Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **422,** 2816–2840 (2012) doi:10.1111/j.1365-2966.2012.20516.x

What shapes the galaxy mass function? Exploring the roles of supernova-driven winds and active galactic nuclei

R. G. Bower,[1⋆] A. J. Benson[2] and Robert A. Crain[3,4]

[1]*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE*
[2]*California Institute of Technology, Mail Code 350-17, Pasadena, CA 91125, USA*
[3]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia*
[4]*Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, the Netherlands*

Accepted 2012 January 6. Received 2012 January 6; in original form 2011 August 8

ABSTRACT

The observed stellar mass function (SMF) is very different to the halo mass function predicted by Λ cold dark matter (ΛCDM), and it is widely accepted that this is due to energy feedback from supernovae and black holes. However, the strength and form of this feedback is not understood. In this paper, we use the phenomenological model GALFORM to explore how galaxy formation depends on the strength and halo mass dependence of feedback. We focus on 'expulsion' models in which the wind mass loading, β, is proportional to $1/v_{\rm disc}^{n}$, with $n =$ 0, 1, 2 and contrast these models with the successful Bower et al. model (B8W7), for which $\beta \propto 1/v_{\rm disc}^{3.2}$. A crucial development is that our code explicitly accounts for the recapture of expelled gas as the system's halo mass (and thus gravitational potential) increases. While models with high wind speed and mass loading result in a poor match to the observed SMF, a model with slower wind speed matches the flat portion of the SMF at $M_\star \sim 10^9$–$10^{11}\,h^{-1}\,{\rm M}_\odot$. When combined with active galactic nucleus feedback, the model provides a good description of the observed SMF above $10^9\,h^{-1}\,{\rm M}_\odot$. In order to explore the impact of different feedback schemes further, we examine how the expulsion models compare with a further range of observational data, contrasting the results with the B8W7 model. In the expulsion models, the brightest galaxies are assembled more recently, and the specific star formation rates of galaxies decrease strongly with decreasing stellar mass. The expulsion models tend to have a cosmic star formation density that is dominated by lower mass galaxies at $z = 1$–3, and dominated by high-mass galaxies at low redshift. These trends are in conflict with observational data, but the comparison highlights some deficiencies of the B8W7 model also. The experiments in this paper not only give us important physical insight into the impact of the feedback process on the formation histories of galaxies, but the strong mass dependence of feedback adopted in B8W7 still appears to provide the most promising description of the observed Universe.

Key words: galaxies: evolution – galaxies: formation.

1 INTRODUCTION

A central issue in the study of galaxy formation is to understand the connection between the mass of galaxies and the mass of their associated dark matter haloes. This problem is not trivial: whilst the mass function of dark matter haloes predicted by the cold dark matter paradigm has a relatively steep slope ($\mathrm{d}\log N/\mathrm{d}\log M \sim -0.9$ over the range of halo masses relevant to the galaxy formation), the observed mass function of galaxies is characterized by a shallow Schechter function. The abundance of galaxies is nearly independent of stellar mass over the interval $M_\star \sim 10^8$–$10^{10.5}\,h^{-1}\,{\rm M}_\odot$, whilst at greater masses the abundance of galaxies declines exponentially. This raises two fundamental questions: (i) why is the stellar mass function (SMF) so flat below $\sim 10^{10.5}\,h^{-1}\,{\rm M}_\odot$, when the abundance of haloes is such a strong function of mass, and (ii) what physical processes are responsible for the exponential suppression of galaxies with mass greater than $\sim 10^{10.5}\,h^{-1}\,{\rm M}_\odot$ (White & Frenk 1991; Benson et al. 2003)?

In this paper we investigate the astrophysical processes that map the halo mass function to the galaxy SMF. There are two common approaches to this problem. Arguably, the most appealing route is to numerically integrate the set of differential equations that describe the rudimentary astrophysical processes (e.g. gravity, hydrodynamics, radiative cooling, star and black hole formation) using a *local* framework (e.g. a set of particles or a grid). This approach aims

⋆E-mail: r.g.bower@durham.ac.uk

to evolve the cosmological density fluctuation power spectrum reflected by the cosmic microwave background using an ab initio description of the problem. This system of equations commonly overpredicts the universal abundance of stars – a problem commonly known as the 'overcooling problem' (e.g. Katz, Weinberg & Hernquist 1996; Balogh et al. 2001; Kereš et al. 2009; Schaye et al. 2010).

In order to get closer to a description of the observed universe, an additional set of processes (collectively known as 'feedback') must be introduced to couple the energy, mass and metals returned by supernovae (SNe) and black holes (active galactic nuclei, AGN) to the surrounding gas. As well as accounting for the stellar mass and gas content of galaxies, feedback enriches the intergalactic medium (IGM) with metals, potentially resulting in an orthogonal set of observational constraints. However, the wide range of scales involved in this problem (from AU scales to tens of Mpc) makes it infeasible to model these processes from first principles, forcing recourse to phenomenological, or 'sub-grid', treatments (e.g. Springel & Hernquist 2003; Springel, Di Matteo & Hernquist 2005; Okamoto, Gao & Theuns 2008b; Schaye & Dalla Vecchia 2008; Booth & Schaye 2009).

Because the properties of model galaxies are remarkably sensitive to the details of sub-grid models, an alternative approach is to establish a set of equations that describes the same astrophysical processes on *macroscopic* scales, typically averaged over the physical scale of a galaxy. This approach is adopted by phenomenological or 'semi-analytic'[1] models, such as the GALFORM code considered here (Cole et al. 2000; Bower et al. 2006, hereafter Bow06; see also Kauffmann, White & Guiderdoni 1993; Hatton et al. 2003; De Lucia et al. 2006; Somerville et al. 2008; Guo et al. 2011). For example, rather than computing the star formation rate at each point in a galaxy, such models typically compute the total rate of star formation over the entire galaxy, and assume that the rate is a simple function of global galaxy parameters (such as the gas mass, disc size and rotation speed). This leads to an alternative set of differential equations that provide a macroscopic description of the physics. A phenomenological model then solves these relatively simple equations within the merging hierarchy of structure formation. Macroscopic treatments are by definition approximate, but their use is ubiquitous in all branches of physics. When applied carefully and within an understood range of validity, the resulting descriptions can lead to significant physical insight.

In this paper, we apply the GALFORM model to seek an understanding of the mass dependence of the equations of feedback from galaxies. We focus initially on the role supernova-driven winds play in establishing the galaxy SMF on scales below $\sim 10^{10.5}\,h^{-1}\,\mathrm{M}_{\odot}$. Although strong suppression of galaxy formation in low-mass haloes is clearly necessary, our understanding of the process remains incomplete, largely because of the difficultly of calculating the effects of feedback from fundamental physical principles. The difficulty primarily arises because the interstellar medium (ISM) in which SNe explode is inherently multiphase (and magnetized), and does not behave like an ideal gas. Thus, accurately calculating the impact of even a solitary supernova is extremely challenging, as the result depends strongly on the initial density of the medium into which energy is injected. As a sequence of SNe explode, a network of low-density channels (or 'chimneys') is established that allow the supernova ejecta to escape into the halo of the galaxy (McKee & Ostriker 1977; Efstathiou 2000; de Avillez & Breitschwerdt 2007). However, without detailed calculations it is impossible to estimate how the mass outflow rate and the specific energy of the outflowing material will depend on the star formation rate and the mass of the host galaxy.

In view of this uncertainty, all currently feasible simulations of galaxy formation parametrize the effect of SNe and include it as a sub-grid calculation (Springel & Hernquist 2003; Dalla Vecchia & Schaye 2008). In hydrodynamical simulations, winds are most commonly modelled either by adding thermal energy to gas particles (or cells) or by giving gas particles a velocity kick. Although adding thermal energy seems a promising route, the energy injected is easily radiated away if the temperature of the particles is too low (Katz et al. 1996). Equally the injection of kinetic energy may drive thermodynamic shocks into the surrounding gas, and this thermal energy may also be radiated. These issues arise because the codes treat the complex multiphase ISM as a single fluid. Several schemes have been developed to circumvent the problem; one approach is to decouple the relevant particles from the numerical scheme for a period of time, either by preventing cooling for a period (Brooks et al. 2007) or by decoupling the particles from hydrodynamical forces (Springel & Hernquist 2003; Okamoto, Nemmen & Bower 2008a; Oppenheimer & Davé 2008, hereafter Op08). Particles therefore retain the energy injected by feedback for a period of time, easing their escape from the dense ISM of the galaxy. An alternative approach is to stochastically heat or kick the particles so that their energy is sufficiently high that their cooling time is long, or the shocks they drive are sufficiently strong that radiative losses are small Dalla Vecchia & Schaye 2008; Creasey et al., in preparation).

Because of the difficulty in interpreting the multiphase nature of the outflow, the mass loading and velocity of winds are not yet strongly constrained by observation (but see Martin 2005; Weiner et al. 2009; Chen et al. 2010; Rubin et al. 2011 for recent progress). Most hydrodynamical calculations have therefore adopted the simplest possible model, in which the mass loading and velocity of winds are independent of the system in which the feedback event is triggered. This also simplifies implementation of the feedback, since there is no requirement to estimate the environment of the ISM on the fly, for example by determining the mass of the dark matter halo or the local gravitational potential. An exception is the momentum scaling model of Oppenheimer & Davé (2006), where the wind parameters are set following a radiatively driven wind model (Murray, Quataert & Thompson 2005), and a number of similar schemes implemented in the Over Whelmingly Large Simulations (OWLS; Schaye et al. 2010).

In contrast, most phenomenological ('semi-analytic') models of galaxy formation assume that the parameters describing feedback adjust to ensure that the specific energy of outflows is matched to the binding energy of the halo. Conservation of total energy thus ensures that the mass loading of outflows is greater in dwarf galaxies than in larger systems (Dekel & Silk 1986). Such models are partially motivated by arguments pertaining to the porosity of the ISM: analytic models (e.g. Efstathiou 2000) consider the formation of channels in the multiphase ISM and suggest that the porosity of the ISM is self-regulating and determined by the gravitational potential of the disc. In general, phenomenological models further assume that material expelled from the disc is recaptured on a time-scale proportional to the dynamical time. This coefficient is

[1] The term 'semi-analytic' commonly used in the literature is misleading. With modern computing power, it is no longer critical that the resulting equations can be solved analytically. The important distinction is that this type of model provides a macroscopic description of the relevant processes. This approach is key to gaining insight into the problem. Such models are usually referred to as 'phenomenological' in other areas of physics.

allowed to be significantly larger than unity in some models in order to approximate the effect of expulsion of gas from the halo.

Current phenomenological models present a coherent picture for the formation of galaxies in a cosmological context, and provide an excellent explanation of many diverse data sets. Although the GALFORM model has been developed to explain the observed properties of galaxies (e.g. Bow06; Font et al. 2008; Lagos et al. 2011), it has also been shown to explain the X-ray scaling relations of groups and clusters (Bower, McCarthy & Benson 2008, hereafter Bow08) and the optical and X-ray emission from AGN (Fanidakis et al. 2011). The models can be used to generate convincing mock catalogues of the observable Universe (Cai et al. 2009) and applied to test the procedures used to derive physical parameters from astronomical observations, and to identify the priorities for the next generation of astronomical instruments. These successes have been driven by the inclusion of two key components of the models: (i) galaxy winds that scale strongly with halo mass and (ii) a 'hot-halo' mode[2] of AGN feedback in which the cooling of gas in quasi-hydrostatic haloes is suppressed. By altering the parametrization of the feedback schemes here, we investigate whether the Bow06 choice is optimal, or whether alternative schemes can similarly reproduce the properties of observed galaxies.

A major advantage of semi-analytic models is that they enable the effects of sub-grid parametrization to be explored quickly and easily (Bower et al. 2010). We exploit this aspect in this study to explore the effect of various descriptions of outflows of material from galaxies. This requires us to generalize the GALFORM model to include the possibility that gas is expelled from the potential of dark matter haloes. Previous efforts to 'calibrate' the semi-analytic method against numerical simulations have included gravity, hydrodynamics and cooling, but not effective feedback (Benson et al. 2001; Helly et al. 2003; De Lucia et al. 2010). Part of the reason for this is the very different treatments of the winds from galaxies. Our extensions of the code allow us to bring the two approaches into closer alignment, and we include a brief comparison with the Galaxies-Intergalactic Medium Interaction Calculation (GIMIC; Crain et al. 2009). This provides a series of relatively high-resolution hydrodynamic simulations, featuring relatively high spatial and mass resolution ($\epsilon = 0.5\,h^{-1}$ kpc and $m_{\rm gas} = 1.4 \times 10^6\,h^{-1}\,{\rm M}_\odot$ for the highest resolution realizations), and that trace a representative cosmological volume (four spherical volumes with comoving radius $18\,h^{-1}$ Mpc and one with comoving radius $25\,h^{-1}$ Mpc). The simulations include radiative cooling (Wiersma, Schaye & Smith 2009a), and sub-grid prescriptions for star formation and the thermodynamics of the ISM (Schaye & Dalla Vecchia 2008), hydrodynamically coupled supernova-driven winds (Dalla Vecchia & Schaye 2008) and metal enrichment resulting from stellar evolution (Wiersma et al. 2009b). Some important successes include the X-ray scaling relations of $L_\star$ galaxies (Crain et al. 2010), and the distribution of satellites and stellar halo properties of the Milky Way (Deason et al. 2011; Font et al. 2011a). Our modified GALFORM scheme implements similar physics, and we show that its behaviour is very similar to GIMIC. This opens a new avenue, allowing us to use GALFORM to better understand how the parametrization of feedback impacts upon the formation and evolution of galaxies, and thus guide both the development of sub-grid treatments in hydrodynamical simulations and the interpretation of observational data.

[2] This mode is often referred to as the 'radio' mode; we prefer the term 'hot halo' as it emphasizes that this type of feedback is assumed to be effective only when the cooling time of the halo is sufficiently long compared to the dynamical time. See Fanidakis et al. (2011) for further discussion.

The structure of this paper is as follows. In Section 2, we discuss the implementation of supernova–wind-driven feedback schemes in semi-analytic models, and introduce a new scheme that allows gas and metals to be expelled from low-mass haloes and later re-accreted when the binding energy of the halo has increased significantly. In contrast to many previous models, we do not assume that material is always re-accreted after a certain time-scale, or that expelled material is always lost from the hierarchy. In Section 3, we present a comparison of different feedback scalings, focusing on the difference between schemes that scale the parameters describing winds with halo mass, and those that adopt fixed parameters. In Section 4, we explore how feedback from SNe can be combined with feedback from black holes in order to generate an exponential break in the mass function. In particular, we compare the effect of 'hot-halo' mode feedback with that of strong quasar-driven winds (similar to those considered by Springel et al. 2005). We consider a number of additional observational constraints in Section 5. While several of the feedback schemes are able to reproduce the high-mass part of the SMF, we show that the specific star formation rate (SSFR) and the downsizing of galaxy formation are important orthogonal constraints. We present a summary of our results in Section 6 Except where otherwise noted, we assume a 7-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*7) cosmology, $\Omega_0 = 0.272$, $\Lambda_0 = 0.728$, $\Omega_{\rm b} = 0.0455$, $h_0 = 0.704$, $n_{\rm s} = 0.967$ and $\sigma_8 = 0.81$ (Komatsu et al. 2011). Throughout, we convert observational quantities to the h scaling of the theoretical model so that stellar masses are quoted in $h^{-1}\,{\rm M}_\odot$ etc.

2 FEEDBACK IN PHENOMENOLOGICAL MODELS

2.1 Feedback as a galactic fountain

We begin by reviewing the conventional approach to feedback in GALFORM, focusing on the implementation used in Bow06. To summarize the key features, gas is expelled from the disc and assumed to circulate in the halo, falling back to the disc on roughly a dynamical time if the cooling time is sufficiently short. If the cooling time is long compared to the dynamical time, the halo is susceptible to a hot-halo mode of feedback if a sufficiently massive central AGN is present. The scheme results in a galactic fountain with material rising from the galaxy disc and later falling back. In the case of haloes with short cooling times, it may be more appropriate to picture the circulating material as cool clouds rather than as material heated to the halo virial temperature.

We parametrize the rate at which gas is expelled from the disc into the halo as

$$\dot{M}_{\rm outflow} = \beta \dot{M}_*, \tag{1}$$

where $\dot{M}_*$ is the star formation rate and the macroscopic mass loading factor, and β is

$$\beta = \left(\frac{v_{\rm disc}}{v_{\rm hot}} \right)^{-\alpha_{\rm hot}}. \tag{2}$$

Here, $v_{\rm disc}$ is the circular speed of the galaxy disc, the parameter $v_{\rm hot}$ sets the overall normalization of the wind loading, and the parameter $\alpha_{\rm hot}$ determines how the mass loading of the wind varies with the disc rotation speed. We will be careful to explicitly distinguish the macroscopic loading factor, β, which represents the loading of the wind as it escapes from the galaxy into the halo, from the sub-grid loading factor η, used in hydrodynamical simulations to represent the amount of ISM material heated or kicked by the supernova

remnant. If winds are hydrodynamically coupled, the macroscopic mass loading is very likely to be significantly larger than η. Furthermore, the physical processes that determine η are themselves highly uncertain and their treatment in numerical models is likely to be resolution dependent.

In the Bow06 implementation, material is modelled as leaving the disc with a specific energy comparable to the binding energy of the halo (i.e. $v_{\rm wind} \sim v_{\rm halo}$). If we assume $v_{\rm disc} \sim v_{\rm halo}$ (where $v_{\rm halo}$ is the circular velocity of the halo at the virial radius), energy conservation then requires that $\alpha_{\rm hot} = 2$ and that, in the standard implementation, no material leaves the halo completely. However, Bow06 found that this scaling did not sufficiently suppress the formation of small galaxies and a stronger scaling, $\alpha_{\rm hot} = 3.2$, was adopted. This gives a good match to the observed K-band luminosity function. The stronger scaling implies that in small haloes SNe couple more efficiently to the cold gas, resulting in a higher mass loading of the wind. Assuming a velocity $\sim v_{\rm halo}/\sqrt{2}$ is sufficient to drive the fountain, the fraction of the total supernova energy needed to power the fountain is

$$f_{\rm fountain} \sim \frac{1}{8}\left(\frac{v_{\rm hot}}{430\,{\rm km\,s^{-1}}}\right)^{3.2} \times \left(\frac{v_{\rm halo}}{200\,{\rm km\,s^{-1}}}\right)^{-1.2}\left(\frac{\epsilon_{\rm SN}}{2.5\times10^{49}\,h\,{\rm erg\,M_\odot^{-1}}}\right)^{-1}, \quad (3)$$

where $\epsilon_{\rm SN}$ is the energy produced by SNe per unit mass of stars formed. Assuming $\epsilon_{\rm SN} = 2.5 \times 10^{49}\,h\,{\rm erg\,M_\odot^{-1}}$ (appropriate for a Chabrier initial mass function, IMF; Dalla Vecchia & Schaye 2008) sufficient energy is, in principle, available to power the fountain in haloes more massive than $v_{\rm halo} \sim 65\,{\rm km\,s^{-1}}$. Interestingly, Font et al. (2011b) find that the properties of Milky Way satellite galaxies are best reproduced if the mass dependence of feedback saturates in such low-mass haloes.

2.2 Allowing for mass loss from the halo

In the revised implementation presented in this paper, we consider the case in which the energy of the gas escaping the disc has systematically greater than the binding energy of the halo. This treatment is necessary if we are to consistently account for winds with outflow speeds that are independent of halo mass. From an observational perspective, this type of wind may be required in order to account for the widespread distribution of metals in the universe. This type of wind was previously considered in Benson et al. (2003), and we briefly review the implementation here.

