A predicted new population of UV-faint galaxies at $z \gtrsim 4$

J. Stuart B. Wyithe, Abraham Loeb and Pascal A. Oesch

1 School of Physics, University of Melbourne, Parkville, Victoria, VIC 3010, Australia
2 Astronomy Department, Harvard University, 60 Garden Street, Cambridge, MA 02138, USA
3 UCO/Lick Observatory, University of California, Santa Cruz, CA 95064, USA

ABSTRACT

We show that a bursty model of star formation explains several puzzling observations of high-redshift galaxies. We begin by showing that because the observed star formation rate integrated over a Hubble time exceeds the observed stellar mass by an order of magnitude, the specific star formation rate requires a duty-cycle of $\sim 10$ per cent. We use the specific star formation rate to calibrate a merger-driven model of star formation regulated by supernova feedback, and reproduce the star formation rate density and stellar mass functions of galaxies at $4 \lesssim z \lesssim 7$. The specific star formation rate is predicted not to evolve rapidly with either mass or redshift at $z \gtrsim 4$, consistent with observation. This is in contrast to expectations from hydrodynamical simulations where star formation closely follows accretion rate, and increases strongly towards high redshift. Bursty star formation explains the observation that there is not enough stellar mass at $z \sim 2–4$ to account for all star formation observed. A duty-cycle of $\sim 10$ per cent implies that there could be 10 times the number of known high-redshift galaxies at fixed stellar mass that have not yet been detected through UV selection. We therefore predict the possible existence of an undetected population of UV-faint galaxies that accounts for most of the stellar mass density at $z \sim 4–8$.

Key words: galaxies: formation – galaxies: high-redshift – cosmology: theory.

1 INTRODUCTION

The galaxy luminosity function represents one of the most important observables for studying the reionization of cosmic hydrogen (e.g. Bouwens et al. 2011; Loeb & Furlanetto 2012; McClure et al. 2013; Oesch et al. 2012, 2013; Ellis et al. 2013). Gaining an understanding of the reionization process therefore requires development of a theoretical picture of the processes involved in star formation at high redshift (e.g. Trenti et al. 2010; Finlator, Oppenheimer & Davé 2011; Muñoz & Loeb 2011; Raicic, Theuns & Lacey 2011; Salvaterra, Ferrara & Dayal 2011; Tacchella, Trenti &Carollo 2013). The observed properties of high-redshift galaxies have been modelled using complex hydrodynamical simulations coupled with sub-grid models for processes including star formation and metal enrichment (e.g. Finlator et al. 2011; Salvaterra et al. 2011; Jaacks, Nagamine & Choi 2012b). These models broadly reproduce the luminosity function evolution as well as the blue colours of the young stellar populations at high redshift. However, while these models are able to reproduce the luminosity function and star formation rate density (SFRD) function, they overproduce the high-redshift stellar mass function, particularly at the low-mass end (Jaacks et al. 2012a).

Recently, Wyithe & Loeb (2013) presented a model for the high-redshift SFRD function, which includes merger-driven star formation regulated by supernova (SN) feedback. This model fits a range of observables, and implies a star formation duty-cycle of only 1–10 per cent, much lower than found in hydrodynamical simulations (Jaacks et al. 2012b). This small duty-cycle is supported by analysis of the spectral energy distributions (SED) of Ly-break galaxies (LBG) including nebular emission, which indicate that star formation histories should be episodic (de Barros, Schaerer & Stark 2012). The model of Wyithe & Loeb (2013) successfully predicts the observed relation between star formation rate and stellar mass at $z \gtrsim 4$.

A puzzling observation in recent high-redshift galaxy research has been that the star formation rate per stellar mass (specific star formation rate; sSFR) does not seem to evolve significantly with either mass or redshift in the range $4 \lesssim z \lesssim 8$ (e.g. Stark et al. 2009; González et al. 2010). While subsequent analyses indicated that the absolute value of the sSFR at $z \geq 4$ might have been underestimated in these first derivations after including updated estimates of dust extinction and accounting for the impact of rest-frame optical emission lines, the current best observational estimates indicate relatively slow evolution across $z \sim 4–7$ (see e.g. de Barros et al. 2012;
Figure 1. The sSFR as a function of redshift calculated based on equation (1) for duty-cycles of 10 per cent and 15 per cent (solid lines) and 100 per cent (dotted lines), in comparison with measurements at stellar masses of 10⁹ and 5 × 10⁹ M⊙. The open triangles are data points from Stark et al. (2013). Other data points are from González et al. (2014), and the labels represent stellar masses based on the assumptions of a constant (CSF) and rising (RSF) star formation history without and with emission lines included (em. line) as described in that paper. We note that the model curves should be compared to the CSF points. We also note that the error bars of González et al. (2014) represent uncertainties on the mean, and are therefore significantly smaller than the ones from Stark et al. (2013), which represent the variance within the sample.

González et al. (2014; Stark et al. 2013). However, the observational debate is far from settled (e.g. Smit et al. 2013).

