Essential physics of early galaxy formation

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 8 September 2014 (MN LATEX style file v2.2)

Essential physics of early galaxy formation

Pratika Dayal[1⋆], Andrea Ferrara[2,3], James S. Dunlop[1] & Fabio Pacucci[2]
[1] *SUPA†, Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, EH9 3HJ, UK*
[2] *Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy*
[3] *Kavli IPMU, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa 277-8583, Japan*

ABSTRACT

We present a theoretical model embedding the essential physics of early galaxy formation ($z \simeq 5-12$) based on the single premise that any galaxy can form stars with a maximal *limiting efficiency that provides enough energy to expel all the remaining gas, quenching further star formation.* This simple idea is implemented into a merger-tree based semi-analytical model that utilises two mass and redshift-independent parameters to capture the key physics of supernova feedback in ejecting gas from low-mass halos, and tracks the resulting impact on the subsequent growth of more massive systems via halo mergers and gas accretion. Our model shows that: *(i)* the smallest halos (halo mass $M_h \leqslant 10^{10}\mathrm{M}_\odot$) build up their gas mass by accretion from the intergalactic medium; *(ii)* the bulk of the gas powering star formation in larger halos ($M_h \geqslant 10^{11.5}\mathrm{M}_\odot$) is brought in by merging progenitors; *(iii)* the faint-end UV luminosity function slope evolves according to $\alpha = -1.75 \log z - 0.52$. In addition, *(iv)* the stellar mass-to-light ratio is well fit by the functional form $\log M_* = -0.38 M_{UV} - 0.13\, z + 2.4$, which we use to build the evolving stellar mass function to compare to observations. We end with a census of the cosmic stellar mass density (SMD) across galaxies with UV magnitudes over the range $-23 \leqslant M_{UV} \leqslant -11$ spanning redshifts $5 < z < 12$: *(v)* while currently detected LBGs contain $\approx 50\%$ (10%) of the total SMD at $z = 5$ (8), the *JWST* will detect up to 25% of the SMD at $z \simeq 9.5$.

Key words: Galaxies: evolution - high-redshift - luminosity function, mass function - stellar content

1 INTRODUCTION

According to the standard cosmological Lambda Cold Dark Matter (ΛCDM) model the early Universe was almost perfectly homogeneous and isotropic, a picture supported by the small temperature anisotropies measured by the BOOMERanG experiment (Lange et al. 2001), and the COBE (Fixsen et al. 1996), WMAP (Hinshaw et al. 2013) and PLANCK satellites (Planck Collaboration et al. 2013). This (now well-established) model predicts that the earliest bound structures were low-mass Dark Matter (DM) halos that formed due to gravitational instability in slightly overdense regions. These low-mass structures acted as building blocks and merged to form successively larger structures in a hierarchical sequence.

The mechanical and radiative energy deposition by stars in the earliest galaxies affected the subsequent star formation history via a number of physical processes collectively referred to as "feedback". These include shock-heating and ejection of gas out of the DM halos, photo-evaporation and molecule dissociation, to mention a few. In general, these processes suppress and, in some situations, can quench further star formation. Feedback is ubiquitously invoked in astrophysics to solve problems ranging from galactic to cosmological scales (Dubinski & Carlberg 1991; Navarro et al. 1996; Moore et al. 1999b; Klypin et al. 1999; Moore et al. 1999a; Springel et al. 2008; White & Frenk 1991; Yoshida et al. 2002; Springel & Hernquist 2003; Gnedin 1998; Aguirre et al. 2001; Tornatore et al. 2007).

Given the low DM masses of the earliest galaxies, a tiny amount of star formation is sufficient to push out most (or indeed all) of the gas from these systems, potentially leading to a complete blow-away (Mac Low & Ferrara 1999). These "feedback-limited" galaxies then have to wait for gas to either be re-accreted from the surrounding intergalactic

⋆ E-mail:prd@roe.ac.uk (PD)
† Scottish Universities Physics Alliance

medium (IGM) or to be brought in by mergers to re-ignite further star formation. As halos build-up mass with time, their DM potential well can sustain much larger star formation rates (SFR) without losing gas. This naturally implies that, at any given time, there is a *limiting star formation efficiency* such that the energy produced by newly-formed stars is sufficient to expel all the remaining gas, quenching further star formation (at least temporarily).

We implement this one simple idea into a semi-analytic model to trace the formation and evolution of galaxies over the first billion years of cosmic time, from redshift $z = 12$ to $z = 5$. Our model follows the assembly of galaxies through the mergers of their DM progenitors. We trace all major baryonic processes including star formation, supernova (SN)-powered gas ejection, gas/stellar mass growth through mergers, and gas accretion from the IGM. In the spirit of maintaining simplicity to isolate the fundamental physics driving galaxy evolution, our model utilises only *two redshift and mass-independent free parameters*: (a) the star formation efficiency threshold, f_*, and (b) the fraction of SN energy that drives winds, f_w.

The construction of such a simple model is very timely, given the immense amount of data on high-z Lyman Break Galaxies (LBGs) that has been acquired over the past few years: surveys with the Wide Field Camera 3 (WFC3) on the *Hubble Space Telescope* (*HST*) have revolutionized our understanding of the faintest LBGs, providing unprecedented constraints on the faint-end of the evolving ultraviolet luminosity function (UV LF; e.g. Bouwens et al. 2007; McLure et al. 2009; Bouwens et al. 2010a; McLure et al. 2010; Oesch et al. 2010b; McLure et al. 2013) while ground-based wide-field surveys such as UltraVISTA are revealing luminous galaxies populating the bright-end of the UV LF at $z \simeq 7$ (e.g. Ouchi et al. 2010; Bowler et al. 2012, 2014). Spectral energy distribution (SED) fitting to broad-band photometry has also been used to build galaxy stellar mass functions (SMF; González et al. 2011), understand the stellar populations in these sources through their UV (β) slopes (Bouwens et al. 2010b; Finkelstein et al. 2012; Wilkins et al. 2011; Dunlop et al. 2012, 2013; Bouwens et al. 2013; Rogers et al. 2014), infer their physical properties (Oesch et al. 2010a; Labbé et al. 2010a; McLure et al. 2011; Labbé et al. 2013) and calculate the growth of stellar mass density (SMD) with redshift (Labbé et al. 2010a; González et al. 2011; Stark et al. 2013). The Cluster Lensing And Supernova survey with Hubble (CLASH; Postman et al. 2012) is now using galaxies magnified by strong-lensing to further constrain the number-density of high-z sources (Zheng et al. 2012; Bradley et al. 2013; Coe et al. 2013) and study their nebular emission (Smit et al. 2014).

A number of earlier works have attempted to model high-z galaxies using cosmological simulations (e.g. Finlator et al. 2007; Dayal et al. 2009; Nagamine et al. 2010; Dayal et al. 2010; Forero-Romero et al. 2011; Salvaterra et al. 2011; Dayal & Ferrara 2012; Jaacks et al. 2012; Dayal et al. 2013). Although these studies can shed light on important physical properties of galaxies (halo/stellar/gas masses, metallicities), their assembly, and their clustering, running simulations necessarily involves a number of assumptions regarding the star-formation density threshold, gas ejection, metal pollution and fraction of SN energy that can power winds, to name a few. Other models have used semi-analytic approaches to reproduce observations invoking multiple free-parameters (Cole 1991; Somerville & Primack 1999; Baugh 2006; Croton et al. 2006; De Lucia et al. 2010; Benson 2012; Lu et al. 2013).

In this work, our aim is instead to build the simplest model, based on only two free parameters than can be readily constrained by existing data [1], in order to isolate the fundamental physics that shapes galaxy evolution in the first billion years. We show that, once the two parameters have been fixed, this model naturally yields the correct mass-to-light ratios, SMF and SMD, without the need to invoke any *ad-hoc* free parameters. We make predictions for the fractional contribution of galaxies of different luminosities to the SMD, which can be tested in the future with the *James Webb Space telescope* (*JWST*).

2 THEORETICAL MODEL

As introduced above, our aim is to build the simplest model for galaxy formation based solely on the balance between the amount of Type II SN (SNII; mass $\geqslant 8\mathrm{M}_\odot$) energy available to drive winds, and the gravitational potential of the host DM halo. Our idea is as follows: if the SNII kinetic energy is larger than the binding energy of a halo, the galaxy will lose all of its gas mass and will be unable to form any more stars. However, halos with a binding energy larger than the SNII kinetic energy will only lose part of their gas and can continue forming stars. This simple idea and its mathematical formulation are detailed in the following.

2.1 Feedback-limited star formation efficiency

Radiative cooling is very efficient in dense low-mass halos at high-z. Left unchecked, this leads to an over-production of stars and too many baryons being locked up in condensed halos (as compared to observations), a problem canonically termed "overcooling" (Benson et al. 2003; Springel & Hernquist 2003). This problem can be alleviated by introducing SN feedback that reduces the star-formation efficiency of small halos by ejecting their gas, quenching further star formation (e.g. Mac Low & Ferrara 1999; Springel & Hernquist 2003; Greif et al. 2007), as formulated below.

The formation of an amount $M_*(z)$ of stars at redshift z can impart the ISM with a total SNII energy E_{SN} given by

$$E_{SN} = f_w E_{51} \nu M_*(z) \equiv f_w v_s^2 M_*(z), \quad (1)$$

[1] The cosmological model used in this work corresponds to the ΛCDM Universe with DM, dark energy and baryonic density parameter values of $(\Omega_\mathrm{m}, \Omega_\Lambda, \Omega_\mathrm{b}) = (0.2725, 0.702, 0.04)$, a Hubble constant $H_0 = 100h = 70\,\mathrm{km\,s^{-1}Mpc^{-1}}$, a primordial spectral index $n_s = 0.96$ and a spectral normalisation $\sigma_8 = 0.83$, consistent with the latest results from the PLANCK collaboration (Planck Collaboration 2013).

where each SNII is assumed to impart an (instantaneous) explosion energy of $E_{51} = 10^{51}$erg to the ISM and $\nu = [134\,\mathrm{M}_\odot]^{-1}$ is the number of Type II SNe per stellar mass formed for a Salpeter IMF between $0.1-100\mathrm{M}_\odot$; we use this IMF in all calculations and simply refer to SNII as SN in what follows. The values of E_{51} and ν yield $v_s = 611$ km s^{-1}. Finally, f_w is the fraction of the SN explosion energy that is converted into kinetic form and drives winds [2].

For any given halo, the energy E_{ej} required to unbind and eject all the ISM gas can be expressed as

$$E_{ej} = \frac{1}{2}[M_{g,i}(z) - M_*(z)]v_e^2, \quad (2)$$

where $M_{g,i}(z)$ is the gas mass in the galaxy at epoch z; the term $M_{g,i}(z) - M_*(z)$ implies that SN explosions have to eject the part of the initial gas mass not converted into stars. Further, the escape velocity v_e can be expressed in terms of the halo rotational velocity,v_c, as $v_e = \sqrt{2}v_c$.