We introduce the parameter $\lambda_{\rm expel}$ to reflect the excess energy of the wind relative to the binding energy of the halo. Specifically, we set the mean specific energy of the wind to

$$E_{\rm mean} = \frac{1}{2}\lambda_{\rm expel} v_{\rm halo}^2. \quad (4)$$

Note that $\lambda_{\rm expel}$ may be a function of halo mass (see Section 2.3). We parametrize the fraction of material that is able to escape using the cumulative energy distribution, $f_E(x)$, where $x = E_{\rm esc}/E_{\rm mean}$ and $E_{\rm esc}$ is a measure of the wind-specific energy needed to escape the halo. We will choose a monotonic function for f_E so that $f_E(0) = 1$, $f_E(1) = 1/e \sim 0.36$ and $f_E \to 0$ for large x. We will set the energy needed to escape the halo to $E_{\rm esc} = v_{\rm halo}^2$ (i.e. we assume that the escape velocity is $\sqrt{2}v_{\rm halo}$). Clearly, this is an oversimplification, since the true escape velocity depends on the details of the potential, the launch radius of the wind, the terminal radius and the ram pressure that the gas encounters. While we adopt this scaling to give a simple interpretation of the wind velocities, there is likely to be a systematic offset when comparing with hydrodynamical simulations. Since it is the ratio of $E_{\rm mean}$ and $E_{\rm esc}$ that determines the result of a model, we could rescale the wind speeds quoted in this paper according to the new pre-factor. For convenience, we can represent $\lambda_{\rm expel}$ as a wind speed, $v_{\rm wind}$, where

$$v_{\rm wind} = {\lambda_{\rm expel}}^{1/2} v_{\rm halo}, \quad (5)$$

and we will refer to models by their wind speed; however, it should be remembered that this is more accurately defined as the specific energy of the wind, and we do not intend to imply that the wind necessarily has a bulk outflow velocity of $v_{\rm wind}$: what really matters is the fraction of the mass of the outflow that escapes from the halo. Combining equations (4) and (5), a significant fraction of the outflow will escape the halo if

$$v_{\rm wind} > \sqrt{2} v_{\rm halo}. \quad (6)$$

We will consider the halo mass dependence of $v_{\rm wind}$ below, but it will be useful to normalize different models at a particular halo mass. For example, a fiducial halo with $v_{\rm halo} = 200\,{\rm km\,s^{-1}}$, for which $\lambda_{\rm expel} \equiv \lambda_{\rm expel,200} = 1$–$4$ corresponds to wind speeds, $v_{\rm wind} = (200, 400, 600, 800)\,{\rm km\,s^{-1}}$.

Material that is not expelled is added to the halo following the scheme described in Bow06. This includes a delay proportional to the dynamical time before the material is allowed to cool again (see Bow06 for details). Material that escapes the halo may be later recaptured as the halo grows in mass (e.g., Oppenheimer et al. 2010). We implement this by scanning through descendant haloes in the dark matter merger tree and adding $f_E(E^i_{\rm esc}/E_{\rm mean}) - f_E(E^{i-1}_{\rm esc}/E_{\rm mean})$ to the reheated gas mass at each step (where $E^i_{\rm esc}$ refers to the escape energy of the descendant halo at timestep i, and i ranges from the step at which the energy is injected to the final output time, $i = 0$). Mass that is added to the halo becomes able to cool on the dynamical time-scale (which we define as $GM_{\rm halo}/v^3_{\rm halo}$). It may not be able to cool if the cooling time is long and AGN feedback is sufficiently effective. The step in $E_{\rm esc}$ may be small if the halo grows only a little by accretion, or may be large if the halo is accreted to become part of a much larger structure.

Note that this scheme differs significantly from the superwind implementation of Baugh et al. (2005), in which expelled material is not considered for recapture. Since the overall baryon fractions of clusters of galaxies are inferred to be close to the cosmic abundance, recapture must be an important part of the feedback process. Finally, we note that some semi-analytic models adopt a feedback scheme in which expelled material becomes available for cooling or star formation on a time-scale that is much longer than the dynamical time. This is an approximation to the superwind scheme that we have described here, but it is not accurate since it does not take the growth rate of the halo into account.

In order to fix on a scheme, we must choose an appropriate form for the (cumulative) distribution functions f_E. Benson et al. (2003) chose an exponential form, $f_E = {\rm e}^{-x}$. This leads to a broad spread of wind particle energies. On the basis of their observational data, Steidel et al. (2011) suggest that wind outflow is more sharply peaked, and we also find that a more sharply peaked distribution better matches the results of hydrodynamical simulations. In the following models, we will assume $f_E = \exp(-x^6)$ in what follows. The precise choice of power is not important, however, and we obtain similar results for $f_E = \exp(-x^2)$.

2.3 A generalized feedback model

In contrast with GALFORM, and largely because of a lack of observational motivation for any particular scaling, hydrodynamical simulations have mostly adopted the simplest case of assuming that the mass loading and velocity of winds are independent of halo properties. We can easily adapt the revised feedback implementation described above to investigate such a scheme in GALFORM, by explicitly including a halo mass dependence in equation (4). We will adopt $v_{\rm halo} = 200\,{\rm km\,s^{-1}}$ as a fiducial halo mass at which to compare the wind mass loading and wind speed for models with different $\alpha_{\rm exp}$ so that we express the mass dependence of the wind speed as $\lambda_{\rm expel} = \lambda_{\rm expel,200}\left(v_{\rm halo}/200\,{\rm km\,s^{-1}}\right)^{\alpha_{\rm exp}}$, where $\alpha_{\rm exp}$ is a dimensionless parameter that differentiates different feedback models. Combining this with equations (4) and (5) gives

$$v_{\rm wind}^2 = v_{\rm wind,200}^2 \left(\frac{v_{\rm halo}}{200\,{\rm km\,s^{-1}}}\right)^{\alpha_{\rm exp}+2}. \quad (7)$$

For the mass loading we have

$$\beta = \beta_{200}\left(\frac{v_{\rm wind,200}}{200\,{\rm km\,s^{-1}}}\right)^{-\alpha_{\rm hot}}. \quad (8)$$

In the case $\alpha_{\rm exp} = 0$ we recover the wind-specific energy scaling with the specific binding energy of the halo. In the case $\alpha_{\rm exp} = -2$, the wind speed is independent of halo mass. Unless otherwise stated, we will set

$$\alpha_{\rm hot} = \alpha_{\rm exp} + 2, \quad (9)$$

so that a fixed fraction of the supernova energy is used to drive winds in haloes of all masses (assuming $v_{\rm disc} \sim v_{\rm halo}$ in equation 2).

Given a wind speed and mass loading normalization, $v_{\rm wind,200}$ and β_{200}, the code parameters are

$$\lambda_{\rm expel,200} = \left(\frac{v_{\rm wind,200}}{200\,{\rm km\,s^{-1}}}\right)^2 \quad (10)$$

and

$$v_{\rm hot} = {\beta_{200}}^{1/\alpha_{\rm hot}}\ 200\,{\rm km\,s^{-1}}. \quad (11)$$

This allows simple comparison to older models. Note that the original GALFORM parameter, $v_{\rm hot}$, expresses the mass loading, and is not a measure of the specific energy of the wind. To avoid this confusion in this paper, we will use the macroscopic mass loading parameter β_{200} to label models in what follows. The maximum available supernova energy sets a limit of $\beta_{200}(v_{\rm wind,200}/200\,{\rm km\,s^{-1}})^2 < 32$.

If $\alpha_{\rm exp} = -2 (\alpha_{\rm hot} = 0)$, the macroscopic mass loading, β, is independent of the halo potential. This would be the case if the wind were completely decoupled until it had escaped from the halo. This would make it impossible to frame the feedback in terms of the standard GALFORM parameters, and for this paper, we will assume that there is always a small coupling between the wind loading and the halo mass. We adopt $\alpha_{\rm exp} = -1.9$ ($\alpha_{\rm hot} = 0.1$) as our minimum value.

While $\alpha_{\rm exp} = 0$ and -2 are natural choices, there is no a priori reason to adopt a particular value of $\alpha_{\rm exp}$, and we will consider $\alpha_{\rm exp} = -1$ ($\alpha_{\rm hot} = 1$) as an intermediate value. For these parameters, the speed of the wind scales with $\sqrt{v_{\rm halo}}$, and its mass loading scales as $1/v_{\rm halo}$. In this case, the material expelled from smaller galaxies is more likely to escape the halo, but this is a weaker function of mass than in the superwind case discussed above. The scaling of the wind mass loading is similar to the momentum-driven model used by Op08, but note that we will assume that the ratio of the total energy of the wind (not its total momentum) to the mass of stars formed is independent of halo mass.

2.4 Parameter values

The parameters of the best-fitting models are given in Table 1. To simplify comparison with previous work, we have translated the feedback parametrization in Bow06 ($v_{\rm hot} = 485\,{\rm km\,s^{-1}}$, $\alpha_{\rm hot} = 3.2$, $\lambda_{\rm expel,200} = 0$ and $\alpha_{\rm cool} = 0.58$) into to the more generalized parameters considered in this paper. Where parameter values are not explicitly given, we adopt those in Bow06 with the exceptions given below. First, we now adopt a background cosmology that is consistent with the *WMAP* 7-year results (Komatsu et al. 2011). Secondly, we use a stellar yield of $p_{\rm yield} = 0.04$ in order to improve the match of galaxy colours as discussed in Font et al. (2008), and a default halo gas distribution with a core radius of $0.025r_{\rm vir}$ as discussed in Bow08. With these revisions, we make small shifts in the standard feedback parameters in order to restore a good fit to the local K-band luminosity function.

The parameters of the baseline model are given in the second row of Table 1. We use $v_{\rm hot} = 430\,{\rm km\,s^{-1}}$ and $\alpha_{\rm cool} = 0.52$, where $\alpha_{\rm cool}$ determines the ratio of free-fall and cooling times at which haloes are taken to be hydrostatic (as opposed to being classified as 'rapidly cooling') so that only when $t_{\rm cool}(r_{\rm cool}) > \alpha_{\rm cool}^{-1} t_{\rm ff}(r_{\rm cool})$ is the AGN feedback effective. These differences have little impact on the properties of sub-L_* galaxies. Since we are initially concerned with the faint end of the luminosity function, we begin by disabling the AGN feedback scheme by setting $\alpha_{\rm cool} = 0$. This allows us to make a simple comparison to hydrodynamical calculations that do not include AGN feedback.

We consider three supernova-driven feedback models following the feedback scheme discussed in the previous section. In what

Table 1. Models considered in this paper and a comparison to the values adopted in Bow06. Updating the cosmological parameters requires that we make small adjustments to restore a good match to the observed mass function. We will refer to this model as B8W7. The following models correspond to the optimal parameters for the different feedback schemes considered in the text. The values in parentheses indicate the values adopted for the AGN 'hot halo' of feedback in Section 4.1.

Model		$\alpha_{\rm hot}$	β_{200}	$v_{\rm wind,200}$ ($\rm km\,s^{-1}$)	$\alpha_{\rm exp}$	$\alpha_{\rm cool}$
Bow06	Bow06	3.2	17	–	–	0.58
Bow08+W7 cosmology	B8W7	3.2	12	–	0.0	0.52
Superwind	pGIMIC	0.1	8	275	−1.9	–
	SW	0.1	8	180	−1.9	(0.35)
Momentum scaling	MS	1.0	8	200	−1.0	(0.45)
Energy scaling	ES	2.0	–	–	0.0	–

follows, we refer to these as 'superwind' (SW), 'momentum scaling' (MS) and 'energy scaling' (ES) models. This nomenclature reflects how the mass loading scales with halo mass, corresponding to $\alpha_{\rm hot} = 0.1, 1$ and 2. For each case, we typically consider six values of the mass loading at $v_{\rm halo} = 200\,{\rm km\,s^{-1}}$, $\beta_{200} = 1, 2, 4, 8, 16$ and 32, and five values of the wind launch velocity at this halo mass, $v_{\rm wind,200} = 100, 200, 400, 600$ and $800\,{\rm km\,s^{-1}}$. For the SW and MS models, optimal values of the feedback parameters have been chosen to provide a reasonable match to the SMF above $M_* \sim 10^9\,h^{-1}\,{\rm M}_\odot$ (a particularly good fit was not possible for the ES scaling) and the corresponding parameters are given in Table 1. We also consider a model intended to replicate the supernova-driven winds implemented in the GIMIC hydrodynamical simulations (Crain et al. 2009). These adopt a fixed sub-grid wind mass loading of $\eta = 4$ and a launch wind velocity of $600\,{\rm km\,s^{-1}}$. This requires ∼80 per cent of the available supernova energy being used to drive the wind. We find that we can best reproduce the resulting SMF produced by the simulation if we adopt a somewhat lower macroscopic wind speed, $v_{\rm wind,200} = 275\,{\rm km\,s^{-1}}$, and higher mass loading, $\beta_{200} = 8$ (although a broad range of parameters give similar results). Since the winds in the simulation are always hydrodynamically coupled, the increased mass loading of the wind is not surprising. We will refer to this 'pseudo-GIMIC' model as pGIMIC below.

While we initially consider the models in the absence of AGN feedback, we will later adjust the AGN feedback parameter $\alpha_{\rm cool}$ in order to set the mass scale of the break in the luminosity function. For this we use the AGN feedback scheme of Bow08. This allows material to be expelled from the central regions of haloes through AGN heating, and results in the baryon fraction in galaxy groups being much lower than the cosmic average. This provides a much improved match to the observed X-ray properties of these systems. The values of $\alpha_{\rm cool}$ that give a good match to the observed luminosity function are given in brackets in the table. The ES model could not be adjusted to give a sufficiently flat faint-end slope to the SMF, and we do not consider the role of AGN in this model.

3 GALACTIC WIND MODELS

3.1 Conventional GALFORM WINDS

We use the galaxy SMF as the starting point for comparing the models and the data. We use the determination of the SMF by Bell et al. (2003) and Li & White (2009), correcting the IMF to the Kennicut parametrization. We convert the observational data to the h^{-1} dependence on the Hubble parameter of the theoretical models. The data are compared with the Bow06 model, and the baseline B8W7 model in Fig. 1 . Both models assume that supernova-driven winds do not escape the parent halo; instead the material ejected from the disc circulates in the halo and returns on the dynamical time-scale if the cooling time is short. The two models are almost indistinguishable, since changes in the background cosmology and the AGN feedback implementation have been intentionally compensated for by small changes in the feedback parameters. As expected, both provide a good description of the observational data.

When we consider alternative wind descriptions, we will not initially consider AGN feedback. To establish a baseline for the comparison in the absence of AGN feedback, we also show the B8W7 model with AGN feedback disabled (by setting $\alpha_{\rm cool} = 0$). This is shown as a red line in the figure. The rollover of the galaxy luminosity function is almost non-existent in this model, while the SMF is largely unaffected below $10^{11}\,{\rm M}_\odot$. This is encouraging, since it shows that the two processes involved in matching the shape of the

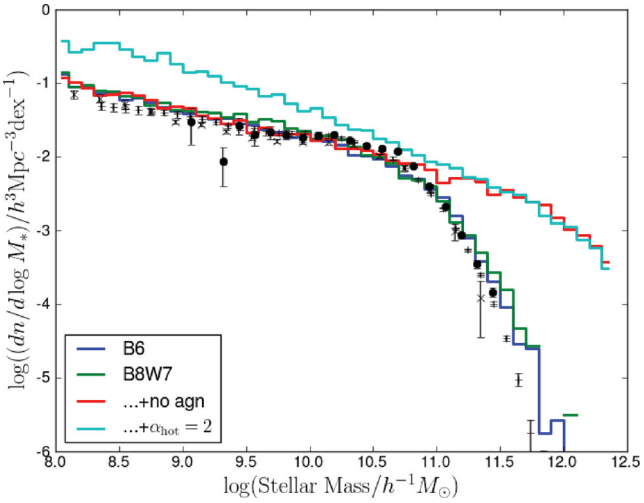

Figure 1. Comparison of the SMF of the Bow06 model (blue line) with the baseline B8W7 model used in this paper (green line). This is based on the *WMAP7* cosmology and includes AGN 'hot-halo' feedback following Bower et al. (2008). The two models are almost indistinguishable. To illustrate the importance of AGN feedback, we show the effect of turning off the AGN feedback (red line). We also show the effect of adopting $\alpha_{\rm hot} = -2$ (cyan line). For comparison, observational data are shown as black points with error bars. The data are taken from Bell et al. (2003) (circles) and Li & White (2009) (crosses).

galaxy luminosity function (i.e. eliminating the overabundance of galaxies of fainter than $\sim 10^{10}\,h^{-1}\,{\rm M}_\odot$ and reducing the abundance of galaxies above the break in the mass function) can be separated. We therefore focus our initial discussion on the modes of supernova-driven feedback, and consider models that do not include an AGN feedback component.

The Bow06 and B8W7 models achieve a good fit to the abundance of low-mass galaxies because of the very strong mass dependence of the feedback ($\alpha_{\rm hot} = 3.2$). For comparison, we show a model with feedback parameters that might be considered to be better motivated by theory. By selecting $\alpha_{\rm hot} = 2$, the wind speed is tuned to the binding energy of the halo. As can be seen, this results in a somewhat steeper faint-end slope and a relatively poor match to the observed SMF. In the following section, we consider in more detail how the choice of feedback scheme affects the mass function, and how the wind parameters can be optimized to improve the fit.

Although the SMF provides a good way to compare the results of different feedback schemes with observations, it is far from simple to interpret the changes in terms of the effect of the different feedback schemes. For example, increasing the effectiveness of the feedback scheme shifts galaxies to lower stellar mass, and so only affects the normalization of the mass function indirectly. The suppression of the normalization arises both because of the lower abundance of the haloes of greater mass and because of the range of halo masses that contribute galaxies of a particular stellar mass. A better way to understand the effect of the schemes is therefore to plot the stellar mass of the central galaxy as a function of halo mass. Since the scatter is not strongly constrained observationally (Moster et al. 2010), we use the relationship found in the Bow06 and B8W7 models as a best guide to the relationship expected in the real Universe. We can then understand how various feedback schemes affect this relationship, and compare with constraints yielded by abundance matching observations with theoretical subhalo mass functions. Of course, the parameters chosen in Bow06 and B8W7 are not unique, and other parameter combinations can give similar quality fits to

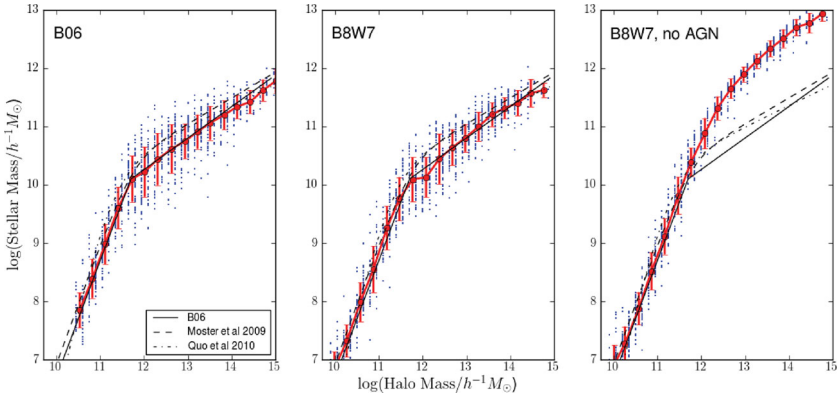

Figure 2. The $M_*-M_{\rm halo}$ relation for central galaxies. Panels from left to right show Bow06, the baseline model B8W7 and the B8W7 model without AGN feedback. Blue points show a random selection of haloes from the model, while the solid red line shows the median and error bars show the 1σ scatter. The solid black line shows an empirical fit to the Bow06 model, while black dashed and dot–dashed lines show halo abundance matching models from Moster et al. (2010) and Guo et al. (2010), respectively. The models in the first two panels include AGN feedback and successfully match the observed SMF. These are well characterized by a broken power law. Although the first two panels have a slightly different background cosmology, the relation between halo mass and stellar mass is similar. Disabling AGN feedback, as shown in the last panel, results in a much weaker rollover of the relationship.

the mass function and other data sets, as we have shown in Bower et al. (2010), but the models do provide a well-documented starting point for our comparison of different feedback schemes.