Most simulations of high-redshift galaxy formation do not produce a plateauing of sSFR at z > 4. This is because simulations generally associate star formation primarily with the accretion of gas. As a result they predict a rapid increase in the sSFR, which can be understood because the specific accretion rate is found to scale as (1 + z)^3.5 (Neistein & Dekel 2008). A possible explanation for the plateauing of sSFR at high redshift was proposed by Pippino, Calura & Matteucci (2013) who suggest that the phase of build-up of stellar mass in spheroids implies high sSFR values at high redshifts, and that the observation of multiple generations of spheroids explains the lack of evolution in sSFR at z > 4. Other possibilities include reduced star formation efficiencies at high redshift due to metallicity dependent star formation (e.g. Krumholz & Dekel 2012). Weinmann, Neistein & Dekel (2011) calculated the specific star formation history within a suite of semi-analytic models. At z > 4, they found that weak evolution of sSFR could be reproduced in the presence of strong SN feedback. In this paper, we find that a SN-regulated model with star formation triggered by mergers and a low duty-cycle naturally reproduces both the large value of specific SFR, and a weak dependence of sSFR on mass and redshift.

The relationship between the observed star formation rate and stellar mass per unit volume has also been an observational focus. Wilkins, Trentham & Hopkins (2008) compiled estimates of stellar mass and star formation rates as a function of redshift in order to investigate whether the integral of star formation rate matches the observed stellar mass. Puzzlingly, at z ~ 2–4, Wilkins et al. (2008) find that there is not enough stellar mass to account for all of the star formation observed. Conversely, at high redshift, Bouwens et al. (2011) and Robertson et al. (2013) find that the observed stellar mass is accounted for by the observed star formation rate (see also Stark et al. 2013). In this paper, we argue that both observations can be understood in the context of a star formation model with a duty-cycle of the order of 10 per cent.

We begin in Section 2 by pointing out the general constraint on the duty-cycle that is provided by observations of the sSFR. Then, in Section 3, we briefly summarize the model for high-redshift star formation presented in Wyithe & Loeb (2013). We next present a comparison of this model with various observables including the SFRD, sSFR, clustering amplitude and stellar mass function in Section 4. We discuss the detectability of a predicted population of low UV luminosity galaxies in Section 5, and finish with a discussion in Section 6. In our numerical examples, we adopt the standard set of cosmological parameters (Komatsu et al. 2011), with values of Ω_m = 0.04, Ω_λ = 0.24 and Ω_k = 0.76 for the density parameters of matter, baryon and dark energy, respectively, h = 0.73, for the dimensionless Hubble constant, and σ_8 = 0.82.

2 THE sSFR OF STAR-FORMING GALAXIES

Before discussing our particular model for SNe-regulated star formation, we begin by looking at the very general constraint on duty-cycle (ε_{duty, tot}) provided by the sSFR. In the simplest model of constantly star-forming galaxies, the sSFR is

\[ \frac{\text{sSFR}}{M_*} = \frac{\text{SFR}}{\text{SFR}} \times (\epsilon_{\text{duty, tot}} H^{-1})^{-1}, \]  

Thus, the sSFR leads to a direct estimate of the duty-cycle of star formation averaged over a Hubble time. This is plotted in the upper two panels of Fig. 1 as a function of redshift for ε_{duty, tot} = 0.1 and 0.15 (solid lines), compared with observations of sSFR at stellar masses of M_* = 10^9 M⊙ and 5 × 10^9 M⊙, respectively. We find weak dependence of the inferred duty-cycle over the range of stellar mass and redshift probed. The dotted lines also show the curves corresponding to a duty-cycle of unity, illustrating the level at which the values of sSFR deviate from expectations of continuous star formation. We note that a rising star formation history in a galaxy will result in a smaller predicted value of stellar mass and hence a larger predicted value of sSFR. The level of this effect, which is much smaller than for the range of duty-cycle shown in Fig. 1, can be seen by comparing the range of sSFR values measured by González et al. (2014) assuming constant and rising star formation histories, respectively. A low duty-cycle has a range of important implications for the properties of the high-redshift galaxy population, and explains several puzzling properties of the observed relation between stellar mass and star formation rate. For the remainder of this paper, we explore these explanations in the context of the merger-driven model of Wyithe & Loeb (2013). However, we stress that the result of low duty-cycle from equation (1) is very general and not dependent on the details of our particular star formation model.

3 MODEL

In this section, we follow the modelling of the SFRD function of high-redshift galaxies presented in Wyithe & Loeb (2013), and
compare this model with the observations of Smit et al. (2012). The reader is referred to Wyithe & Loeb (2013) for details of this model.

The star formation rate in a galaxy halo of mass $M$ that turns a fraction $f_\text{s}$ of its disc mass $m_\text{d}M$ into stars over a time can be estimated as $\text{SFR}$

$$\text{SFR} = 0.15 M_\odot \text{yr}^{-1} \left( \frac{m_\text{d}}{0.17} \right) \left( \frac{f_\text{s}}{0.1} \right) \left( \frac{M}{10^8 M_\odot} \right) \left( \frac{t_\text{SF}}{10^7 \text{yr}} \right)^{-1} \text{.}$$

(2)

In our model, bursts of star formation are assumed to be triggered by major mergers, yielding an estimate of the SFRD (i.e. galaxies per Mpc$^{-3}$ per unit of SFR) described by

$$\Phi(\text{SFR}) = \epsilon_{\text{duty}} \left( \frac{\Delta M}{t_H} \right) \left( \frac{dN_{\text{merge}}}{dM} \right) \left( \frac{dt}{dM} \right)^{-1} \text{,}$$

(3)

where $\epsilon_{\text{duty}}$ is the fraction of the Hubble time ($t_H$) over which each burst lasts, and $dN/dM$ is mass function of dark matter haloes (Press & Schechter 1974; Sheth & Tormen 1999). To calculate the rate of major mergers ($dN_{\text{merge}}/dt$), we find the number of haloes per logarithm of mass $\Delta M$ per unit time that merge with a halo of mass $M_1$ to form a halo of mass $M$ (Lacey & Cole 1993). As in Wyithe & Loeb (2013), we define major mergers to have a 2:1 mass ratio (i.e. $M_1 = \frac{1}{2} M$ and $\Delta M = M/3$).

