

The physics driving the cosmic star formation history

Joop Schaye,^{1*} Claudio Dalla Vecchia,¹ C. M. Booth,¹ Robert P. C. Wiersma,¹
Tom Theuns,^{2,3} Marcel R. Haas,¹ Serena Bertone,⁴ Alan R. Duffy,^{1,5} I. G. McCarthy⁶
and Freeke van de Voort¹

¹*Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, the Netherlands*

²*Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE*

³*Department of Physics, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerpen, Belgium*

⁴*Santa Cruz Institute for Particle Physics, University of California at Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA*

⁵*Jodrell Bank Centre for Astrophysics, Alan Turing Building, The University of Manchester, Manchester M13 9PL*

⁶*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA*

Accepted 2009 November 10. Received 2009 November 10; in original form 2009 September 28

ABSTRACT

We investigate the physics driving the cosmic star formation (SF) history using the more than 50 large, cosmological, hydrodynamical simulations that together comprise the Overwhelmingly Large Simulations project. We systematically vary the parameters of the model to determine which physical processes are dominant and which aspects of the model are robust. Generically, we find that SF is limited by the build-up of dark matter haloes at high redshift, reaches a broad maximum at intermediate redshift and then decreases as it is quenched by lower cooling rates in hotter and lower density gas, gas exhaustion and self-regulated feedback from stars and black holes. The higher redshift SF is therefore mostly determined by the cosmological parameters and to a lesser extent by photoheating from reionization. The location and height of the peak in the SF history, and the steepness of the decline towards the present, depend on the physics and implementation of stellar and black hole feedback. Mass loss from intermediate-mass stars and metal-line cooling both boost the SF rate at late times. Galaxies form stars in a self-regulated fashion at a rate controlled by the balance between, on the one hand, feedback from massive stars and black holes and, on the other hand, gas cooling and accretion. Paradoxically, the SF rate is highly insensitive to the assumed SF law. This can be understood in terms of self-regulation: if the SF efficiency is changed, then galaxies adjust their gas fractions so as to achieve the same rate of production of massive stars. Self-regulated feedback from accreting black holes is required to match the steep decline in the observed SF rate below redshift 2, although more extreme feedback from SF, for example in the form of a top-heavy initial stellar mass function at high gas pressures, can help.

Key words: stars: formation – galaxies: evolution – galaxies: formation – cosmology: theory.

1 INTRODUCTION

The cosmic history of star formation (SF) is one of the most fundamental observables of our Universe. Measuring the global star formation rate (SFR) density as a function of redshift has therefore long been one of the primary goals of observational astronomy (e.g. Lilly et al. 1996; Madau et al. 1996; Steidel et al. 1999; Ouchi et al. 2004; Arnouts et al. 2005; Schiminovich et al. 2005; Hopkins & Beacom 2006; Bouwens et al. 2007).

Modelling the cosmic star formation history (SFH) is not an easy task because it depends on a complex interplay of physical processes

and because a large range of halo masses contributes. To predict the SFH within the context of the cold dark matter cosmology, one must first get the dark matter halo mass function right. These days this is the easier part, as the cosmological parameters are relatively well constrained. Then one must model the rate at which gas accretes, cools, collapses and turns into stars. Even if one does not attempt to model the cold, interstellar gas phase and uses empirical SF laws to estimate the rate at which gas is converted into stars on kpc scales, there are a host of feedback processes that need to be taken into account. Stars produce radiation which can heat gas, exert radiation pressure and change its ionization balance and hence its cooling rate. Massive stars explode as supernovae (SNe) which can drive both small-scale turbulence and large-scale outflows. Stars also produce heavy elements and dust which change the rate at which gas cools.

*E-mail: schaye@strw.leidenuniv.nl

Accretion on to supermassive black holes (BHs) in the centres of galaxies also results in radiative, thermal and mechanical feedback. Finally, magnetic fields and cosmic rays may be important.

Despite this complexity, many authors have used semi-analytic models (e.g. White & Frenk 1991; Hernquist & Springel 2003; Bower et al. 2006; Croton et al. 2006; Somerville et al. 2008; Fontanot et al. 2009) or hydrodynamical simulations (e.g. Nagamine, Cen & Ostriker 2000; Pearce et al. 2001; Ascasibar et al. 2002; Murali et al. 2002; Sommer-Larsen, Götz & Portinari 2003; Springel & Hernquist 2003b; Kereš et al. 2005; Nagamine et al. 2006; Di Matteo et al. 2008; Oppenheimer & Davé 2008; Choi & Nagamine 2009; Crain et al. 2009; Booth & Schaye 2009a) to study the SFH of the Universe.

Modern semi-analytic models use cosmological dark matter simulations to generate halo mass functions and merger trees. These are then combined with simple prescriptions for the baryonic physics, which include a large number of free parameters, to predict galaxy SFRs. The parameters of the models are tuned to match particular sets of observations, after which the model is used to make predictions for other properties. While this approach has been very productive, it can be difficult to understand what physics is driving the results and there is a danger that model predictions may be correct even when the underlying prescriptions do not reflect the real world (because they were tuned to match particular observations). Furthermore, current semi-analytic models are only of limited use for the study of the gas around and in between galaxies.

While hydrodynamical simulations attempt to model much more physics from first principles than semi-analytic models, they do make extensive use of subgrid prescriptions for the physics that they cannot resolve directly. Some of these prescriptions are physically motivated and relatively well understood. One example is the radiative cooling rate which is determined by atomic physics. However, even here there are significant uncertainties such as the effects of photoionization, non-equilibrium, dust and relative abundance variations. Other subgrid prescriptions, such as the stellar initial mass function (IMF), are empirically motivated. Most subgrid models contain, however, a mixture of physical and empirical ingredients. For example, prescriptions for SF, thermal/kinetic feedback from stars and active galactic nuclei (AGN) and for the growth of BHs typically use physically motivated frameworks whose parameters are calibrated using observations.

When it comes to predictions for SFRs in galaxies, the subgrid models used in cosmological hydrosimulations play a critical role. Differences in the subgrid prescriptions used by different authors are likely to be much more important than differences in the codes used to model gravity and hydrodynamics. The crucial role played by subgrid models implies that hydrodynamical simulations have some of the same weaknesses as semi-analytic models. On the other hand, direct simulation enables one to probe much deeper, e.g. by following gas flows in three dimensions, or by studying the interactions of galaxies with their environment and the intergalactic medium (IGM).

While simulations offer the tantalizing possibility of laboratory-like control, the huge dynamic range required makes it computationally very expensive to explore parameter space. The Overwhelmingly Large Simulations (OWLS) project, which was made possible by the temporary availability of the supercomputer that serves as the correlator for the Low Frequency Array (LOFAR) telescope, aims to use the potential of simulations to gain insight into the physics that determines the formation of galaxies and the evolution of the IGM. The OWLS project consists of a large suite of cosmological, smoothed particle hydrodynamics (SPH) simulations with varying

box sizes and resolutions. Each production simulation uses 2×512^3 particles, which places them among the largest dissipative simulation ever performed. The real power of the project stems, however, not from the size of the simulations, but from the fact that they are repeated many (more than 50) times, each time varying a subgrid prescription, most of which were newly developed for this project. Although we have not attempted to fine-tune the subgrid parameters to match particular data sets, we do hope that our investigations will help trigger future work in directions that will improve agreement with the observations.

While the OWLS project aims to study a variety of problems in astrophysical cosmology, we will limit ourselves here to the cosmic SFH. To further limit the scope of the paper, we will postpone a discussion of the SFR as a function of halo mass to another paper (Haas et al., in preparation). We note, however, that this latter function is in some ways physically more interesting than the SFH. This is because the cosmic SFR can be thought of as a convolution of the halo mass function, which depends only on cosmology and redshift, and the SFR as a function of halo mass and redshift, which depends on poorly understood astrophysics.

We will also not compare to observations of the build-up of the cosmic stellar mass density. The cumulative SFR must of course equal the stellar mass (after taking stellar mass loss into account). While this will be true by construction for *ab initio* models, it is not necessarily the case for observationally inferred quantities. This is because observational probes of SF measure only the rate of formation of massive stars, since those dominate the electromagnetic output of young stellar populations. To obtain the total SFR, it is necessary to extrapolate to low masses using an assumed IMF. Since the IMF is uncertain and may not be universal, measurements of the total stellar mass are of great interest. However, it should be noted that massive stars are the ones that matter from a cosmological perspective because they dominate the chemical, radiative and mechanical feedback processes that regulate the formation of stars and galaxies. As the objective of this paper is to explore the effects of varying physical prescriptions rather than to fit the observations, we will for simplicity limit ourselves to comparisons with measurements of the evolution of the SFR density. We emphasize, however, that it should be kept in mind that these are often inconsistent with observations of the stellar mass density (e.g. Wilkins, Trentham & Hopkins 2008; Cowie & Barger 2008, but see also Reddy et al. 2008; Reddy & Steidel 2009) and subject to large systematic uncertainties (e.g. Conroy, Gunn & White 2009).

This paper is organized as follows. In Section 2 we introduce the numerical techniques that are common to all OWLS runs, while Sections 3 and 4 describe the physics modules employed in the reference model and the other OWLS runs, respectively. As this paper also serves to introduce the OWLS project, we will discuss the ingredients of the different simulations in some detail. The SFHs predicted by the models are presented and discussed in the sections that describe them. Finally, we discuss and summarize our main findings in Section 5.

2 SIMULATIONS

In this section, we describe the numerical techniques that are common to all simulations that make up the OWLS project. The modifications and additions to the subgrid modules are discussed in Section 3 for our reference model and in Section 4 for the other runs.

The simulations were performed with a significantly extended version of the N -body Tree-PM, SPH code GADGET3 (last described

Table 1. List of simulations run assuming the reference model.

