii) If the redshift of reionization is $z_{re} < 10$, the $z = 0$ HOF depends sensitively on the minimum halo mass that can host a luminous galaxy before reionization. For example, if gas cooled mainly by molecular (rather than atomic) hydrogen processes in haloes of virial temperature $T_{200} \sim 10^3$ K, then the mass scale at which 50 per cent of the haloes are luminous at $z = 0$ can vary by more than a factor of 10 compared to the atomic hydrogen cooling case (see e.g. Fig. 13).

(iv) The potential well of dark matter haloes less massive than $M_{200} \lesssim 10^6 \, M_\odot \, h^{-1}$ at $z = 0$ is not deep enough to overcome the pressure force produced by the external Haardt & Madau (2001) UVB radiation field. The gas completely evaporates from these haloes which are therefore gas free at $z = 0$. They will also be star-free unless galaxy formation occurred in haloes of mass $M_{200} \sim 10^4 \, M_\odot \, h^{-1}$ prior $z_{re} > 10$, which is unlikely.

(v) The gas content of dark matter haloes in the mass range $10^6 \lesssim M_{200}/M_\odot \, h^{-1} \lesssim 5 \times 10^9$ at $z = 0$ is drastically reduced (but not completely removed) by the presence of the UVB radiation field. A small fraction of the remaining gas is prevented from collapsing by the external radiation field, with which it is in thermal equilibrium, and is confined, in hydrostatic equilibrium, within the potential well of the halo, without forming stars. These are the objects we previously identified in cosmological hydrodynamics simulations and called RELHICs (Reionization limited H I clouds; Benítez-Llambay et al. 2017). The gas density profiles of RELHICs are flat at large distances from the halo centre and have a well-defined maximum value at the centre; this latter property is a direct consequence of the gas being in equilibrium in the cuspy dark matter distribution of ΛCDM haloes.

(vi) The amount of gas required to overcome the gravitational acceleration of dark matter haloes more massive than $M_{200} \sim 5 \times 10^9 \, M_\odot \, h^{-1}$ exceeds the universal baryon fraction (Ω_b/Ω_m) by orders of magnitude. This gas cannot remain in hydrostatic equilibrium and will therefore collapse to make a luminous galaxy.

We tested the results of our analytical model by comparing with cosmological hydrodynamics simulations. We find that our predicted HOF is reproduced remarkably well in the simulations. Similarly, the gas mass and internal structure of RELHICs predicted by our model are reproduced very well in the simulations.

Our results apply strictly to field (or central) dark matter haloes, not to satellites of more massive systems. Unlike field haloes whose mass can only grow with time, satellites can lose mass by tidal stripping. Note, however, that our model applies to satellites before they fell into their host halo.

Our simulation indicates that there should be fewer than 30 field luminous galaxies inhabiting dark matter haloes of mass $M_{200} \lesssim 10^9 \, M_\odot \, h^{-1}$ in the Local Volume.[10] Finding many more nearby field dwarfs in such low-mass haloes would require either a revision of our current understanding of how (and when) reionization occurred or of how galaxy formation proceeded before reionization. Interestingly,

Jethwa, Erkal & Belokurov (2018) and Graus et al. (2019) have recently claimed, based on simulations, that such a population of galaxies might be required to explain the sheer number and radial distribution of ultrafaint galaxies observed around the Milky Way but Bose et al. (2020) argue that these results are affected by numerical resolution and that, when these effects are taken into account, the data are in good agreement with the observed number of ultrafaint satellites. In any case, our model provides a theoretical framework to address these issues and interpret that data without the need for cosmological simulations and thus, independent of computational cost or resolution limitations.

In summary, if the onset of galaxy formation is largely determined by atomic hydrogen cooling in small haloes before reionization, then we expect a cutoff in the galaxy mass function at a halo mass, $M_{200} \sim 10^8 \, M_\odot \, h^{-1}$, at $z = 0$. Finding an isolated luminous galaxy in a dark matter halo below this mass scale would be a remarkable discovery that would challenge not only the current cosmological model, but also current understanding of galaxy formation.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

---

[10]We take the Local Volume to be sphere of 10 Mpc centred on the Milky Way, and assume that our simulation is representative of the Local Volume.