The duty-cycle of the starburst is defined as

$$\epsilon_{\text{duty}} = \frac{t_\text{SF} + t_H}{t_H} \text{,}$$

(4)

where we have assumed that the most massive stars fade after $t_\text{SF} \sim 3 \times 10^6$ yr (Barkana & Loeb 2001), and defined the starburst lifetime as $t_\text{SF}$. We define

$$\Psi(\text{SFR}) = \ln 10 \times \text{SFR} \times \Phi \text{,}$$

(5)

which has units of Mpc$^{-3}$ dex$^{-1}$, for direct comparison with published observations.

Our model includes a prescription for SN feedback in which a fraction $F_{\text{SN}}$ of each SN energy output, $E_{\text{SN}}$, heats the galactic gas mechanically. The total stellar mass is given by $M_\star = m_\star M_{\text{f, int}}$, where $f_{\text{int}} = N_{\text{merge}} f_\text{s}$, and $N_{\text{merge}}$ is the number of major mergers per Hubble time. The feedback model assumes a self-gravitating exponential disc with scale radius $R_\text{d}$, mass fraction relative to the halo $m_\text{d}$ and spin parameter $\lambda \sim 0.05$ (Mo, Mao & White 1998). Wyithe & Loeb (2013) found

$$f_\text{s} = \frac{0.008}{N_{\text{merge}}} \left( \frac{M}{10^{10} M_\odot} \right)^{2/3} \left( \frac{1 + z}{10} \right) \left( f_\text{d} F_{\text{SN}} \right)^{-1} \text{.}$$

(6)

The parameter $f_\text{s}$ describes the fraction of SN energy that contributes to feedback owing to the finite time-scale of the SN feedback, while $f_\text{d}$ describes the fraction of SN energy that contributes owing to the finite disc scaleheight. These values are calculated based on the model of Clarke & Oey (2002), in which clusters of $N_\text{c}$ SN produce superbubbles with a radius $R_\text{s}$. The characteristic time-scale associated with the feedback is set by the lifetime of the lowest mass SN progenitor ($t_\text{e} \sim 4 \times 10^7$ years). Wyithe & Loeb (2013) found

$$f_\text{d} = 0.85 \left( \frac{N_\text{c}}{10} \right)^{-1/3} \left( \frac{E_{\text{SN}}}{10^{51} \text{erg}} \right)^{-1/3} \left( \frac{\lambda}{0.05} \right)^{2/3} \left( \frac{m_\text{d}}{0.17} \right)^{-1/3} \times \left( \frac{M}{10^8 M_\odot} \right)^{-1/9} \left( \frac{1+z}{10} \right)^{-2/3} \left( \frac{c_\text{s}}{10 \text{km s}^{-1}} \right)^2 \text{,}$$

(7)

as long as $f_\text{d} < 1$ and $f_\text{d} = 1$ otherwise. Here, $c_\text{s}$ is the sound speed in the gas, which we assume to have a temperature of $10^4 \text{ K}$. The break corresponds to a superbubble radius that is larger than the scaleheight of the disc. In cases where $t_\text{SF} < t_\text{e} \sim 4 \times 10^7$ yr, only

$$f_\text{d} \equiv \left( t_\text{SF}/t_\text{e} \right)^2$$

(8)

of the overall SN energy output is generated by the time the starburst concludes. In cases where $t_\text{SF} > t_\text{e}$, we have $f_\text{d} = 1$.

We utilize equations (6)–(8) with equation (3) as a function of the parameters $t_\text{SF}$ and $f_\text{max}$.

4 RESULTS

Wyithe & Loeb (2013) used the merger-driven model described above to argue that the duty-cycle for star formation in high-redshift galaxies was a few per cent. However, while much lower than unity the model-independent estimate based on the sSFR shown in Fig. 1 is significantly larger, with a value of $\sim 0.1$. To understand the origin of this difference we note that there is a degeneracy in the model between the duty-cycle $\epsilon_{\text{duty}}$ and the parameter $F_{\text{SN}}$ which governs the fraction of SN energy that is harnessed for feedback. Wyithe & Loeb (2013) arbitrarily chose a value for this parameter since it is unconstrained by just the SFRD function. However, for the analysis in this paper the inclusion of the constraint on sSFR allows us to constrain the value of $F_{\text{SN}}$ in addition to the values of $f_{\text{max}}$ and $t_\text{SF}$.