The first panel of Fig. 2 illustrates the dependence of central galaxy stellar mass on halo mass for the Bow06 model. The red line shows the mean relation. The dashed black line shows a broken power-law approximation to the model, described by

$$\begin{aligned}\log M_* &= 10.1 + 0.54(\log(M_{\rm halo}) - 11.7) \\ &\qquad \text{if } \log(M_{\rm halo}) \geq 11.7 \\ &= 10.1 + 2(\log(M_{\rm halo}) - 11.7) \\ &\qquad \text{if } \log(M_{\rm halo}) < 11.7. \end{aligned} \quad (12)$$

At the break in the curve, 14 per cent of the baryons in the halo have been converted into stars. Note that there is considerable scatter about these relations in the model. We repeat this relation in all plots so that they can be compared easily. The error bars indicate the $\pm1\sigma$ range of the model galaxies. Close to the break, the scatter in the model exceeds an order of magnitude. We supplement the empirical approximation with relations from Moster et al. (2010) and Guo et al. (2010). The relations shown are derived from matching the abundance of subhaloes in N-body simulations to observational data, assuming that the scatter in the relation is negligible. The models are based on $(\Omega_{\rm m}, \Omega_\Lambda, h, \sigma_8, n_{\rm s}) = (0.26, 0.74, 0.72, 0.77, 0.95)$ and $(0.25, 0.75, 0.73, 0.9, 1)$ cosmologies, respectively, but note that the differences in the predicted abundance of 10^{10}–$10^{12}\,h^{-1}\,{\rm M}_\odot$ haloes are small. Thus, the relations are similar for stellar masses below $10^{10}\,h^{-1}\,{\rm M}_\odot$ but are offset from the Bow06 relation at high masses due to the large scatter about the mean relation. Because of the steep break in the mass function, scatter boosts the abundance of massive galaxies relative to a relation without scatter (see Moster et al. 2010 for further discussion), and thus the scatter and the normalization of the high-mass $M_* - M_{\rm halo}$ are tightly correlated.

The second panel of Fig. 2 shows the baseline B8W7 model. The scatter in M_* at a given halo mass is reduced compared to Bow06, although it is still larger ($\sigma = 0.3$ dex) than the scatter explored by Moster et al. (2010) (up to 0.15 dex), particularly around the break in the relation. The final panel shows the effect of *disabling* AGN feedback in the B8W7 model. The power-law relation now extends to higher mass before slowly rolling over as the result of the increasing cooling times of massive haloes. The scatter in the relation around $M_{\rm halo} \sim 10^{12}\,h^{-1}\,{\rm M}_\odot$ is now much reduced. This arises because the efficacy of AGN feedback in this model has a strong dependence on the accretion history of haloes (see Bow08 for further discussion).

In summary, the B8W7 model provides a match to the observed SMF due to the very strong halo mass dependence of the wind mass loading and the suppression of cooling in haloes with relatively long cooling times. This paper investigates whether models with more general feedback schemes can achieve a similar success.

3.2 Superwind models with fixed wind speed and mass loading

3.2.1 Effect of wind parameters

We now consider SW models, in which the mass loading and velocity of winds are (almost) independent of the halo mass. This mimics the schemes that are commonly adopted in hydrodynamical simulations. We begin by contrasting the results with the baseline B8W7 model.

Fig. 3 shows the effect of varying the mass loading for a fiducial wind speed of $v_{\rm wind,200} = 400\,{\rm km\,s^{-1}}$. For comparison, the blue curve shows the effect of low mass loading and low wind speed such that supernova-driven feedback is ineffective. Increasing the mass loading reduces the normalization of the mass function below M_*, but the power-law dependence at low galaxy luminosities becomes steep, increasing the discrepancy with observations. As the mass loading increases above $\beta_{200} = 4$, feedback takes a 'bite' out of the mass function. This can be understood as a transition in the effectiveness of feedback. In high-mass haloes ($v_{\rm halo} > v_{\rm wind,200}$), material falls back on to the central galaxy on the dynamical time-scale, while in low-mass haloes it is expelled and ceases to be available to fuel star formation. The time-scale for the return of expelled material therefore makes a transition when the two speeds are equal (see Op08). (The effect is most clearly seen by plotting the stellar mass of galaxies against their halo mass, as we discuss below.) Although increasing the mass loading tends to suppress the abundance of M_* galaxies, the faint-end slope of the mass function is always much steeper than the observations. It is not possible to improve the fit to the mass function by adjusting this parameter.

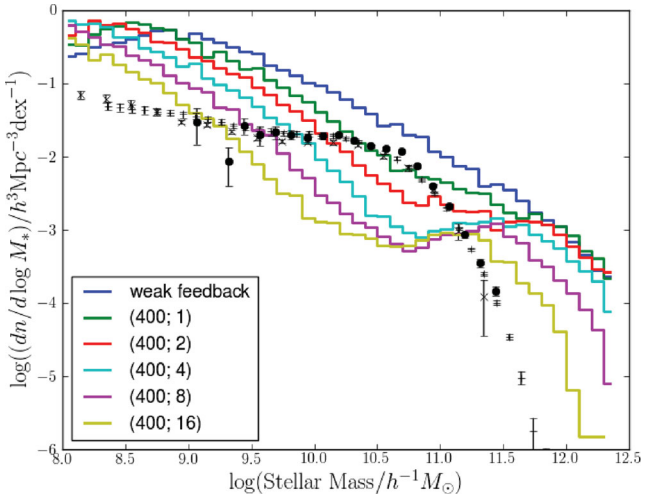

Figure 3. Effect of mass loading in the SW model. These models (solid lines) have a wind velocity of $v_{\rm wind,200} = 400\,{\rm km\,s^{-1}}$ and mass loading of $\beta_{200} = 1, 2, 4, 8$ and 16. We also include a model with low mass loading ($\beta_{200} = 1$) and low speed ($v_{\rm wind,200} = 100\,{\rm km\,s^{-1}}$) to show the effect of suppressing the feedback channel (blue line). With low mass loading, the faint end of the function is extremely steep. Increasing the mass loading suppresses the normalization of the low-mass part of the mass function, but tends to introduce a noticeable dip in the SMF if the loading is high. A high mass loading also suppresses the abundance of the highest mass galaxies, but the shape of the mass function does not match the observational data. Since AGN feedback has not been included, the cut-off seen here is the result of the long cooling times in larger haloes.

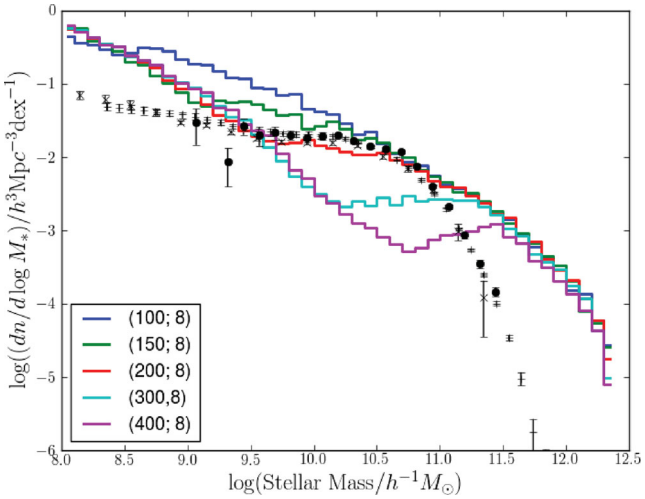

Figure 4. The effect of changing the wind speed in the SW model. These models have a fixed mass loading of $\beta_{200} = 8$ and a varying wind speed of $v_{\rm wind} = 100$–$400\,{\rm km\,s^{-1}}$. The figure shows how changing the speed of the wind adjusts the location and depth of the dip in the SMF, although the faintest and most massive galaxies are almost unaffected. An intermediate wind speed has the effect of producing a flat portion of the luminosity function.

Although the abundance of bright galaxies greatly exceeds the observations, a rollover is evident when the high mass loading occurs. As was the case where AGN were disabled in the B8W7 model, the rollover is driven by the long cooling times of high-mass haloes. In this situation, the bottleneck is the cooling time of the material in the halo, and the star formation rate is proportional to the inverse of the wind mass loading.

The effect of changing the wind speed is explored in Fig. 4, where we show the mass function obtained by varying $v_{\rm wind,200} = 100$–$400\,{\rm km\,s^{-1}}$ at a fixed mass loading of $\beta_{200} = 8$. At a wind speed greater than $200\,{\rm km\,s^{-1}}$, the mass function begins to resemble the observational data with a flat slope around the knee of the luminosity function. This is encouraging: we infer that a suitable choice of wind parameters enables a match to the properties of $10^{9.5}$–$10^{11}\,h^{-1}\,{\rm M}_\odot$ galaxies to be obtained. However, while the model matches the observational data down to $M_* \sim 10^{9.5}$, the number of galaxies rises rapidly at lower masses, and an additional feedback mechanism would need to be introduced to explain the low abundance of dwarf galaxies. Since the halo masses of these galaxies are sufficiently high that they are unlikely to be affected by photoheating (Crain et al. 2007; Okamoto et al. 2008b), the only option would be to explore winds that scale with halo mass. On the other hand, a high abundance of faint galaxies (at $z > 6$) would provide an abundant source of ionizing photons to drive the reionization of the universe (Benson et al. 2006; Jaacks et al. 2012). If the mass-dependent scheme suppresses the formation of pre-reionization of small galaxies too dramatically, there will not be sufficient photons to reionize the universe. The constraint is quite weak, however. Even with the strong halo mass dependence of winds in the Bow06 model, we find that it is sufficient to assume that feedback saturates when $v_{\rm halo} < 65\,{\rm km\,s^{-1}}$ in order to provide the necessary ionizing flux (Font et al. 2011).

In summary, with a suitable choice of parameters, the SW scheme offers an attractive explanation for the flat portion of the SMF in the range of $10^{9.5}$–$10^{11}\,h^{-1}\,{\rm M}_\odot$. The model, however, predicts that lower mass galaxies will be more abundant than observed. In contrast, the strong B8W7 halo mass dependence of feedback in the B8W7 model results in a flat SMF down to below $10^{8}\,h^{-1}\,{\rm M}_\odot$.

3.2.2 *Comparison with hydrodynamical simulations*

The SW feedback scheme is similar to the approach adopted in many hydrodynamic simulations, and it is interesting to briefly compare the results in order to gain confidence that the phenomenological description we use appropriately represents the physics of a full hydrodynamical treatment. We first compare with the GIMIC (Crain et al. 2009). The highest resolution realizations of these simulations have gas particles with mass $1.45 \times 10^6\,h^{-1}\,{\rm M}_\odot$, and a softening length of $0.5\,h^{-1}$ kpc. This is sufficient to resolve the onset of the Jeans instability in galactic discs while at the same time allowing reconstruction of a representative cosmological volume. Feedback is implemented by imparting kinetic energy to stochastically chosen neighbouring particles of newly formed stars. The kicked particles remain hydrodynamically coupled at all times. The simulations adopt sub-grid wind parameters $\eta = 4$ and $v_{\rm wind,200} = 600\,{\rm km\,s^{-1}}$. This choice was motivated by observations suggesting that the wind speed was independent of halo mass (see Martin 2005 for discussion) and by requiring that the total stellar mass density matched the observed universe. Alternative choices are explored in the OWLS simulations (Schaye et al. 2010). These parameters determine the input properties of particles. Since the simulation is fully hydrodynamic (so that wind particles remain hydrodynamically coupled to the surrounding gas particles), we should not expect them to directly translate into the macroscopic wind parameters used in GALFORM.

In Fig. 5 we compare the SMF from GIMIC with the SW model (a full comparison of individual galaxy merger trees will be presented in a future paper). It is important to note that the GIMIC simulations did not include AGN feedback, and were constrained to matching the observed stellar mass density, rather than the portion of the

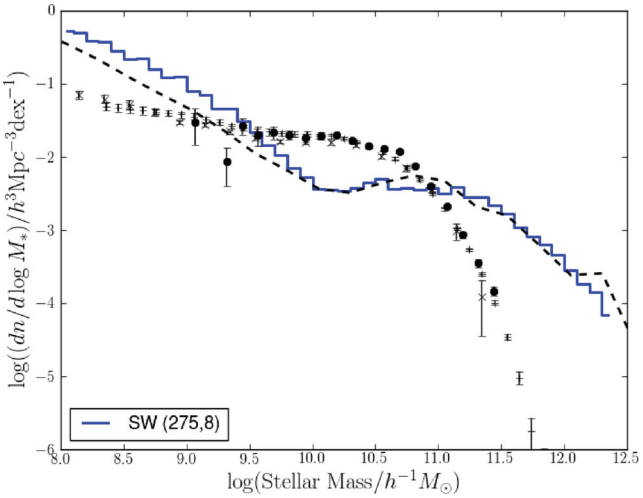

Figure 5. Comparison with the SMF of the GIMIC hydrodynamic simulation (dashed line) and the GALFORM model (pGIMIC) with $(v_{\rm wind,200}, \beta_{200}) = (275, 8)$ (solid blue histogram). The GALFORM parameters have been chosen to reproduce the position and amplitude of the dip seen in the hydrodynamic simulation. The GIMIC wind parameters are (600, 4), but hydrodynamical coupling is expected to result in additional macroscopic mass loading and lower effective wind speeds. Observational data are reproduced from Fig. 1.

mass function below stellar masses of $10^{11}\,h^{-1}\,{\rm M}_\odot$. In order to run the GALFORM model, we revert to the cosmological parameters used in Bow06 (both models are based on the Millennium simulations; Springel et al. 2005). With a suitable choice of wind normalization ($v_{\rm wind,200} = 275\,{\rm km\,s^{-1}}$ and $\beta_{200} = 8$), the GALFORM code reproduces the hydrodynamic mass function well. Although the GALFORM wind has lower speed and somewhat higher mass loading, it should be remembered that these are the effective macroscopic wind parameters. As shown by Dalla Vecchia & Schaye (2008), the ram pressure induced by the hydrodynamic coupling of winds tends to slow the outflow and increases its mass loading as it leaves the disc of the galaxy.

Op08 also present similar models with which we can compare. These simulations have significantly lower mass resolution than GIMIC, with gas particle masses of up to $1.5 \times 10^7\,h^{-1}\,{\rm M}_\odot$ and a softening length of $1.9\,h^{-1}\,{\rm kpc}$. Consequently, these hydrodynamic simulations do not attempt to resolve physics within galaxies. Star formation is implemented following the sub-grid scheme of Springel & Hernquist (2003): winds are implemented kinetically, but the affected particles are decoupled from hydrodynamic forces until the surrounding density is low.

Op08 consider three distinct models: a high wind speed model, a model with much lower wind speed and another in which the wind speed (and mass loading) scales (inversely) with the local halo velocity dispersion. Their strong wind model ($v_{\rm wind,200}$, $\beta_{200} = 680$, 2) produces results that are similar to those of the GIMIC simulations. However, the 'slow wind' model ($v_{\rm wind,200}$, $\beta_{200} = 340$, 2) provides a good match to the observed mass function over the range plotted in their paper. There are three regimes of the SMF for the slow wind: a steep slope at low mass, flat around $10^{10}\,h^{-1}\,{\rm M}_\odot$ and then steep again at higher masses. As we have seen, GALFORM can reproduce this behaviour if the wind velocity is low ($200\,{\rm km\,s^{-1}}$) and the mass loading somewhat higher (between 4 and 8). These have comparable total wind energy to the Op08 models. In low-mass haloes, even the high wind loading considered does not sufficiently suppress star formation, compared to the observations. At intermediate masses, the slope is roughly flat as the wind becomes less effective and eventually stalls. Then at very high mass, cooling becomes inefficient and the slope steepens. Obviously, as with GIMIC, the match to the observed SMF at such high masses is poor because the simulations do not include AGN feedback (but see Gabor et al. 2011). We will consider models in which the wind parameters scale with the properties of the halo in Section 3.3.

In summary, this brief comparison shows that the expulsion scheme implemented in our phenomenological model describes the effects seen in hydrodynamic simulations well. By using these models to better explore the parameter space of galaxy feedback, we can create a closer connection between phenomenological models and fully hydrodynamic simulations. A comparison with the observational data highlights two important issues: first, the steep slope of the mass function below $M_* = 10^{10}\,h^{-1}\,{\rm M}_\odot$ and, secondly, the overabundance of galaxies more massive than $M_* = 10^{11.5}\,h^{-1}\,{\rm M}_\odot$. In the following sections, we will explore how these discrepancies can be resolved by introducing more feedback schemes that scale with halo mass, and including feedback from AGN.

3.2.3 *Stellar mass as a function of halo mass*

In order to better understand how the feedback scheme can shape the SMF, it is useful to examine the relation between the halo mass and the stellar mass of central galaxies. We present the relation for the Bow06 model in Fig. 2 and show that the relationship can be characterized by a broken power law. Fig. 6 illustrates the effect of changing the feedback scheme to the SW model.

The first panel shows the effect of including only minimal feedback ($v_{\rm wind,200}$, $\beta_{200}) = (100, 1)$. The discrepancies compared to the empirical relationship (black lines) are evident. At all halo masses, the associated stellar mass is too high, and the relation shows little change of slope. This model corresponds to the weak feedback (blue) line in Fig. 3. At a given stellar mass, the galaxies are formed in lower mass haloes than indicated by the observed relation. These haloes are much more abundant, and thus the SMF is normalized too high. The slope of the relation is also in clear disagreement and this results in the overly steep faint-end slope of the predicted mass function.

The remaining panels illustrate the effect of increasing the wind speed at a fixed mass loading. In the second panel, $(v_{\rm wind,200}, \beta_{200}) = (100, 8)$. This model is shown as the blue line in Fig. 4. The outflow has a low speed, so little mass escapes the halo, but the high mass loading results in effective suppression of galaxy stellar mass. As a result, the model matches the normalization of the knee of the SMF well, and this is reflected by the $M_*-M_{\rm halo}$ relation coming close to the kink of the observed relationship. However, several discrepancies from the observed relationship remain clear. In particular, the relation shows little change of slope. While the relationship at high stellar mass can be improved with AGN feedback, the relation at lower stellar masses is too shallow. As a result, galaxies of a given stellar mass are overabundant compared to the observed relation, as is evident in Fig. 4.

The third and fourth panels show the effect of increasing the mass loading in the model, $(v_{\rm wind,200}, \beta_{200}) = (200, 8)$ and $(400, 8)$. These are shown as the green and purple lines in Fig. 4. The increasing wind speed creates a kink in the M_*–$M_{\rm halo}$ relation, with the stellar mass formed in haloes around $10^{11}\,h^{-1}\,{\rm M}_\odot$ being very strongly suppressed. In the kinked region, the steepening of the $M_*-M_{\rm halo}$ relation means that a particular halo mass contributes to a wide spread of stellar masses, resulting in a suppression of the mass function normalization. By suitable adjustment of the parameters,

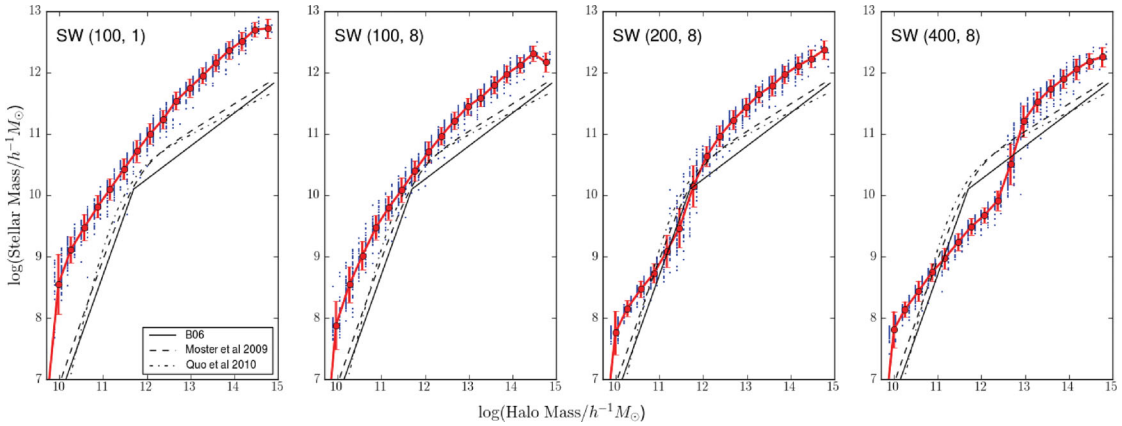

Figure 6. Comparison of the $M_*-M_{\rm halo}$ relation for SW models with various parameters. The first panel shows a model with weak feedback ($v_{\rm wind,200}$, β_{200}) = (100, 1). The remaining panels show ($v_{\rm wind,200}$, β_{200}) = (100, 8), (200, 8) and (400, 8) (from left to right), to illustrate the effect of increasing the wind speed at a fixed mass loading. The SMFs corresponding to these models can be seen in Fig. 4, and the effects seen in the SMF are more readily interpreted in terms of the amplitude and slope of the $M_*-M_{\rm halo}$ relation. Black lines show the expected relationship derived from the observational data (see Fig. 2).

the suppression can be tuned to create a flat portion of the SMF. We will exploit this in Section 4.1. Below the kink, however, the slope of the M_*–$M_{\rm halo}$ relation is much shallower than that seen in the B8W7 model (Fig. 1) and the slope of the mass function is therefore inevitably steeper than the observational data.