We then define the *ejection efficiency*, f_*^{ej}, as the fraction of gas that must be converted into stars to "blow-away" the remaining gas from the galaxy (i.e. $E_{ej} \leqslant E_{SN}$). This can be calculated as

$$f_*^{ej}(z) = \frac{v_c^2(z)}{v_c^2(z) + f_w v_s^2}. \quad (3)$$

The *effective efficiency* can then be expressed as

$$f_*^{eff} = min[f_*, f_*^{ej}]. \quad (4)$$

This represents the maximum fraction of gas that can be converted into stars in a galaxy without expelling all the remaining gas. Since v_c scales with the halo mass (M_h), efficient star formers (hosted by large DM halos) can continuously convert a fraction f_* of their gas into stars, while feedback-limited systems can form stars with a maximum efficiency dictated by f_*^{ej} that decreases with decreasing halo mass. Matching the bright and faint ends of the evolving UV LF requires $f_* = 0.03$ and $f_w = 0.1$ as explained in Sec. 3.1 below.

Galaxies of a given M_h value are more compact (i.e. have deeper potential wells) and rotate faster with increasing redshift as $v_c \propto (1+z)^{1/2}$. Using Eqn. 3 this implies that a given f_*^{ej} value is reached for progressively lower M_h values with increasing redshift, as shown in Fig. 1. Given that $f_*^{eff} = min[f_*, f_*^{ej}]$, this means that f_*^{eff} saturates to f_* for lower M_h values with increasing redshift. In other words, galaxies of a given halo mass are more efficient at holding on to their gas with increasing redshift as a result of their deeper potential wells. This feedback function (behaviour of f_*^{ej} as a function of halo mass) is shown in Fig. 1 for $z = 5$ to $z = 20$. Quantitatively, while galaxies with masses as low as $M_h \simeq 10^{8.45}\mathrm{M}_\odot$ saturate to $f_* = 0.03$ and become efficient star formers at $z = 20$, galaxies have to be as massive as $M_h = 10^{9.25}\mathrm{M}_\odot$ at $z = 5$ to achieve the same f_* value.

[2] Using the lifetime function of Padovani & Matteucci (1993), stars with masses between $9.98-100\mathrm{M}_\odot$ would contribute to the total SN energy for our model time step of 20 Myrs as opposed to the range $8-100\mathrm{M}_\odot$ used in this work. This would lead to a value of $\nu = [184\,\mathrm{M}_\odot]^{-1}$ and $v_s = 521$ km s^{-1}. However, this decrease in v_s by a factor of 1.17 does not affect any of our model results, confirming that neglecting the lifetimes of stars that result in SNII is a valid approximation.

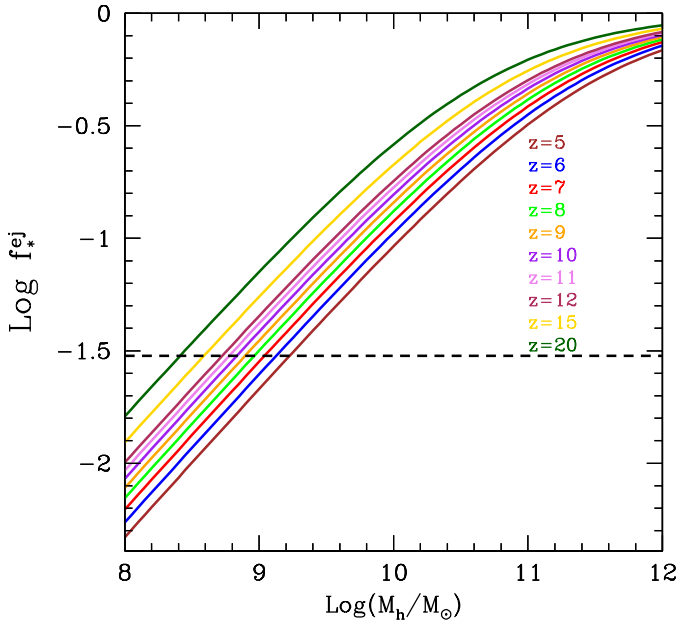

Figure 1. The ejection efficiency (f_*^{ej}) as a function of halo mass for $z \simeq 5-20$; this is the star-formation efficiency required to eject all the gas from the galaxy and quench further star formation. The horizontal line shows $f_* = 0.03$. Since $f_*^{eff} = min[f_*, f_*^{ej}]$, galaxies with $f_*^{ej} > 0.03$ saturate at an effective efficiency of $f_*^{eff} = f_* = 0.03$. We assume each SN imparts an explosion energy of $E_{51} = 10^{51}$erg, of which a fraction $f_w = 0.1$ drives winds and the SN rate (ν) is calculated assuming a Salpeter IMF between $0.1-100\mathrm{M}_\odot$. These values of E_{51} and ν lead to $v_s = 611$ km s^{-1} (see Sec. 2.1 for details).

2.2 Merger tree physics

We implement the above simple physical ideas into standard DM halo merger trees tracing the formation of increasingly larger systems from the mergers of smaller progenitors as shown in Fig. 2 (White & Frenk 1991; Lacey & Silk 1991; Cole et al. 1994). We build merger trees for 800 $z = 4$ galaxies equally spaced in log M_h between $10^{8-13}\mathrm{M}_\odot$ using the modified binary merger tree algorithm with accretion presented in Parkinson et al. (2008). In brief, the merger tree for each simulated DM halo starts at $z = 4$ and runs backward in time up to $z = 20$, with each halo fragmenting into its progenitors. At any given time-step, a halo of mass M_0 can either lose a part of its mass (i.e. fragment into halos below the mass resolution limit M_{res}) or fragment into two halos with masses $M_{res} < M < M_0/2$. The mass below the resolution limit then accounts for "smooth-accretion" from the IGM, in which the halo is embedded. We run our merger tree using 70 steps equally spaced in time (by 20 Myrs) and with a resolution mass $M_{res} = 10^8\mathrm{M}_\odot$. Each of the simulated $z = 4$ halos is associated with the correct number density

by matching its halo mass to the Sheth-Tormen mass function (Sheth & Tormen 1999). Then, at any redshift, every progenitor is given the same number density as its $z = 4$ successor. The use of the conditional mass function given by the extended Press-Schechter theory (Bower 1991; Lacey & Cole 1993) and the modifications introduced by Parkinson et al. (2008) ensure that progenitor halo mass function matches the Sheth-Tormen mass function at any z (see Sec. 2.1 Dayal et al. 2014).

Once the merger tree for each galaxy has been constructed, we implement our baryonic physics model within it. Given that the baryonic properties of parent halos depend on those of their progenitors at earlier times, we now proceed forward in time from $z = 20$ and follow the joint halo/galaxy evolution. We start from the first DM progenitor (with halo mass M_0) along a branch of the merger tree and assume that it has an initial gas mass $M_{g,i}(z) = (\Omega_b/\Omega_m)M_0(z)$. A fraction of this gas mass gets converted into a (newly formed) stellar mass $M_*(z)$, such that

$$M_*(z) = f_*^{eff} M_{g,i}(z). \quad (5)$$

In the spirit of maintaining simplicity, we assume that every stellar population has a fixed metallicity of $0.05\mathrm{Z}_\odot$ and each newly-formed stellar population has an age of 2 Myr. Using these parameters with the population synthesis code STARBURST99 (Leitherer et al. 1999), the UV luminosity at $\lambda = 1500$ Å produced by a newly-formed stellar mass can be expressed as

$$L_{UV} = 10^{33.077}\left(\frac{M_*}{\mathrm{M}_\odot}\right) \mathrm{erg\,s^{-1}\AA^{-1}}. \quad (6)$$

This star-formation episode must then result in a certain amount of gas $M_{g,ej}(z)$ being ejected from the galaxy at the given z-step. The fraction of gas mass ejected depends on the ratio of the SN kinetic energy available (E_{SN}) and the potential energy required to unbind the gas not turned into stars (E_{ej}) which can be expressed as (see Eqns. 1 and 2)

$$\frac{E_{SN}}{E_{ej}} = \frac{f_w v_s^2 f_*^{eff} M_{g,i}(z)}{M_{g,i}(z)(1-f_*^{ej})v_c^2} = \frac{f_w v_s^2 f_*^{eff}}{(1-f_*^{ej})v_c^2}. \quad (7)$$

Substituting f_*^{ej} from Eqn. 3, we obtain

$$\frac{E_{SN}}{E_{ej}} = \frac{f_*^{eff}}{f_*^{ej}}. \quad (8)$$

The value of $M_{g,ej}(z)$ therefore depends on whether $f_*^{eff} = f_*$ or $f_*^{eff} = f_*^{ej}$: in the former case the galaxy is an "efficient star-former" that can support a large amount of stellar mass being formed without losing much of its gas, while it is a "feedback-limited" system in the latter case with all of its ISM gas being blown-away (see also Fig. 2). Using Eqn. 8, $M_{g,ej}$ can be mathematically expressed as

$$M_{g,ej}(z) = [M_{g,i}(z) - M_*(z)]\frac{f_*^{eff}}{f_*^{ej}}, \quad (9)$$

since the initial gas mass is reduced by the amount that is converted into stars. The final gas mass, $M_{g,f}(z)$, remaining in the galaxy at that time-step can then be expressed as

$$M_{g,f}(z) = [M_{g,i}(z) - M_*(z)]\left[1 - \frac{f_*^{eff}}{f_*^{ej}}\right]. \quad (10)$$

On the other hand, a galaxy inherits a certain amount of stars and gas from its progenitors following merging events. In addition, this galaxy also obtains a part of its DM (and gas) mass through "smooth-accretion" from the IGM. Consider, for example, a galaxy of halo mass M_0 at redshift z that has progenitors with halo masses M_1 and M_2 at redshift $z + \Delta z$. The difference between the sum of the progenitor masses and M_0 then yields the unresolved halo mass that is smoothly-accreted from the IGM, such that $M_{h,acc}(z) = M_0 - (M_1 + M_2)$. We then make the simple (and reasonable) assumption that the smoothly-accreted DM pulls in a cosmological ratio of gas mass with it such that the accreted gas mass is $M_{g,acc}(z) = (\Omega_b/\Omega_m)M_{h,acc}(z)$. Thus, the total initial gas mass in the galaxy at z is the sum of the newly accreted gas mass, as well as that brought in by its merging progenitors, i.e.

$$M_{g,i}(z) = M_{g,acc}(z) + \sum M_{g,f}(z + \Delta z). \quad (11)$$

This $M_{g,i}(z)$ value is then used to calculate the new stellar mass formed in the galaxy as described by Eqn. 5. The total stellar mass in this galaxy is now the sum of mass of the newly-formed stars, and that brought in by its progenitors such that

$$M_{*,tot}(z) = M_*(z) + \sum M_*(z + \Delta z). \quad (12)$$

Eqs. 9 and 10 are again used to obtain the ejected, and final gas masses at the given z-step.