REFERENCES

Babul A., Rees M. J., 1992, MNRAS, 255, 346
Barkana R., Loeb A., 1999, ApJ, 523, 54
Benítez-Llambay A., Navarro J. F., Abadi M. G., Gottlöber S., Yepes G., Hoffman Y., Steinmetz M., 2015, MNRAS, 450, 4207
Benítez-Llambay A. et al., 2017, MNRAS, 465, 3913
Benítez-Llambay A., Frenk C. S., Ludlow A. D., Navarro J. F., 2019, MNRAS, 488, 2387
Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002a, MNRAS, 333, 156
Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002b, MNRAS, 333, 177
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bose S., Deason A. J., Belokurov V., Frenk C. S., 2020, MNRAS, 495, 743
Bower R. G., 1991, MNRAS, 248, 332
Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517
Couchman H. M. P., Rees M. J., 1986, MNRAS, 221, 53

Crain R. A. et al., 2015, MNRAS, 450, 1937
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Efstathiou G., 1992, MNRAS, 256, 43P
Ferrero I., Abadi M. G., Navarro J. F., Sales L. V., Gurovich S., 2012, MNRAS, 425, 2817
Fitts A. et al., 2017, MNRAS, 471, 3547
Gnedin N. Y., 2000, ApJ, 542, 535
Graus A. S., Bullock J. S., Kelley T., Boylan-Kolchin M., Garrison-Kimmel S., Qi Y., 2019, MNRAS, 488, 4585
Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays, p. 64
Haardt F., Madau P., 2012, ApJ, 746, 125
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Han J., Jing Y. P., Wang H., Wang W., 2012, MNRAS, 427, 2437
Han J., Cole S., Frenk C. S., Benitez-Llambay A., Helly J., 2018, MNRAS, 474, 604
Hoeft M., Yepes G., Gottlöber S., Springel V., 2006, MNRAS, 371, 401
Hu C.-Y., Naab T., Walch S., Glover S. C. O., Clark P. C., 2016, MNRAS, 458, 3528
Ikeuchi S., 1986, Ap&SS, 118, 509
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Jethwa P., Erkal D., Belokurov V., 2018, MNRAS, 473, 2060
Jorgenson R. A., Murphy M. T., Thompson R., Carswell R. F., 2014, MNRAS, 443, 2783
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Krumholz M. R., Ellison S. L., Prochaska J. X., Tumlinson J., 2009, ApJ, 701, L12
Lacey C., Cole S., 1993, MNRAS, 262, 627
Ludlow A. D., Navarro J. F., Angulo R. E., Boylan-Kolchin M., Springel V., Frenk C., White S. D. M., 2014, MNRAS, 441, 378
Michałowski M. J. et al., 2015, A&A, 582, A78
Munshi F., Brooks A. M., Christensen C., Applebaum E., Holley-Bockelmann K., Quinn T. R., Wadsley J., 2019, ApJ, 874, 40
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Okamoto T., Frenk C. S., 2009, MNRAS, 399, L174
Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920
Planck Collaboration et al., 2014, A&A, 571, A1
Planck Collaboration et al., 2020, A&A, 641, A6
Press W. H., Schechter P., 1974, ApJ, 187, 425
Quinn T., Katz N., Efstathiou G., 1996, MNRAS, 278, L49
Rees M. J., 1986, MNRAS, 218, 25P
Sawala T. et al., 2016, MNRAS, 456, 85
Schaye J., 2001, ApJ, 562, L95
Schaye J. et al., 2015, MNRAS, 446, 521
Shen S., Madau P., Conroy C., Governato F., Mayer L., 2014, ApJ, 792, 99
Somerville R. S., 2002, ApJ, 572, L23
Springel V., 2005, MNRAS, 364, 1105
Sykes C., Fumagalli M., Cooke R., Theuns T., Benítez-Llambay A., 2019, MNRAS, 487, 609
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Thoul A. A., Weinberg D. H., 1996, ApJ, 465, 608
Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, Nature, 585, 39
Wiersma R. P. C., Schaye J., Smith B. D., 2009, MNRAS, 393, 99
Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, MNRAS, 442, 2560
Xu H., Wise J. H., Norman M. L., 2013, ApJ, 773, 83
Xu H., Wise J. H., Norman M. L., Ahn K., O'Shea B. W., 2016, ApJ, 833, 84
Yoshida N., Abel T., Hernquist L., Sugiyama N., 2003, ApJ, 592, 645
Yoshida N., Omukai K., Hernquist L., Abel T., 2006, ApJ, 652, 6