Simulation	L (h^{-1} Mpc)	N	m_b ($h^{-1} M_\odot$)	m_{dm} ($h^{-1} M_\odot$)	ϵ_{com} (h^{-1} kpc)	ϵ_{prop} (h^{-1} kpc)	z_{end}
<i>REF_L006N128</i>	6.25	128^3	1.4×10^6	6.3×10^6	1.95	0.50	2
<i>REF_L012N256</i>	12.50	256^3	1.4×10^6	6.3×10^6	1.95	0.50	2
<i>REF_L012N512</i>	12.50	512^3	1.7×10^5	7.9×10^5	0.98	0.25	2
<i>REF_L025N128</i>	25.00	128^3	8.7×10^7	4.1×10^8	7.81	2.00	0
<i>REF_L025N256</i>	25.00	256^3	1.1×10^7	5.1×10^7	3.91	1.00	2
<i>REF_L025N512</i>	25.00	512^3	1.4×10^6	6.3×10^6	1.95	0.50	1.45
<i>REF_L050N256</i>	50.00	256^3	8.7×10^7	4.1×10^8	7.81	2.00	0
<i>REF_L050N512</i>	50.00	512^3	1.1×10^7	5.1×10^7	3.91	1.00	0
<i>REF_L100N128</i>	100.00	128^3	5.5×10^9	2.6×10^{10}	31.25	8.00	0
<i>REF_L100N256</i>	100.00	256^3	6.9×10^8	3.2×10^9	15.62	4.00	0
<i>REF_L100N512</i>	100.00	512^3	8.7×10^7	4.1×10^8	7.81	2.00	0

Note. From left to right, the columns show simulation identifier, comoving box size, number of dark matter particles (there are equally many baryonic particles), baryonic particle mass, dark matter particle mass, comoving (Plummer-equivalent) gravitational softening, maximum physical softening and final redshift.

in Springel 2005), a Lagrangian code used to calculate gravitational and hydrodynamical forces on a system of particles. Most of the simulations were run in periodic boxes of size $L = 25$ and 100 comoving h^{-1} Mpc and each of the production runs uses 512^3 dark matter and equally many baryonic particles (representing either a collisionless star or collisional gas particles). The particle masses in the 2×512^3 particle $25 h^{-1}$ Mpc ($100 h^{-1}$ Mpc) box are $6.34 \times 10^6 h^{-1} M_\odot$ ($4.06 \times 10^8 h^{-1} M_\odot$) for dark matter and $1.35 \times 10^6 h^{-1} M_\odot$ ($8.66 \times 10^7 h^{-1} M_\odot$) for baryons. Note, however, that baryonic particle masses change during the course of the simulation due to mass transfer from star to gas particles. The $25 h^{-1}$ Mpc simulation volumes were run as far as redshift 2 and the $100 h^{-1}$ Mpc volumes were run to redshift 0. Comoving gravitational softenings were set to $1/25$ of the initial mean inter-particle spacing but were limited to a maximum physical scale of $0.5 \text{ kpc } h^{-1}$ ($2 \text{ kpc } h^{-1}$) for the $25 h^{-1}$ Mpc ($100 h^{-1}$ Mpc) simulations. The switch from a fixed comoving to a fixed proper softening happens at $z = 2.91$ in all simulations. We used $N_{\text{ngb}} = 48$ neighbours for the SPH interpolation.

In order to assess the effects of the finite resolution and box size on our results, we have run a suite of cosmological simulations, all using the same physical model, using different box sizes (ranging from 6.25 to $100 h^{-1}$ Mpc) and particle numbers (ranging from 128^3 to 512^3). The particle masses and gravitational softenings for each of these simulations are listed in Table 1.

The initial conditions were generated with CMBFAST (version 4.1; Seljak & Zaldarriaga 1996) and evolved to redshift $z = 127$, where the simulations were started, using the Zel'Dovich (1970) approximation from an initial glass-like state (White 1996).

Before discussing each of the variations made to the subgrid models in Section 4, we first turn our attention to the physics included in the reference model.

3 THE REFERENCE MODEL

3.1 Description of the model

When simulations lack the required numerical resolution or the physics to accurately model a physical process, we must resort to subgrid models. In this section, we describe the ‘reference’ physical model (*REF*) that is used as a base from which all further investigations of the behaviour of the simulations can be launched. As

discussed in Section 4, we will do this by varying one physical process at a time and comparing the resulting SFH to the one predicted by the reference model.

We emphasize that the *REF* model should not be viewed as our ‘best’ model. As its name implies, it functions as a reference point for our systematic variation of the parameters. In fact, we will show in future papers that some variations yield much better agreement with particular types of observations. For example, the inclusion of AGN feedback dramatically improves the agreement with observations of groups of galaxies at redshift 0 (McCarthy et al. 2009).

To illustrate the dynamic range in the simulations, Figs 1 and 2 show two factors of 10 zooms into the tenth most massive haloes in, respectively, the *REF_L025N512* run at $z = 2$ and model *REF_L100N512* at $z = 0$. Note, however, that the left-hand panels still show only a small fraction of the simulation volume.

3.1.1 Cosmology

We assume the values for the cosmological parameters derived from the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) 3-year results (Spergel et al. 2007), $\{\Omega_m, \Omega_b, \Omega_\Lambda, \sigma_8, n_s, h\} = \{0.238, 0.0418, 0.762, 0.74, 0.951, 0.73\}$, which are consistent with the *WMAP* 5-year data (Komatsu et al. 2009).¹ The primordial baryonic mass fraction of helium is assumed to be 0.248.

3.1.2 Radiative cooling and heating

Radiative cooling is central to simulations of the formation of galaxies as it enables baryons to dissipate their binding energy which allows their collapse to proceed within virialized structures. Photo-heating by the ionizing background radiation also plays a key role because it strongly increases pressure forces in low-density gas, thereby smoothing out small-scale baryonic structures.

Previous cosmological simulations have typically included radiative cooling assuming primordial abundances (e.g. Springel & Hernquist 2003b). Some recent studies have included metal-line cooling (e.g. Scannapieco et al. 2005; Romeo et al. 2006;

¹The most notable difference is in σ_8 , which is 1.6σ lower in *WMAP3* than in *WMAP5*.

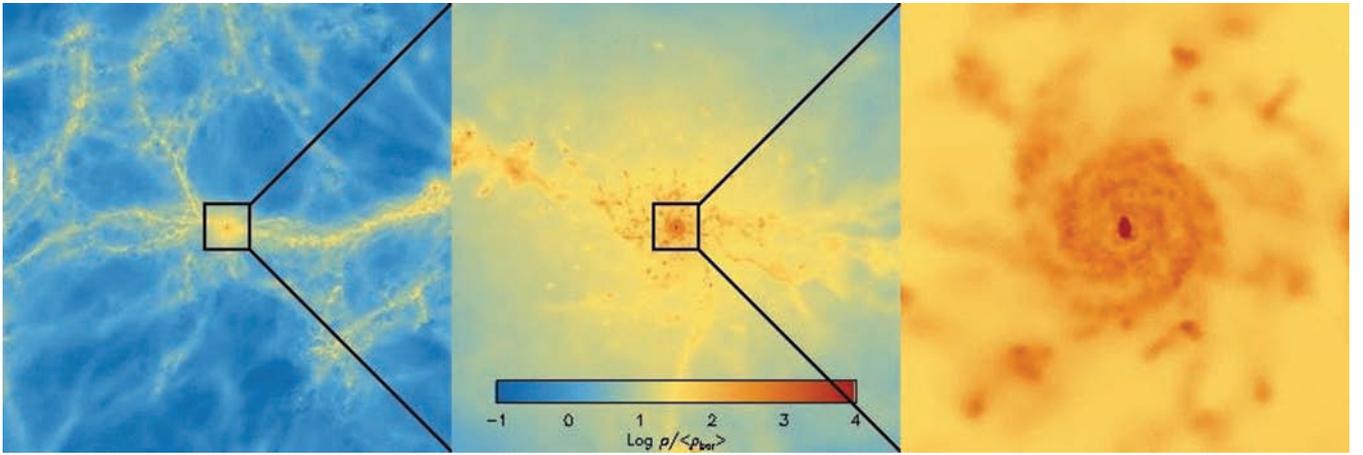


Figure 1. Zoom into a $M_{200} = 10^{12.2} M_{\odot}$ halo at $z = 2$ in the *REF_L025N512* simulation. From left to right, the images are 10 , 1 and $0.1 h^{-1}$ Mpc on a side. All slices are $1 h^{-1}$ Mpc thick. Note that the first image shows only a fraction of the total simulation volume, which is cubic and $25 h^{-1}$ Mpc on a side. The colour coding shows the projected gas density, $\log_{10} \rho / (\rho_b)$, and the colour scale ranges from -1 to 4 (which is lower than the true maximum of the image). The coordinate axes were rotated to show the galaxy face-on. This halo is the tenth most massive in the simulation. About half of the haloes in this mass range host extended disc galaxies, while the other half have highly disturbed morphologies due to ongoing mergers.