Finally, the total UV luminosity of the galaxy is a sum of the new luminosity as well as that brought by its progenitors,

$$L_{UV,tot} = L_{UV}(z) + \sum L_{UV,*}(z + \Delta z). \quad (13)$$

Using STARBURST99, we find that the UV luminosity for a burst of stars (normalized to a mass of $1\mathrm{M}_\odot$ and metallicity $0.05\,\mathrm{Z}_\odot$) decreases with time as

$$\log\left(\frac{L_{UV}(t)}{\mathrm{erg\,s^{-1}\AA^{-1}}}\right) = 33.0771 - 1.33\log(t/t_0) + 0.462, \quad (14)$$

where t is the age of the stellar population (in yr) at z and $\log(t_0/\mathrm{yr}) = 6.301$; we remind the reader that by construction, each newly-formed stellar population has an age of 2 Myr.

These simple ideas of halo mass/stellar growth, star formation and its associated feedback and gas mass ejection/accretion/merged are implemented into our merger tree, tracing the growth of galaxies from $z = 20$ to $z = 4$. As shown in Fig. 2, in this model low-mass galaxies are *feedback-limited star formers*: star formation with a low efficiency is sufficient to eject all the gas from these systems, quenching further star formation. These feedback-limited galaxies then have to wait for gas to either be accreted from their surrounding IGM or for gas to be brought in by mergers to re-ignite further star formation. On the other hand, massive galaxies are efficient star formers that can sustain much larger star formation rates at a fixed efficiency f_*. As

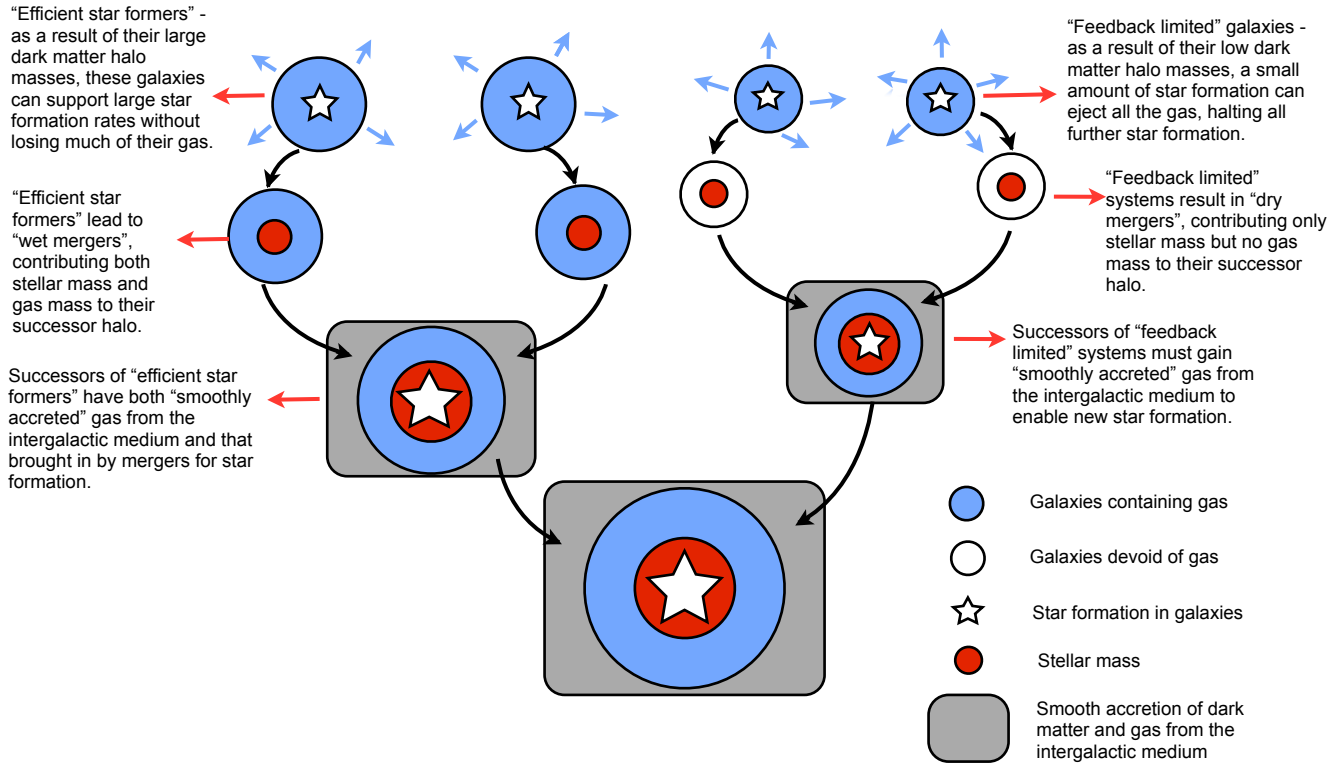

Figure 2. Merger tree showing the assembly of a galaxy in the hierarchical structure formation model. Low-mass systems are "feedback-limited" as SN winds can expel most/all of their gas, quenching further star formation. On the other hand, larger-mass systems are "efficient star-formers" and can convert gas into stars at a fixed efficiency (f_*) at any given time (see Sec. 2.1). Galaxies build up in mass hierarchically from the mergers of smaller progenitors which can be a combination of feedback-limited systems and efficient star-formers. In addition to gaining gas from the mergers of their progenitors, galaxies also "smoothly accrete" both DM and gas from the IGM.

expected, both low and high-mass galaxies assemble from building blocks that are a combination of "feedback-limited" and "efficient" systems, with galaxies of progressively lower masses becoming efficient star formers with increasing redshift, as explained in Sec. 2.1.

As seen from the equations in this Section, our fiducial model only has two free parameters - the threshold SF efficiency (f_*) and the fraction of SN energy that drives winds (f_w), both of which are *independent of redshift and halo mass*.

3 UV LUMINOSITY FUNCTION

Once we have implemented the baryonic physics described in Sec. 2.2 into the merger tree and inferred the UV luminosity of each galaxy, we can construct the evolving LBG UV LF to be compared with observations, and shed light on the physics that shapes it.

3.1 Understanding the UV LF evolution

We start by discussing the behaviour of the UV LF inferred from the halo mass function (HMF). We can attempt to construct a simple (and likely unphysical) model of the UV LF by multiplying the HMF with a star-formation efficiency value η (column 4 of Table 1), chosen to match to the knee of the observed UV LF. While η is redshift-dependent, its value is constant (i.e. independent of halo mass) at any given redshift. This is equivalent to a model without feedback wherein every galaxy contains a cosmological ratio of baryons to DM equal to Ω_b/Ω_m, leading to the observed UV LF tracing the same shape as the HMF. Quantitatively, it transpires from the required efficiency scaling with redshift, that galaxies of the same UV luminosity reside in halos that are twice as massive at $z \simeq 5$ as at $z \simeq 8$ leading to a steepening in the predicted faint-end slope ($\alpha_{HMF} = -2.11 \rightarrow -2.32$ from $z = 5 \rightarrow 8$). This simple comparison serves to emphasize the important role of feedback; a viable model requires that DM halos form stars with an efficiency that progressively decreases with halo mass

Given that galaxies of a given luminosity are hosted in lower-mass halos at earlier times, one might then expect them to enter the feedback-limited regime producing a flattening of the LF faint-end towards high-z compared to the underlying evolving HMF. It is therefore intriguing to see that the faint-end slope of the theoretical UV LF also steepens with increasing redshift as shown in column 5 of Table 1; within errors the faint end slope of the theoretical UV LF

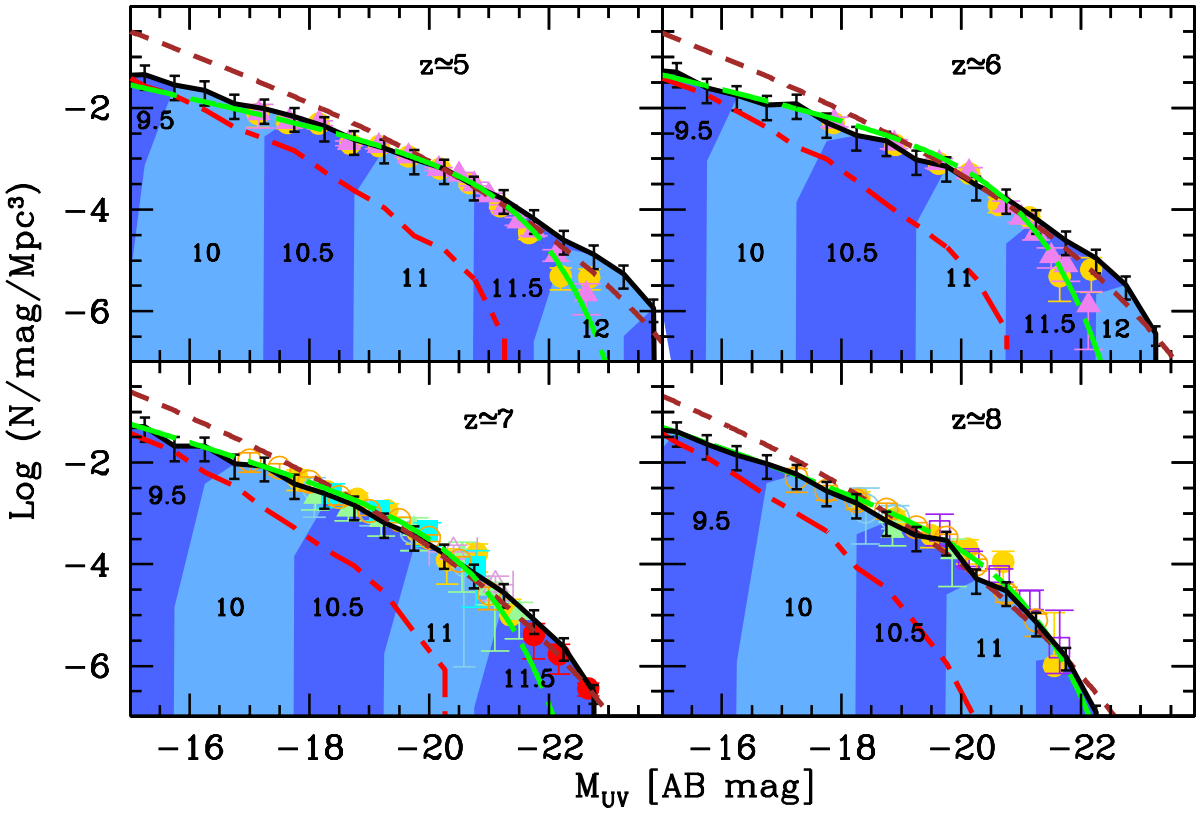

Figure 3. The evolving LBG UV LF at $z \simeq 5-8$. In all panels, the black solid line shows the results using our fiducial model, i.e. including gas accretion from progenitors and the IGM, and SN-powered gas ejection; the black error bars show poissonian errors arising from the luminosity dispersion in each bin. The dashed red line shows that the UV LF would have been severely under-estimated (with increasing brightness) had we not considered the gas that is brought in by mergers. The brown line shows the results obtained by multiplying the halo mass function with a constant star-formation efficiency appropriate for the redshift considered (column 4 of Table 1). In all panels, the dashed green line shows the observationally inferred best-fit Schechter function (McLure et al. 2009, 2013) and points show observational results: (a) $z \simeq 5$: Bouwens et al. (2007, filled circles) and McLure et al. (2009, filled triangles); (b) $z \simeq 6$: Bouwens et al. (2007, filled circles) and McLure et al. (2009, filled triangles); (c) $z \simeq 7$: Oesch et al. (2010b, filled squares), Bouwens et al. (2010a, empty blue circles), Bouwens et al. (2011, filled circles), Castellano et al. (2010, empty triangles), McLure et al. (2010, filled triangles), McLure et al. (2013, empty orange circles) and Bowler et al. (2014, filled red circles); (d) $z \simeq 8$: Bouwens et al. (2010a, empty blue circles), Bouwens et al. (2011, filled circles), McLure et al. (2010, filled triangles), Bradley et al. (2012, empty squares) and McLure et al. (2013, empty red circles). The numbers in the shaded areas under the UV LF show the central value of the DM halo mass bin hosting the galaxy as obtained from the fiducial model.

remains constant from $z \simeq 8$ to $z \simeq 5$ with respect to the underlying HMF, as shown in Fig. 3.