4.1 Comparison with observations

To constrain the parameters $t_\text{SF}$ and $f_{\text{max}}$, we fit our model to both the SFRD function (Smit et al. 2012), and the sSFR (González et al. 2014). We fit our model separately at the four different redshifts $z \sim 4$–7 considered. To find the allowed regions of parameter space, we calculate the $\chi^2$ of the model as

$$\chi^2(f_{\text{max}}, t_\text{SF}) = \sum_{i=0}^{N_{\text{obs}}} \left( \frac{\log \Psi(\text{SFR}_i, f_{\text{max}}, t_\text{SF}, z) - \log \Psi_{\text{obs}}(\text{SFR}_i, z)}{\sigma_{\text{SFR}}(\text{SFR}_i, z)} \right)^2 + \left( \frac{\log s\text{SFR}(f_{\text{max}}, t_\text{SF}, z) - \log s\text{SFR}_{\text{obs}}(z)}{\sigma_{\text{SFR}}(z)} \right)^2 \text{.}$$

(9)

Here, $\Psi_{\text{obs}}(\text{SFR}_i, f_{\text{max}}, t_\text{SF}, z)$ is the observed SFRD at redshift $z$. The uncertainty in SFRD (in dex) is given by $\sigma_{\text{SFR}}(\text{SFR})$.

While our best-fitting model provides an excellent description of the SFRD functions as shown in Wyithe & Loeb (2013), the reduced $\chi^2$ based on the quoted errors does not suggest a formally acceptable best fit. This results in an artificially small region of parameter space that is allowed around the best fit. To show a more realistic estimate of the constraints of allowed values of the physical input parameters, we have increased the quoted error bars on the SFRD function by factors of 3 and 2, respectively, when calculating likelihoods at $z \sim 4$ and $z \sim 5$, so as to obtain a reduced $\chi^2$ of order unity. We note that the best-fitting regions without this error bar inflation lie within the ranges quoted.

The value of sSFR is evaluated at a stellar mass of $M_\star = 10^8 M_\odot$. Given the systematic uncertainties on measurements of sSFR which arise from the unknown star formation history and impact of emission lines on stellar mass estimates, we choose an uncertainty corresponding to the range of estimates for the two assumed star formation histories and the two cases of with/without emission lines considered in González et al. (2014).
UV-faint galaxies at $z \gtrsim 4$

Figure 2. Constraints on the model parameters $f_{\star,\text{max}}$ and $t_{\text{SF}}$ at four different redshifts (constraints are independent at each redshift). In each case, three contours are shown corresponding to differences in $\chi^2$ relative to the best-fitting model of $\Delta \chi^2 = \chi^2 - \chi^2_{\text{min}} = 1, 2.71$ and 6.63. Projections of these contours on to the axes provide the 68.3, 90 and 99 per cent confidence intervals on individual parameter values. The vertical grey regions represent time-scales longer/shorter than the lifetime of the highest/lowest mass SN progenitor ($3 \times 10^6$yr/$4 \times 10^7$yr). Because our model does not provide a best fit that is formally consistent with the data (see discussion in text), we have chosen to present examples for model parameters $t_{\text{SF}}$ and $f_{\star,\text{max}}$ that are close to the best fit, and which all have the same value of $t_{\text{SF}}$. The symbols refer to the parameter locations of models shown in Figs 3–8.

Finally, the SFRD function and sSFR are sensitive to the value of $F_{\text{SN}}$, and we therefore integrate the likelihood over a range of values uniformly distributed within the range $-2.5 < \log_{10} F_{\text{SN}} < 0$:

$$\mathcal{L}(f_{\star,\text{max}}, t_{\text{SF}}) \propto \int_{-2.5}^{0} \text{d}(\log_{10} F_{\text{SN}}) e^{-\chi^2/2}.$$  \hfill (10)

As noted in Wyithe & Loeb (2013), the relation between SFR and $M$ in equation (2) is not perfect. To account for scatter in this relationship, we therefore convolve the model SFRD function from equation (3) with a Gaussian of width 0.5 dex in SFR (González et al. 2011).

4.2 Parameter constraints

In Fig. 2, we show constraints on models. Given the sSFR-driven constraints on $F_{\text{SN}}$, we find that the shape of the SFRD function requires starburst durations of a few tens of Myr at $z \sim 5$, $z \sim 6$ and $z \sim 7$, with a few per cent of the gas turned into stars per burst. This value can be compared with the lifetimes shorter than that of the most massive stars ($t_s \sim 3 \times 10^6$ yr, left-hand grey region), as well as with times in excess of the lifetime of the least massive stars that produce SNe (right-hand grey region). Overall, we find that star formation in high-redshift galaxies is terminated on the same time-scale as feedback from SNe can be produced, as was shown in Wyithe & Loeb (2013).

We show the comparison between observed and modelled SFRD functions in Fig. 3. The four different redshifts are illustrated with examples having values of $F_{\text{SN}} = 0.03$ at $z = 4$ and $F_{\text{SN}} = 0.1$ at $z = 5–7$ (cf. $F_{\text{SN}} = 0.1$ and 0.3 in Wyithe & Loeb 2013). The four curves shown correspond to model parameters $t_{\text{SF}}$ and $f_{\star,\text{max}}$ labelled by the symbols in Fig. 2. We note that the burst lifetime was chosen to be the same in each case with $t_{\text{SF}} = 2.5 \times 10^7$ yr, and we have chosen common values for $f_{\star,\text{max}}$ across several redshifts. The thick solid lines shown represent models close to the best fit to the observational data. The other three values were chosen to occupy different parts of the $t_{\text{SF}} - f_{\star,\text{max}}$ parameter space in order to illustrate the dependence of the predicted SFRD function on the different parameters.