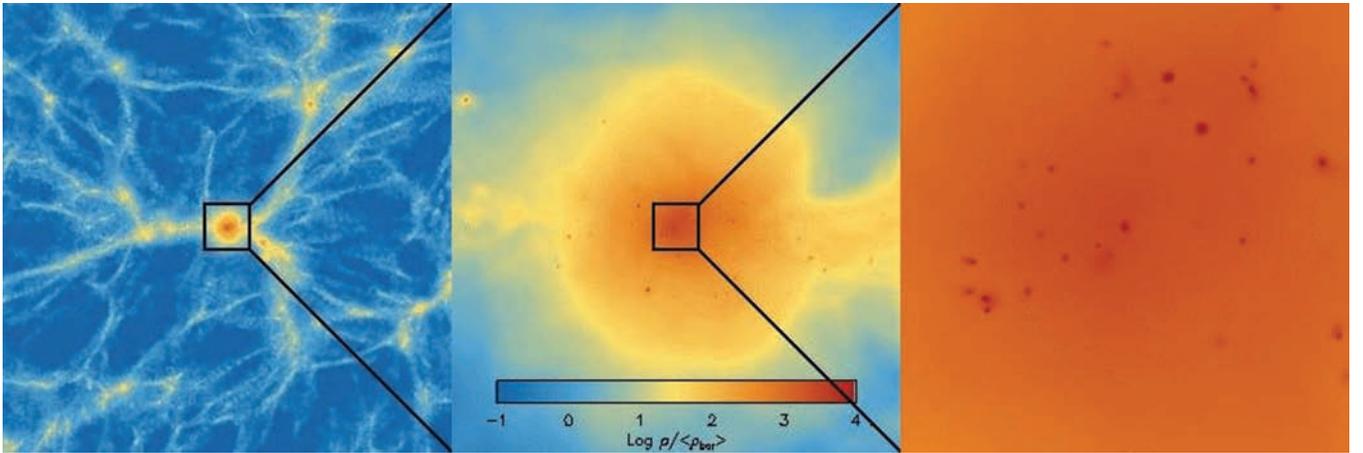


Figure 2. Zoom into a $M_{200} = 10^{14.2} M_{\odot}$ halo at $z = 0$ in the *REF_L100N512* simulation. From left to right, the images are 40 , 4 and $0.4 h^{-1}$ Mpc on a side. All slices are $1 h^{-1}$ Mpc thick. Note that the first image shows only a fraction of the total simulation volume, which is cubic and $100 h^{-1}$ Mpc on a side. The colour coding shows the projected gas density, $\log_{10} \rho / (\rho_b)$, and the colour scale ranges from -1 to 4 (which is lower than the true maximum of the image). This halo is the tenth most massive in the simulation.

Oppenheimer & Davé 2006; Tornatore et al. 2007; Choi & Nagamine 2009), but under the assumption of collisional ionization equilibrium and fixed relative abundances. Photoionization by the UV background radiation does not only provide a source of heat, it also reduces the cooling rates for both primordial and metal-enriched plasmas (Efstathiou 1992; Wiersma, Schaye & Smith 2009a). Wiersma et al. (2009a) emphasized the importance of including this effect as well as variations in the relative abundances of the elements.

We implemented radiative cooling and heating using the method and tables of Wiersma et al. (2009a).² In brief, net radiative cooling rates are computed element by element in the presence of the cosmic microwave background and the Haardt & Madau (2001) model for the UV and X-ray background radiation from quasars and galax-

ies. Hence, variations in relative abundances and photoionization of heavy elements are both taken into account. The contributions of the 11 elements, hydrogen, helium, carbon, nitrogen, oxygen, neon, magnesium, silicon, sulphur, calcium and iron, are interpolated as a function of density, temperature and redshift from tables that have been pre-computed using the publicly available photoionization package *CLOUDY*, last described by Ferland et al. (1998), assuming the gas to be optically thin and in (photo)ionization equilibrium.

The simulations model hydrogen reionization by ‘switching on’ the Haardt & Madau (2001) background at $z = 9$. Prior to reionization, the cooling rates are computed in the presence of the cosmic microwave background and a photodissociating background which we obtain by cutting off the $z = 9$ Haardt & Madau (2001) spectrum at 1 Ryd. Note that the presence of a photodissociating background suppresses H_2 cooling at all redshifts. Reionization has the effect of rapidly heating all of the gas to temperatures of $\sim 10^4$ K. The assumption that the gas is optically thin is likely to lead to an

²We used their equation (3) rather than (4) and *CLOUDY* version 05.07 rather than 07.02.

underestimate of the gas temperature shortly after reionization (e.g. Abel & Haehnelt 1999). For the case of helium reionization, we correct for this effect by heating the gas by a total amount of 2 eV per atom. This extra helium reionization heating takes place at a central redshift 3.5, with the heating spread with a Gaussian filter with $\sigma(z) = 0.5$ in redshift. This prescription was chosen to match observations of the temperature history of the IGM (Schaye et al. 2000) as shown in fig. 1 of Wiersma et al. (2009b).

3.1.3 Star formation

Cosmological simulations such as ours miss both the resolution and the physics to model the cold interstellar medium (ISM), let alone the formation of stars within molecular clouds. SF is therefore implemented stochastically by converting gas particles into collisionless star particles, which represent simple (or single) stellar populations. We convert entire particles, because the spawning of multiple star particles per gas particle affects the efficiency of feedback from SF (Dalla Vecchia & Schaye 2008). Hence, the particle number is conserved in our simulations.

Gas with densities exceeding the critical density for the onset of the thermogravitational instability (hydrogen number densities $n_{\text{H}} = 10^{-2}$ – 10^{-1} cm^{-3}) is expected to be multiphase and to form stars (Schaye 2004). We therefore impose an effective equation of state (EOS) with pressure $P \propto \rho^{\gamma_{\text{eff}}}$ for densities³ $n_{\text{H}} > n_{\text{H}}^*$ where $n_{\text{H}}^* = 0.1 \text{ cm}^{-3}$, normalized to $P/k = 1.08 \times 10^3 \text{ cm}^{-3} \text{ K}$ at the threshold. We use $\gamma_{\text{eff}} = 4/3$, for which both the Jeans mass and the ratio of the Jeans length to the SPH kernel are independent of the density, thus preventing spurious fragmentation due to a lack of numerical resolution (Robertson & Kravtsov 2008; Schaye & Dalla Vecchia 2008). Only gas on the EOS is allowed to form stars. Schaye & Dalla Vecchia (2008) demonstrated that our choice of threshold reproduces the threshold surface density for $\text{H}\alpha$ emission from SF that is observed in nearby galaxies.

Previous cosmological simulations used Schmidt-type (i.e. power laws of the volume density) SF laws and tuned one or more free parameters to fit the observed Kennicutt–Schmidt SF law, which is a surface density law. This approach is unsatisfactory as the parameters would really need to be retuned if disc scaleheights change as a result of varying abundances (and hence cooling rates), SN feedback or changes in the assumed EOS of the ISM.

Schaye & Dalla Vecchia (2008) showed that because the surface density in a self-gravitating system is directly related to the pressure, the Kennicutt–Schmidt SF law can be rewritten as a pressure law. This enables one to reproduce arbitrary input Kennicutt–Schmidt laws independently of the assumed EOS. Moreover, because the parameters are observables, no tuning is required. Following Schaye & Dalla Vecchia (2008), we thus compute the SFR for star-forming gas particles using

$$\dot{m}_* = m_{\text{g}} A (1 \text{ M}_{\odot} \text{ pc}^{-2})^{-n} \left(\frac{\gamma}{G} f_{\text{g}} P \right)^{(n-1)/2}, \quad (1)$$

where m_{g} is the mass of the gas particle, $\gamma = 5/3$ is the ratio of specific heats (not to be confused with the effective EOS imposed on to the ISM), f_{g} is the mass fraction in gas (which we assume to be unity) and P is the total pressure. The parameters A and n are,

respectively, the amplitude and slope of the observed Kennicutt (1998) law, $\dot{\Sigma}_* = A (\Sigma_{\text{g}}/1 \text{ M}_{\odot} \text{ pc}^{-2})^n$ with $A = 1.515 \times 10^{-4} \text{ M}_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$ and $n = 1.4$. The amplitude of this relation has been renormalized by a factor⁴ of 1/1.65 to account for the fact that the original analysis of Kennicutt (1998) assumed the Salpeter (1955) IMF whereas we use the Chabrier (2003) IMF.

3.1.4 Stellar evolution and chemodynamics

Our implementation of stellar evolution and chemical enrichment is discussed in detail in Wiersma et al. (2009b). Here, we will provide only a brief summary.

Each star particle represents a single stellar population that is specified by its initial mass, age and its chemical composition (which it inherits from its progenitor gas particle). We follow the timed release, by both massive stars (Type II SNe and stellar winds) and intermediate-mass stars [Type Ia SNe and asymptotic giant branch (AGB) stars], of all 11 elements that contribute significantly to the radiative cooling rates. During each time-step, star particles distribute the mass they eject over their neighbouring gas particles⁵ using the SPH interpolation scheme.

We assume a Chabrier (2003) IMF spanning the range of 0.1–100 M_{\odot} and use the metallicity-dependent stellar lifetimes of Portinari, Chiosi & Bressan (1998) and the complete set of nucleosynthetic yields of Marigo (2001) and Portinari et al. (1998) along with the SN Type Ia (SN Ia) yields of the W7 model of Thielemann et al. (2003). Since SN Ia are thought to result from binary evolution, a single stellar population will produce SN Ia over an extended period. We implement the release of mass and energy (which we inject in thermal form) by SN Ia using empirically derived rates normalized to the observed cosmic SN Ia rate (see fig. A6 of Wiersma et al. 2009b). For reference, our assumed rate implies that a fraction of about 0.025 of stars with initial mass between 3 and 8 M_{\odot} end their lives as SNe Ia.

For the purpose of both radiative cooling and stellar evolution, we define the abundance of a particular element as the ratio of the SPH estimates of its mass density to the total gas mass density. Wiersma et al. (2009b) showed that the use of such ‘smoothed abundances’ for the cooling rates significantly increases the SFR compared to the standard approach of using ‘particle metallicities’, i.e. the ratio of the elemental mass to the total mass of a particle. While the use of smoothed abundances reduces the effects of the lack of metal mixing inherent to SPH, it does not solve the problem.