The physical reason for the constant faint-end slope difference between the UV LF and the HMF can be explained as follows. As seen from the shaded regions in Fig. 3, the host halo masses increase by a factor ≈ 3 from $M_h = 10^{9.5}\,(10^{11})\mathrm{M}_\odot$ for $M_{UV} = -16\,(-21)$ at $z = 8$ to $10^{10}\,(10^{11.5})\mathrm{M}_\odot$ for $M_{UV} = -16\,(-21)$ at $z = 5$. However, a galaxy of a given halo mass has half the time to assemble by $z = 8$ when the age of the Universe is about 0.66 Gyr compared to the 1.21 Gyr available by $z = 5$. As a result of the shorter time available, galaxies of a given halo mass assemble at twice the rate leading to their SFRs (and luminosities) being twice as large at $z = 8$ compared to $z = 5$. So, although the HMF shifts to progressively lower masses with increasing redshift, the fact that galaxies of a given mass are more efficient at forming stars with increasing redshift helps maintain a constant slope offset between the UV LF and underlying HMF, as shown in Fig. 3.

In summary, our model reproduces both the slope and amplitude of the UV LFs as measured by a number of surveys remarkably well, as shown in Fig. 3 and quantified in Table 1. The strength of our model lies in the fact that it yields UV LFs over 9.5 (7) UV magnitudes at $z = 5$ (8) using only two time- and mass-independent free parameters $f_* = 0.03$, $f_w = 0.1$. These two parameters together shape the UV LF: while f_w affects the faint-end slope of the UV LF where feedback is most effective, f_* determines the amplitude and normalization at the bright-end where galaxies can form stars at the maximum allowed f_* value. Of course, since $f_*^{eff} = min[f_*, f_*^{ej}]$, its value depends both on the halo

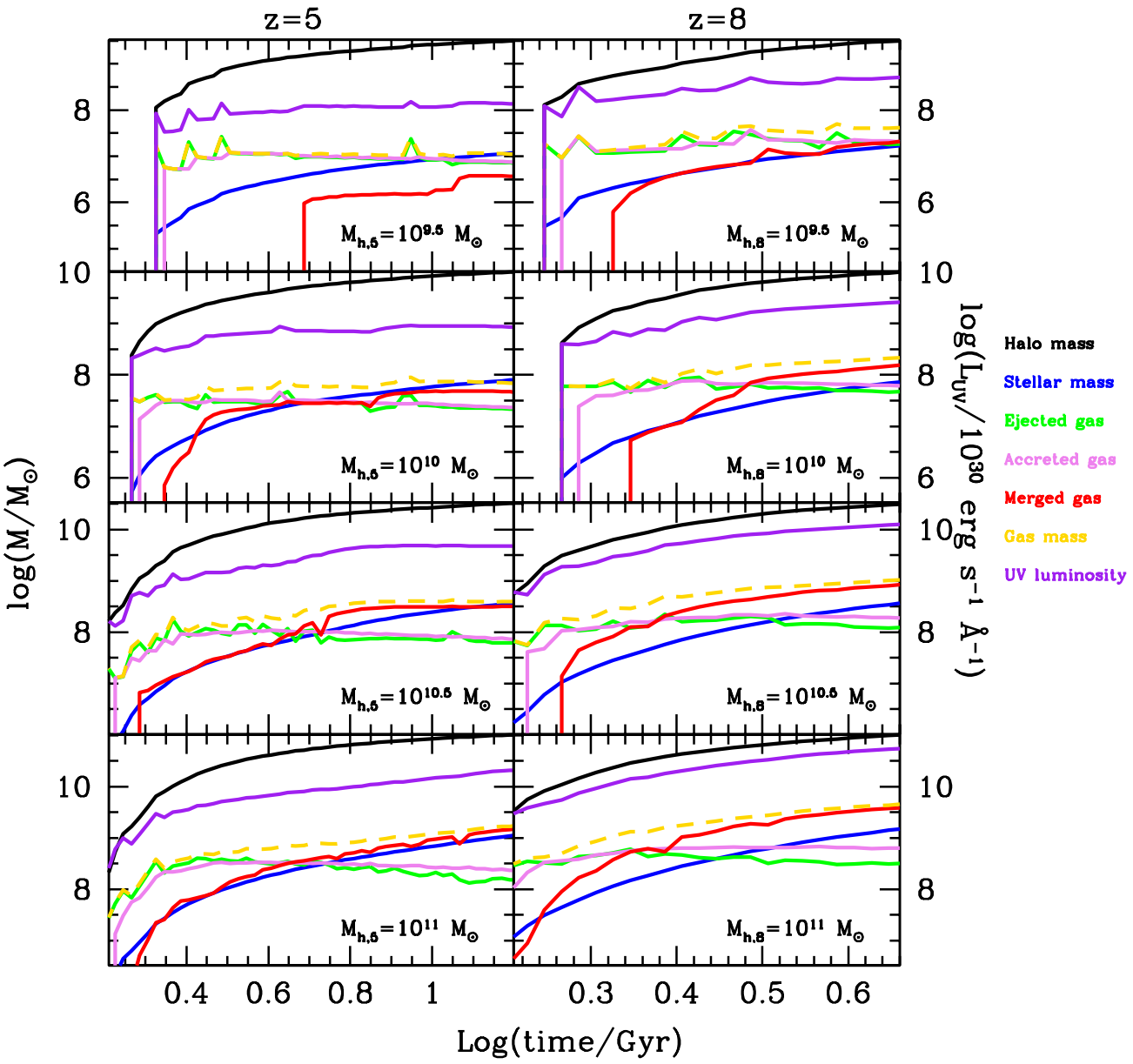

Figure 4. Galaxy assembly as a function of cosmic time. For the halo mass value given in each panel, we show the assembly of galaxies with halo masses in the range $10^{9.5}-10^{11}\mathrm{M}_\odot$ at $z=5$ (left column) and $z=8$ (right column). Each panel shows the total halo and stellar mass (in $\mathrm{M}_\odot$) built-up by a certain time; the ejected/accreted/merged and final gas mass values (in $\mathrm{M}_\odot$) are instantaneous quantities. We also show the instantaneous UV luminosity in units of $\log(L_{UV}/\mathrm{erg\,s^{-1}\AA^{-1}})$ which can be treated as a proxy for the SFR; the UV luminosity has been uniformly scaled down by 10^{30} for display reasons. Note that the time-axis is different in the left and right-columns: the Universe is twice as old at $z=5$ with an age of 1.21 Gyr as compared to $z=8$ where the age is 0.66 Gyr.

mass and redshift through the dependence of f_*^{ej} on these two quantities as explained in Sec. 2.1.

Further, our model naturally predicts that the bright-end of the UV LF at $z\simeq 7$ must be flatter than the usually-assumed Schechter function and is compatible with both the HMF and the double power-law (DPL) slope estimated from the widest-area survey carried out so far at this redshift (Bowler et al. 2014). Our model slightly over-predicts the number of bright galaxies at $z\simeq 5$ and $z\simeq 6$, as can be seen from the same figure. Whether this discrepancy is due to physical effects that have been ignored (e.g. dust attenuating the luminosity from these massive galaxies (Dayal et al. 2009), halo mass quenching (Peng et al. 2010) and/or AGN feedback), or is in fact due to remaining issues with the data analysis (e.g. the application of inadequate aperture corrections when undertaking photometry of the brightest high-redshift galaxies) remains a matter for further study. By virtue of reproducing the observed UV LF, the redshift

Table 1. For the redshift shown in column 1, we show the observed faint-end slope of the UV LF (McLure et al. 2009, 2013) in column 2. Columns 3 shows the faint-end slope of the UV LF obtained by scaling the HMF using the star-formation efficiency shown in column 4. Column 5 shows the faint-end UV LF slope values obtained from our fiducial model. The faint-end slopes for the scaled HMF and theoretical UV LF have been computed over the absolute magnitude range $-18 \leqslant M_{UV} \leqslant -12$.

z	α_{obs}	α_{HMF}	η	α_{th}
5	$-1.66^{+0.66}_{-0.66}$	-2.11 ± 0.07	0.006	-1.74 ± 0.48
6	$-1.71^{+0.11}_{-0.11}$	-2.19 ± 0.10	0.0075	-1.89 ± 0.71
7	$-1.90^{+0.14}_{-0.15}$	-2.25 ± 0.13	0.009	-2.02 ± 0.74
8	$-2.02^{+0.22}_{-0.23}$	-2.32 ± 0.15	0.011	-2.10 ± 0.62
9	–	–	–	-2.19 ± 0.23
10	–	–	–	-2.25 ± 0.48
11	–	–	–	-2.32 ± 0.34
12	–	–	–	-2.61 ± 0.83

evolution of the star formation rate density (SFRD) predicted by our model is in excellent agreement with high-z SFRD observations (Ellis et al. 2013).

Finally, as shown in Dayal et al. (2013), we clarify that the evolution of the UV LF is a *combination* of luminosity and density evolution that depends on the luminosity range probed: the evolution at the bright end is genuine luminosity evolution, driven by the brightest galaxies continuing to brighten further with time; the evolution at the faint end is a mix of positive and negative luminosity and density evolution as these tiny systems brighten and fade in luminosity, and continually form and merge into larger systems.