The kink is created by the wind stalling at a particular halo mass so that material no longer escapes the halo and falls back on the dynamical time-scale. Increasing the speed of the wind shifts the region of the kink, but leaves the relations at high and low halo masses unchanged. This is reflected in the SMF, with the wind speed effecting a transition region between unchanging abundances of high- and low-mass galaxies. The transition between the regimes appears to become steeper for higher wind energies, resulting in a noticeable dip in the mass function.

3.3 Momentum scaling models

While Op08 find that the slow wind model fits the observed data over a range from $10^{9.5}$ to $10^{11}\,h^{-1}\,{\rm M}_\odot$, their preferred model is one in which the feedback parameters vary with halo mass. Their preferred scheme is intended to mimic the effect of a momentum-driven wind (Murray et al. 2005), and the ratio of total momentum to mass of stars formed is held fixed. We adopt a similar scheme, scaling the wind mass loading inversely with the circular speed of the disc (i.e. $\beta \propto 1/v_{\rm disc}$). In contrast to Op08, however, we scale the wind speed so that the total wind energy (per stellar mass formed) is independent of halo mass. (The total of the wind momentum in Op08 exceeds the momentum available from photon by almost an order of magnitude.) The effect of using this 'momentum scaling' ($\alpha_{\rm hot}$ = 1) in the GALFORM model is shown in Fig. 7. The first panel shows a model with relatively modest wind speed normalization ($v_{\rm wind,200}$ = 200 km s^{-1}), considering a range of mass loading normalizations β_{200} = 2−32. Note that in low-mass haloes, the mass loading will be higher and the wind speed lower. The panel shows that normalization of the SMF steadily decreases as the wind speed increases. The overall SMF is flatter than that seen in the SW case (where the wind properties are independent of halo mass, Fig. 4), and we see that some models compare favourably with the data above a stellar mass of $10^9\,h^{-1}\,{\rm M}_\odot$.

The second panel shows the effect of increasing the wind speed at a fixed mass loading, β_{200} = 8. As the wind speed increases, the abundance of M_* galaxies is suppressed. As we have seen in the SW models, a high mass loading can create a dip in the mass function (where $v_{\rm wind} \sim v_{\rm halo}$). The dip tends to be more smeared

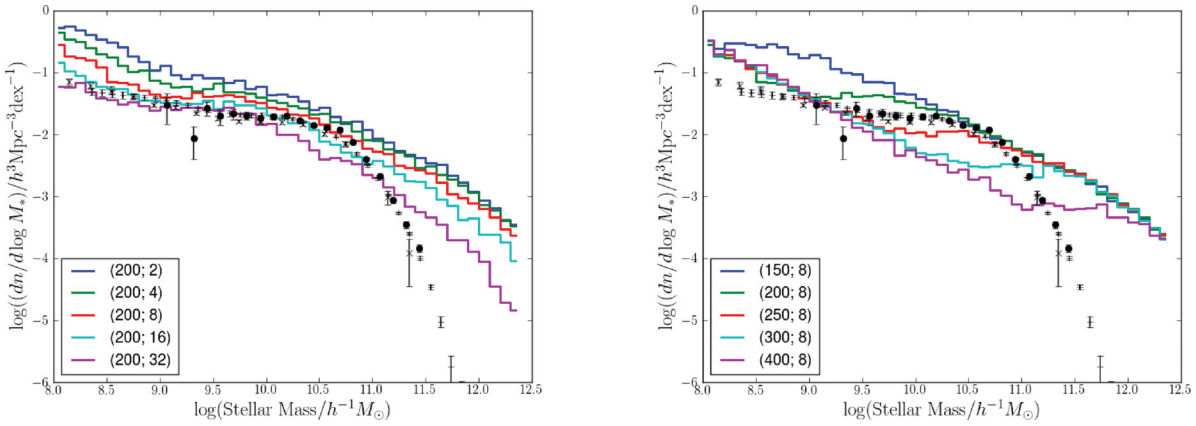

Figure 7. The effect of using the 'momentum' scaling (MS) model ($\alpha_{\rm hot}$ = 1). The first panel shows a wind speed normalization of $v_{\rm wind}$ = 200 km s^{-1}, with varying mass loading, β_{200} = 2−32. For a low wind speed, a high normalization of the mass loading effectively suppresses the formation of low-mass galaxies, leading to an encouraging match to a large portion of the mass function. The second panel shows a fixed wind loading of 8 and varying normalization of the wind speed, $v_{\rm wind}$ = 100–400 km s^{-1}. Increasing the wind speed has the undesirable effect of creating a dip in the mass function.

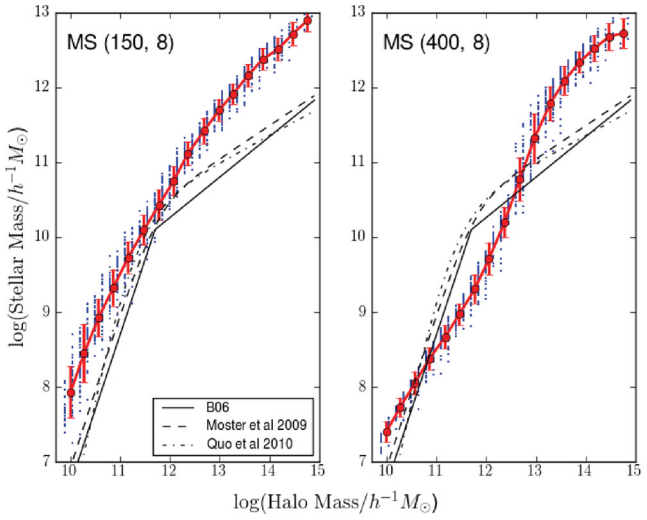

Figure 8. Comparison of the $M_*-M_{\rm halo}$ relation for two momentum scaling models ($\alpha_{\rm hot} = 1$). The first panel shows $(v_{\rm wind,200}, \beta_{200}) = (150, 8)$ and the second panel shows (400, 8). Comparing this plot with Fig. 6 shows that the principal effect of the halo mass dependence of the wind is to smooth out the kink in the $M_*-M_{\rm halo}$ relation. See Fig. 2 for explanation of lines and symbols.

out for the MS wind, however. The origin of this feature is shown clearly in Fig. 8, where the M_*–$M_{\rm halo}$ relation is shown for an MS model with parameters $(v_{\rm wind,200}, \beta_{200}) = (150, 8)$ and (400, 8). The effect of the MS feedback scheme is to introduce a kink into this relation, with the location of the kink depending on the wind speed normalization. The effect is similar to that seen previously in the SW models, but the kink is more diffuse, resulting in a smoother transition between the high- and low-mass regimes. The second panel of this figure should be compared with the last panel of Fig. 6 as the feedback is the same in 200 km s^{-1} haloes in both cases. The kink in the M_*–$M_{\rm halo}$ relation is clearly much smoother in the momentum-driven model, resulting in a less prominent dip in the mass function. Below $M_* = 10^9\,h^{-1}\,{\rm M}_\odot$, the slope of the M_*–$M_{\rm halo}$ relation is again relatively shallow, leading to an overabundance of low-mass galaxies.

The difference between the MS and SW models is also seen in the behaviour of the high-mass end of the SMF. At high masses, little of the feedback material is able to escape the halo, but the mass loading of the two models differs. As a result, for models with equal normalization, the stellar mass associated with high-mass haloes is higher in the MS model than in the SW model (compare the last panels of Figs 8 and 6, for example). This is reflected in a greater abundance of high-mass galaxies in the MS model versus SW (compare Figs 7 and 4). We will show, later, that this has important consequences for the abundance of the massive galaxies at high redshift.

In summary, the dependencies on mass loading and wind speed in the MS model show similar trends to the SW model. However, the rise in the abundance of the faintest galaxies is shallower, and the dip in the SMF tends to be smoothed out. With a suitable choice of wind parameters, this model is able to match the observed SMF over a greater range of galaxy mass.

3.4 Energy scaling models

Finally, we consider models in which the wind speed is a fixed multiple of the halo circular velocity, such that a fixed fraction of the wind escapes regardless of the halo mass. In the left-hand panel of Fig. 9, we fix $\beta_{200} = 4$ and allow the velocity of the wind to increase, $v_{\rm wind,200} = 100$–400. The results for normalized wind speeds between 100 and 200 (blue and green lines) are identical because little material has sufficient specific energy to escape the halo. Further increases in wind speed change the normalization dramatically as material leaves the halo and takes longer to become available for cooling again. However, the shape of the SMF changes little, and it is not possible to recreate the dip in the mass function that was seen in previous models. In the right-hand panel, we show the effect of varying the mass loading of the wind. For this model, the effect is similar to that of varying the wind speed. Since there is no characteristic mass at which the wind stalls, the loss of material from the halo is similar regardless of whether a relatively small fraction of baryons are expelled with high specific energy (and thus remain outside of the halo for an extended period), or a large fraction of material is expelled with lower specific energy. Finally, we note that the abundance of high-mass galaxies continues the increasing

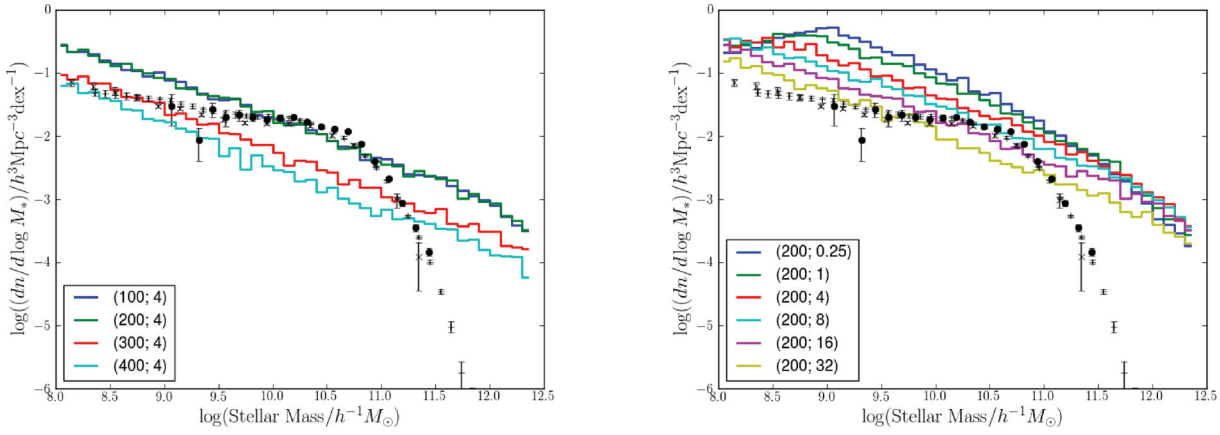

Figure 9. Left-hand panel: the effect of varying the wind speed in the energy scaling (ES) model ($\alpha_{\rm hot} = 2$). These models have increasing wind speed normalization, $v_{\rm wind,200} = 100, 200, 300, 400$; all the models shown have $\beta_{200} = 4$. Increasing the wind speed normalization results in further suppression of the SMF. However, the effect is strongest around the knee of the mass function, and the resulting function does not match the observational data. Right-hand panel: the effect of varying the wind loading at a fixed wind speed. For this model, the effects of varying mass loading and wind speed is quite similar and the mass function cannot be shaped to match the observational data.

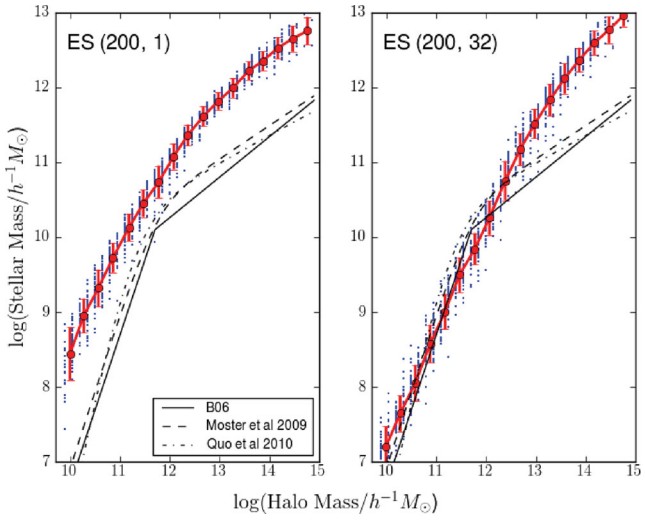

Figure 10. Comparison of M_* versus $M_{\rm halo}$ for a model with energy scaling of the wind (ES, $\alpha_{\rm hot} = 2$). The two panels show models with ($v_{\rm wind,200}$, β_{200}) = (200, 1) and (200, 32). These span the range of models shown in the right-hand panel of Fig. 9. See Fig. 2 for the explanation of lines and symbols.

trend seen in Figs 7 and 4 due to the decrease in mass loading in high-mass haloes (for equal normalization).

Fig. 10 shows the effect of this type of feedback on the M_*–$M_{\rm halo}$ relation. As can be seen, the kink that enabled us to produce a flat component of the mass function in the first two feedback schemes is absent. The scaling of the wind speed with halo mass means that winds do not stall at a particular halo mass. Instead, the feedback steepens the overall slope of the M_*–$M_{\rm halo}$ relation. However, although it is closer to the observed relation, the difference in slope leads to a significant mismatch with the SMF, as shown in Fig. 9. This clearly illustrates the way in which the normalization of the mass function is strongly dependent on the slope of the M_*–$M_{\rm halo}$ relation as well as its normalization.

Overall, the effect of this type of feedback is less encouraging, and we do not consider this model further. The effects of mass loading and wind speed are very similar, and the primary effect of both is to suppress the normalization of the SMF rather than to alter its shape. While including AGN feedback induces a break at the bright end of the mass function by suppressing cooling in hydrostatic haloes, the faint-end slope is not affected. In contrast, the B8W7 model achieves a much improved match to the mass function by adopting a stronger halo mass dependence of the wind mass loading.

4 THE ROLE OF AGN FEEDBACK

4.1 The hot-halo (or 'radio') mode

None of the models discussed so far is able to match the abrupt turn over of the SMF. In this section, we consider the 'hot-halo' mode of AGN feedback, associated with the heating of the group and cluster diffuse material. Bow06 argued that the AGN feedback loop can only be established if the cooling time is longer than the dynamical time (or sound crossing time) of the halo, and showed that allowing AGN to suppress cooling in hydrostatic haloes resulted in a good description of many galaxy formation properties. Bow08 extended this by allowing the AGN to expel material from hydrostatic haloes (rather than simply replacing the energy radiated). They showed that this model was able to match the observed X-ray scaling relations of groups and clusters as well as many of the observed properties of galaxies. Typically, the heat input is assumed to be associated with low-excitation radio sources (Croton et al. 2006; Best et al. 2007); however, the exact heating mechanism is not important to the scheme. The crucial distinction is that only hydrostatic haloes are affected and that the cold gas disc of the galaxy is affected only indirectly because of the reduction in the supply of cooling gas (van de Voort et al. 2011 discuss these effects in the context of hydrodynamic simulations). It is important to note that the effectiveness of AGN feedback is not a threshold imposed at a fixed mass, but is the result of dynamically tracking the relative cooling and dynamical times of the halo as it evolves.

The results of including this type of AGN feedback in the SW and MS models are shown in Fig 11. In each model we choose the wind speed and mass loading to achieve a good match to the abundance of galaxies with stellar mass between $10^{9.5}$ and 10^{11} h^{-1} M$_\odot$

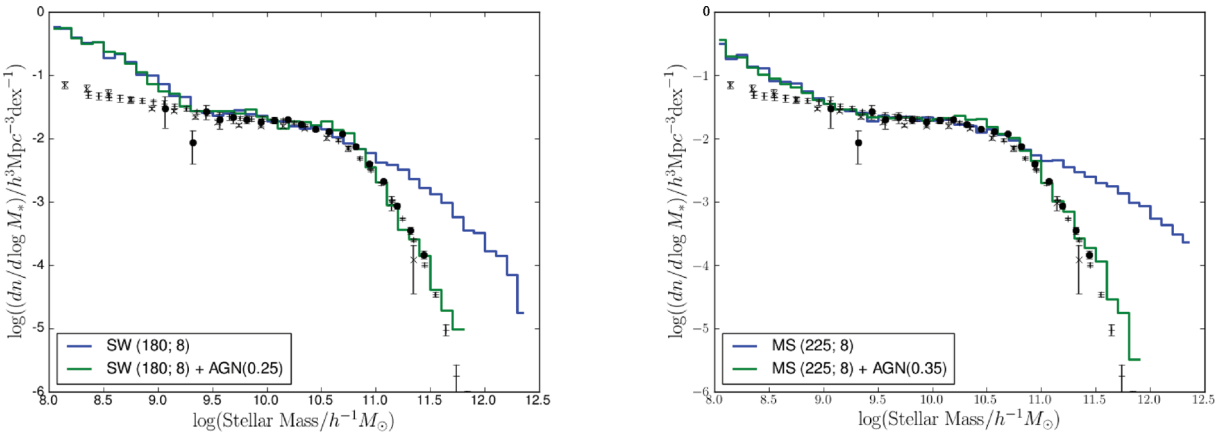

Figure 11. Left-hand panel: the effect of 'hot-halo' AGN feedback on the SW model. Both models have ($v_{\rm wind,200}$, β_{200}) = (180, 8), and we compare $\alpha_{\rm cool}$ = 0.0 and 0.25. Increasing $\alpha_{\rm cool}$ adjusts the ratio of cooling and free-fall times at which AGN are assumed to become effective. The increase shifts the break in the mass function to lower stellar mass so that the model matches the break in the mass function well. Right-hand panel: the effect of AGN feedback on the MS ($\beta \propto 1/v_{\rm disc}$) model. In this case, we compare models with ($v_{\rm wind,200}$, β_{200}) = (225, 8) for $\alpha_{\rm cool}$ = 0.0, 0.35. In both panels, black points show observational data. We have supplemented the Bell et al. (2003) (circles) and Li & White (2009) (pluses) with preliminary data from the GAMA survey (crosses; Baldry et al., 2011) to provide an independent assessment of the observational uncertainties at $M_* < 10^9$ h^{-1} M$_\odot$. The upturn in the model mass functions at low M_* is inconsistent with both recent data sets.

and have then adjusted the $\alpha_{\rm cool}$ parameter to achieve a good match to the observed SMF. Increasing $\alpha_{\rm cool}$ adjusts the ratio of cooling and free-fall times at which the AGN is assumed to become effective. Larger values make AGN feedback more effective and shift the break in the mass function to lower stellar mass. With a suitable value for $\alpha_{\rm cool}$ the models match the observed mass function well. For the SW model, we find that $(v_{\rm wind,200},\, \beta_{200}) = (180,\, 8)$ gives a reasonable match to the observed SMF if combined with $\alpha_{\rm cool} = 0.25$. For the MS model, we find that $(v_{\rm wind,200},\, \beta_{200}) = (225,\, 8)$ combined with $\alpha_{\rm cool} = 0.35$ gives a good description of the observational data for stellar masses above $\sim 10^9\, h^{-1}\, {\rm M}_\odot$. Both models overpredict the abundance of the lowest mass galaxies, although this problem is reduced in the MS model. To emphasize this point, we have superposed preliminary data from the GAMA survey (Baldry et al. 2011). This provides an independent assessment of the mass function for low-mass galaxies. Although the detailed shape of the mass function differs slightly from Li & White (2009), the differences are much smaller than the discrepancy between the models and the observational data. In contrast, the B8W7 model keeps a shallow mass function slope down to the faintest galaxies plotted (see Fig. 1). The difference in behaviour arises from the steeper slope of the stellar mass – halo mass relation in B8W7.