4.3 Specific star formation rate

Fig. 4 shows the sSFR as a function of mass for the models listed in Fig. 3 at $z = 4–7$, illustrating the broad success of the model in reproducing the observed sSFR for the constrained parameters (particularly $F_{\text{SN}}$). Beyond the narrow range of observed stellar mass values, the model predicts that the sSFR remains quite insensitive to stellar mass (or star formation rate). In this figure, the wiggles result from the simple assumption of an abrupt change in efficiency with halo mass in equation (6). Fig. 5 shows the sSFR as a function of redshift for these models, illustrating the model prediction that
the sSFR does not evolve with redshift. This finding is in agreement with observations at $z \gtrsim 4$.

While our model is fit to the data of González et al. (2014), it is also in reasonable agreement with the measurements of Stark et al. (2013), as shown in Fig. 5. As noted earlier, the true value of the sSFR at $z \gtrsim 4$ is still under debate. In particular, current measurements are significantly limited by the unknown intrinsic dispersion in mass to light ratios as a function of UV luminosity. The main difference between the Stark et al. (2013) and the González et al. (2014) measurements is the treatment of this dispersion. We note that constraining our model using the measurements of Stark et al. (2013) rather than González et al. (2014) would result in lower predicted duty-cycles.

4.4 Star formation efficiency and average duty-cycle

To understand the properties of high-redshift galaxies produced by our model with multiple merger induced burst episodes, we evaluate the total star formation efficiency $f_{\star, \text{tot}} = H^{-1} dN_{\text{merge}}/dt f_{\star}$, as well as the overall duty-cycle $\epsilon_{\text{duty, tot}} = N_{\text{merge}} sSFR / t_H = f_{\phi} dN_{\text{merge}} / dt$. These quantities are plotted in Fig. 6 for parameter choices corresponding to the examples in Fig. 3. In highly star-forming galaxies of SFR $\sim 1-100 M_\odot$ yr$^{-1}$, we find that $\sim 5-10$ per cent of the gas forms stars, whereas we find smaller fractions, down to a per cent in galaxies with lower star formation rates. We find duty-cycles of $\sim 10-20$ per cent. The duty-cycles are larger at higher redshift, and for systems of higher star formation rate, in agreement with the observational estimate of Lee et al. (2009), who infer a duty-cycle at $z \sim 4$ of $15-60$ per cent (at 1$\sigma$).

4.5 The stellar mass function

The next observable that we consider is the stellar mass function.Observationally, the stellar mass function at $z \gtrsim 4$ is generally derived for star-forming, UV-bright, Ly-break-selected galaxies. This can be estimated from our model as

$$\Phi(M_\star) = \Phi(\text{SFR}) \times \left( \frac{d\text{SFR}}{dM_\star} \right) = \Phi(\text{SFR}) \times s\text{SFR}.$$  \hspace{1cm} (11)

The resulting stellar mass function is plotted as the grey curves in Fig. 7 for the models shown in Fig. 3–8. The data points are from González et al. (2014). Since our model produces both the correct sSFR and the correct SFRD function, it is no surprise that the agreement is good. This agreement is in contrast to many hydrodynamical models of galaxy formation, in which the constant accretion leads to high duty-cycles, and hence an SFR to halo mass ratio which is too low. This low mass-to-light ratio in turn results in a stellar mass function that is too steep.

The low duty-cycle of our bursty model predicts the possible existence of many galaxies of large stellar mass that are not actively star forming, and which may therefore be missing from a
Ly-break-selected stellar mass function. Using our model, we therefore also calculate the predicted mass function in the case where the sample was selected based on stellar mass rather than on SFR. This is

\[
\Theta_{\text{all}}(M_\star) = \frac{1}{\epsilon_{\text{duty,tot}}} \Phi(\text{SFR}) \times \left( \frac{d\text{SFR}}{dM_\star} \right)
\]

\[
= \frac{1}{\epsilon_{\text{duty,tot}}} \Phi(\text{SFR}) \times \text{sSFR}
\]

\[
= \frac{1}{\epsilon_{\text{duty,tot}}} \times \Theta(M_\star).
\]

(12)

In Fig. 7 we show the resulting predicted stellar mass functions for the models shown in Figs 3–6 (black curves). Owing to the low duty-cycle, these curves are a factor of \(\sim 5–10\) higher than the Ly-break-selected case, indicating that high-redshift surveys may currently be missing most of the stellar mass produced at early times.

4.6 Clustering of star-forming galaxies

Finally, to check whether the relationship between halo mass and SFR is correctly reproduced, in our model we calculate the
correlation length in samples above a limiting SFR for the four models in this paper. The results are plotted in Fig. 8. The correlation length is calculated for a sample above the limiting star formation rate by averaging over correlation functions weighted by the SFR density function. For comparison, we also include clustering measurements from Lee et al. (2009) and Overzier et al. (2006). To convert from an apparent magnitude limit to an intrinsic star formation rate, we assumed a flat SED with $\beta = -2$ for computation of a $K$-correction, and a conversion from UV luminosity to SFR using Kennicutt (1998). We find that the clustering length increases rapidly towards high SFRs, in agreement with observations. Our model yields a clustering length in the best-fitting model which is consistent with clustering measurements at $z \sim 5$–6, but underestimates observations at $z \sim 4$. This underestimate may indicate that a smooth contribution to the star formation is required at lower redshifts which would require a population of more massive, biased haloes.