3.1.1 Faint end galaxies: starving or inefficient?

A natural question that arises at this point is whether the faint-end of the UV LF lies below that which would be inferred from the HMF because the fainter galaxies are fuel-supply limited ("starving") as a result of their progenitors having ejected most/all of their gas content, or because they themselves are star-forming efficiency-limited (i.e. $f_*^{eff} < f_*$) due to their low masses. This question can easily be answered using Fig. 1: as shown there, galaxies with $M_h \geqslant 10^9\,, (10^{9.25}) \mathrm{M}_\odot$ at $z = 8\,(5)$ can form stars at the maximum allowed efficiency of $f_*^{eff} = f_* = 3\%$. Given that, from our fiducial model, galaxies brighter than $M_{UV} = -15$ are hosted in halos more massive than $10^{9.5}\mathrm{M}_\odot$ at all $z = 5$ to $z = 8$ (Fig. 3), galaxies *on the currently observable UV LF are not efficiency limited*; their luminosities are depressed as a result of their progenitors having lost most/all of their gas content, resulting in a gas mass that is lower than the cosmological baryon fraction, i.e. starvation. The observed UV LF thus holds imprints of the entire past gas build-up history of its progenitors.

3.1.2 Gas supply: mergers or accretion?

This bring us to the question of how these galaxies build up their gas and stellar content: is their assembly dominated by the gas brought in by mergers or that accreted from the IGM? A first hint can be obtained from Fig. 3 where, in the dashed red lines, we show the theoretical UV LF that results from assuming each galaxy loses all its gas content at every $z-$step resulting in dry mergers (although we use the same f_*^{eff} values as the fiducial model). As can be seen, while similar to the fiducial model at the faint end, this UV LF drops off very steeply with increasing luminosity. This implies that gas brought in by mergers is not very important at the faint end since the tiny progenitors of these galaxies are feedback limited even in the fiducial model, resulting in largely dry mergers. However, the progenitors of increasingly luminous galaxies are not feedback limited, and contribute to building up a large gas reservoir that powers star formation. Indeed, the LF at the bright end drops to 10% of the fiducial model value (M_{UV} decreases by ≈ 2 magnitudes) if the gas brought in by mergers is not taken into account.

To further elucidate the above point, we now examine in detail four representative galaxies with halo masses $M_h = 10^{9.5,10,10.5,11}\mathrm{M}_\odot$ at two epochs corresponding to $z = 5$ and $z = 8$ (see Fig. 4). As expected from the hierarchical model, the earlier a galaxy starts building up, the larger its final mass can become (Dayal et al. 2013), i.e. the progenitors of the most massive galaxies start assembling earliest with the progenitors of progressively less massive galaxies forming later. As shown in Fig. 4, while the progenitors of $M_h = 10^{11}\mathrm{M}_\odot$ galaxies start forming about 0.2 Gyr after the Big Bang, the progenitors of $10^{9.5}\mathrm{M}_\odot$ galaxies appear only ≈ 0.2 Gyr later than that. Moreover, the progenitors of galaxies as small as $10^{9.5}\mathrm{M}_\odot$ are feedback-limited and bring in negligible amounts ($\lesssim 10\%$ of the gas acquired by accretion) of gas when the galaxy starts assembling. As a result, the initial stellar mass build up is dominated by self-accreted gas, with mergers becoming important only at the later stages. Gas brought in by mergers starts dominating progressively earlier for more massive halos; while episodes of star formation are essentially driven by minor mergers, massive galaxies build up their gas reservoir from both accretion and mergers. Finally, it is interesting to note that galaxies of similar halo mass produce a similar amount of stars at both $z = 5$ and $z = 8$. As galaxies at $z = 8$ have about half the time to assemble as compared to galaxies at $z = 5$, it follows that the specific star formation rate correspondingly decreases with time.

To summarize, the *faint-end of the LF is feedback-suppressed* below the HMF since star formation in the tiny progenitors of these galaxies led to complete gas ejection, reducing their gas mass below the cosmological baryon fraction. Smooth accretion from the IGM dominates over mergers in assembling the gas mass of the faintest galaxies on the UV LF. Mergers (smooth-accretion) become progressively more (less) important with increasing luminosity, with *mergers supplying most of the gas mass for galaxies at the bright end* of the LF. Given their halo masses, while $z = 5 - 8$ galaxies brighter than $M_{UV} = -15$ can form stars at the

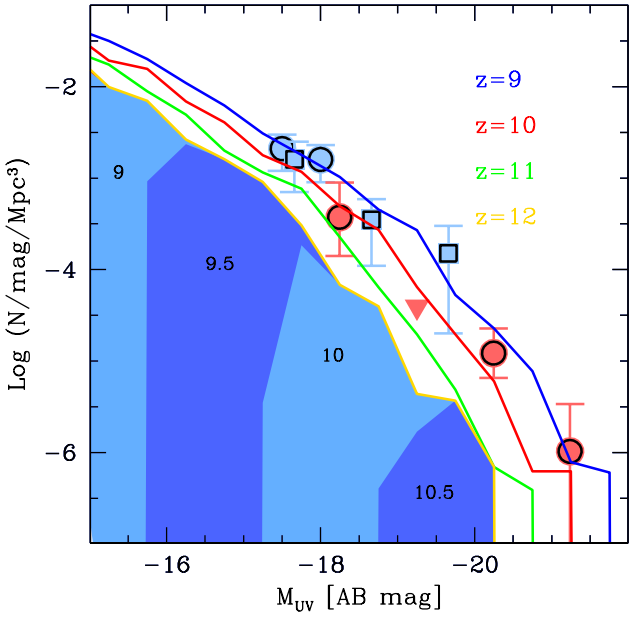

Figure 5. The evolving LBG UV LF at $z \simeq 9-12$: lines show our model results with data points showing observational results. The UV LF from our fiducial model is in excellent agreement with observations at $z \simeq 9$: (filled blue circles McLure et al. 2013) and (filled blue squares Oesch et al. 2013) and at $z \simeq 10$: (filled red circles Bouwens et al. 2014); the downward pointing triangle represents the upper-limit of the $z \simeq 10$ data at $M_{UV} \simeq -19.25$. We present our predictions for $z = 11, 12$ that should be testable with the *JWST*. The numbers in the shaded area under the $z = 12$ UV LF show the mass of the DM halo hosting the galaxy. With $M_h \geqslant 10^9 M_\odot$, galaxies as faint as $M_{UV} = -15$ at $z = 12$ are not star-formation efficiency limited; see Sec. 3.2.

maximum allowed efficiency, their lower luminosities arise as a result of their low baryon fraction (since their progenitors are feedback-limited), i.e. these galaxies are *not star-formation efficiency limited, but starved due to limited fuel supply.*

3.2 Faint-end slope evolution

Buoyed by the success of our model in reproducing the observed UV LFs from $z = 5$ to 8, we extend our results to redshifts as high as $z = 9-12$ (Fig. 5), and provide a functional form for the redshift-evolution of the faint-end slope, α (also see Table 1).

It is encouraging to see that the UV LF predicted by our fiducial model is in agreement with the sparsely-sampled UV LF at $z = 9$ (McLure et al. 2013; Oesch et al. 2013). We remind the reader that we use the same values of our two free parameters at all z such that $f_* = 0.03$ and $f_w = 0.1$, i.e. we do not invoke any *redshift-dependent ad-hoc free parameters*. Further, the $z = 10$ theoretical UV LF is also consistent with the UV LF inferred by Bouwens et al. (2014). However, our predictions are slightly higher than the observational upper limits inferred by these authors at $M_{UV} = -19.5$. Improved data are needed at this epoch to determine whether this apparent drop in the UV LF is real.

Galaxies at $z = 9-12$ more luminous than $M_{UV} = -15$ are hosted in halos with mass $M_h \geqslant 10^9 M_\odot$ as shown by the shaded regions (for $z = 12$) in Fig. 5. From Fig. 1, we see that galaxies with masses as low as $M_h = 10^{8.75} M_\odot$ start forming stars with the maximum allowed efficiency of $f_* = 0.03$. This implies that the galaxies on the UV LF at $z = 9-12$ are not themselves efficiency-limited but fuel-supply limited; the progenitors of galaxies at the faint-end were feedback limited, reducing their gas mass below the cosmological baryon fraction.

A by-product of our model is the prediction of the redshift-evolution of the faint-end slope of the UV LF. The exact value of this slope and its redshift-evolution is important, given that it is the galaxies occupying the faint-end of the UV LF that provide most of the photons for cosmic reionization (e.g. Choudhury & Ferrara 2007; Salvaterra et al. 2011). As explained in Sec. 3.1, the UV LF steepens with increasing redshift from $z = 5$ to 12, mirroring the behaviour of the scaled HMF which shifts to progressively lower masses. From our fiducial model we provide a simple functional form for our predicted redshift evolution of α (which can be used in reionization calculations):

$$\alpha = -1.75 \log z - 0.52. \quad (15)$$

This relation is valid over two orders of magnitude in luminosity, for the magnitude range $-18 \leqslant M_{UV} \leqslant -12$ and over the redshift range $z = 5$ to $z = 12$.

4 STELLAR MASS FUNCTION AND DENSITY

Having checked that our fiducial model predicts the correct evolving luminosity function for high-z galaxies, we now use it to study the evolving stellar mass function (SMF). The measurement of the SMF at extreme redshifts from observations has remained difficult because a robust estimate of M_* ideally requires rest-frame near-infrared data. In the absence of such data of the required depth, various indirect approaches can be used to get an estimate the SMF at a given z: (a) scaling the HMF with a constant factor assuming a certain $M_* - M_h$ relation, (b) scaling the theoretical UV LF assuming a $M_* - M_{UV}$ relation, (c) convolving the best-fit functional form (Schechter or DPL) of the observed UV LF with a mass-to-light (M/L) ratio including its scatter (e.g. González et al. 2011), and (d) once stellar masses have been derived by SED fitting to broad-band colours (albeit without the benefit of deep rest-frame near-infrared data), binning up the observed masses can yield an estimate of the SMF (e.g. Stark et al. 2009; Labbé et al. 2010a). This last approach is most similar to theoretical models that produce the SMF by binning galaxies on the basis of their stellar mass (e.g. Nagamine et al. 2010; Dayal & Ferrara 2012; Dayal et al. 2013; Hutter et al. 2014), albeit with the data the completeness of the result SMF is inevitably limited by the depth of the data and the wavelength of the selection band. We present the SMFs obtained by using all of the above four approaches in Sec. 4.2 below.