In Fig. 12, we compare the evolution of the mass function for the models discussed above. The figure also shows the B8W7 model, and recent observational data from Drory et al. (2005), Bundy, Ellis & Conselice (2005), Marchesini et al. (2009) and Mortlock et al. (2011) (plus, circle, cross and triangle, respectively). All of the models include AGN feedback following the Bow08 scheme. The key issue that we wish to test with this plot is whether the models generate sufficient massive galaxies at high redshifts, and we focus on the brightest galaxies at each epoch. Compared with B8W7 and the observational data, the new models show a rapid decrease in the abundance of the most massive galaxies at higher redshift. The discrepancy is worst for the SW model.

The differences in the behaviour of the models can be traced to the differences in wind mass loading in high-mass haloes (see Section 3.2.1). The effect arises since the 'hot-halo' mode of feedback is *not* a simple cut-off in cooling at high halo mass, but explicitly compares the halo cooling time and dynamical time, taking into account the halo formation history. In practice the effective halo mass threshold increases slowly with redshift. Moreover, since none of the models can eject material from massive haloes, the efficiency of star formation is inversely proportional to the mass loading. The net result is that massive galaxies appear at higher redshifts in the model with the strongest halo mass dependence of the mass loading. Since $\beta \propto v_{\rm halo}^{-3.2}$ in the B8W7 model, this model provides the best match to the observational data, followed by the MS model ($\beta \propto v_{\rm halo}^{-1}$).

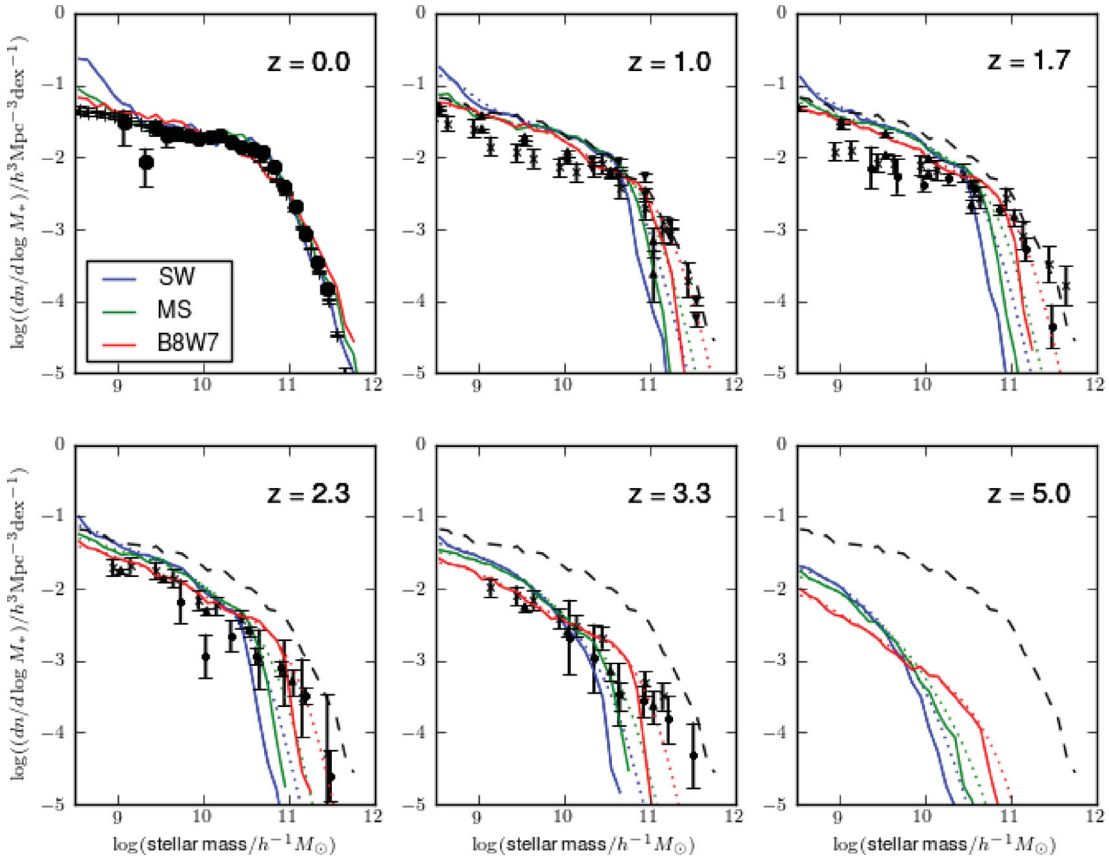

Figure 12. The evolution of the SMF compared between three models that match the $z = 0$ SMF well: B8W7 (red), SW (blue) and MS (green). All three models include AGN feedback, and the $z = 0$ SMF of B8W7 is repeated as a black dashed line in each panel. The effect of observational errors is indicated by the dotted lines which show the effect of convolving the models with a stellar mass error of 0.2 dex. For comparison, we show as black symbols the observational data for the high-redshift mass function from a variety of sources (see text for details). Out to $z = 2$, all three models result in similar evolution of the SMF. The break in the mass function evolves more quickly in the SW and MS models than in the B8W7 model. However, the deficit of massive galaxies in the MS model could be accounted for if the random stellar mass errors are greater than 0.2 dex. All of the models show an excess of $M_* \sim 10^{10}\, h^{-1}\, {\rm M}_\odot$ galaxies at $z = 1{-}2$ compared to most observational data sets.

We should, however, note that there are considerable random uncertainties in determining the stellar masses of high-redshift galaxies. Applying this convolution will tend to smear the model predictions, resulting in a tail of higher mass galaxies (see discussion in Marchesini et al. 2009). Thus, while this data set picks out the B8W7 model, a careful analysis of the observational errors is required before reaching a definitive conclusion. In order to illustrate the effect of random uncertainties in the mass determination, dotted lines show the result of convolving the model with a random error of 0.2 dex. This has a pronounced effect on the abundance of the most massive galaxies, as a small population of galaxies that are mistakenly assigned low stellar mass can easily overwhelm the true population. The comparison still favours the B8W7 model, but assigning larger mass errors would make it difficult to exclude the MS model with high confidence.

Focusing on lower mass galaxies, we see that all of the models appear to overpredict the observed normalization of the mass function at $z = 1$–2 (Marchesini et al. 2009). Although there is considerable scatter between data sets, and the survey volumes are relatively small, this does appear to be a persistent problem, and only the data from Drory et al. (2005) are consistent with the evolution seen in the models. This discrepancy is also evident if the K-band luminosity functions are compared directly (e.g. Cirasuolo et al. 2010). Pozzetti et al. (2010) suggest that the problem lies with the mass dependence of the SSFRs of the model galaxies, and we will examine this in Section 5.1. Our preferred interpretation is that the data require a stronger redshift dependence of feedback. We have already shown that the GIMIC model provides a good description of the mass function at $z > 1$, so a promising route would be to vary the wind speed parameter between 180 at $z = 0$ and 275 at $z = 1$. Alternatively, the required variation in the fraction of the wind escaping would naturally arise if we were to choose a criterion for wind escape based on the halo mass rather than circular velocity, at least at low redshift. It is unclear why this choice should be physically motivated, however. Perhaps a better explanation could be the greater gas content of high-redshift discs, and thus the tendency for star formation to occur in more massive star-forming complexes (Jones et al. 2010; Genzel et al. 2011).

In summary, introducing a hot-halo mode of feedback creates a break in the stellar mass function in all three models. As a result, all three provide a good match to the observed SMF above a stellar mass of $\sim 10^{9.5}\,h^{-1}\,\mathrm{M}_{\odot}$. At lower stellar masses, the mass function of the SW model rises steeply, and is inconsistent with the observational data. This trend is less pronounced in the MS model, while the B8W7 model has a flat SMF to much lower masses. The models predict different evolution of the SMF, with B8W7 showing the highest abundance of massive galaxies at $z = 1$ and above. All three models predict an abundance of $10^{10}\,h^{-1}\,\mathrm{M}_{\odot}$ galaxies at $z = 1$ that appears to be at odds with the data, and suggest that the effective wind speed should be higher at $z > 1$ than at the present day.

4.2 The 'starburst' (or 'QSO') mode

Another channel of AGN feedback, often referred to as the 'QSO' or 'starburst' mode, also has the potential to be important because a large fraction of the black hole mass is accreted in this way. The GALFORM model assumes that black hole growth is triggered when gas is transported to the centre of a galaxy by disc instabilities or galaxy mergers. Most of the cold gas fuels a burst of star formation, but a small fraction is accreted by the black hole (e.g. Springel et al. 2005). Since most mergers are gas rich, this results in a strong correlation between black hole mass and bulge mass very like that observed. We will use the term 'QSO mode' and 'starburst mode' interchangeably, perhaps the term 'starburst' should be preferred since it makes it clear that this channel only occurs during such events. The key distinction is that the 'starburst' mode acts on the cold gas of the host galaxy, rather than acting through the heating of hot gas in the haloes of galaxy groups and clusters. In the starburst mode, feedback may lead to explosive winds that blow cold gas out of the host galaxy. If cold gas is removed from the system at sufficiently high specific energy it will suffer a long delay before it is able to cool once again. This type of feedback has been explored in idealized numerical simulations which have shown that the energetics of the black hole can plausibly remove the whole ISM of the merging galaxies (Springel et al. 2005; Hopkins et al. 2006) (although higher resolution simulations suggest that the geometry of the central outflow may play an important role; Hopkins & Elvis 2010). However, this channel expels only the cold material from the system, and does not prevent further accretion. As the halo grows, it accretes new satellite galaxies, together with their gas, so that (in practice) star formation quickly re-establishes itself.

Fig. 13 illustrates the effect of the QSO mode of expulsion. All of the models we consider reproduce the observed correlation between the mass of the black hole and the mass of the galaxy bulge. We start from the SW model [with parameters $(v_{\mathrm{wind},200},\,\beta_{200}) = (180,\,8)$, blue line]. In this model, black holes grow strongly as a result of galaxy mergers and disc instabilities (see Bow06), but this results in no effective feedback. The default model assumes that all the energy generated in black hole events is radiated without doing significant mechanical work. In order to explore what would happen if this radiation coupled effectively to the surrounding gas (or if the quasar accretion disc produced a high speed wind), we

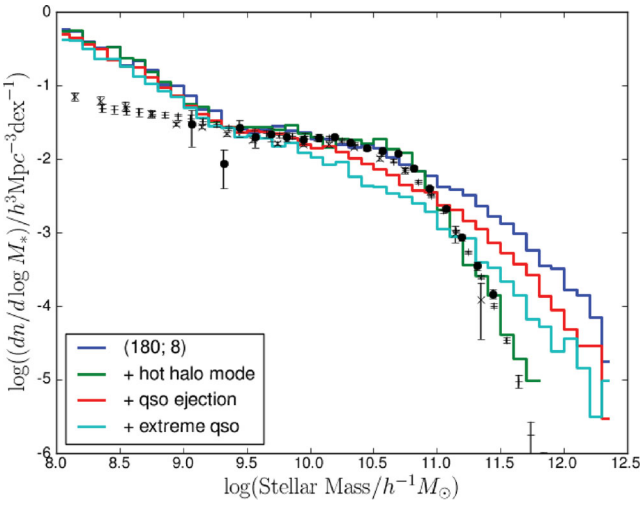

Figure 13. Left-hand panel: starting from the SW model [$(v_{\mathrm{wind},200},\beta_{200}) = (180,\,8)$, blue line], we contrast the effect of the AGN 'hot-halo' mode feedback scheme (implemented by setting $\alpha_{\mathrm{cool}} = 0.25$, green line) with the effect of inducing a wind with very high mass loading ($\beta_{200} = 16$) and wind speed ($v_{\mathrm{wind},200} = 1130\,\mathrm{km\,s^{-1}}$) during bursts of star formation (red line). We assume that the energy to drive such powerful outflows comes from the QSO phase of AGN growth. In order to show the effect of more frequent outbursts, we also show an 'extreme QSO' model in which discs are much more unstable than in B8W7 (cyan line), so that black hole fuelling occurs more frequently. The main effect of the QSO mode is to suppress the formation of galaxies near the break of the mass function and to drive the mass function towards a power-law shape. The sharp break in the mass function is only created if AGN feedback is only effective in hydrostatic haloes.

implement a 'QSO mode' of feedback by using much stronger feedback during starbursts (compared to quiescent star formation events). We illustrate the effect by using $(v_{\rm wind,200},\ \beta_{200}) = (1130,\ 16)$ during bursts (the results are similar for other parameter choices) so that the energy of the wind during the burst is 80 times larger than that during quiescent star formation. The figure shows that even winds of this strength have a modest effect on the mass function. Furthermore, their effect is to suppress the abundance of M_* galaxies rather than to create an exponential break in the SMF. We can produce a stronger effect on the mass function by increasing the frequency of starbursts. A simple way to achieve this is to tighten the disc stability criterion so that discs more frequently become unstable. The effect is illustrated by the 'extreme QSO' model in the plot (cyan line). The model has been shifted further from the observed SMF, giving the mass function an almost power-law form.

We can compare these models with Gabor et al. (2011) who modify the hydrodynamical models of Op08 to investigate the effect of quenching star formation after galaxy mergers. They contrast this form of feedback with a model in which star formation is suppressed in hot haloes. The implementation of their schemes is similar to that adopted here (although their quasar-mode feedback scheme is triggered only by mergers, while it is triggered by both mergers and disc instabilities in our model), and the results are qualitatively similar. In particular, the merger model tends to have a relatively weak effect on the overall mass function, and fails to imprint a characteristic scale on the SMF.

We have experimented with using other models as a starting point. If we start from a model with much weaker quiescent feedback, a strong 'starburst' mode feedback fails to reproduce the shape of the observed SMF, again tending to drive the mass function towards a power law. The 'starburst' mode does not have the required effect because it does not scale strongly with halo mass (as is the case for the 'hot-halo' mode). In summary, while the starburst/QSO channel might supplement the feedback from supernova during star bursts, it does not provide a scheme for creating a break in the SMF.

5 FURTHER CONSIDERATIONS

5.1 The star-forming sequence

We have compared the different schemes on the basis of the SMF and the central stellar mass. In this section, we compare the models with the star formation rates of the central galaxies. We focus on B8W7 and the best-fitting models with constant wind and $\beta \propto 1/v_{\rm disc}$ feedback scalings, SW+AGN $[(v_{\rm wind,200},\ \beta_{200},\ \alpha_{\rm cool}) = (180,\ 8,\ 0.25)]$ and MS+AGN $[(v_{\rm wind,200},\ \beta_{200},\ \alpha_{\rm cool}) = (225,\ 8,\ 0.35)]$. The SMFs of these models are shown in Fig. 11.

Fig. 14 shows the logarithm of the $z = 0$ SSFR as a function of stellar mass. We focus on the properties of star-forming galaxies (which we define as having SSFR $>0.01\,{\rm Gyr}^{-1}$). In the model we restrict attention to central galaxies to avoid uncertainties in the treatment of satellite galaxies. In all the models, there is a clear sequence of star-forming galaxies that can be cleanly compared to the observed star-forming sequence. There is also a large population of galaxies which are not seen on this diagram because their star formation rates are extremely low. These are satellite galaxies or central galaxies in haloes with effective AGN feedback.

We compare the theoretical models to observational data from Brinchmann et al. (2004, updated to DR7). The observational data are shown as black contour lines in the figure. The observed relation is almost flat (with low-mass galaxies having slightly higher SSFR) up to stellar masses of $10^{11}\,h^{-1}\,{\rm M}_\odot$ (above which star formation

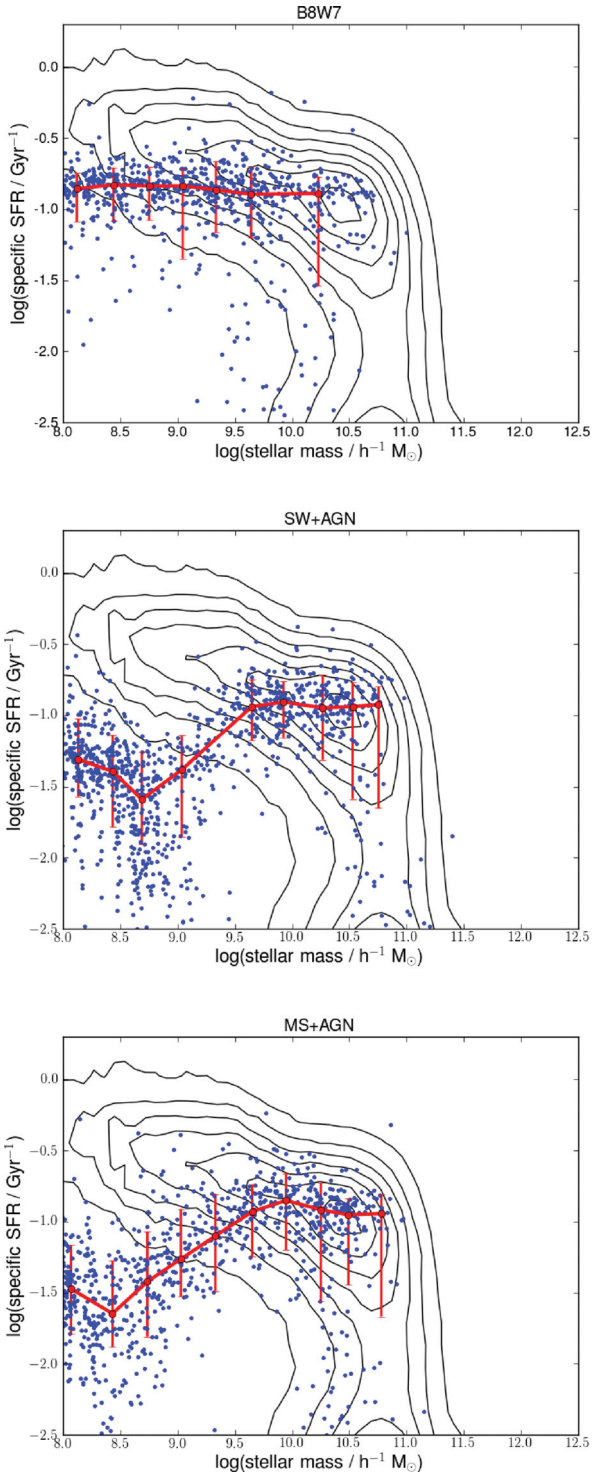

Figure 14. Comparison of the SSFRs of central galaxies for three models which provide encouraging fits to the SMF. In separate panels, we show the baseline model, B8W7, SW+AGN $[(v_{\rm wind,200},\ \beta_{200},\ \alpha_{\rm cool}) = (180,\ 8,\ 0.25)]$ and the MS+AGN model $[(v_{\rm wind,200},\ \beta_{200},\ \alpha_{\rm cool}) = (250,\ 8,\ 0.35)]$. The star formation sequence for local galaxies is shown as black contours (see text for details). A random sample of model galaxies are shown as blue points, with the median relation and 10th and 90th percentiles shown as a red line with error bars. The B8W7 model has a constant SSFR regardless of stellar mass. In contrast, the SW and MS models produce relations that are strongly dependent on system mass and are incompatible with the observational data: instead of a near-constant SSFR, the SW and MS model have SSFRs that are lower in small galaxies.

is suppressed by AGN feedback). Given the age of the universe, the location of the sequence at SSFR $\sim 10^{-0.8}\,\mathrm{Gyr}^{-1}$ implies that the SSFRs of all galaxies have been steady (or slowly rising) over the history of the universe. Note that the uncertainties due to dust obscuration would tend to increase the star formation rates of the most massive galaxies. However, the data presented already include an extinction correction based on the Balmer decrement, and agree well with SSFRs based on SED fitting (McGee et al. 2011).