3.1.5 Energy feedback from core collapse supernovae

The use of an effective EOS with a pressure that exceeds that of the warm, neutral ISM can be considered a form of weak feedback that reflects the fact that energy injected by massive stars and SNe drives small-scale turbulence. However, as we will show explicitly in Section 4.4, this form of feedback does not lead to a significant suppression of SF. Indeed, observations show that starburst galaxies drive large-scale winds (e.g. Veilleux, Cecil & Bland-Hawthorn

³Gas particles are only placed on the EOS if their temperature is below 10^5 K when they cross the density threshold and if their density exceeds 57.7 times the cosmic mean. These criteria prevent SF in intracluster gas and in intergalactic gas at very high redshift, respectively (Schaye & Dalla Vecchia 2008).

⁴This normalization factor is calculated from the asymptotic ratio (which is reached after only 10^8 yr) of the numbers of ionizing photons predicted from models of stellar populations with a constant SFR (Bruzual & Charlot 2003).

⁵As discussed in Wiersma et al. (2009b), we do not change the entropy of the receiving particles.

2005) which may, over time, eject large amounts of gas and may therefore dramatically reduce the SFR.

As discussed in Dalla Vecchia & Schaye (2008), thermal energy from SNe is quickly radiated away in simulations like ours because the ratio of the heated mass to that of the star particle is too large. In Dalla Vecchia & Schaye (2009) and Section 4.8.3 we show that it is possible to overcome this overcooling problem, without ad hoc suppression of the radiative cooling rates, by decreasing this ratio. However, for our reference model we use the more standard approach of injecting SN energy in kinetic form using the prescription of Dalla Vecchia & Schaye (2008), which is a variation of the recipe of Springel & Hernquist (2003a).

After a short delay of 30 Myr, corresponding to the maximum lifetime of stars that end their lives as core-collapse SNe, newly formed star particles inject kinetic energy into their surroundings by kicking a fraction of their SPH neighbours in random directions. Wind particles are not allowed to form stars for a period of 15 Myr in order to avoid high velocity star particle ejection (a numerical artefact we observed occasionally in high-resolution simulations of isolated galaxies). These time delays are not important for the results presented here.

Each SPH neighbour i of a newly formed star particle j has a probability of $\eta m_j / \sum_{i=1}^{N_{\text{ngb}}} m_i$ of receiving a kick with a velocity v_w . Thus, if all baryonic particles had equal mass, each newly formed star particle would kick, on average, η of its neighbours. Our reference model uses the default parameters of Dalla Vecchia & Schaye (2008), i.e. $\eta = 2$ and $v_w = 600 \text{ km s}^{-1}$. Assuming that each star with initial mass in the range of 6–100 M_{\odot} injects 10^{51} erg of kinetic energy, these parameter values imply that the total wind energy accounts for 40 per cent of the available kinetic energy for our IMF [if we ignore the electron capture SNe predicted by models with convective overshoot (e.g. Chiosi, Bertelli & Bressan 1992) and consider only stars in the mass range of 8–100 M_{\odot} , this works out to be 60 per cent]. The value $\eta = 2$ was chosen in part because it roughly reproduces the peak in the cosmic SFR, as we will show later.

Note that contrary to the widely used kinetic feedback recipe of Springel & Hernquist (2003a), the kinetic energy is injected *locally* and the wind particles are *not* decoupled hydrodynamically. As discussed by Dalla Vecchia & Schaye (2008) and as we will show in Section 4.8.2, these differences have important consequences.

3.1.6 Black hole growth and feedback from AGN

The reference model does not include a prescription for the growth of supermassive BHs and feedback from AGN. However, AGN are included in the OWLS project as an optional model that is switched on for a subset of the simulations (Section 4.1.0).

3.2 Convergence tests

Before trying to interpret the results of numerical simulations, we must check whether they have converged numerically. For cosmological simulations, this means checking whether the box size was sufficiently large and whether the resolution was sufficiently high. To isolate the effects of the size of the simulation volume and the resolution, it is necessary to vary one while holding the other fixed.

To test for numerical convergence, we have run a suite of simulations of the reference model. The main numerical parameters of these runs are listed in Table 1. The simulation names contain strings of the form $L_{xxx}N_{yyy}$, where xxx is the simulation box size

in comoving $h^{-1} \text{ Mpc}$ and yyy is the cube root of the number of particles per species (dark matter or baryonic).

For readers not interested in the details, we first give the conclusions so that they can skip to Section 4. Even our $25 h^{-1} \text{ Mpc}$ boxes are sufficiently large to obtain a converged prediction for the cosmic SFH. The resolution of the $L025N512$ and $L100N512$ runs suffices for redshifts $z < 7$ and $z < 3$, respectively. While the SFR typically increases if the resolution is improved, the situation reverses at low redshift before convergence has been attained.

The discussion below is in parts similar to the one we presented in Wiersma et al. (2009b), where we considered the convergence of the cosmic metal distribution in the reference simulations.

3.2.1 Box size

The results of cosmological simulations may depend on the size of the simulation volume for at least two reasons. First, because the mean density in the box is fixed by the cosmological parameters, the box size determines what types of objects can be sampled. The distribution of density fluctuations can only be modelled correctly on scales that are much smaller than the box. Secondly, because the Fourier modes of the density field only evolve independently in the linear regime, the box must be large compared to the scales on which the density contrast is non-linear. Otherwise the missing power on scales greater than the box size will decrease the power on the scales that are sampled by the simulation.

Fig. 3 shows the SFR per unit comoving volume as a function of redshift for two sets of simulations of the reference model. The solid curves show our two fiducial box sizes and particle numbers: 100 comoving $h^{-1} \text{ Mpc}$ ($REF_L100N512$, black) and 25 comoving $h^{-1} \text{ Mpc}$ ($REF_L025N512$, red), both using 2×512^3 particles. The data points show the observations as compiled by Hopkins & Beacom (2006). To facilitate easy comparisons, we will show these

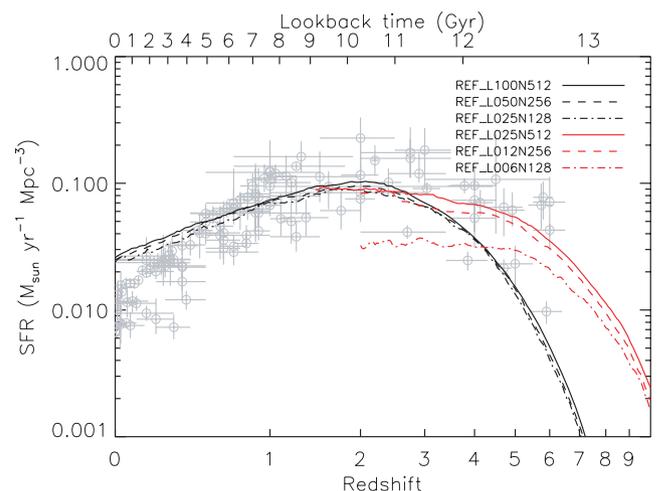


Figure 3. The effect of the box size on the cosmic SFH. The curves show the cosmic SFR density as a function of redshift (bottom x -axis) and lookback time (top x -axis) for different simulations of the reference model. The data points show the compilation of observations from Hopkins & Beacom (2006), converted to our IMF and cosmology. The black curves extending down to $z = 0$ are for simulations that use the same numerical resolution as model $REF_L100N512$. The red curves, which do not continue to $z = 0$, correspond to runs with the resolution of $REF_L025N512$. A $25 h^{-1} \text{ Mpc}$ box is sufficiently large to obtain a converged prediction for the SFH down to $z = 0$.

data points and at least one of the two fiducial runs in all subsequent figures.

We caution the reader that the data are subject to large systematic uncertainties due to, for example, the assumed IMF and dust correction. Also observe that the scatter is clearly too large compared with the error bars, despite the fact that Hopkins & Beacom (2006) applied a uniform dust correction and that the same IMF was assumed for all observations. Given these uncertainties, models whose predictions are discrepant with respect to these observations cannot automatically be ruled out.

Focusing first on the *L100* run (black, solid curve) we see a sharp rise at high redshift, a peak at $z \approx 2$ followed by a steady decline to $z = 0$. Qualitatively this matches the data, although the simulations appear to underestimate the SFR beyond the peak as well as the steepness of the decline below $z = 1$. Note that the height of the peak, i.e. the maximum SFR, is sensitive to the fraction of the SN energy that is used to generate galactic winds. While we fixed the wind velocity to 600 km s^{-1} based on other considerations (see Dalla Vecchia & Schaye 2008), the mass loading $\eta = 2$ (which corresponds to 40 per cent of the SN energy for $v_w = 600 \text{ km s}^{-1}$) was chosen partly because it gives roughly the right maximum SFR. We will show later that the underestimate of the SFR at $z > 3$ can be attributed to a lack of numerical resolution, while the overestimate of the SFR at $z < 0.5$ reflects the fact that our galactic winds cannot suppress SF in massive galaxies.

Comparing the three black curves extending to $z = 0$, which correspond to box sizes of 25, 50 and $100 h^{-1} \text{ Mpc}$, we see that even a $25 h^{-1} \text{ Mpc}$ box is large enough to obtain a converged estimate for the cosmic SFH. This is perhaps surprising, as there clearly exist structures with sizes that are of the same order or greater than this. Apparently, rare objects like clusters of galaxies do not contribute significantly to the mean SFR. This is consistent with Crain et al. (2009), who used zoomed simulations to show that while the SFR in different $25 h^{-1} \text{ Mpc}$ regions varies by up to an order of magnitude, the SFH in a region of this size whose mean density equals the cosmic mean closely tracks the global SFH.

Comparing the three red curves that end at higher redshifts, which correspond to box sizes of 6.125, 12.5 and $25 h^{-1} \text{ Mpc}$ and particle masses that are 64 times smaller than those used for the black curves, we see that while a $12.5 h^{-1} \text{ Mpc}$ box is nearly sufficiently large for $z > 2$ (recall that we already established that $25 h^{-1} \text{ Mpc}$ is sufficiently large using the lower resolution simulations), $6.125 h^{-1} \text{ Mpc}$ is clearly insufficient to obtain a converged estimate of the SFH.