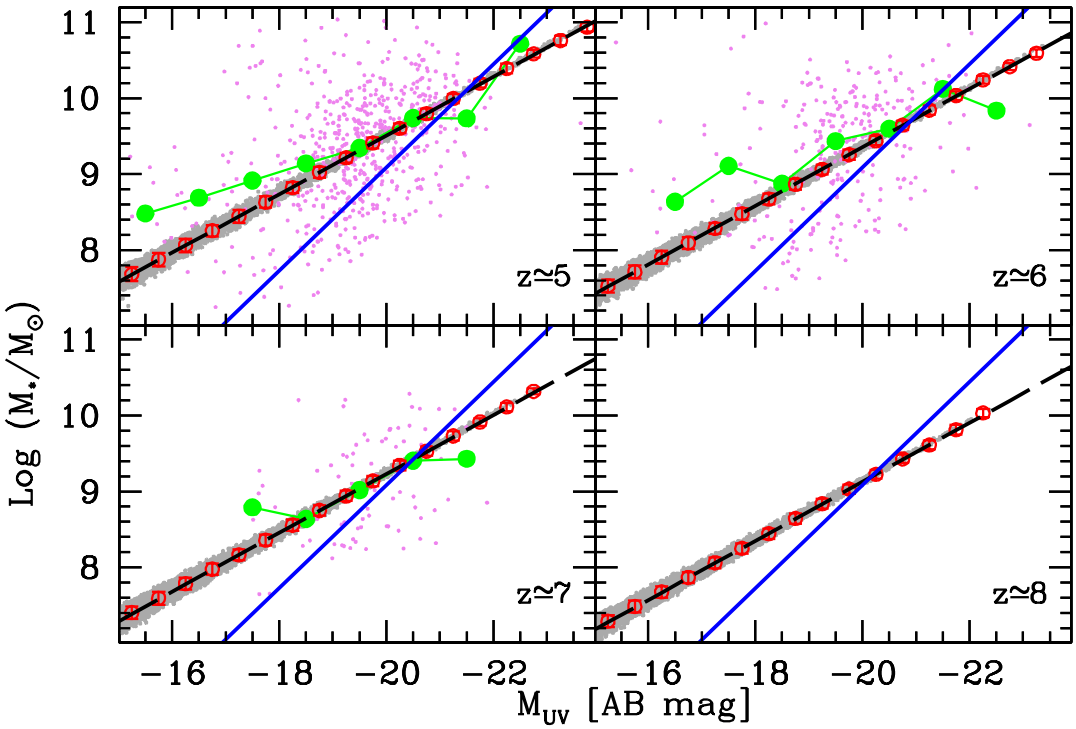

Figure 6. Mass-to-light relation showing galaxy stellar mass as a function of UV magnitude. Red points show the predicted average M_* value in each UV bin together with the 1σ error, and gray points show the predicted values for all galaxies brighter than $M_{UV} = -15$ at that redshift from the theoretical model. Violet points show the values for real galaxies in the CANDELS and HUDF fields as inferred by Grazian et al. (A&A submitted), with green points showing the observed medians in each UV bin. The black line shows our best-fit theoretical power-law relation, while the blue line shows the relation previously inferred from data by González et al. (2011) at $z = 4$ (and applied unchanged at higher redshifts).

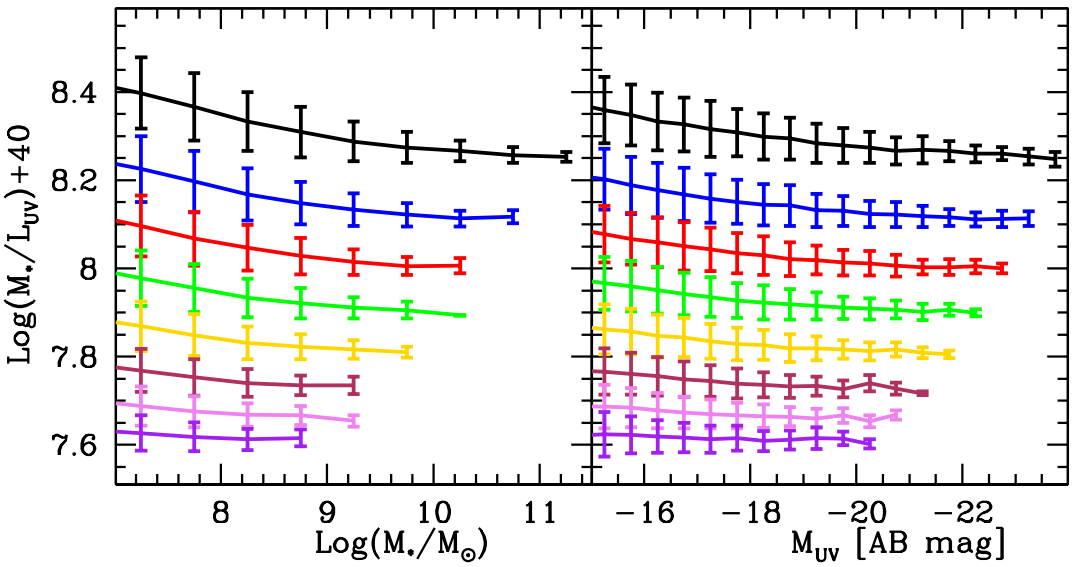

Figure 7. The theoretical mass-to-light ratio as a function of stellar mass (left panel) and UV absolute magnitude (right panel) for $z \simeq 5 - 12$ (from top to bottom); M_* and L_{UV} are in units of $M_\odot$ and erg/s/Å, respectively and we show this ratio arbitrarily scaled up by a factor of 10^{40}.

4.1 Mass-to-light relation through time

We start from the M/L relation that links the total stellar mass M_* and the UV luminosity L_{UV}, as obtained from our model. As noted in Eqn. 14, the UV luminosity from a burst of star formation declines rapidly with time as $L_{UV} \propto t^{-1.3}$. This implies that the UV luminosity of a starburst galaxy is typically dominated by the most recent burst. Combining this with the fact that increasingly massive galaxies have larger gas masses available for instantaneous star formation and have built up a larger stellar mass over their history, we would expect L_{UV} to scale with M_*. As shown in Fig. 6, this is indeed the behaviour shown by our model. We find that the $M_* - M_{UV}$ relation evolves smoothly with redshift with the form

$$\log M_* = \gamma M_{UV} + \delta, \tag{16}$$

where the slope of the relation has a constant value of $\gamma \simeq -0.38$ at all $z = 5$ to 12. However, the zero-point of the relation changes with redshift such that

$$\delta = -0.13\,z + 2.4, \tag{17}$$

over the redshift range $z = 5-12$. Thus, the normalisation of the M/L relation decreases by about 0.4 dex for an increase in redshift by $\Delta z = 3$; in other words, galaxies of a given luminosity are associated with lower M_* (or M_h) values with increasing redshift. As shown in Sec. 3.1, this increase in luminosity arises because galaxies form stars faster with increasing redshift, given the shorter cosmic time available for them to assemble a certain halo mass.

As shown in Fig. 6, the mass-to-light relation recently inferred from the CANDELS and HUDF data by Grazian et al. (in preparation; also Duncan et al., in preparation) differs considerably to that previously deduced by González et al. (2011). Grazian et al. find $\gamma = -0.4$ and $\delta = 1.6$ at $3.5 < z < 4.5$, which is in excellent agreement with our model predictions at $z = 5$ as shown in Fig. 6; a combination of observational factors including low galaxy numbers and an increase in the errors associated with the SED fitting to the broad band colours of faint objects ($M_{UV} \gtrsim -17$) probably leads to the slight mismatch between their observational and our theoretical values at the faint end (A. Grazian, Private communication). However, the slope (-0.68) and normalization (-4.51) inferred by González et al. (2011) at $z = 4$ (and applied to redshifts as high as $z = 7$) are very different both from our model and from the new Grazian et al. results.

The reason for the change in the mass-to-light relation inferred from the observations is not completely clear, but is almost certainly in part due to the improved near-infrared data (*HST* WFC3/IR Y_{105}, J_{125}, H_{160}, and VLT Hawk-I K-band) now available in the relevant deep survey fields. For example, while González et al. (2011) selected galaxies as LBGs based on their UV colours, Grazian et al. have been able to use H_{160} as the primary selection band, and even for those galaxies which would still be easily selected as LBGs have been able to use near-infrared detections to better constrain ages and hence stellar masses (where previously only near-infrared upper limits were available for many objects). Whatever the origin of this change, it is clear that new Grazian et al. results are in better agreement with the predicted mass-to-light relation which results from our model (with the median stellar masses inferred by González et al. (2011) being apparently under-estimated by about an order of magnitude). As shown in fig. 6, while the median values inferred by González et al. (2011) slowly converge towards the values found here (and by Grazian et al.) at brighter magnitudes, the steeper slope of their M/L relation inevitably affects the inferred SMF, as discuss below in Sec. 4.2.

We also show our M/L ratios as a function of M_* and M_{UV} in Fig. 7. As discussed above, galaxies of a given mass assemble faster with increasing redshift, i.e. they show higher luminosities for a given M_*, leading to a decrease in the amplitude of the M/L ratio. As expected from the evolution of this relation (see Eqn. 17), the M/L ratio decreases by about 0.3-0.4 dex for an increase in redshift of $\Delta z = 3$, both as a function of M_* and M_{UV}. Further, while the M/L ratio is flat for $z \geqslant 9$, it becomes a decreasing function of luminosity at later times. This is because the rate at which massive galaxies build up and form stars increases with time as they gain progressively larger gas masses (both from accretion and mergers), leading to a smaller M/L than low-mass systems (see Fig. 4).

4.2 Building stellar mass functions

We now show the evolving SMF obtained directly from our model, and compare it to those obtained by scaling the UV LF/HMF, and convolving functional fits (Schechter function and double power-law) to the UV LF with our M/L ratios (including associated scatter) as shown in Fig. 8. The evolving (fiducial) SMF is obtained by binning the number of theoretical galaxies in M_* bins at $z = 5 - 8$ for galaxies brighter than $M_{UV} = -18$ (the approximate limit of current observations) and extending to magnitudes as faint as -15 (as expected to be detectable by future instruments such as the *JWST*. Firstly, our model predicts that the faint-end of the SMF continues to rise to masses as low as $M_* = 10^7 \mathrm{M}_\odot$ at $z = 8$. This lower limit increases to $M_* \simeq 10^{7.8} \mathrm{M}_\odot$ at $z = 5$ as galaxies typically become more massive.

We now compare the SMF from our model to that built by González et al. (2011) integrating down to $M_{UV} = -18$. Starting with the low-mass end, we find the theoretical SMF is shifted towards higher M_* values (by about an order-of-magnitude) at all $z = 5 - 8$, since González et al. (2011) have under-estimated the median M/L ratio at this end (see sec. 4.1). At the massive-end, our SMF samples much higher M_* values since we integrate up to magnitudes as high as $M_{UV} = (-23.75, -22.5)$ at $z = (5, 8)$, compared to the value of $M_{UV} = -21$ used by González et al. (2011). However, the shape and amplitude of the theoretical and observationally-inferred SMFs are quite similar in the mass range $10^{8.4-10} \mathrm{M}_\odot$ due to the underlying function being compatible with a Schechter function in both cases; we remind the reader that while such functional shape arises naturally in our model, González et al. (2011) have obtained the SMF by explicitly convolving a Schechter function with their M/L ratio.