5 THE OBSERVED ABILITY OF A PASSIVE HIGH-REDSHIFT POPULATION

As outlined above, due to the low effective duty-cycle, our model predicts a significant population of galaxies that is not star forming at a given point in time and might therefore be missed by current high-redshift surveys, which are only sensitive to the rest-frame UV. An accurate estimate for the extent of this missing population should be derived through detailed SED modelling of the whole galaxy population, which is beyond the scope of this paper. However, we briefly outline the likely impact of a low duty-cycle and bursty star formation history on the observability of the full galaxy population at high redshift.

There are two ways for a non-star-forming galaxy at high redshift to be missed in a Lyman-break-selected sample. First, because the UV luminosity could dim below the detection limit of a survey, and secondly, the UV continuum colour could evolve too far to the red where the effective survey volume of a typical LBG selection drops significantly for sources with UV continuum slopes $\beta \gtrsim -1$. In practice, however, the drop in the UV luminosity is likely the main cause for a galaxy not being selected. Based on simple star formation histories with short star formation times ($t_{SF} < 100$ Myr) and using Bruzual & Charlot (2003) stellar population models, one finds that a galaxy dims by an order of magnitude (i.e. 2.5 mag) in the rest-frame UV light after $\lesssim 100$ Myr of a passive phase. Given the star formation lifetime of $t_{SF} \sim 10^7$ yr and duty-cycle of $\sim 10$ percent, this value of $\sim 100$ Myr is comparable to, but shorter than the average time between bursts in our model at $z > 4$, which may lead to sources being missed from current surveys. Furthermore, after 100 Myr, a galaxy’s rest-frame UV

Figure 6. The values of total star formation efficiency $f_{\text{tot}}$ (i.e. the sum of $f_\star$ over all mergers), and the overall duty-cycle (i.e. the fraction of a Hubble time during which a galaxy is star bursting) as a function of SFR. The four curves shown correspond to the SFRD functions shown in Fig. 3, with model parameters $t_{SF}$ and $f_{\text{max}}$ designated by the symbols in Fig. 2.
continuum slope would have reddened by $\Delta \beta > 1$, which would further diminish its chance to be selected as a robust high-redshift source.

Our model therefore suggests the existence of a significant population of UV-faint galaxies with red UV continuum slopes of $\beta \gtrsim -1$. Such sources could in principle be searched for based on Spitzer/IRAC imaging which samples rest-frame optical wavelengths of $z \sim 4$–8 sources. In the rest-frame optical, a galaxy would only dim by $\sim 1.5$ mag after 100 Myr without star formation. Such galaxies would be detectable in the deepest current IRAC images over the Hubble Ultra Deep Field (reaching down to $\sim 27$ mag$_{AB}$; see Oesch et al. 2013) out to $z \sim 6$ if their stellar masses are a few times $10^9 \, M_\odot$. However, the selection of such galaxies is complicated by possible contamination from very dusty low-redshift sources (see e.g. Wilkland et al. 2008; Caputi et al. 2012), as well as by confusion due to the broad IRAC point spread function. The advent of JWST will, therefore, greatly facilitate the identification of passive high-redshift sources due to its higher resolution and much better sensitivity in several filters sampling the observed $\sim 2 \, \mu m$ regime, and will for the first time enable a full census of high-redshift galaxies.

6 DISCUSSION

In this paper, we have described a bursty model for SN-regulated high-redshift star formation. Our model successfully reproduces the star formation-selected stellar mass function because it predicts both the SFRD and the correct sSFR of star-forming galaxies. However, if we are considering the stellar mass function of the whole galaxy population, then there is stellar mass missing from the observed census. Indeed, based on our model, we argue that surveys currently find only $\sim 10$–20 per cent of the total stellar mass density at $z \gtrsim 4$. We argue that the low duty-cycle of star formation in this model produces possible solutions to two observed puzzles in high-redshift galaxy formation.

The first puzzle relates to the observation that the sSFR does not evolve significantly with redshift or mass at $z \geq 4$, in contrast to theoretical expectation. Our bursty model naturally produces this behaviour. We point out that the value of the sSFR, and its observed evolution at high redshift directly constrain the duty-cycle (averaged over a Hubble time) of high-redshift star formation to be approximately 10 per cent, independent of a specific model for star formation.