Now that we have established that our fiducial box sizes are sufficiently large, we turn our attention to the convergence with respect to resolution.

3.2.2 Numerical resolution

While the simulation box limits the maximum sizes of the structures that can form, numerical resolution may even affect the properties of common objects. For example, hydrodynamical simulations that do not resolve the Jeans scales may underestimate the fraction of mass in collapsed structures and thus the SFR. Moreover, we cannot expect to form dark matter haloes whose masses are comparable to or smaller than the particle mass.

Before showing the results of the convergence tests, it is useful to consider what to expect. The *L025N512* model has a dark matter particle mass of $m_{\text{dm}} \approx 6 \times 10^6 h^{-1} M_{\odot}$. From a comparison of the mass functions of dark-matter-only simulations, we find that $\sim 10^2$ particles are needed to robustly define a halo. Thus, we expect the

halo mass function to be converged for $M > 6 \times 10^8 h^{-1} M_{\odot}$. Comparing this to the mass corresponding to the virial temperature below which photoheating is expected to suppress SF, $M \sim 0.7 \times 10^8 h^{-1} M_{\odot} (T_{\text{vir}}/10^4 \text{ K})^{3/2} ((1+z)/10)^{-3/2}$, we see that even after reionization we are missing haloes expected to form stars. However, haloes of such low mass are only expected to dominate the cosmic SFR at very high redshift. The particle mass for *L100N512* is 64 times greater, and we can thus only probe the mass function down to $M \sim 4 \times 10^{10} h^{-1} M_{\odot}$. This will lead us to underestimate the SFR at $z \gtrsim 3$ (see e.g. fig. 5 of Crain et al. 2009).

Let us now consider the hydrodynamics. The Jeans scales depend on the density and the temperature of the gas. The temperature of substantially overdense⁶ gas is $\gtrsim 10^4 \text{ K}$ in our simulations. In reality, gas at interstellar densities ($n_{\text{H}} \gtrsim 10^{-1} \text{ cm}^{-3}$) is sufficiently dense and self-shielded to form a cold ($T \ll 10^4 \text{ K}$) gas phase, allowing it to form stars (Schaye 2004). However, our simulations impose an effective EOS for gas with densities exceeding our SF threshold of $n_{\text{H}} = 10^{-1} \text{ cm}^{-3}$. For our EOS ($P \propto \rho^{4/3}$), the Jeans mass is independent of the density. Hence, if we resolve the Jeans mass at the SF threshold, we resolve it everywhere. The Jeans mass is given by

$$M_{\text{J}} \approx 1 \times 10^7 h^{-1} M_{\odot} f_{\text{g}}^{3/2} \left(\frac{n_{\text{H}}}{10^{-1} \text{ cm}^{-3}} \right)^{-1/2} \left(\frac{T}{10^4 \text{ K}} \right)^{3/2}, \quad (2)$$

where f_{g} is the (local) fraction of the mass in gas. Thus, we do not expect convergence unless the gas particle mass $m_{\text{g}} \ll 10^7 M_{\odot}$. To achieve convergence, a simulation will, however, also need to resolve the Jeans length L_{J} . This implies that the maximum, proper gravitational softening, ϵ_{prop} , must be small compared with the Jeans length

$$L_{\text{J}} \approx 1.5 h^{-1} \text{ kpc} f_{\text{g}}^{1/2} \left(\frac{n_{\text{H}}}{10^{-1} \text{ cm}^{-3}} \right)^{-1/2} \left(\frac{T}{10^4 \text{ K}} \right)^{1/2}. \quad (3)$$

Note that since L_{J} scales as $L_{\text{J}} \propto \rho^{-1/3}$ for our EOS, ϵ_{prop} will always exceed L_{J} for sufficiently high densities. However, since the Jeans mass does not decrease with density, we do not expect runaway collapse for star-forming gas.

Comparing the above equations with the gas particle mass and softening scales for our fiducial simulations (see Table 1), we see that while *L025N512* marginally resolves the Jeans scales for $f_{\text{g}} \approx 1$, this is not the case for the simulations that go down to $z = 0$, although *L050N512* has $m_{\text{g}} \approx M_{\text{J}}$ and $\epsilon_{\text{prop}} \approx L_{\text{J}}$ and is therefore not far off. Note, however, that none of our simulations comes close to resolving the Jeans scales prior to reionization, when SF in haloes with virial temperatures less than 10^4 K may have been important.

Fig. 4 compares the SFHs predicted by simulations with varying resolutions. The left-hand and right-hand panels test the convergence of the *L025N512* and *L100N512* models, respectively. Focusing first on the solid and dashed black curves in the left-hand panel, we see that a particle mass eight times greater than our fiducial value (and a softening twice our fiducial value) is sufficient for $z < 4$. For $z < 2$, even a particle mass that is a factor of 64 smaller appears to be sufficient. Comparison of our fiducial run with the higher resolution run *L012N512* (red, dot-dashed) indicates that the former is likely nearly converged for $z < 7$. Note that we do not expect perfect agreement when comparing *L025* and *L012* runs even if they have converged in terms of resolution since the two runs

⁶Gas with very low overdensities can have temperatures substantially below 10^4 K due to the adiabatic Hubble expansion, but the Jeans scales corresponding to these low densities are nevertheless large.

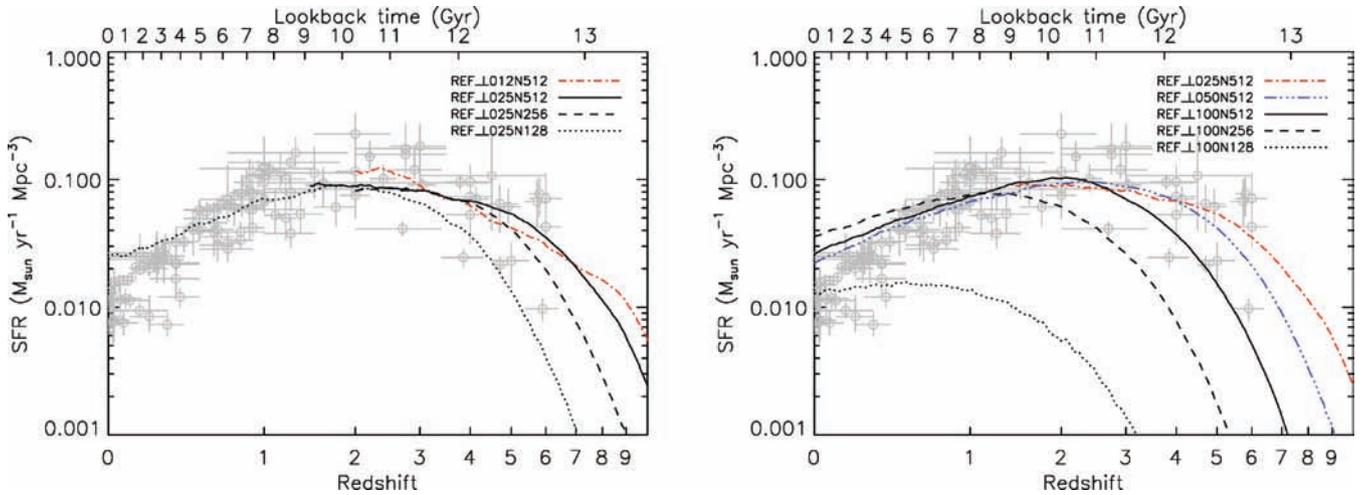


Figure 4. As Fig. 3, but comparing the SFHs for simulations of the reference model that differ in terms of their numerical resolution. The left-hand and right-hand panels test the convergence of the SFH predicted by the $L025N512$ and $L100N512$ runs, respectively. The SFH in the $L025N512$ run is clearly fully converged for $z < 4$ (compare with the lower resolution $L025N256$), but comparison with the higher resolution $L012N512$ (which, however, has a box size that is slightly too small to be converged; see Fig. 3) suggests that it is nearly converged for $z < 7$. Comparing $L100N512$ with $L050N512$ (which use a box size sufficiently large to provide a converged result; see Fig. 3), we see that the former has nearly converged for $z < 3$. Interestingly, increasing the resolution beyond that of $L100N256$ decreases the SFR at late times.

necessarily have different initial conditions and a $12.5 h^{-1}$ Mpc box is not fully converged in terms of box size (see Fig. 3).

The kink at $z = 9$ in the SFH of $L012N512$ is due to the negative feedback associated with reheating during reionization. While this effect is easily visible for $L012N512$, it is only just detectable in $L025N512$. This is expected. Reionization will suppress SF in haloes with virial temperatures of $\lesssim 10^4$ K, which corresponds to halo masses of $\sim 10^8 M_{\odot}$ at this redshift. Such haloes are resolved with $\sim 10^2$ particles in $L012N512$, but contain only ~ 10 particles in $L025N512$.

Comparing the three black curves in the right-hand panel, which show the SFH for runs $L100N128$, $L100N256$ and $L100N512$, we see no evidence for full convergence, although the difference between the two highest resolution runs is small for $z < 1.5$. Indeed, comparison with $L050N512$ and $L025N512$, whose box sizes we have already shown to be sufficiently large, reveals that the fiducial run $L100N512$ has nearly converged for $z < 3$ (and $L050N512$ for $z \lesssim 4$).

An increase in the numerical resolution typically increases the SFR, particularly at high redshift, when it is dominated by haloes near the resolution limit. Observe, however, that the opposite happens at low redshift once the resolution is increased beyond that of $L100N256$. This reduction reflects the fact that, in simulations with higher resolution, the gas that would otherwise be available or SF has already been used up or ejected by lower mass progenitors at higher redshifts.