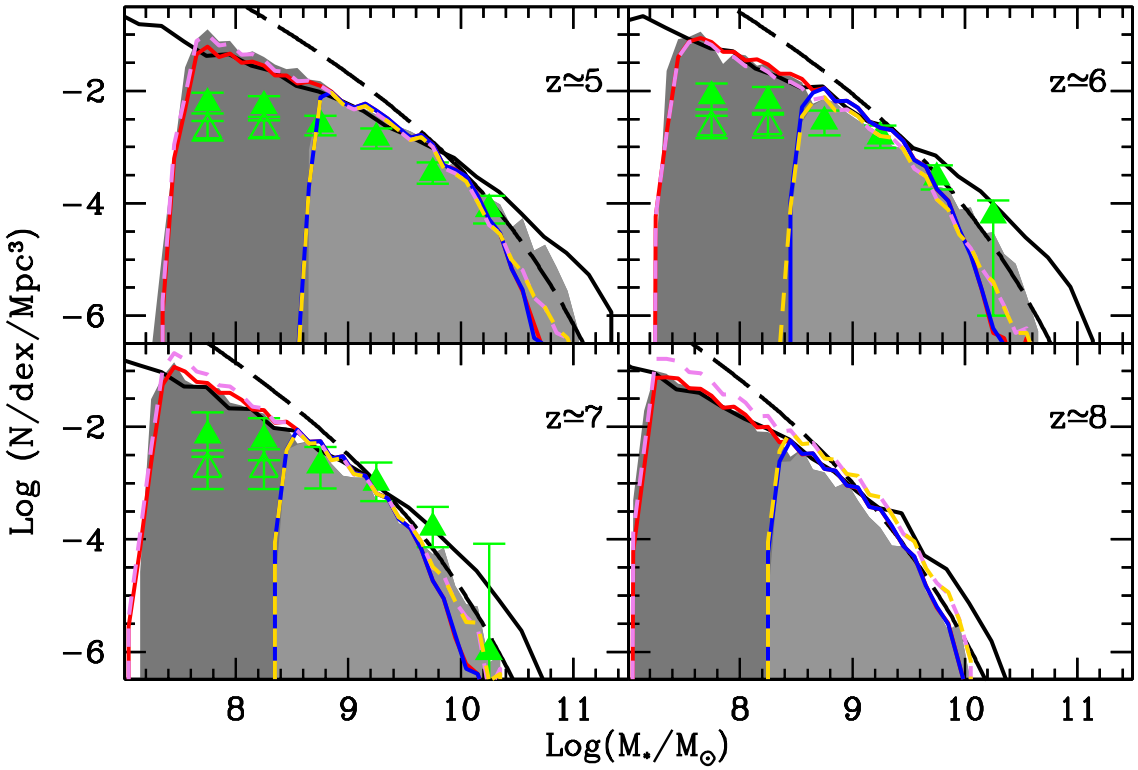

Figure 8. The evolving stellar mass function (SMF) for the redshift marked in each panel. In each panel, the dark (light) shaded regions show the SMF obtained by directly binning model galaxies brighter than $M_{UV} - 15$ and -18, respectively. The solid and dashed black lines have been obtained by scaling the theoretical UV LF (Fig. 3)and DM Halo mass functions at the appropriate redshift to match to the massive-end of the theoretical SMFs. The red (blue) lines show the SMF obtained by convolving the theoretical M/L ratio (Fig. 6) with the observationally-inferred best-fit Schechter function (McLure et al. 2009, 2013) for galaxies brighter than $M_{UV} = -15$ and -18, respectively. The violet (gold) lines show the SMF obtained by convolving the theoretical M/L ratio with the observationally-inferred best-fit double power law (DPL) function (Bowler et al. 2014) for galaxies brighter than $M_{UV} = -15$ and -18, respectively; we have used the $z = 7$ DPL at $z = 8$. Filled (empty) points show the corrected (uncorrected) stellar mass functions inferred observationally by González et al. (2011).

We now present theoretical SMFs obtained by convolving the M/L ratio with two different functional forms of the UV LF: the Schechter function (parameters from McLure et al. 2009, 2013) and the DPL (parameters from Bowler et al. 2014); we use the same DPL parameters at $z = 8$ as at $z = 7$. For each function, we sample the M_* value in each M_{UV} bin shown in Fig. 6 assuming a Gaussian error distribution to build the boot-strapped SMFs shown in Fig. 8. Starting at the massive-end, we find that the fiducial SMF is in better agreement with that obtained from the DPL at $z = 7$, while the SMF from the Schechter function under-estimates the SMF. This is as expected given that the fiducial UV LF is compatible with the slow drop-off shown by a double power-law for $z = 7$ at the bright-end (Sec. 3.1). Mirroring the fiducial UV LF that is over-estimated with respect to observations at $z = 5, 6$, the fiducial SMF is also over-estimated at the massive-end compared to both the Schechter function and the DPL. The UV LF is compatible with both the Schechter function and the DPL at the faint end, leading to the less-massive end of the fiducial SMF being in reasonable agreement with the SMFs from both these functions; as expected, the SMF inferred by González et al. (2011) is flatter than that obtained by either function given the different M/L provided by this work.

Interestingly, while the SMF obtained by arbitrarily scaling the UV LF matches the fiducial SMF at the faint end, it over-predicts the mass (by about 0.15 dex) for a given number density for $M_* \geqslant 10^{9.5}$M$_\odot$. As shown in Fig. 7, this is because the M/L ratios of galaxies decrease with increasing M_* (or M_{UV}), i.e. the SMF would be *over-estimated* at the bright-end assuming a constant M/L ratio across the whole UV LF.

As for the SMF obtained by scaling the HMF to match the bright-end, we find it over-estimates the faint-end. This is as expected, given that the galaxies at the faint-end are feedback-limited due to their progenitors having ejected all their gas; this reduces their M_* (and M_{UV}) value below that for larger galaxies.

Interestingly, while the HMF scaling required to reproduce the UV LF (see Table 1 and Fig. 3) is a function of redshift, the HMF scaling required to reproduce the bright end of the SMF (Fig. 8) is independent of redshift and has

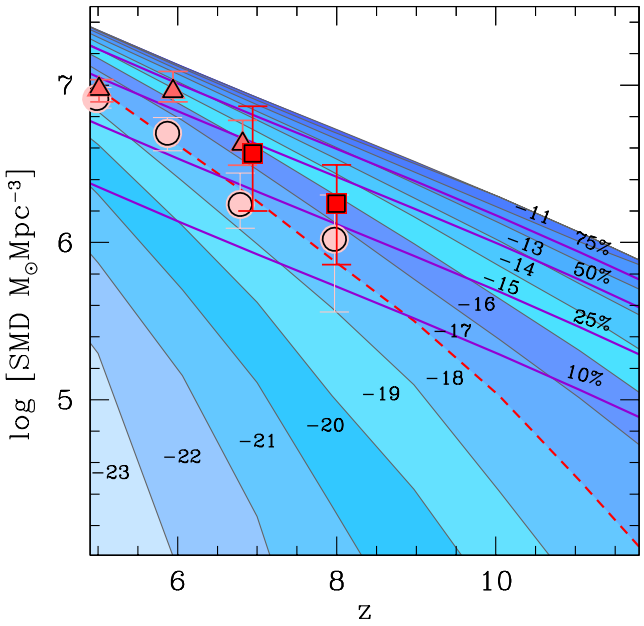

Figure 9. The stellar mass density (SMD) as a function of redshift. The different coloured contours show the contribution to the SMD from galaxies brighter than the magnitude value marked in the contour. The solid purple lines show fractions of the SMD to allow estimates of the magnitude limits which reveal the galaxies providing 10%, 25%, 50% and 75% of the total SMD at any redshift. The dashed red line shows the SMD from galaxies that have already been detected ($M_{UV} \lesssim -18$) to allow comparison with the data points: González et al. (2011, filled triangles), Stark et al. (2013, filled circles) and Labbé et al. (2010a,b, filled squares). Our model predicts that most of the stellar mass in the Universe at $5 \leqslant z \leqslant 12$ is locked up in systems too faint to be detected by the *HST*, but deep surveys with the *JWST* should reveal over a half (a fourth) of the total stellar mass in the Universe at redshifts as high at $z \simeq 5$ (9.5).

the physically reasonable value of $M_* = 0.016 M_h$. Essentially these two figures demonstrate that our fiducial model delivers the correct LF shape which cannot be produced by simple scaling of the HMF, while at the same time producing a stellar mass function which, at the high mass end, does indeed mirror the shape of the HMF.

To summarize, we find that assuming a constant M/L ratio across the whole UV LF would lead to an overestimation of the massive end of the SMF, while a scaled HMF overpredicts the faint end. A good match to the fiducial SMF can then be obtained from an M_* *that traces the UV luminosity at the low-mass end and the halo mass at the high-mass end.*

4.3 Cosmic stellar mass census

Our model also yields the stellar mass density (SMD; stellar mass per unit comoving volume) as a function of redshift, which can be directly compared to observations. Since both the halo mass function and the host halos of equally-luminous galaxies shift to progressively lower masses with increasing redshift, a comparison of the total theoretical SMD with observations is clearly a non-trivial test of our model. Encouragingly, the results of the model are in extremely good agreement with observationally inferred SMD values as shown in Fig. 9. Conducting a census of the total stellar mass we find that, as a result of their enormous numbers, small, faint galaxies contain most of the stellar mass in the Universe at $5 \leqslant z \leqslant 12$. Indeed, galaxies brighter than current observational limits ($M_{UV} \leqslant -18$) contain about 50% of the total stellar mass at $z \simeq 5$. This value then steadily decreases with redshift such that observed galaxies contain a quarter of the total stellar mass at $z \simeq 6.5$ and only 10% of the total stellar mass at $z \simeq 9$ (i.e. at redshifts $z \geqslant 6.5$, most of the stellar mass of the Universe is locked up in galaxies too small to have been detected so far). The next generation of space instruments such as the *JWST*, along with future developments in the use of gravitational lensing (e.g. in the Frontier Fields), will play an important role in revealing about half (a quarter) of the stellar mass in the Universe up to $z \simeq 8$ ($z \simeq 9.5$).

5 CONCLUSIONS

Recent deep *HST* and ground-based surveys have enabled the construction of statistically-significant samples of galaxies at $z \geqslant 5$, providing key new information on the evolving galaxy ultraviolet luminosity function (UV LF), mass-to-light ratios (M/L), stellar mass functions (SMF) and the integrated stellar mass Density (SMD) out to $z \simeq 10$. In this study we have endeavoured to isolate the essential physics driving early galaxy formation/evolution by building a very simple semi-analytical model wherein DM merger trees are implemented with the key physics of star formation, SN feedback and the resulting gas ejection, and the growth of progressively more massive systems (via halo mergers and gas accretion). In the spirit of maintaining simplicity, our model utilises a total of two *redshift- and mass-independent free-parameters*: the star formation efficiency threshold, f_*, and the fraction of SN energy that drives winds, f_w. Our model is based on the single premise that any galaxy can form stars with a maximal *effective efficiency* (f_*^{eff}) *that provides enough energy to expel all the remaining gas, quenching further star formation.* The value of $f_*^{eff} = min[f_*, f_*^{ej}]$ where f_*^{ej} is the star-formation efficiency required to eject all gas from a galaxy. In this model, low-mass galaxies form stars with an effective efficiency that is sufficient to eject all gas from a galaxy $f_*^{eff} = f_*^{ej}$, while massive systems can form stars with a larger fixed efficiency value of $f_*^{eff} = f_*$.