In each panel, a random sample of model galaxies is shown as blue points, while the red line with error bars shows the median SSFR and the 10th and 90th percentiles of the distribution. We include only central star-forming galaxies (with SSFR $>0.01\,\mathrm{Gyr}^{-1}$) in this calculation, but very similar results are obtained if we include star-forming satellite galaxies as well. The observed relationship is reproduced fairly well by B8W7. The SSFR in the model is flat over a wide range in stellar mass, from below $10^7\,h^{-1}\,\mathrm{M}_\odot$ to almost $\sim 10^{11}\,h^{-1}\,\mathrm{M}_\odot$ (where the relation dips as the supply of fuel for star formation is suppressed by the hot-halo feedback).

In contrast, the SW and MS models predict a relation with a noticeable *decline* in SSFR towards lower stellar masses. Although the declining trend is less abrupt in the MS model (bottom panel) than that in the SW model (middle panel), both relations are clearly inconsistent with the observational data. We infer from this figure that a dramatic change of feedback efficiency cannot be responsible for the flattening of the SMF. This problem is also seen in the hydrodynamical models of GIMIC and Op08 (see Davé, Oppenheimer & Finlator 2011), although the limited mass resolution of those simulations limits the comparison to galaxies more massive than $\sim 10^9\,h^{-1}\,\mathrm{M}_\odot$, and consequently the mass dependence is not so clearly evident.

It is interesting to understand the origin of the dip. In the SW and MS models, the flat region of the SMF is created by a transition between the two feedback regimes: for low halo masses, feedback is extremely effective at expelling gas from the halo and the star formation rate is strongly suppressed. At higher masses, however, the wind velocity is no longer sufficient to escape the halo and the cold gas mass and star formation rate increase. However, because the division between the two regimes occurs at a fixed escape velocity, we expect that the halo mass of the transition evolves rapidly with redshift, $M_{\mathrm{halo,kink}} \sim (1+z)^{-3/2}$. Allowing for the dependence of stellar mass on halo mass, $M_* \propto {M_{\mathrm{halo}}}^2$ (e.g. Fig. 2, the relation evolves slowly with redshift) we expect the stellar mass at which the transition occurs to evolve as $M_{*,\mathrm{kink}} \sim (1+z)^{-3}$. Thus, the transition mass evolves more quickly than the mass of an individual galaxy. Thus, the transition mass is much smaller at high redshift. Over time, an individual galaxy makes a transition from the regime in which ejection is ineffective to the one in which it is. Consequently, galaxies somewhat below the transition mass at $z = 0$ have low current star formation rates compared to their past average. At the very lowest stellar masses, the SSFR in the SW and MS models begins to increase. Galaxies that lie well below the kink in the M_*-M_{halo} relation have experienced similar feedback during their formation history and the rise is thus to be expected.

Although the B8W7 model provides the best description of the observational data, it does not reproduce the weak trend for the SSFR to increase as the stellar mass decreases (SSFR $\propto {M_*}^{-0.22}$) that is seen in the data. The strength of this trend is controversial, but does not appear to be an observational selection effect. It is seen regardless of the star formation diagnostic that is applied, and is apparent at higher redshifts as well as locally (but this depends critically on the definition of the sample; for a recent overview, see Karim et al. 2011). In Section 4.1, we noted that the surprisingly rapid evolution of the normalization of the observed mass function suggested that the effective wind speed should scale with redshift. This change would also have implications for the SSFR, since the present-day star formation rate would increase relative to the past average. Varying the escape speed rather than the mass loading could create a tilt in the SSFR $-M_*$ relation since the effect will be strongest around the kink in the M_*-M_{halo} relation but result in little change in the star formation histories of the most massive star-forming galaxies.

In summary, the SSFR of galaxies provides an important additional discriminant of the models. The B8W7 model comes closest to matching the observed data, with the characteristic SSFR that is almost independent of stellar mass. In contrast, the SW and MS models show SSFRs that decline with decreasing stellar mass, while the observational data show a slightly increasing trend. Further exploration of feedback schemes that scale systematically with redshift is required to identify a model which produces a better match to the observational data.

5.2 The star formation history of the universe

We have seen that all the models reproduce the observed build-up of stellar mass reasonably well, and another way to tackle this question is to investigate the star formation rates of galaxies directly. In Fig. 15 we show the evolution of the cosmic star formation rate. For the models this is calculated by integrating the contribution of galaxies down to stellar masses of $10^8\,h^{-1}\,\mathrm{M}_\odot$. As a result of including feedback from AGN in order to suppress the formation of the most massive galaxies, the behaviour of all three models

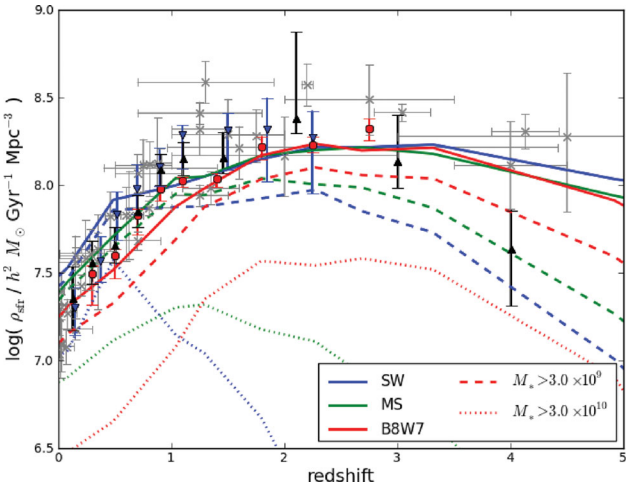

Figure 15. The evolution of the cosmic star formation rate density as a function of redshift. Solid lines show the total star formation rate density. Line colours distinguish different feedback models. All three models include hot-halo AGN feedback that suppresses the formation of stars in high-mass haloes. While the models all show similar star formation rates above $z = 2$, the decline in the star formation rate between $z = 1$ and the present day differs markedly. The solid lines should be compared with the observational data, shown as black crosses (from the compilation of Hopkins 2004), red circles [from the stacked Very Large Array (VLA) analysis of Karim et al. 2011], blue triangles from Rodighiero et al. (2010) and black triangles from Cucciati et al. (2011). The dashed and dotted lines show the contribution from galaxies more massive than $10^{9.5}$ and $10^{10.5}\,h^{-1}\,\mathrm{M}_\odot$, respectively. Despite the overall similarity of the star formation histories in the three models, the way the star formation rate is divided between stellar masses varies greatly.

is broadly similar in this plot. Above $z = 2$, the integrated star formation rates of the models are very similar. However, the strength of the decline in the star formation rate differs between the models, being much stronger for the B8W7 model than for the SW scheme. As expected, the MS scheme lies in between.

We compare the models with the compilation of observational data from Hopkins (2004) and the more recent data of Karim et al. (2011) (radio, red circles), Rodighiero et al. (2010) [mid-infrared (IR), blue triangles] and Cucciati et al. (2011) [rest-frame ultraviolet (UV)]. The radio and mid-IR based measurements have the advantage of that the obscured star formation is accounted for (see Karim et al. 2011 for further discussion). None of the models matches the trends in the observational data perfectly. Given the scatter in the observational data sets, there is little reason to choose between the MS and B8W7 models on the basis of this plot. However, the uncertainties in the observational data arise in large part because of the large extrapolation required to correct the observed star formation rate density for galaxies that are too faint to be directly detected.

An important question is therefore to examine how the star formation rate density depends on the stellar mass of galaxies. The mass dependence of the cosmic star formation rate in the models is illustrated by dashed and dotted lines in Fig. 15. The line styles show the integrated contribution from galaxies more massive than $10^{9.5}$ and $10^{10.5}\,h^{-1}\,\mathrm{M}_\odot$, respectively. Although the total star formation rates of the models are similar, the way the star formation is distributed between galaxy masses is different, with the SW and MS models showing a shift from star formation that is dominated by low-mass ($<10^{9.5}\,h^{-1}\,\mathrm{M}_\odot$) galaxies at high redshift (note the large difference between the solid and dashed curves) to a dominance by lower mass galaxies at low redshift. In contrast, in the B8W7 model, massive galaxies make a similar contribution to the total star formation rate density at all redshifts (i.e. there is a constant offset between the solid, dashed and dotted lines). Thus, although all three models show a similar total star formation rate density at $z > 2$, the contribution of different mass galaxies is very different in the models.

Given the different contributions of high- and low-mass galaxies in the models, it is very dangerous to quantitatively compare to observational data for the total star formation rate density since the observations are usually based on extrapolation of the properties of high-mass galaxies. Therefore, in Fig. 16 we show the mass dependence of the cosmic star formation rate, separating the redshift dependence into separate panels. The different models are distinguished by the coloured lines. The solid lines show the distribution measured for the model galaxies, and the dashed lines illustrate the effect of a 0.2-dex error in stellar mass assignment. Using this plot, it is then possible to directly compare to observational measurements based on mass-complete samples (Gilbank et al. 2011; Karim et al. 2011). Historically, such plots have been used to infer that the star formation density is dominated by large galaxies (at high redshifts), and by lower mass galaxies at low redshift. In practice, the more complete and sufficiently deep data sets show that the increase in the star formation rate density with redshift is similar for all galaxy masses; however, the limitations of the observed data sets, particularly at high redshift, are evident.

A comparison of the models and observational data in this logarithmic plot allows the contribution from the tails of the mass distribution to be clearly seen. At low redshift, star formation in the SW and MS models is more concentrated towards high-mass galaxies than suggested by the data. In contrast, star formation in the B8W7 model is somewhat too flat. Also, while a significant contribution to the star formation rate comes from low-mass

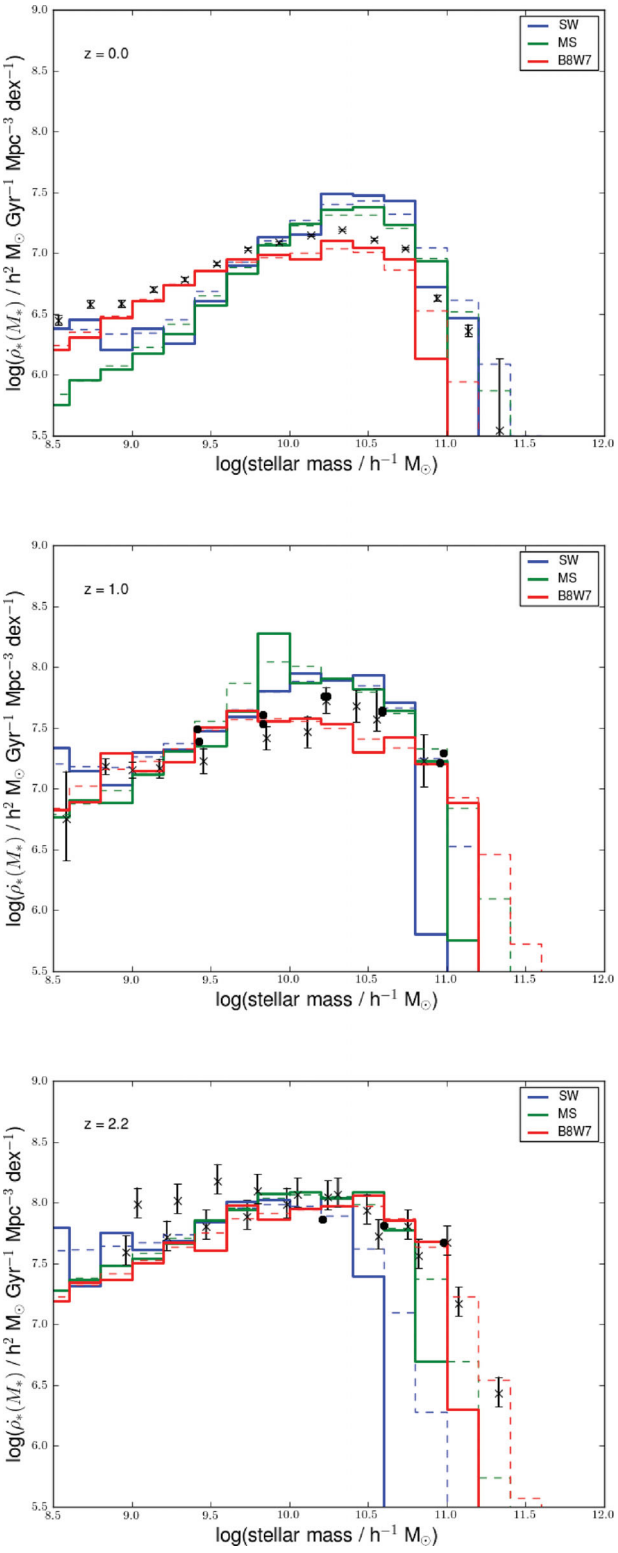

Figure 16. The evolution of the cosmic star formation rate density as a function of the stellar mass. The coloured lines differentiate different feedback schemes, while the panels compare to data at $z = 0.0$, 1.0 and 2.2. Solid lines show the distribution obtained from the models, dashed lines illustrate the effect of a random error of 0.2 dex in the stellar mass estimates. Observational data collected by Gilbank et al. (2011) and Karim et al. (2011) are shown as black symbols. The figure gives a more accurate comparison of the observed star formation rate density than that predicted by the models (see text for discussion).

galaxies (as suggested by the observations), the model fails to generate sufficient massive star-forming galaxies. This occurs because the AGN feedback parameters need to match the observed SMF set an effective halo mass threshold that is lower in B8W7 than the SW and MS models. This is driven by the larger scatter in galaxy mass at fixed halo mass in B8W7. Thus, the model contains galaxies of the stellar mass required to match the observed SMF, but star formation has been suppressed in too many of them. A stellar mass error of 0.2 dex improves the match (as shown by the dashed line), but this probably overestimates the uncertainty in the local data. The difference in behaviour at $z = 0$ tallies with the differences in the models' SSFR–M_* relations seen in Fig. 14.

The behaviour of the models is similar at $z = 1$: B8W7 results in a distribution that is flatter than the data, while the MS and SW models are slightly too peaked. The order of the models at the massive end is reverse, with the B8W7 model suggesting the greater population of massive star-forming galaxies. The differences are small, however, and likely to be masked by the uncertainties in the stellar mass determination (see dashed lines). At $z = 2$, a deficit of massive star-forming galaxies is apparent for the SW model, and the rapid drop at high masses is not reconciled with the observational data if an error of 0.2 dex is assigned to the stellar mass. This deficit is to be expected from our comparison with the SMF of this model: the SW model lacks sufficient high-mass galaxies.

In summary, while all models result in similar predictions for the evolution of the integrated cosmic star formation rate, the mass dependence of the cosmic star formation rate exposes important differences between the models. Although none of the trends is conclusive, the comparison highlights different issues with the models. The SW and MS models tend to under-represent star formation from low-mass galaxies at low redshift. At high redshift, the SW model under-represents the contribution to star formation from high-mass galaxies. In contrast, the B8W7 model tends to under-represent the contribution from high-mass galaxies at lower redshift. As the consistency and completeness of the observational data improves, this approach has great potential for discriminating between feedback models.

5.3 Where are the baryons?

As we have seen, in the successful models, only a small fraction of baryons are locked up into stars. It is of interest to examine the phase of the remaining baryons in the model, and this is illustrated in Fig. 17. We focus on the same three models that provided a good description of the SMF above $M_* \sim 10^9\,h^{-1}\,\mathrm{M}_\odot$: B8W7, SW $(v_{\mathrm{wind},200}, \beta_{200}) = (200, 4)$ and MS $(v_{\mathrm{wind},200}, \beta_{200}) = (200, 16)$. The SMFs of these models are shown in Fig. 11.

The distribution of baryonic matter between phases is plotted as a function of halo mass. We include all baryons associated with the halo by the model, and define the baryon fraction as the mass of baryons, M_b, divided by the cosmic average (i.e. $f_\mathrm{b} = M_\mathrm{b}/\Omega_\mathrm{b} M_\mathrm{halo}$). In the absence of feedback, the baryons collapse with the dark matter, and all haloes would be baryonically 'closed' with fractions of the order of unity (Crain et al. 2007). The lines are cumulative so that the mass in stars is shown as a blue line. The green line adds the contribution of cold gas, and the red line includes the contribution from diffuse halo gas (i.e. 'hot' plus 'reheated' components in the terminology of Bow06). This gives the total baryon mass associated with the halo. Note that, because of the expulsion feedback considered in these models, the total mass of baryons in the halo need not reach the cosmic value.

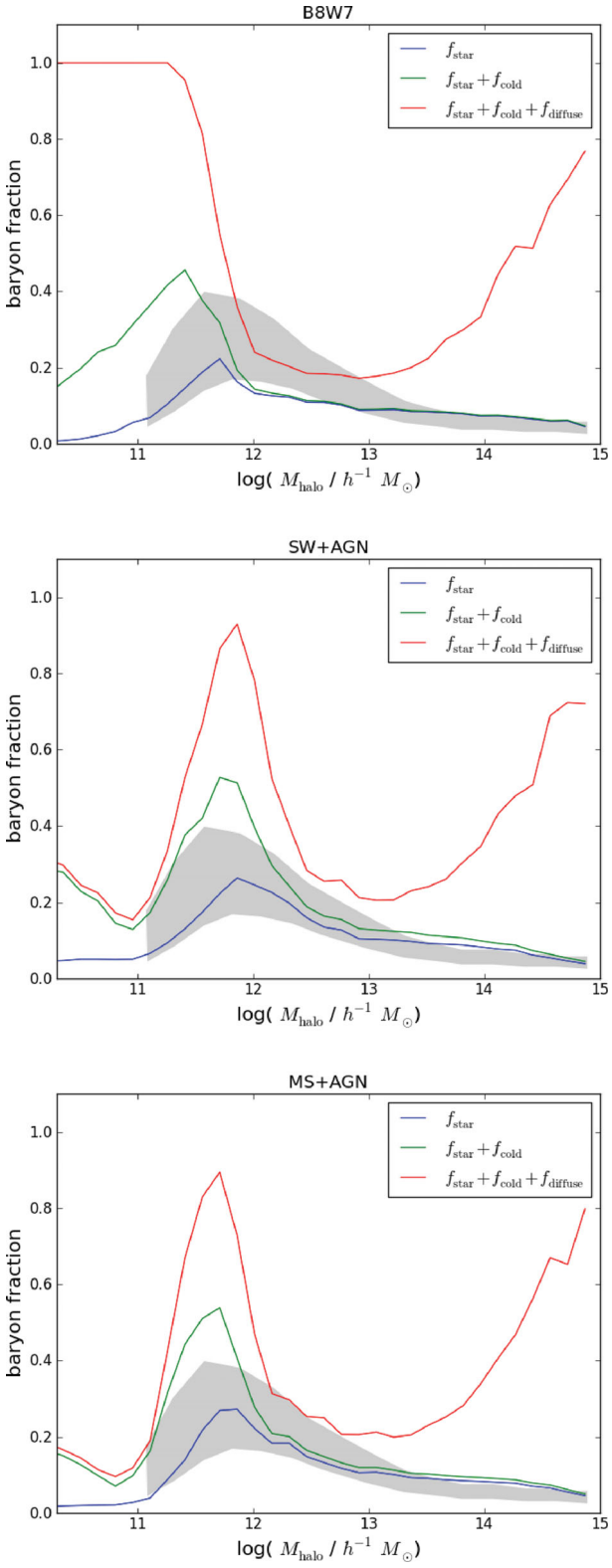

Figure 17. The fraction of baryons (relative to $M_\mathrm{halo}\Omega_\mathrm{b}/\Omega_0$) in different phases as a function of halo mass. The lines are cumulative showing, blue: stars, green: adding cold gas, red: adding hot and reheated gas. The three panels show the B8W7, SW and MS models (all including the 'hot-halo' mode of AGN feedback). The models include the effect of hot-halo AGN feedback, as shown in Fig. 11. The shaded region indicates the total stellar mass content of haloes measured in observational data sets (see text for details).