Single simulations currently lack the dynamic range to obtain a converged result for the SFH over a wide range of redshifts. One strategy that has been used to overcome this limitation (e.g. Springel & Hernquist 2003b) is to combine a suite of simulations with different box sizes and to use, for each redshift, the one that yields the highest SFR. This procedure comes down to plotting all SFH curves and using the envelope that encompasses all as the best estimate for the SFH. However, our finding that, at low redshift, the predicted SFR decreases as it approaches convergence indicates that this procedure may overestimate the SFR at late times. Because the low-redshift SFR depends on the amount of gas that was consumed

or ejected at earlier times, one really does need a large dynamic range to model the SFH down to $z = 0$.

Summarizing, the convergence tests are consistent with our expectations based on our estimates of the minimum resolved dark halo mass and a comparison of the mass and length resolutions with the Jeans scales. For our purposes, the resolution of the $L100N512$ run suffices for $z < 3$ and that of $L025N512$ for $z \lesssim 7$.

Before investigating the physics driving the predicted SFH, we caution the reader that convergence of our reference model does not automatically imply that other physics variations are also converged at the same resolution. On the other hand, it would be surprising if resolution effects would change qualitative conclusions drawn from comparisons in the regime for which *REF* is converged. This situation would change, however, if we did not impose an effective EOS on the ISM because in that case, the Jeans scales could become much smaller than in the models considered here.

4 VARIATIONS ON THE REFERENCE MODEL

In this section, we describe the full set of OWLS runs. Most simulations differ from the reference model in only the choice of a single parameter or the presence of a certain aspect of the subgrid physics. In this way, cross-comparison of different simulations allows us to isolate the effects of different physical processes and the importance of different numerical parameters. The full list of simulations is shown in Table 2, which lists the simulation identifier, indicates whether or not a given simulation was run in the 25 and $100 h^{-1}$ Mpc boxes and gives a reference to the section that discusses the simulation. Except for the *MILL* runs, all simulations that were run in the same box size used identical initial conditions.

The order in which we present the different variations roughly parallels the order in which the subgrid models were discussed in Section 3.1. The different subsections can be read independently of each other. We begin by comparing our *WMAP-3* cosmology to the *WMAP-1* cosmology, which was for example assumed in the

Table 2. List of main physics variations employed in the OWLS project.

Simulation	L025	L100	Section	Description
<i>AGN</i>	✓	✓	4.1.0	Includes AGN
<i>DBLIMFCONTSFV1618</i>	✓	✓	4.7.2	Top-heavy IMF at high pressure, cont. SF law, extra SN energy in wind velocity
<i>DBLIMFV1618</i>	✓	✓	4.7.2	Top-heavy IMF at high pressure, extra SN energy in wind velocity
<i>DBLIMFCONTSFML14</i>	✓	✓	4.7.2	Top-heavy IMF at high pressure, cont. SF law, extra SN energy in mass loading
<i>DBLIMFML14</i>	✓	✓	4.7.2	Top-heavy IMF at high pressure, extra SN energy in mass loading
<i>EOS1p0</i>	✓	✓	4.4	Slope of the effective EOS changed to $\gamma_{\text{eff}} = 1$
<i>EOS1p67</i>	✓	–	4.4	Slope of the effective EOS changed to $\gamma_{\text{eff}} = 5/3$
<i>IMFSALP</i>	✓	✓	4.7.1	Salpeter (1955) IMF
<i>IMFSALPML1</i>	✓	–	4.7.1	Salpeter (1955) IMF; wind mass loading $\eta = 2/1.65$
<i>MILL</i>	✓	✓	4.1	Millennium simulation cosmology, $\eta = 4$ (twice the SN energy of <i>REF</i>)
<i>NOAGB_NOSNIa</i>	–	✓	4.6	No mass loss from AGB stars and SN Ia
<i>NOHeHEAT</i>	✓	–	4.3	No extra heat input around helium reionization
<i>NOREION</i>	✓	–	4.3	No hydrogen reionization
<i>NOSN</i>	✓	✓	4.8	No SN energy feedback from SNe
<i>NOSN_NOZCOOL</i>	✓	✓	4.2	No SN energy feedback from SNe and cooling assumes primordial abundances
<i>NOZCOOL</i>	✓	✓	4.2	Cooling assumes primordial abundances
<i>REF</i>	✓	✓	3	Reference model
<i>REIONZ06</i>	✓	–	4.3	Hydrogen reionization occurs at $z = 6$
<i>REIONZ12</i>	✓	–	4.3	Hydrogen reionization occurs at $z = 12$
<i>SFAMPLx3</i>	✓	–	4.5.2	Normalization of the Kennicutt–Schmidt SF law increased by a factor of 3
<i>SFAMPLx6</i>	✓	–	4.5.2	Normalization of the Kennicutt–Schmidt SF law increased by a factor of 6
<i>SF Slope1p75</i>	✓	–	4.5.2	Slope of the Kennicutt–Schmidt SF law increased to 1.75
<i>SFTHRESZ</i>	✓	–	4.5.1	Critical density for onset of SF is a function of metallicity (equation 4)
<i>SNIaGAUSS</i>	–	✓	4.6	Gaussian SN Ia delay function
<i>WDENS</i>	✓	✓	4.8.1	Wind mass loading and velocity depend on gas density (SN energy as <i>REF</i>)
<i>WHYDRODEC</i>	✓	–	4.8.2	Wind particles are temporarily hydrodynamically decoupled
<i>WML1V848</i>	✓	✓	4.8.1	Wind mass loading $\eta = 1$, velocity $v_w = 848 \text{ km s}^{-1}$ (SN energy as <i>REF</i>)
<i>WML4</i>	✓	✓	4.8	Wind mass loading $\eta = 4$ (twice the SN energy of <i>REF</i>)
<i>WML4V424</i>	✓	–	4.8.1	Wind mass loading $\eta = 4$; wind velocity $v_w = 424 \text{ km s}^{-1}$ (SN energy as <i>REF</i>)
<i>WML8V300</i>	✓	–	4.8.1	Wind mass loading $\eta = 8$; wind velocity $v_w = 300 \text{ km s}^{-1}$ (SN energy as <i>REF</i>)
<i>WPOT</i>	✓	✓	4.9	Wind mass loading and vel. vary with grav. potential (‘momentum-driven’)
<i>WPOTNOKICK</i>	✓	✓	4.9	Same as <i>WPOT</i> except that no extra velocity kick is given to winds
<i>WHERMAL</i>	✓	–	4.8.3	SN energy injected thermally
<i>WVCIRC</i>	✓	✓	4.9	Wind mass loading and vel. vary with halo circ. vel. (‘Momentum-driven’)

Note. From left to right, the columns show the simulation name, whether or not the simulation was run in the 25 and 100 Mpc h^{-1} boxes, respectively, the section number containing the description of the model and a very brief description of the changes in the model relative to the *REF* simulation. Except for the *MILL* runs, all simulations that were run using the same box size used identical initial conditions.

widely used ‘Millennium simulation’ (Springel et al. 2005b). In Sections 4.2 and 4.3, we investigate aspects of the radiative cooling and heating by turning off metal-line cooling and varying the redshift of reionization, respectively. We study the importance of our treatment of the unresolved ISM in Section 4.4 by varying the EOS. We vary the subgrid model for SF in Section 4.5, where we try a metallicity-dependent SF threshold as well as a range of Kennicutt–Schmidt SF laws. In Section 4.6, we investigate the effect of intermediate-mass stars by turning off mass loss from AGB stars and by varying the time-delay function for SNe Ia. Section 4.7 investigates the effect of using a Salpeter IMF and an IMF that becomes top-heavy at high pressures. Many aspects of our prescription for kinetic SN feedback are varied in Section 4.8. We not only try a range of parameter values, but also check the effect of temporarily decoupling the hydrodynamical forces on wind particles. This section also investigates a promising way of injecting SN feedback in thermal form. Another form of feedback from young stars is studied in Section 4.9, where we discuss the results of various simplified implementations of radiatively driven winds. Finally, we study the effect of AGN feedback in Section 4.1.0.

4.1 Cosmology

To investigate the dependence on cosmology and in order to facilitate comparisons to earlier work, we change the cosmological parameters from our fiducial *WMAP* 3-year values (Spergel et al. 2007) to the cosmology used in many studies including the Millennium simulation (Springel et al. 2005b). We refer to this latter set of cosmological parameters, which were chosen to be consistent with a combined analysis of the 2-degree Field Galaxy Redshift Survey and the first-year *WMAP* data, as the ‘Millennium cosmology’ and denote the models *MILL*. The Millennium cosmology uses the cosmological parameter values $\{\Omega_m, \Omega_b, \Omega_\Lambda, \sigma_8, n_s, h\} = \{0.25, 0.045, 0.75, 0.9, 1.0, 0.73\}$. Note that because of the change in the values of Ω_m and Ω_b , the dark matter and the (initial) baryonic particle masses are, respectively, 4.5 and 7.7 per cent higher for the *MILL* run than for the *REF* model.