APPENDIX A: EFFECTIVE DENSITY–TEMPERATURE RELATION OF STAR-FREE GASEOUS HALOES

Table A1 summarizes the density–temperature relation used in this work which is derived from the Wiersma et al. (2009) cooling tables.

Table A1. Temperature–density relation at different redshifts derived from the Wiersma et al. (2009) cooling tables, which assume an optically thin gas in ionization equilibrium, illuminated by the Haardt & Madau (2001) UVB radiation field. For low-density gas the relation is such that the photoheating time-scale equals the Hubble time. For high-density gas, the temperature is such that the cooling rate equals the photoheating rate. Entries in the table are $\log_{10}(T/K)$.

$\log_{10}\left(n_{\mathrm{H}}/\mathrm{cm}^{-3}\right)$	−8.0	−7.5	−7.0	−6.5	−6.0	−5.5	−5.0	−4.5	−4.0	−3.5	−3.0	−2.5	−2.0	−1.5	−1.0	−0.5	0.0
$z = 0.0$	2.89	3.16	3.45	3.73	4.02	4.28	4.49	4.54	4.43	4.30	4.20	4.13	4.08	4.03	4.00	3.97	3.94
$z = 1.1$	2.83	3.00	3.23	3.49	3.76	4.03	4.29	4.51	4.63	4.53	4.40	4.29	4.20	4.14	4.09	4.04	4.00
$z = 2.0$	2.79	2.91	3.10	3.33	3.59	3.86	4.13	4.37	4.59	4.57	4.47	4.36	4.26	4.18	4.12	4.07	4.03
$z = 3.0$	2.81	2.89	3.03	3.24	3.49	3.75	4.02	4.27	4.51	4.54	4.44	4.33	4.23	4.17	4.11	4.06	4.02
$z = 4.1$	2.78	2.84	2.96	3.15	3.39	3.65	3.93	4.20	4.43	4.49	4.39	4.29	4.21	4.15	4.09	4.05	4.01
$z = 5.2$	2.70	2.76	2.88	3.07	3.31	3.57	3.84	4.13	4.35	4.43	4.35	4.26	4.18	4.13	4.08	4.03	4.00
$z = 6.1$	2.60	2.69	2.84	3.03	3.27	3.53	3.79	4.05	4.27	4.35	4.31	4.23	4.16	4.11	4.06	4.02	3.99
$z = 7.2$	2.44	2.54	2.72	2.93	3.19	3.45	3.71	3.95	4.15	4.29	4.27	4.21	4.15	4.10	4.05	4.01	3.98
$z = 8.0$	2.28	2.42	2.60	2.85	3.11	3.37	3.63	3.85	4.03	4.21	4.25	4.19	4.14	4.09	4.04	4.01	3.98
$z = 9.0$	2.04	2.24	2.48	2.73	3.03	3.27	3.51	3.71	3.91	4.11	4.22	4.18	4.12	4.08	4.03	4.00	3.97

APPENDIX B: IS GALAXY FORMATION TRIGGERED WHEN GASEOUS HALOES EXCEED THE CRITICAL MASS?