As we should expect, given that the models reproduce the $M_*-M_{\rm halo}$ relations deduced from subhalo abundance matching, all the models have a similar peak in the fraction of baryons locked into stars around $M_{\rm halo} \sim 10^{11.5}\,h^{-1}\,{\rm M}_\odot$ (solid lines). In the B8W7 model, this declines rapidly to lower masses, and declines to an almost constant value ($\sim$0.08) in higher mass systems. In the SW model, the excess abundance of small galaxies means that the decline to lower mass haloes is less steep, reaching a constant value of 0.05 in haloes smaller than $10^{11}\,h^{-1}\,{\rm M}_\odot$. The MS model is intermediate between the two, as we should expect. The amplitude of the peak is smaller in the B8W7 case than in the other models. These differences are largely due to the contribution from very faint satellite galaxies ($M_* < 10^9\,h^{-1}\,{\rm M}_\odot$) in the SW and MS models, although it is also noticeable that the peak in stellar abundance is broader in these systems. This is consistent with all of the models giving a similar match to the mass function because of the interaction between the halo abundance and the scatter in the M_*–$M_{\rm halo}$ relation.

The most striking difference between the plots lies in the behaviour of the total baryon contributions below $M_{\rm halo} \sim 10^{11.5}\,h^{-1}\,{\rm M}_\odot$. In the B8W7 model (in this regime), all of the baryons reside in the host halo, and the diffuse gas makes up the majority of the baryons. It is questionable whether such a large halo component (which we refer to as the 'circumgalactic medium', CGM) is compatible with observational data. Even though recent observations from the Cosmic Origins Spectrograph (COS) on *Hubble Space Telescope (HST)* suggest that there is a significant CGM (Tumlinson et al. 2011), it is likely to account for 'only' a similar fraction to the mass in the galaxy's stars (i.e. $\sim$10 per cent of $\Omega_{\rm b}$; see Prochaska et al. 2011 for discussion). In contrast, the SW and MS models are able to eject much of the material from the halo so that low-mass galaxies contain little diffuse gas in their haloes, and most of the baryons are locked into cold gas. As cold gas is ejected from these galaxies, it escapes from the halo. Nevertheless, the feedback scheme in these models has been tuned to match the observed SMF, so that the outflowing wind stalls as the halo mass approaches $10^{12}\,h^{-1}\,{\rm M}_\odot$, resulting in a peak in the abundance of diffuse gas that matches the peak in the stellar fraction. At masses above $\sim 10^{12}\,h^{-1}\,{\rm M}_\odot$, material is driven out of the halo by AGN feedback in all three models, resulting in a second dip in the abundance of diffuse gas. As discussed in Bow08, this results in a good match to the X-ray luminosities of groups and clusters. The rapidly rising gas fractions account for the tilt in X-ray scaling relations compared to the 'self-similar' predictions.

In order to compare the models with observational data, we show measurements of the stellar mass fraction of haloes from Leauthaud et al. (2012) as a shaded region in Fig. 17. This analysis is based on combining a subhalo occupation distribution model and integrated stellar mass measurements for a sample of COSMOS X-ray-detected groups, for which halo masses have been determined using a weak lensing analysis. In contrast to some previous results (e.g. Gonzalez, Zaritsky & Zabludoff 2007; Giodini et al. 2009), Leuthaud's careful analysis shows that the stellar mass content of high-mass haloes is low and comparable to the mean stellar mass fraction of the universe as a whole (as would be expected on quite general grounds; Balogh et al. 2008). The shaded region indicates the plausible systematic uncertainties in the analysis. The survey includes only directly detected galaxies, and could underestimate the stellar mass content of some haloes with significant intracluster light. However, the correction for intracluster light is likely to be small and certainly less than 50 per cent (Zibetti et al. 2005; McGee & Balogh 2010). As Leauthaud et al. (2012) discuss, the differences from previous work arise from the treatment of stellar populations and satellite galaxies, not from the contribution of the intracluster light.

All of the models compare with these observational data reasonably well. The SW and MS models fit particularly well, with a peak in the stellar mass fraction close to $\sim 10^{12}\,h^{-1}\,{\rm M}_\odot$. At higher masses, the observations show a shallow decline, which is also seen in the models. Differences between the model and the data could certainly be accounted for by the missing intracluster light distribution. The B8W7 model shows a somewhat sharper decline than is evident in the data, together with a peak in the stellar mass fraction that is shifted to lower masses. However, this part of the observational region is dominated by the scatter in the $M_*-M_{\rm halo}$ relation: all of the models presented reproduce the observed SMF, as seen in Fig. 11.

In summary, the different feedback schemes result in significant differences in the distribution of mass between the phases of baryonic matter. In the B8W7 model, haloes below $\sim 10^{11.5}\,h^{-1}\,{\rm M}_\odot$ retain the cosmic baryon fraction, mostly in the form of a diffuse gas halo. In the SW and MS models, most baryons have been ejected from low-mass haloes. In higher mass haloes, all three models suggest that the stellar fraction should quickly converge to the average value. The shape of the models fits well with observational measurements.

5.4 Metal enrichment of the intergalactic medium

Although the models presented have all been developed in order to explain the galaxy population of the universe and the X-ray scaling relations of galaxy clusters, the metal enrichment of the IGM could offer an important discriminator between the models. Indeed, a major motivation for considering powerful winds is to explain the widespread metal enrichment of the universe. Recent observations have shown that metals are widely distributed in the IGM. For example, Prochaska et al. (2011) show that the low-redshift metal absorption line cross-section is compatible with $0.3\,L^*$ galaxies being surrounded by a metal-enriched halo extending out to $\sim$300 kpc (i.e. well beyond the virial radius of the galaxy). Quantitative comparison with these data is fraught with difficulties, however, since the observations must be corrected for the ionization state of the absorbing clouds. This issue is further complicated by the multiphase nature of the absorbers. The inherent correlation of bright and faint galaxies must also be taken into account, particularly if smaller galaxies dominate the enrichment. In addition, the phenomenological models include a number of simplifying assumptions that make quantitative comparison with the data difficult. In particular, the models we have presented here assume that enrichment and metal recycling is instantaneous and that metals are distributed uniformly throughout the galaxy halo. These limitations must be carefully taken into account in a detailed comparison, and we will not attempt this here. Nevertheless, it is instructive to briefly examine the differences between one model and another.

All of the models result in significant ejection of metals from the galaxy disc, and there is little to distinguish between the models on the basis of the mass of metals ejected from the disc. Rather, the important discriminator is the distribution of the metals. In the B8W7 models, the metals remain trapped in the dark matter potential of the galaxy. In the SW and MS models, metals ejected from small galaxies will be more widely distributed. At face value, the observed widespread distribution of metals would favour the SW or MS model.

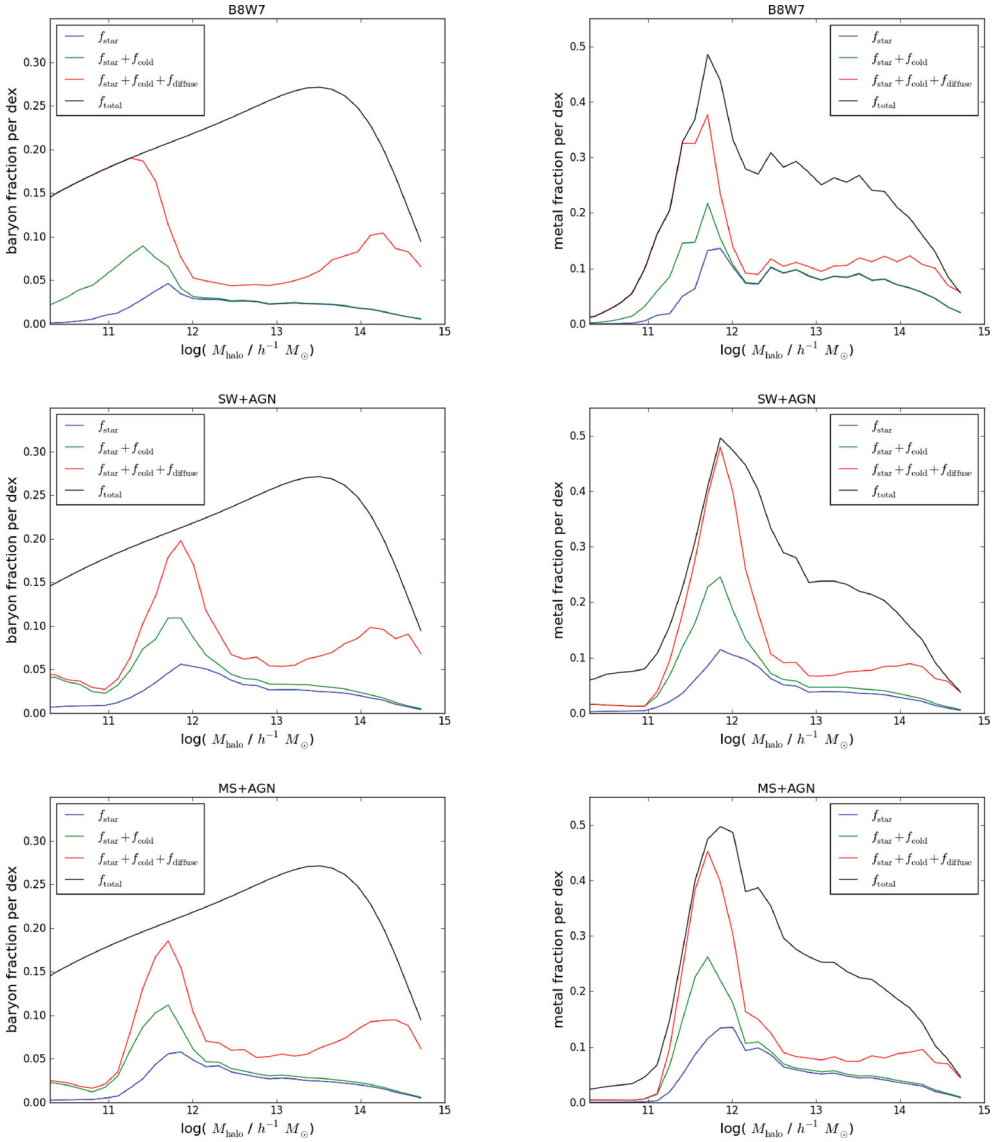

Figure 18. Left-hand panels: the contribution of baryons in different phases to the total baryon content of the universe at $z = 0$, plotted as a function of halo mass. The lines are cumulative showing, blue: stars, green: adding cold gas, red: adding hot and reheated gas, and can be compared to the mass fraction (per halo mass bin) in different phases shown in Fig. 17. The solid black line shows the total contribution associated with haloes of a given mass in the absence of any mass loss. The difference between the red and black lines indicates the contribution to the total intergalactic baryon mass that has been expelled from haloes. Right-hand panels: the contribution of haloes and phases to the total metal mass content of the universe at $z = 0$. The black line indicates the metal mass that would be associated with haloes in the absence of mass ejection, so that the difference between the black and red lines indicates the contribution to intergalactic metals. The figure shows that the contribution to intergalactic metals is dominated by the mass expelled from galaxy groups by AGN feedback in all models. This dominates over the metals that are expelled from the haloes of low-mass galaxies.

In practice, however, the situation is more complex. At low redshift, in all of these models, most of the intergalactic metals are associated with material that has been ejected from galaxy groups by the action of AGN feedback. The situation at low redshift is illustrated in Fig. 18. The panels on the left-hand side show the distribution of the baryon mass between haloes of different mass, such that the integral under the curve gives the total baryonic mass of the universe in haloes more massive than $10^{10}\,h^{-1}\,{\rm M}_\odot$. The coloured lines illustrate the division of the baryons between different phases for the different models. The blue, green and red lines show the cumulative contribution from star, cold gas and diffuse gas, respectively, as previously discussed. The black line shows the contribution to the baryon mass in the absence of expulsion. It is noticeable that this is widely distributed, with all haloes less massive than $10^{14}\,h^{-1}\,{\rm M}_\odot$ making a similar contribution to the total mass content. Because the distribution of mass between different haloes is rather flat, the coloured lines bear close resemblance to those in Fig. 17, and attention can be focused on the difference between the red and black lines. Since the scale is linear, this is the mass contribution from material that has been ejected from haloes. In the B8W7 model, this contribution only comes from the material expelled by AGN in haloes more massive than $10^{12}\,h^{-1}\,{\rm M}_\odot$, while in the SW and MS models, expulsion feedback means that there is a second important contribution from haloes less massive than $10^{11.5}$. The integral of the difference between the red and black curves measures the total mass fraction ejected from haloes. We find that

the fractions are 52, 61 and 66 per cent for the B8W7, SW and MS models, respectively (for haloes more massive than $10^{10}\,h^{-1}\,\mathrm{M}_{\odot}$).

The panels on the right show the equivalent plot for the contribution to the metal content of the universe. The line colouring follows that used in the left-hand panels. The black line shows the contribution from stars that are now in haloes whose masses are given on the horizontal axis. Of course, when the metals were produced, the stars may have been in lower mass haloes. The metals locked into stars, cold gas and diffuse gas are shown by the coloured lines, and the area under the curve gives the total contribution. In contrast to the left-hand panels, the metal contributions are much more strongly weighted to haloes more massive than $10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$, which reflects the low efficiency of star formation in smaller haloes, and thus the low rate of metal production. As a result, the contribution to the global metal budgets from low-mass haloes is strongly downweighted compared to the contribution from galaxy-group haloes (driven by AGN feedback, see below). The fraction of metals that have been lost from haloes is given by the difference between the red and black lines. To be clear, the difference between the lines corresponds to the fraction of metals that are absent from haloes of the mass given on the x-axis at $z = 0$. Some of these metals may have been ejected from progenitor haloes of smaller mass. We find that the models have total intergalactic metal fractions of 42, 45 and 46 per cent for the B8W7, SW and MS models, respectively. In the B8W7 model, the black and red curves are superposed below a halo mass of $10^{11.5}\,h^{-1}\,\mathrm{M}_{\odot}$, but the results are similar for all the models, with by far the largest contribution coming from the group-scale haloes. Above $10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$, the behaviour of the models is similar since all the models share a common AGN feedback scheme. In the absence of AGN feedback, very few metals are ejected from such haloes. Focusing on lower mass haloes, greater differences are apparent. In particular, the SW and MS models have an extended tail making a useful contribution to the intergalactic metals. This tail is not present in the B8W7 model, both because the material remains trapped in the halo potential and because the SMF is much flatter in the B8W7 model than in either the SW or MS model. Nevertheless, low-mass haloes make a relatively little contribution to the total metal content of the universe in all the models.

At higher redshifts, the fraction of intergalactic metals drops rapidly in all the models, reflecting the lower abundance of high-mass haloes with effective 'hot mode' feedback. At these redshifts, a larger fraction of metals come from dwarf galaxies in low-mass haloes and this is reflected in greater differences in the intergalactic metal fraction between the models. In Fig. 19 , we show the evolution of the fraction of metals that have been ejected from galaxies. To be specific, we show the fraction of all metals produced at higher redshifts, associating ejected metals with the halo in which the stars that produced the metals are located at the specified redshift. Metals produced by high-redshift galaxies may be recaptured by larger haloes at low redshift, just as the original stars of the early galaxy may have been merged into a larger system. In principle, metals that are missing from haloes of mass greater than $10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$ may have been ejected at earlier time from a progenitor halo of much lower mass; however, by re-running the models without AGN feedback, we have shown that this process is negligible. In haloes more massive than $10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$, the loss of metals from the halo is entirely due to the effect of AGN feedback.

The black line shows the total fraction of metals that are outside galaxy haloes, while the coloured lines show how this fraction depends on halo mass: the blue (red) line shows the material that has been lost from haloes less massive than $10^{11}\,h^{-1}\,\mathrm{M}_{\odot}$ ($10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$)

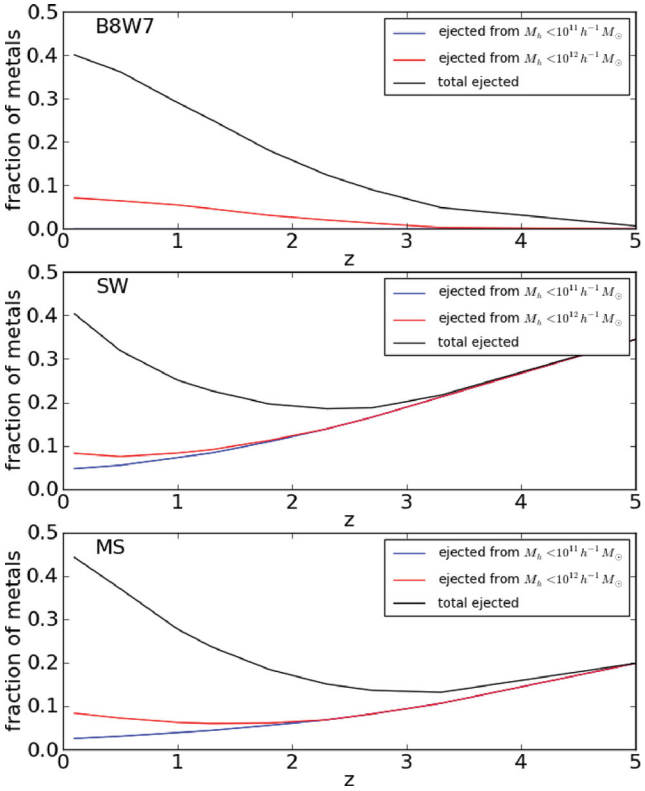

Figure 19. The fraction of metals that have been expelled from haloes as a function of redshift for the three different models. The black line shows the total fraction of ejected metals, while the red and blue lines show the fraction ejected from haloes less massive than 10^{12} and $10^{11}\,h^{-1}\,\mathrm{M}_{\odot}$, respectively. At low redshift, the intergalactic metal content is dominated by ejection from galaxy group and cluster haloes by the 'hot-halo' mode of feedback. At high redshift, there are few such massive haloes and this ejection mechanism is unimportant. As a result, there are few intergalactic metals in the B8W7 model at $z > 3$. In the SW and MS models, the fraction of metals ejected from haloes is much higher at this redshift, primarily as a result of ejection from dwarf galaxies.

so that the contribution from material ejected from small (large) galaxies can be seen. The blue line is not visible for the B8W7 model because material is not expelled from galaxy haloes (the small contribution from the green line is due to hot-halo mode feedback in $M_{\mathrm{halo}} \sim 10^{12}\,h^{-1}\,\mathrm{M}_{\odot}$ systems). The area above the red line shows the contribution from metals ejected by the 'hot-halo' mode of feedback. Although this fraction declines quickly with increasing redshift, it dominates the galaxy contribution out to $z \sim 2$, even in the MS and SW models. At $z > 3$, the IGM metal distribution reflects the ejection of metals from dwarf galaxy haloes in these models, and the observed widespread enrichment of the high-redshift IGM (Aguirre et al. 2008; Frank et al. 2010; Steidel et al. 2011) seems to favour the expulsion models (SW and MS) over B8W7 (Bertone, Stoehr & White 2005; Cen & Chisari 2011; Booth et al. 2012).

Although these figures make for an interesting comparison between models, a quantitative comparison to the observed abundance of intergalactic metals is extremely difficult. In particular, the current model does not specify how far outside the group haloes the metal distribution will extend, and does not begin to address its temperature and ionization state. This issue can only be satisfactorily addressed if we account for the non-spherical nature of halo

accretion and outflow. Equally, the extended diffuse gas distribution within haloes will also make a contribution to the metals seen in random sight lines. It would therefore be premature to rule out any of the models without careful consideration of the effects of galaxy clustering. This task is beyond the scope of the present paper but is clearly an important avenue for future work.