The second puzzle lies in the relation between the observed growth of stellar mass and the observed instantaneous star formation rate. The stellar mass density that is directly observed in samples at high redshift is the stellar mass density in the population of star-forming galaxies (González et al. 2011). This quantity may be different from the total stellar mass density in the Universe. There seems to be disagreement between the relationship of star formation rate to stellar mass observed at $z \sim 6$ (Bouwens et al. 2011) and at $z \sim 2$–4 (Wilkins et al. 2008). Specifically, at $z \sim 2$–4,
Figure 8. The correlation length in samples above a limiting SFR. The data points are from Lee et al. (2009) and Overzier et al. (2006). The correlation length was calculated for each sample above the limiting star formation rate by averaging over correlation functions weighted by the SFR density function. Clearly, our model predictions are in good agreement with current observational clustering measurements at \( z \sim 5 - 6 \), providing an independent verification of the viability of the model. At \( z \sim 4 \), our models underestimate the clustering strength somewhat, which may indicate that a smooth contribution to the star formation is required at lower redshifts in a population of more massive, biased haloes. The four curves shown correspond to the SFRD functions shown in Fig. 3, with model parameters \( t_{SF} \) and \( f_{\star,\text{max}} \) designated by the symbols in Fig. 2.

Wilkins et al. (2008) find that there is not enough growth in stellar mass to account for all of the star formation observed. Wilkins et al. (2008) calculate the stellar mass density \( (\rho_{\star,\text{obs}}) \) using fits to the stellar mass function, and take the derivative across a redshift interval \( \Delta z \sim 0.5 \) to find an inferred star formation rate \( \dot{\rho}_{\star,\text{inf}} \) in units of mass per time per Mpc\(^3\). Comparing with the observed star formation rate \( \dot{\rho}_{\star,\text{obs}} \) at \( z \sim 2 - 4 \), Wilkins et al. (2008) found that \( \rho_{\star,\text{obs}} > \dot{\rho}_{\star,\text{inf}} \) with a difference of \( \sim 0.6 \) dex. However, with a duty-cycle smaller than unity, the stellar mass in the star-forming galaxies was built up over a time shorter than the survey depth, meaning that \( \dot{\rho}_{\star,\text{inf}} \) is an overestimate relative to the observed stellar mass. In this case, only a fraction \( \epsilon_{\text{duty}} \) of galaxies with stellar mass \( M_{\star} \) are observed in a particular survey, but all galaxies would have starbursts during a time corresponding to the survey depth (note this does not imply that the instantaneous SFRD is underestimated). Thus, if a Ly-break is needed to identify the galaxies in which stellar mass is observed, much of the stellar mass at a particular time would be missed by the survey since it is contained in non-star-forming galaxies. The difference is a factor of inverse the duty-cycle, explaining the disagreement between observed and inferred quantities found by Wilkins et al. (2008). We note that if the galaxy sample were selected on stellar mass rather on UV luminosity, the estimates of SFRD based on instantaneous SFRD and the derivative of stellar mass density would agree.

In a complementary analysis, Bouwens et al. (2011) have taken the stellar mass function at \( z \sim 6 - 8 \) determined by González et al. (2011) and differentiated to get the star formation rate in a survey at \( z > 6 \). However, in contrast to the results of Wilkins et al. (2008) at \( z \sim 2 - 4 \), in this higher redshift case the resulting stellar mass is found to agree well with the stellar mass inferred directly. At first sight, this is a failure for our model, which predicts that these estimates should differ by a factor of inverse duty-cycle as they do at lower redshift. The solution to this apparent contradiction lies in the fact that at \( z \sim 6 - 8 \) the survey depth of \( \Delta z \sim 0.5 \) corresponds to a time difference across the survey that is similar to the starburst lifetime (but shorter than the time between major mergers), in contrast to the case at \( z \sim 2 - 4 \) where \( \Delta z \sim 0.5 \) corresponds to a time that is longer than the starburst lifetime.

Thus, in difference to observations at \( z \sim 2 - 4 \), at \( z \gtrsim 6 \) we expect that the observed galaxies had star formation episodes for
a time that is similar to the survey depth, meaning that the stellar mass census within star-forming galaxies does include all the stellar mass that was generated during the survey depth time interval. As a result, in the $z \gtrsim 6$ samples, integrating the observed star formation rate between the upper and lower redshifts of the survey gives a stellar mass that approximately equals the mass observed in those $z \sim 6$ galaxies, in agreement with the comparison of Bouwens et al. (2011) (see also Robertson et al. 2013). However, this equivalence is a coincidence, and does not correspond to a large duty-cycle. Rather, galaxies that are not star forming, and therefore not seen in the survey did not form stars during the survey interval. Thus, as in the $z \sim 2-4$ case, there is additional stellar mass in quiescent galaxies that is not accounted for in the observed stellar mass function (e.g. Choi & Nagamine 2012).

We note that this finding is in contrast to recent work at $z \lesssim 3$ that has used the UltraVISTA survey to construct a $K$-selected catalogue covering masses $M_* \gtrsim 10^{11}$ $M_\odot$ (Ilbert et al. 2013; Muzzin et al. 2013). These authors find stellar mass functions selected by stellar mass to agree reasonably well with UV-selected samples, and that star-forming galaxies dominate the stellar mass density at low redshift. This would indicate that $z \lesssim 3$ UV selections are not necessarily missing too many galaxies at the very massive end of the population. At $z \gtrsim 4$ our model is not consistent with this behaviour, which is indicative of a larger duty-cycle or a component of continuous star formation. Full SED modelling of a galaxy population based on merger trees will be required to understand all the observational selection effects on the observability of the predicted non-star-forming galaxy population from our model in detail.