The main differences with respect to the reference model are the values of Ω_b and σ_8 which are, respectively, 8 and 22 per cent higher for *MILL* than for *REF*. Both changes are expected to increase the SFR. The higher value of σ_8 has a particularly large

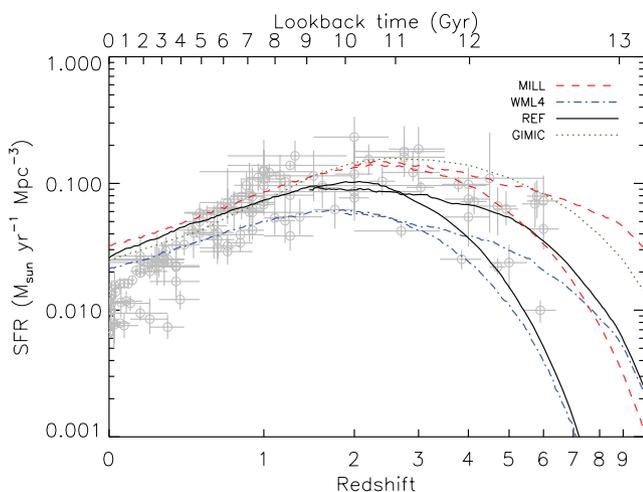


Figure 5. The effect of cosmology on the cosmic SFH. The curves show the cosmic SFR density as a function of redshift (bottom x -axis) and lookback time (top x -axis) for models *MILL* (red, dashed) and *WML4* (blue, dot-dashed). For comparison, the results for the reference model are also shown (black, solid). Results are shown for both the 25 and the $100 h^{-1}$ Mpc boxes, with the smaller box predicting higher SFRs at high redshift. All simulations used 2×512^3 particles. The data points show the compilation of observations from Hopkins & Beacom (2006), converted to our IMF and our fiducial cosmology. For comparison, we also show the SFH predicted by GIMIC (Crain et al. 2009, olive, dotted). Models *MILL* and *WML4* differ only in terms of the assumed cosmology. In particular, model *MILL* assumes values for Ω_b and σ_8 that are, respectively, 8 and 22 per cent higher than our fiducial values. Both *MILL* and *WML4* assume that the wind mass loading, and hence the fraction of the SN energy that is injected in the form of winds, is twice as high as for model *REF*. Note that the lookback time axis (top x -axis) applies only to the *REF* cosmology. Although the data points assume the *REF* cosmology, they would be very similar for the *MILL* cosmology. Compared to our fiducial *WMAP-3* cosmology, the *MILL* cosmology predicts higher SFRs, particularly at high redshift.

effect at high redshift, because structure formation proceeds faster in the *MILL* cosmology. In order to roughly match the peak in the observed SFH, we doubled the mass loading factor to $\eta = 4$ for the SN-driven winds. Hence, the winds account for 80 per cent of the available energy from SNe. To isolate the effect of cosmology, we therefore compare the *MILL* simulation to model *WML4* which employs the same wind parameters, but is otherwise identical to the reference model.

Fig. 5 compares the SFHs in the *MILL* (red, dashed) and *WML4* (blue, dot-dashed) runs. The change from the *WMAP-3* to the *MILL* cosmology strongly boosts the SFR. The difference increases with redshift from about 0.2 dex at $z = 0$ to 0.34 dex at $z = 2$ (for both box sizes). By $z = 9$, the difference has increased to 1.0 dex. Clearly, for a quantitative comparison with observations, it is important to use the correct cosmology. At high redshift, when the haloes that dominate the SF in the simulation correspond to rare fluctuations, the predicted cosmic SFR becomes extremely sensitive to the value of σ_8 . We will show in Haas et al. (in preparation) that the differences are much smaller for haloes of a fixed mass, which implies that the change in the halo mass function accounts for most of the differences in the SFHs predicted for the two cosmologies.

The olive, dotted curve in Fig. 5 shows the SFH predicted by the Galaxies-Intergalactic Medium Interaction Calculation (GIMIC; Crain et al. 2009). GIMIC consists of a series of hydrodynamical simulations that zoom in on 25 Mpc subvolumes of the $500 h^{-1}$ Mpc

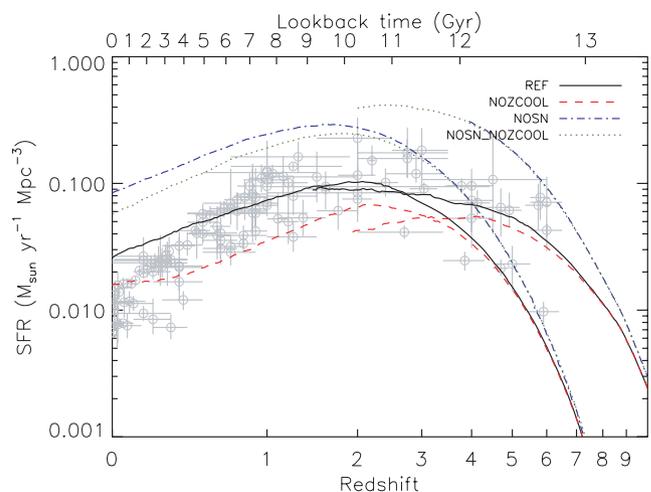


Figure 6. As Fig. 5, but comparing the SFHs for models with and without metal-line cooling, both in the presence and absence of SN-driven winds. Except at very high redshift, metal-line cooling strongly increases the SFR. The boost due to metal-line cooling is greater when SN feedback is included, which implies that metals radiate away a significant fraction of the energy injected by SNe.

dark matter only Millennium simulation and was run using the same code and parameter values as *MILL*. Fig. 5 shows the SFH computed from the weighted average of the five GIMIC subvolumes. The particle mass (gravitational softening) used for the GIMIC runs⁷ is eight (two) times larger than that of our $25 h^{-1}$ Mpc box and thus eight (two) times smaller than that of our $100 h^{-1}$ Mpc box. At very high redshift, the SFR in the GIMIC run is intermediate between that of our two *MILL* box sizes. For $z < 7$ it is very close to that of the $25 h^{-1}$ Mpc box and at $z < 3$ it agrees with the $100 h^{-1}$ Mpc run, although the GIMIC SFR falls off somewhat more steeply below $z = 2$. These differences are exactly what is expected from resolution effects as can be seen by comparing the SFHs for the 25, 50 and $100 h^{-1}$ Mpc boxes shown in the right-hand panel of Fig. 4. The excellent agreement confirms that our box sizes are sufficiently large to obtain a converged prediction for the cosmic SFH.

4.2 Metal-line cooling

Simulations without any radiative cooling are of interest for the study of hot gas in groups of clusters of galaxies (we have run such a simulation for this purpose in the $100 h^{-1}$ Mpc box), but in order to form stars, the gas must be able to radiate away its binding energy. Despite the importance of cooling, most cosmological studies still use highly simplified prescriptions, ignoring metal-line cooling or including it under the assumption of collisional ionization equilibrium and fixed relative abundances. Our simulations are the first to compute the cooling rates element by element and the first high-resolution simulations of cosmological volumes that include the effect of photoionization on the heavy elements.

Fig. 6 compares the reference simulations with runs that ignored metal-line cooling (*NOZCOOL*; red, dashed). As expected, the two agree at very high redshift where there has not been enough time to enrich the gas significantly and where much of the gas falls in

⁷These are the numbers for the intermediate resolution GIMIC runs. The high-resolution runs use the same particle masses and force resolution as our $25 h^{-1}$ Mpc box, but they end at $z = 2$ and do not include the highest density subvolume.

cold (e.g. White & Frenk 1991; Birnboim & Dekel 2003; Kereš et al. 2005). At late times, the runs without metal-line cooling consistently predict lower SFRs. For our high-resolution *L025N512* runs, the difference increases with cosmic time to 0.3 dex at $z = 2$. While the SFR increases to $z = 2$ for *REF*, it peaks at $z = 4$ when metal-line cooling is ignored. Interestingly, for the *L100* runs the difference decreases after peaking at about 0.4 dex around $z = 0.4$.

Also shown in Fig. 6 are four⁸ runs without SN-driven winds (but still including metal production and mass loss from SNe), both with and without metal-line cooling. Clearly, SN feedback strongly suppresses the SF, a point that we will come back to in Section 4.8.

Interestingly, while metal-line cooling also enhances the SFR in the absence of SN feedback, its effect is smaller than when SN feedback is included. Put another way, the factor by which SN feedback reduces the SFR is smaller when metal-line cooling is included. There are two possible explanations for this effect, which may both be right. First, metal-line cooling may reduce the efficiency of SN feedback, probably because it increases radiative losses in gas that has been shock-heated by the wind. Secondly, SN feedback may increase the effect of metal-line cooling, probably because it increases the fraction of the gas that is enriched. The former explanation is likely to be most relevant for galaxy groups, as we will show elsewhere that galactic winds do not dominate the enrichment of the intragroup medium.

Recently, Choi & Nagamine (2009) have also investigated the effect of metal-line cooling on the SFH. While they also used *GADGET*, their simulations used about an order of magnitude fewer particles and were stopped at higher redshifts. They did not include stellar evolution; they assumed a different cosmology and different sub-grid prescriptions for SF and SN feedback. Their cooling rates were taken from Sutherland & Dopita (1993) and assumed fixed relative abundances and collisional ionization equilibrium. Hence, there are many differences compared to our implementation of cooling, as we do include stellar evolution and compute the cooling rates element by element, including the effects of photoionization by the meta-galactic UV/X-ray background. For their simulations with resolutions similar to our *L025N512* runs (but a much smaller box), they found that metal-line cooling increases the SFR by less than 0.1 dex at $z = 3$, whereas we find 0.16 dex. In simulations with resolution comparable to our *L100N512*, they found an increase of about 0.25 dex by $z = 1$ whereas we find 0.3 dex. Thus, the results are broadly consistent although we find a somewhat larger boost due to metal cooling.