In summary, all three models that we have considered eject the majority of metals from the galaxy disc. The metals locked into stars and cold gas represent a relatively minor component of the universe. In the expulsion models, SW and MS, the majority of metals completely escape from haloes smaller than $10^{11.2}\,h^{-1}\,\mathrm{M}_{\odot}$. Thus, these metals are widely distributed across intergalactic space. In contrast, the B8W7 model retains metals within lower mass haloes, creating a circumgalactic medium. However, the dominant contribution to intergalactic metals in all three models comes from the material ejected from group-scale haloes by the action of AGN feedback. While current observational data suggest that the metal haloes of galaxies extend well beyond their virial radii (and thus that the haloes are more extended than the B8W7 model would suggest for individual galaxies), a more detailed study of the effects of clustering (and the observed clustering of metal lines) is required before the models can be distinguished on this basis.

6 CONCLUSIONS

'Feedback' is a fundamental component of galaxy formation models, allowing us to understand the marked differences between the dark matter halo mass function and the galaxy SMF. In order to explain the observational data, galaxy winds must strongly suppress the formation of low-mass galaxies. Phenomenological (or 'semi-analytic') models, such as GALFORM, have been shown to provide a good description of the galaxy mass function together with its evolution, the SSFRs of galaxies and the contribution of galaxies of different masses to the cosmic star formation rate density. In the standard GALFORM model, these successes are achieved by adopting a feedback parametrization that varies strongly with system mass, such that the speed of the outflow tracks the mass of the halo while the mass loading of the wind decreases strongly with halo mass. In contrast, most hydrodynamical models, including the GIMIC simulations, adopt a wind that is independent of the system mass.

In this paper, we have modified the semi-analytic code to implement feedback schemes similar to those usually adopted in hydrodynamic models. Our scheme allows the wind speed and mass loading to be adjusted and the consequent effect on the SMF of galaxies, its evolution and other galaxy properties to be determined. We focus on three particular models. Two models have fixed wind parameters: the pGIMIC model has high wind speed and modest mass loading similar to those of the GIMIC hydrodynamical simulations (Crain et al. 2009); the SW model has a slower wind speed (at the same mass loading). We also consider a model (MS) in which the wind mass loading scales with the inverse of the disc circular velocity. This is similar to the momentum scaling models of Op08, although we set the wind speed by keeping fixed the total energy of the wind (rather than its total momentum) across halo masses. Initially, we also considered a model with a $1/v_{\mathrm{disc}}^2$ dependence of mass loading (ES), but found that this could not reproduce the observed SMF. We contrast these models with a model based on Bow06 and Bow08, but adapted to the *WMAP*7 cosmology used in this paper (B8W7). For each of these models, we consider the effect of including a 'hot-halo' mode of AGN feedback following the gas expulsion scheme of Bow08.

The main results are as follows.

(i) We find that the phenomenological description of feedback that we use in our code reproduces many of the trends seen in hydrodynamic simulations. This gives us confidence that the phenomenological approach captures the key physical effects of galactic winds well and allows us to explore the parameter space of wind properties quickly and efficiently. A comparison of the phenomenological model and hydrodynamic simulation also allows us to better understand the key physical processes in galaxy formation. A model which uses a high wind speed and moderate mass loading (and does not include AGN feedback) reproduces the form of the SMF seen in the GIMIC numerical simulations well. Although the strong feedback suppresses the stellar masses of galaxies strongly, the effect is most prominent around $10^{10.5}\,h^{-1}\,\mathrm{M}_{\odot}$, so that the resulting SMF has a dip in the abundance of galaxies at these masses. At lower masses, the SMF rises steeply since star formation is too weakly suppressed, while at higher masses the wind is not sufficiently energetic to escape the halo. Using the phenomenological model, we are able to explore how the dip in the mass function depends on the assumed wind speed and mass loading.

(ii) The observed mass function is better matched by adopting a relatively slow wind, as seen in the hydrodynamic simulations of Op08. The SW model, with a wind speed of $180\,\mathrm{km\,s^{-1}}$ and a wind mass loading of 8, results in a good match to the observed galaxy mass function over the range 10^9–$10^{11}\,h^{-1}\,\mathrm{M}_{\odot}$. Introducing the AGN scheme developed in Bow08 suppresses the formation of galaxies from cooling in hot gas haloes, resulting in a good match to the high-mass turnover of the mass function. The combined (AGN plus supernova-driven wind) model provides a good description of galaxy abundance down to $10^9\,h^{-1}\,\mathrm{M}_{\odot}$ (but rises much more steeply than the observational data at lower masses). Although this is encouraging, we find that the model performs poorly compared to B8W7 in several other aspects. In particular, when we examine the SSFRs in this model, we see that the fixed feedback scheme results in a kink in the SSFRs of galaxies at around a stellar mass of $10^{9.5}\,h^{-1}\,\mathrm{M}_{\odot}$, such that low-mass galaxies have a factor of $\sim$5 lower star formation rates than their massive counterparts. This is not seen in observational data. The model also tends to underrepresent the abundance of massive ($M_* > 10^{11}\,h^{-1}\,\mathrm{M}_{\odot}$) galaxies at $z > 1$ compared to the B8W7 model and the observational data.

(iii) We also consider a model in which the wind mass loading scales with the inverse of the circular velocity of the disc. This model results in a more subtle transition between the regime where material easily escapes the halo and that in which it stalls. With suitable choice of parameters, and the inclusion of the AGN feedback scheme, the model reproduces the observed galaxy mass function above $M_* \sim 10^9\,h^{-1}\,\mathrm{M}_{\odot}$. However, although the slope of the SSFR$-M_*$ relation is weaker than that seen in the SW model, the trend is reversed compared to the slightly rising relation seen in observational data. The model tends to underpredict the abundance of massive, $z > 1$ galaxies compared to B8W7. The deficit is, however, relatively small and the MS model could be compatible with the observational data if the random uncertainties in the stellar mass estimates are greater than 0.2 dex.

(iv) Feedback from AGN may have two very distinct forms. The 'hot-halo' (or 'radio') mode feedback suppresses the supply of gas from cooling haloes. This is a key component of the successful models we present. However, it is interesting to investigate if this can be replaced by a 'starburst' or 'QSO' mode of feedback. We implement this by enhancing the wind speed and mass loading during starburst events. We find that while this mode may enhance

the feedback from quiescent star formation, it does not introduce a characteristic break in the galaxy mass function and cannot be seen as an alternative to the 'hot-halo' mode feedback. Similar conclusions are reached by Gabor et al. (2011), when they introduce a comparable feedback scheme in hydrodynamical models.

(v) We have also investigated the star formation history of the each of the models, and its dependence on the stellar mass. Looking at the integrated star formation rate, the main difference lies in the strength of the decline between $z = 2$ and the present day. This is strongest in the B8W7 model. Nevertheless, given the uncertainties in the observational estimates, both the MS and B8W7 models seem to provide a reasonable description. However, the models show a much greater variety in behaviour when the results are broken down by stellar mass. For this reason, comparison of the total star formation rate may be misleading since the observational results assume large extrapolations to account for galaxies that are not directly observed. It is therefore cleaner to compare the contribution to the total star formation rate density at each epoch as a function of stellar mass. None of the models provides a perfect description of the data. At low redshift, the distribution is broader in the B8W7 model than is the case for the SW or MS models. Observed contribution from low-mass galaxies is better described by B8W7, but the model contains too few high star formation rate galaxies at low redshift. At $z \sim 2$, the situation is reversed, with all the models showing a similar contribution from low-mass galaxies, and the primary difference being the paucity of high stellar mass galaxies in the MS and SW models.

(vi) The differences between the models are emphasized by comparing the mass fractions of baryons in different phases as a function of halo mass. The conventional model, B8W7, assumes that haloes less massive than $10^{11.5}\,h^{-1}\,\mathrm{M}_\odot$ are baryonically closed. Almost all of the baryons in low-mass haloes are placed in cool clouds in the halo (because of the short halo cooling time). In contrast, the SW and MS models expel much of their baryonic content from the halo. The fraction of baryons retained in the halo has a maximum of 80 per cent at $M_{\rm halo} = 10^{11.7}\,h^{-1}\,\mathrm{M}_\odot$ and drops rapidly to a minimum of $\sim$10 per cent at $M_{\rm halo} = 10^{11}\,h^{-1}\,\mathrm{M}_\odot$. At the minimum, the majority of baryons are predominantly stars and cold gas. At masses above $10^{12}\,h^{-1}\,\mathrm{M}_\odot$, the 'hot-halo' mode of feedback takes over expelling the halo material, resulting in low baryon fractions as discussed in Bow08, and in good agreement with X-ray observations of groups and clusters. The halo baryon fractions of the MS and SW models offer a good explanation of the low abundance of H I clouds around M31 (Sancisi et al. 2008). Further work is required to compare the model predictions to such data and to the circumgalactic gas haloes inferred from absorption line systems at low redshift (Tumlinson et al. 2011) since the ionization state of the extragalactic gas must be carefully computed and combined with X-ray limits on the emission from galactic haloes (Crain et al. 2010). In all of these models, the recycling of previously ejected gas plays an important role.

Another potential discriminant between the models is the distribution of metals in the ISM. All the models predict that a major fraction of the metals produced by stars will reside outside galaxies. In the absence of AGN feedback, the B8W7 model predicts that the metals will be confined within galaxy haloes, while it will be more widespread in the expulsion models (SW and MS). We find, however, that at $z < 2$ the major contributor of extragalactic metals is the AGN powered hot-halo feedback that expels diffuse gas from galaxy groups. This material is highly enriched compared to the winds from low-mass galaxies: combined with the high fraction of the total baryon budget that has been expelled from groups, this component dominates the diffuse metal content of the low-redshift Universe. At $z > 3$, AGN feedback makes little contribution to the intergalactic fraction of metals and the observed widespread distribution of metals at this epoch favours the expulsion feedback schemes of the MS and SW models. However, the nature of our phenomenological model makes it difficult to predict the dispersal of the expelled material, and it is unclear how far from the parent galaxy groups the metals will be spread (Booth et al. 2012). Careful consideration of the dynamics of the outflow is needed to make a meaningful comparison with observational data on metal absorption lines.

In summary, although we have introduced a feedback scheme that reproduces the results of hydrodynamical simulations well, we find that the original wind scheme of Bow06 produces a better match to observational data on the stellar mass content of haloes, the SSFRs of galaxies and the evolution of these quantities. Comparison of the models with the observational data highlights several aspects of the observational data that are not consistent with any of the models considered. First, none of the models reproduces the rapid drop in the normalization of the observed SMF between $z = 0$ and $z = 1$–2. Secondly, although the B8W7 model reproduces the observed dependence of the SSFR on stellar mass better than the SW or MS model, none reproduces the observed tendency for lower mass galaxies to have higher SSFR than their massive counterparts. Both of these trends suggest that the effective wind speed (relative to the halo escape velocity) should increase with redshift. One possibility is that this might arise because of the more clumpy concentration of star-forming regions in high-redshift galaxies. This is an interesting possibility for further investigation. There is also no intrinsic reason why galaxy winds should be a simple power law, and the observational data suggest we should explore feedback schemes in which the exponent of the mass dependence varies with system. For example, the problem of the SW and MS models overproducing galaxies below $10^9\,h^{-1}\,\mathrm{M}_\odot$ could be solved by introducing a strong mass dependence to the feedback below this limit. Rather than introducing phenomenological modifications to the feedback scheme, an alternative approach would be to simulate idealized galaxies at high (10 pc) or ultra-high (0.1pc) resolution and to extract suitable parametrizations for the outflow (Hopkins, Quataert & Murray 2011; Stringer et al. 2011; Creasey et al., in preparation).

In future, we will combine the expulsion mode of feedback considered here with developments of the GALFORM code (Lagos et al. 2011) to more accurately trace the cold gas content of the universe. We will also make a one-to-one comparison of the formation histories of galaxies from GIMIC, and the forthcoming suite of EAGLEsimulations, with the GALFORM code. Our aim is *not only* to develop the GALFORM model as an emulator of numerical simulations (e.g. Bower et al. 2010) but also to use it as a fundamental tool for understanding the key components of successful galaxy formation models. Only by representing the evolution of galaxies as a simple set of coupled differential equations can we claim to have understood the galaxy formation problem, and to have separated the key processes from the details.

ACKNOWLEDGMENTS

We thank our collaborators who have helped to develop the GALFORM project, Shaun Cole, Carlton Baugh, Cedric Lacey, Carlos Frenk, Claudia Lagos and Nikos Fanidakis. We thank Ivan Baldry for allowing us access to pre-release data on the galaxy mass function from the GAMA survey, and thank the anonymous referee for thoughtful comments. AJB acknowledges the support of the

Gordon & Betty Moore Foundation, RAC is supported by the Australian Research Council via a Discovery Project grant and RGB thanks the STFC for support through the rolling grant scheme.

REFERENCES

Aguirre A., Dow-Hygelund C., Schaye J., Theuns T., 2008, ApJ, 689, 851
Baldry I. K. et al., 2011, preprint (arXiv:1111.5707)
Balogh M. L., Pearce F. R., Bower R. G., Kay S. T., 2001, MNRAS, 326, 1228
Balogh M. L., McCarthy I. G., Bower R. G., Eke V. R., 2008, MNRAS, 385, 1003
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289
Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, MNRAS, 320, 261
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
Benson A. J., Sugiyama N., Nusser A., Lacey C. G., 2006, MNRAS, 369, 1055
Bertone S., Stoehr F., White S. D. M., 2005, MNRAS, 359, 1201
Best P. N., von der Linden A., Kauffmann G., Heckman T. M., Kaiser C. R., 2007, MNRAS, 379, 894
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Booth C. M., Schaye J., Delgado J. D., Dalla Vecchia C., 2012, MNRAS, 420, 1053
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645 (Bow06)
Bower R. G., McCarthy I. G., BensonA. J., 2008, MNRAS, 390, 1399 (Bow08)
Bower R. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, MNRAS, 407, 2017
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Brooks A. M., Governato F., Booth C. M., Willman B., Gardner J. P., Wadsley J., Stinson G., Quinn T., 2007, ApJ, 655, L17
Bundy K., Ellis R. S., Conselice C. J., 2005, ApJ, 625, 621
Cai Y.-C., Angulo R. E., Baugh C. M., Cole S., Frenk C. S., Jenkins A., 2009, MNRAS, 395, 1185
Cen R., Chisari N. E., 2011, ApJ, 731, 11
Chen Y.-M., Tremonti C. A., Heckman T. M., Kauffmann G., Weiner B. J., Brinchmann J., Wang J., 2010, AJ, 140, 445
Cirasuolo M., McLure R. J., Dunlop J. S., Almaini O., Foucaud S., Simpson C., 2010, MNRAS, 401, 1166
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Crain R. A., Eke V. R., Frenk C. S., Jenkins A., McCarthy I. G., Navarro J. F., Pearce F. R., 2007, MNRAS, 377, 41
Crain R. A. et al., 2009, MNRAS, 399, 1773
Crain R. A., McCarthy I. G., Frenk C. S., Theuns T., Schaye J., 2010, MNRAS, 407, 1403
Croton D. J. et al., 2006, MNRAS, 365, 11
Cucciati O. et al., 2011, preprint (arXiv:1109.1005)
Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431
Davé R., Oppenheimer B. D., Finlator K., 2011, MNRAS, 415, 1
de Avillez M. A., Breitschwerdt D., 2007, ApJ, 665, L35
De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
De Lucia G., Boylan-Kolchin M., Benson A. J., Fontanot F., Monaco P., 2010, MNRAS, 406, 1533
Deason A. J. et al., 2011, MNRAS, 415, 2607
Dekel A., Silk J., 1986, ApJ, 303, 39
Drory N., Salvato M., Gabasch A., Bender R., Hopp U., Feulner G., Pannella M., 2005, ApJ, 619, L131
Efstathiou G., 2000, MNRAS, 317, 697
Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, MNRAS, 410, 53
Font A. S. et al., 2008, MNRAS, 389, 1619
Font A. S., McCarthy I. G., Crain R. A., Theuns T., Schaye J., Wiersma R. P. C., Dalla Vecchia C., 2011a, MNRAS, 416, 2802
Font A. S. et al., 2011b, MNRAS, 417, 1206
Frank S., Mathur S., Pieri M., York D. G., 2010, AJ, 140, 835
Gabor J. M., Davé R., Oppenheimer B. D., Finlator K., 2011, MNRAS, 417, 2676
Genzel R. et al., 2011, ApJ, 733, 101
Gilbank D. G. et al., 2011, MNRAS, 414, 304
Giodini S. et al., 2009, ApJ, 703, 982
Gonzalez A. H., Zaritsky D., Zabludoff A. I., 2007, ApJ, 666, 147
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Guo Q. et al., 2011, MNRAS, 413, 101
Hatton S., Devriendt J. E. G., Ninin S., Bouchet F. R., Guiderdoni B., Vibert D., 2003, MNRAS, 343, 75
Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., Pearce F. R., 2003, MNRAS, 338, 913
Hopkins A. M., 2004, ApJ, 615, 209
Hopkins P. F., Elvis M., 2010, MNRAS, 401, 7
Hopkins P. F., Hernquist L., Cox T. J., Robertson B., Springel V., 2006, ApJS, 163, 50
Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950
Jaacks J., Choi J.-H., Nagamine K., Thompson R., Varghese S., 2012, MNRAS, 420, 1606
Jones T. A., Swinbank A. M., Ellis R. S., Richard J., Stark D. P., 2010, MNRAS, 404, 1247
Karim A. et al., 2011, ApJ, 730, 61
Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
Kauffmann G., White S. D. M., Guiderdoni B., 1993, MNRAS, 264, 201
Kereš D., Katz N., Davé R., Fardal M., Weinberg D. H., 2009, MNRAS, 396, 2332
Komatsu E. et al., 2011, ApJS, 192, 18
Lagos C. d. P., Lacey C. G., Baugh C. M., Bower R. G., Benson A. J., 2011, MNRAS, 416, 1566
Leauthaud A. et al., 2012, ApJ, 746, 95
Li C., White S. D. M., 2009, MNRAS, 398, 2177
Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, ApJ, 701, 1765
Martin C. L., 2005, ApJ, 621, 227
McGee S. L., Balogh M. L., 2010, MNRAS, 403, L79
McGee S. L., Balogh M. L., Wilman D. J., Bower R. G., Mulchaey J. S., Parker L. C., Oemler A., 2011, MNRAS, 413, 996
McKee C. F., Ostriker J. P., 1977, ApJ, 218, 148
Mortlock A., Conselice C. J., Bluck A. F. L., Bauer A. E., Grützbauch R., Buitrago F., Ownsworth J., 2011, MNRAS, 413, 2845
Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569
Okamoto T., Nemmen R. S., Bower R. G., 2008a, MNRAS, 385, 161
Okamoto T., Gao L., Theuns T., 2008b, MNRAS, 390, 920
Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265
Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577 (Op08)
Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, MNRAS, 406, 2325
Pozzetti L. et al., 2010, A&A, 523, A13
Prochaska J. X., Weiner B., Chen H., Mulchaey J. S., Cooksey K. L., 2011, ApJ, 740, 91
Rodighiero G. et al., 2010, A&A, 518, L25
Rubin K. H. R., Prochaska J. X., Ménard B., Murray N., Kasen D., Koo D. C., Phillips A. C., 2011, ApJ, 728, 55
Sancisi R., Fraternali F., Oosterloo T., van der Hulst T., 2008, A&AR, 15, 189
Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210
Schaye J. et al., 2010, MNRAS, 402, 1536
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776

Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, ApJ, 736, 160
Stringer M. J., Bower R. G., Cole S., Frenk C. S., Theuns T., 2011, preprint (arXiv:1111.2529)
Tumlinson J. et al., 2011, Sci, 334, 948
van de Voort F., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 2782
Weiner B. J. et al., 2009, ApJ, 692, 187
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Wiersma R. P. C., Schaye J., Smith B. D., 2009a, MNRAS, 393, 99
Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574
Zibetti S., White S. D. M., Schneider D. P., Brinkmann J., 2005, MNRAS, 358, 949

This paper has been typeset from a TEX/LATEX file prepared by the author.