Galaxy formation in our simulation typically takes place soon after the RELHICs exceed the critical mass for gas to lose pressure support. We show this in Fig. B1, which displays the distribution of the difference, $t_{\rm cross} - t_{\rm sf}$, i.e. the difference between the cosmic time at which the simulated dark matter haloes that host a luminous galaxy at $z = 0$ first crossed the critical mass for galaxy formation to take place ($t_{\rm cross}$), and the time at which the first star was formed in the halo ($t_{\rm sf}$). The difference is close to 0, indicating that the process of galaxy formation is indeed linked to the event of exceeding the critical mass. The small bias towards negative values (i.e. $t_{\rm sf} \gtrsim t_{\rm cross}$) is likely due to the time that it takes for the gas to undergo gravitational collapse before turning into stars, which is typically a few hundred Myr.
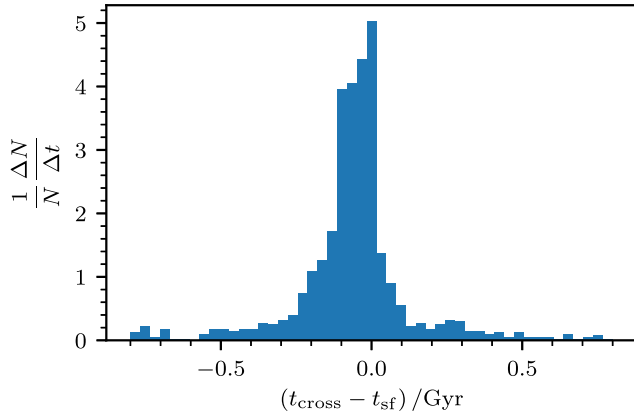

Figure B1. Distribution of $t_{\rm cross} - t_{\rm sf}$, i.e. the difference of the time at which a simulated luminous dark matter halo first crossed the critical mass for galaxy formation to take place ($t_{\rm cross}$), and the time at which the first star was formed in the halo ($t_{\rm sf}$). As expected, most haloes form their stars soon after they become more massive than the critical mass. There is, however, a large population of haloes for which star formation occurs slightly after they cross the critical mass. This delay, which is a negligible fraction of the Hubble time at $z = 0$, is likely due to the time that it takes for the gas to collapse after it becomes self-gravitating.

APPENDIX C: EPS MASS ASSEMBLY HISTORIES

Fig. C1 shows the comparison between the EPS assembly histories (orange dashed lines) and those constructed directly from the simulation using HBT+ (blue dots) for four bins of present-day halo mass. The shaded regions indicate the (10–90)th percentiles of the EPS assembly histories and the error bars the corresponding quantity for the simulated haloes. The EPS assembly histories agree well with those measured in the simulation down to the resolution limit of the simulation ($M_{200} \sim 10^6 \, {\rm M_\odot} \, h^{-1}$). Furthermore, as shown in Fig. 10, the HOF derived using the EPS assembly histories (green dot-dashed line) agrees remarkably well with the HOF measured in the simulation.
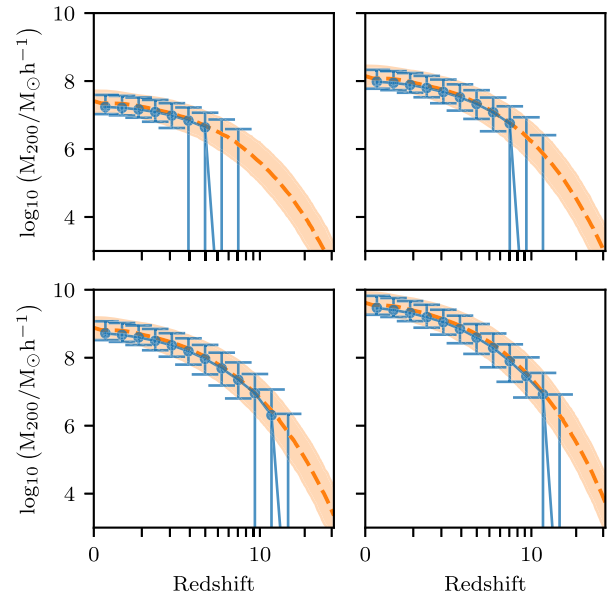

Figure C1. Comparison of synthetic assembly histories constructed using the Extended-Press-Schechter (EPS) formalism (orange dashed line) and assembly histories measured directly in our simulation (blue dots), for four different present-day mass bins. The shaded region shows the (10–90)th percentiles around the median for the EPS assembly histories; the error bars show the same quantities for the simulated dark matter haloes.

This paper has been typeset from a T$_{\rm E}$X/L^{A}T$_{\rm E}$X file prepared by the author.