This simple model reproduces both the slope and amplitude of the evolving UV LF at $z = 5$ to $z = 10$ with parameter values $f_* = 0.03$ and $f_w = 0.1$ (i.e. a model wherein at most 3% of the gas can form into stars and 10% of SNe energy is converted into kinetic form to drive outflows). Although the Halo Mass Function (HMF) scaled to match the knee of the observed UV LF shifts to progressively lower masses at high z, our model correctly reproduces the evolving faint-end slope (α) of the UV LF; the physical explanation is the faster assembly of galaxies with increasing

redshift. Additionally, the model naturally predicts that the bright-end slope of the UV LF is somewhat flatter than the steep exponential drop-off provided by the Schechter function, and is better described as either following a double power-law Bowler et al. (2014), or by the high-mass shape of the HMF.

Galaxies with $M_{UV} \leqslant -15$ are hosted in halos with mass $M_h \geqslant 10^{9.5}(10^9)\mathrm{M}_\odot$ for $z = 5-8\,(12)$, a mass scale above the feedback-limited regime at these redshifts. Hence, the faint-end of the UV LF lies below the HMF because the progenitors of these galaxies have ejected all their gas, resulting in dry mergers. These systems are therefore *fuel-starved*, rather than feedback-limited. We also show that the physics driving galaxy growth is halo mass-dependent: while the smallest halos build up their gas (and stellar) mass by accreting gas from the IGM, the bulk of the gas powering star formation in the largest halos is brought in by merging progenitors. From our model we have provided a simple functional form for the redshift evolution of the faint-end slope, $\alpha = -1.75\log z - 0.52$, a key ingredient for reionization calculations since faint galaxies provide the bulk of hydrogen ionizing photons ($E \geqslant 1$ Ryd).

The stellar mass-to-light relation from our model is well-fit by the functional form $\log M_* = -0.38M_{UV} - 0.13\,z + 2.4$. This relation is in excellent agreement with the recent observational results obtained by Grazian et al. (A&A submitted), but is significantly flatter (i.e. higher average masses at low luminosities) than the relation previously deduced from observations of LBGs by González et al. (2010). We also show the SMFs obtained by (a) binning up M_* from the fiducial model, (b) convolving the M/L relation with a Schechter function and DPL, (c) scaling the fiducial UV LF, and (d) scaling the HMF. A good match to the fiducial SMF can be obtained by using a M_* tracing the UV luminosity at the faint-end and the HMF at the high-mass end.

The census of the cosmic SMD implies that while 50% of the SMD is in detected LBGs at $z \simeq 5$, this fraction drops steadily to 10% at $z \simeq 9$. While the next generation of instruments such as the *JWST* should be capable of revealing up to 25% of the SMD at $z \simeq 9.5$, most of the SMD at high-z is predicted to be locked up in galaxies that will likely remain "invisible" for the immediate future.

We end with a few caveats. First, it appears that the fiducial model slightly over-predicts the bright-end of the UV LF at $z = 5, 6$. This could either arise due to physical effects that have been ignored in order to maintain simplicity (e.g. dust attenuation, AGN feedback, mass quenching) which may become increasingly important with reducing redshift, or due to current data limitations at the brightest UV luminosities (a situation that should be resolved soon by Bowler et al. in preparation). Second, our model assumes a maximally-efficient feedback scenario wherein SN-powered kinetic winds in small galaxies can sweep up all the gas and eject it out of the system. Although this is justifiable based on energy balance arguments, it might be possible that a fraction of the gas remains within the halo potential well, albeit in a heated/dynamically perturbed state not allowing stars to readily form (Mori et al. 2002; Fangano et al. 2007). Third, we have ignored feedback mechanisms including (i) the energy injected by massive stars before the onset of SN explosions (Hopkins et al. 2011; Stinson et al. 2013) and, (ii) the photo-evaporation of gas on the outskirts of galaxies due to a UV background (UVB) created by reionization. Point (i) is justified by our assumption of instantaneous star formation and ISM SN energy injection. However, making the simplifying assumption that stellar feedback only adds to the total energy injected into the ISM gas (Eqn. 1), f_*^{ej} can be left unchanged by decreasing the value of f_w by the same amount. As for (ii), we show that SN feedback always dominates over the effect of the UVB (Dayal et al. 2014); hence including the decrease in the baryon fraction due to a UVB would leave our results unchanged. However, these details of feedback are highly model-dependent and progress can be made only by comparing such predictions with actual data as those imminently expected from the Frontier Fields and ground-based observatories (Yue et al. 2014, in prep.)

In the future, our aim is to include the effects of AGB dust, AGN feedback and the UVB created by both Hydrogen and helium reioniozation. These are some of the main physical effects whose inclusion will allow our model to be extended all the way to $z = 0$.

ACKNOWLEDGMENTS

JSD and PD acknowledge the support of the European Research Council via the award of an Advanced Grant. JSD also acknowledges the support of the Royal Society via a Wolfson Research Merit award, and the contribution of the EC FP7 SPACE project ASTRODEEP (Ref. No: 312725). The authors thank A. Grazian and co-authors for allowing us to use their results and for their constructive comments. PD thanks A. Mazumdar and the anonymous referee for their insightful comments.

REFERENCES

Aguirre A., Hernquist L., Schaye J., Katz N., Weinberg D. H., Gardner J., 2001, ApJ, 561, 521
Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101
Benson A. J., 2012, New A, 17, 175
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, ApJ, 670, 928
Bouwens R. J. et al., 2010a, ApJ, 725, 1587
Bouwens R. J. et al., 2011, ApJ, 737, 90
Bouwens R. J. et al., 2013, ArXiv:1306.2950
Bouwens R. J. et al., 2014, ArXiv:1403.4295
Bouwens R. J. et al., 2010b, ApJ, 708, L69
Bower R. G., 1991, MNRAS, 248, 332
Bowler R. A. A. et al., 2012, ArXiv e-prints
Bowler R. A. A. et al., 2014, MNRAS, 440, 2810
Bradley L. D. et al., 2012, ApJ, 760, 108
Bradley L. D. et al., 2013, ArXiv:1308.1692
Castellano M. et al., 2010, A&A, 524, A28

Choudhury T. R., Ferrara A., 2007, MNRAS, 380, L6
Coe D. et al., 2013, ApJ, 762, 32
Cole S., 1991, ApJ, 367, 45
Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781
Croton D. J. et al., 2006, MNRAS, 365, 11
Dayal P., Dunlop J. S., Maio U., Ciardi B., 2013, MNRAS, 434, 1486
Dayal P., Ferrara A., 2012, MNRAS, 421, 2568
Dayal P., Ferrara A., Saro A., 2010, MNRAS, 402, 1449
Dayal P., Ferrara A., Saro A., Salvaterra R., Borgani S., Tornatore L., 2009, MNRAS, 400, 2000
Dayal P., Mesinger A., Pacucci F., 2014, ArXiv e-prints
De Lucia G., Boylan-Kolchin M., Benson A. J., Fontanot F., Monaco P., 2010, MNRAS, 406, 1533
Dubinski J., Carlberg R. G., 1991, ApJ, 378, 496
Dunlop J. S., McLure R. J., Robertson B. E., Ellis R. S., Stark D. P., Cirasuolo M., de Ravel L., 2012, MNRAS, 420, 901
Dunlop J. S. et al., 2013, MNRAS, 432, 3520
Ellis R. S. et al., 2013, ApJ, 763, L7
Fangano A. P. M., Ferrara A., Richter P., 2007, MNRAS, 381, 469
Finkelstein S. L. et al., 2012, ApJ, 756, 164
Finlator K., Davé R., Oppenheimer B. D., 2007, MNRAS, 376, 1861
Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576
Forero-Romero J. E., Yepes G., Gottlöber S., Knollmann S. R., Cuesta A. J., Prada F., 2011, MNRAS, 415, 3666
Gnedin N. Y., 1998, MNRAS, 294, 407
González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, ApJ, 735, L34
González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., Brammer G. B., 2010, ApJ, 713, 115
Greif T. H., Johnson J. L., Bromm V., Klessen R. S., 2007, ApJ, 670, 1
Hinshaw G. et al., 2013, ApJS, 208, 19
Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950
Hutter A., Dayal P., Partl A. M., Müller V., 2014, ArXiv:1401.2830
Jaacks J., Choi J.-H., Nagamine K., Thompson R., Varghese S., 2012, MNRAS, 420, 1606
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Labbé I. et al., 2010a, ApJ, 716, L103
Labbé I. et al., 2010b, ApJ, 708, L26
Labbé I. et al., 2013, ApJ, 777, L19
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lacey C., Silk J., 1991, ApJ, 381, 14
Lange A. E. et al., 2001, Phys. Rev. D, 63, 042001
Leitherer C. et al., 1999, ApJS, 123, 3
Lu Y. et al., 2013, ArXiv:1312.3233
Mac Low M.-M., Ferrara A., 1999, ApJ, 513, 142
McLure R. J., Cirasuolo M., Dunlop J. S., Foucaud S., Almaini O., 2009, MNRAS, 395, 2196
McLure R. J. et al., 2013, MNRAS, 432, 2696
McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D. P., Targett T. A., Ellis R. S., 2010, MNRAS, 403, 960
McLure R. J. et al., 2011, MNRAS, 418, 2074
Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999a, ApJ, 524, L19
Moore B., Quinn T., Governato F., Stadel J., Lake G., 1999b, MNRAS, 310, 1147
Mori M., Ferrara A., Madau P., 2002, ApJ, 571, 40
Nagamine K., Ouchi M., Springel V., Hernquist L., 2010, PASJ, 62, 1455
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Oesch P. A. et al., 2010a, ApJ, 709, L21
Oesch P. A. et al., 2010b, ApJ, 709, L16
Oesch P. A. et al., 2013, ApJ, 773, 75
Ouchi M. et al., 2010, ApJ, 723, 869
Padovani P., Matteucci F., 1993, ApJ, 416, 26
Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557
Peng Y.-j. et al., 2010, ApJ, 721, 193
Planck Collaboration ., 2013, ArXiv e-prints:1303.5076
Planck Collaboration et al., 2013, ArXiv e-prints:1303.5062
Postman M. et al., 2012, ApJS, 199, 25
Rogers A. B. et al., 2014, MNRAS, 440, 3714
Salvaterra R., Ferrara A., Dayal P., 2011, MNRAS, 414, 847
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Smit R. et al., 2014, ApJ, 784, 58
Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V. et al., 2008, MNRAS, 391, 1685
Stark D. P., Ellis R. S., Bunker A., Bundy K., Targett T., Benson A., Lacy M., 2009, ApJ, 697, 1493
Stark D. P., Schenker M. A., Ellis R., Robertson B., McLure R., Dunlop J., 2013, ApJ, 763, 129
Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013, MNRAS, 428, 129
Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Wilkins S. M., Bunker A. J., Stanway E., Lorenzoni S., Caruana J., 2011, MNRAS, 417, 717
Yoshida N., Stoehr F., Springel V., White S. D. M., 2002, MNRAS, 335, 762
Zheng W. et al., 2012, Nature, 489, 406