7 CONCLUSION

We have shown that a bursty model of high-redshift star formation reproduces a range of observations of the high-redshift galaxy population. In particular, we point out that the observed sSFR requires a duty-cycle of $\sim 10$ percent, which follows directly from the fact that the observed star formation rate in galaxies integrated over a Hubble time exceeds the observed stellar mass by an order of magnitude. We use this observational constraint to calibrate the efficiency of feedback in a model for the high-redshift star formation rate which includes merger-driven star formation regulated by SN feedback. This model reproduces the SFRD function and the stellar mass function of galaxies at $4 \lesssim z \lesssim 7$.

Most current observations select high-redshift galaxies by their UV luminosity, and hence only detect star-forming galaxies without too much extinction. Since galaxies dim in the UV within 100 Myr after the end of a starburst, the finding of an $\sim 10$ per cent duty-cycle therefore implies that there may be 10 times the number of known galaxies at fixed stellar mass that have not yet been detected. Thus, pending more detailed future work including full modelling of the stellar populations, our model suggests the possibility of a large undetected population of UV-faint galaxies that accounts for most of the stellar mass density at $z = 4-8$.

Unfortunately, at $z > 4$ there are currently no good constraints on non-star-forming galaxies. These UV-faint galaxies would be detectable through their rest-frame optical emission with the Spitzer Space Telescope or JWST. However, it is difficult to define selection criteria that select such sources without significant contamination from lower redshift dusty galaxies (Wiklind et al. 2008; Caputi et al. 2012). The existence of a large population of undetected galaxies which are not forming stars would not affect the global star formation rate history or inferences about the reionization of the intergalactic medium, but would affect the estimated cumulative stellar mass as a function of redshift and the number density of passive galaxies at each redshift.

ACKNOWLEDGEMENTS

JSBW acknowledges the support of an Australian Research Council Laureate Fellowship. AL was supported in part by NSF grant AST-0907890 and NASA grants NNX08AL43G and NNA09DB30A. PO acknowledges support by NASA through Hubble Fellowship grant HF-51278.01 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555.

REFERENCES

Barkana R., Loeb A., 2001, ApJS, 349, 125
Bouwens R. J. et al., 2011, ApJ, 737, 90
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Caputi K. I. et al., 2012, ApJ, 750, L20
Choi J.-H., Nagamine K., 2012, MNRAS, 419, 1280
Clarke C., Oey M. S., 2002, MNRAS, 337, 1299
de Barros S., Schaerer D., Stark D. P., 2012, preprint (arXiv:1207.3663)
Ellis R. S. et al., 2013, ApJ, 763, L7
Finlator K., Oppenheimer B. D., Davé R., 2011, MNRAS, 410, 1703
González V., Labbé I., Bouwens R. J., Illingworth G. F., Franx M., Kriek M., Brammer G. B., 2010, ApJ, 713, 115
González V., Labbé I., Bouwens R. J., Illingworth G. F., Franx M., Magee D., 2014, ApJ, 781, 34
Ilbert O. et al., 2013, A&A, 556, A55
Jaacks J., Choi J.-H., Nagamine K., Thompson R., Varghese S., 2012a, MNRAS, 420, 1606
Jaacks J., Nagamine K., Choi J. H., 2012b, MNRAS, 427, 403
Kennicutt R. C. Jr, 1998, ApJ, 498, 541
Komatsu E. et al., 2011, ApJS, 192, 18
Krumholz M. R., Dekel A., 2012, ApJ, 753, 16
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lee K.-S., Giavalisco M., Conroy C., Wechsler R. H., Ferguson H. C., Somerville R. S., Dickinson M. E., Urry C. M., 2009, ApJ, 695, 368
Loeb A., Furlanetto S., 2012, The First Galaxies in the Universe. Princeton Univ. Press, Princeton, NJ, p. 337
McLure R. J. et al., 2013, MNRAS, 432, 2696
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Muñoz J. A., Loeb A., 2011, ApJ, 729, 99
Muzzin A. et al., 2013, ApJ, 777, 18
Neistein E., Dekel A., 2008, MNRAS, 383, 615
Oesch P. A. et al., 2012, ApJ, 759, 135
Oesch P. A. et al., 2013, ApJ, 772, 136
Oesch P. A. et al., 2013, ApJ, 773, 75
Overzier R. A., Bouwens R. J., Illingworth G. D., Franx M., 2006, ApJ, 648, L5
Pipino A., Calura F., Matteucci F., 2013, MNRAS, 432, 2541
Press W. H., Schechter P., 1974, ApJ, 187, 425
Raičević M., Theuns T., Lacey C., 2011, MNRAS, 410, 775
Robertson B. E. et al., 2013, ApJ, 768, 71
Salvato M., Ferrara A., Dayal P., 2011, MNRAS, 414, 847
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Smit R., Bouwens R. J., Franx M., Illingworth G. D., Labbé I., Oesch P. A., van Dokkum P. G., 2012, ApJ, 756, 14
Smit R. et al., 2013, preprint (arXiv:1307.5847)
Stark D. P., Ellis R. S., Bunker A., Bundy K., Targett T., Benson A., Lacy M., 2009, ApJ, 697, 1493
Stark D. E., Schenker M. A., Ellis R., Robertson B., ure R., Dunlop J., 2013, ApJ, 763, 129

UV-faint galaxies at $z \gtrsim 4$