Clearly, except at very high redshift, metal-line cooling is very important. It boosts the SFRs by allowing more of the gas that accretes on to haloes to cool and by reducing the efficiency of SN feedback. Without metal cooling, predictions for the total number of stars formed could easily be low by a factor of 2. Note, however, that we may have overestimated the effect of metal-line cooling for massive galaxies because we have not included any feedback processes capable of stopping cooling flows in such systems. This results in an overestimate of the SFRs and the metallicities in the central regions of groups and clusters of galaxies. On the other hand, as discussed in Wiersma et al. (2009b), the fact that SPH underestimates small-scale metal-mixing (because metals are stuck on particles) causes us to underestimate the total mass that has been enriched, while overestimating the metallicity of the particles that

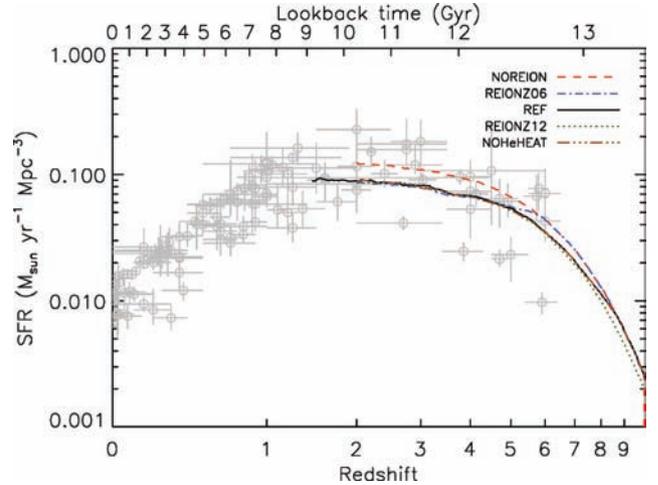


Figure 7. As Fig. 5, but comparing the SFHs for models in which hydrogen was reionized – by switching on a uniform ionizing background – at redshifts $z_r = 12$ (*REIONZ12*), 9 (*REF*), 6 (*REIONZ06*) or not at all (*NOREION*). All simulations use a $25 h^{-1}$ Mpc box and 2×512^3 particles. After the ionizing background is switched on, the SFR quickly changes to the rate predicted by the model with the highest redshift of reionization. Note that the factor by which reionization suppresses the SFR is limited by the resolution of the simulations and will be more severely underestimated at higher redshifts. Turning off the extra heat input of 2 eV per atom around helium reionization ($z \approx 3.5$, model *NOHeHEAT*) has no discernible effect on the SFH.

have received metals. Wiersma et al. (2009b) found that the net effect of increased metal mixing is to boost the SFR.

4.3 Reionization

As our simulations do not include radiative transfer, we need to assume that the background radiation is uniform and that the gas is optically thin. Hydrogen reionization is thus implemented in our simulations by switching on the Haardt & Madau (2001) model for the ionizing background radiation at some redshift z_r , corresponding to the epoch of reionization. Note, however, that we assume that a photodissociating background is already present at $z > z_r$, which effectively suppresses molecular cooling at all redshifts. As described in Wiersma et al. (2009a), switching on the ionizing radiation results in a sudden increase in the radiative heating rate and a sudden decrease in the radiative cooling rate above 10^4 K. As a result, cold gas is quickly heated to $T \sim 10^4$ K, removing gas from haloes with virial temperatures of $< 10^4$ K (e.g. Couchman & Rees 1986; Okamoto, Gao & Theuns 2008).

Fig. 7 compares *L025N512* runs with $z_r = 12, 9$ (i.e. the reference model), 6 and 0 (i.e. no reionization).⁹ After reionization, the SFR deviates from the curve corresponding to the simulation without an ionizing background (*NOREION*) and quickly asymptotes to the model with the highest redshift of reionization (*REIONZ12*). Apparently, the gas in the simulation rapidly loses memory of the time of reheating, as was also found by Pawlik, Schaye & van Scherpenzeel (2009). This is expected, as the sound-crossing time-scale is only 10^8 yr ($l/1$ kpc) for 10^4 K gas.

Naively, one would have expected the suppression of the SFR due to photoionization to decrease at late times, as haloes with $T_{\text{vir}} \gg 10^4$ K start to dominate the cosmic SFR. Interestingly, we

⁸Note that *NOSN_NOZCOOL_L025N512* was stopped earlier.

⁹The run with $z_r = 12$ uses the $z = 9$ Haardt & Madau (2001) model for $z = 9-12$.

do not find this. If anything, the suppression keeps increasing with time, reaching 0.15 dex (a 30 per cent reduction) by $z = 2$. While this is probably mostly a resolution effect, it does indicate that photoionization also reduces the SFR in haloes with higher virial temperatures, either because of the reduction of the cooling rates (Wiersma et al. 2009a) or because it makes the cold gas more susceptible to galactic winds (Pawlik & Schaye 2009).

We emphasize that because of our limited resolution (Haas et al., in preparation, show that we underestimate the SFR in haloes with masses less than $10^{10} M_{\odot}$) and because we assume the presence of a photodissociating background at all redshifts, it is likely that we have strongly underestimated the reduction of the SFR due to photoheating and that this underestimate becomes more severe at higher redshifts. Moreover, our assumption that the gas is optically thin results in an underestimate of the heating rates during reionization.

Helium is thought to have been reionized around $z \approx 3.5$, and the increase in the photoheating rates associated with this event can explain the relatively high temperature of the IGM inferred from observations of quasar absorption spectra (e.g. Schaye et al. 2000). As described in Wiersma et al. (2009b), by injecting 2 eV per atom at $z \approx 3.5$, we are able to match the observationally inferred temperatures. Fig. 7 shows that omitting this extra heat (*NOHeHEAT*) does not yield any noticeable changes in the SFH. This is not surprising, as the temperature increase is confined to low-density gas, far away from galaxies, for which adiabatic cooling dominates over radiative cooling.

4.4 The equation of state of the ISM

Our simulations have neither the resolution nor the physics to model the multiphase ISM. As discussed in Section 3.1.3, we therefore impose a polytropic EOS with slope $\gamma_{\text{eff}} = 4/3$ for gas with densities that exceed our SF threshold of $n_{\text{H}} = 0.1 \text{ cm}^{-3}$. This slope was chosen because it results in a constant Jeans mass and thus suppresses artificial fragmentation. In this section, we will check the effect of varying the slope of the EOS. Schaye & Dalla Vecchia (2008) showed that changes in the EOS can significantly alter the morphology of galaxies. A softer EOS results in tighter spiral arms, thinner discs and increased fragmentation.

Different groups use different prescriptions for the ISM. For example, Springel & Hernquist (2003a) use a complicated function that results from a semi-analytic model of the multiphase ISM. They interpret the pressure implied by their EOS, which is steeper than $4/3$ at densities similar to our SF threshold, as a form of SN feedback. In the past, many cosmological simulations have been run that do not impose an EOS, but which also do not include the physics necessary to model the cold interstellar gas phase (e.g. radiative transfer and molecule formation). Such simulations effectively use an isothermal EOS.

Fig. 8 compares the SFHs of runs with $\gamma_{\text{eff}} = 1$ (i.e. isothermal), $4/3$ (*REF*) and $5/3$ (i.e. adiabatic). Clearly, changes in the EOS do not have a significant effect on the predicted SFH. This may be surprising, given that the EOS can strongly affect the structure of galaxies.

One reason why the results are insensitive to the EOS is that we use the prescription for SF of Schaye & Dalla Vecchia (2008). As discussed by these authors, for self-gravitating systems such as galaxies, the observed Kennicutt–Schmidt surface density law is in effect a pressure law. By implementing it as a pressure law, we can thus reproduce the observed SF law independently of the assumed EOS of the star-forming gas. Previous cosmological simulations have, however, used volume density laws, in which case

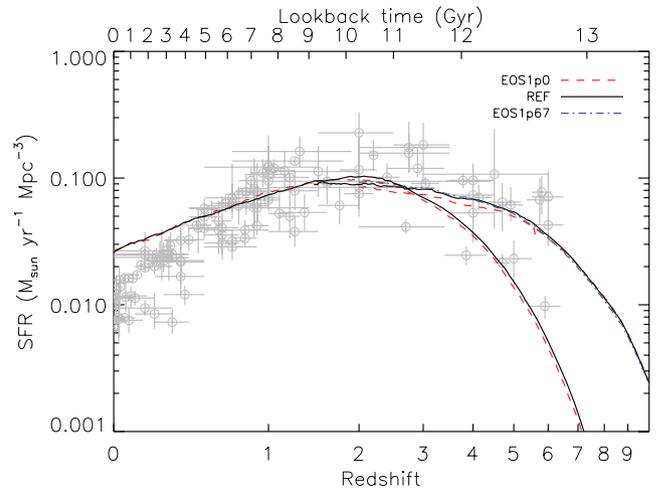


Figure 8. As Fig. 5, but comparing the SFHs for models assuming different EOSs for the unresolved ISM. The slope of the polytropic EOS imposed on the ISM is 1.0 (i.e. isothermal) for model *EOS1p0* (red, dashed), $4/3$ for *REF* (solid, black) and $5/3$ (i.e. adiabatic) for *EOS1p67* (blue, dot-dashed). The SFH is insensitive to the assumed slope of the polytropic EOS that is imposed on to the ISM.

the predicted Kennicutt–Schmidt law must depend on the assumed EOS because the latter sets the scaleheight of the disc. If the EOS is changed, then the same surface density corresponds to a different volume density, but the relation between surface density and pressure will remain unchanged. It is therefore not clear whether the results of previous simulations are as insensitive to the imposed EOS as we find here. However, we will show in the next section that the results are in fact insensitive to the gas consumption time-scale, because SN feedback enables galaxies to regulate their SFRs.

4.5 Star formation

4.5.1 The star formation threshold

Schaye (2004) argued that there is a critical density for the formation of a cold, interstellar gas phase and that the transition from the warm to the cold gas phase triggers gravitational instability on a wide range of length-scales. Gas with densities below the threshold is kept warm ($T \sim 10^4 \text{ K}$) and stable by the presence of a UV background. The predicted critical gas surface density $\Sigma_{\text{g}} \sim 3\text{--}10 M_{\odot} \text{ pc}^{-2}$, which agrees well with SF thresholds inferred from $H\alpha$ observations of nearby galaxies, corresponds to $n_{\text{H}} \sim 10^{-2}\text{--}10^{-1} \text{ cm}^{-3}$ for a self-gravitating disc at 10^4 K . Our reference model uses $n_{\text{H}} = 10^{-1} \text{ cm}^{-3}$. The Schaye (2004) model predicts that the critical density for SF is a weakly decreasing function of metallicity. We have therefore run a simulation, *SFTHRESZ*, that uses the predicted scaling (equations 19 and 24 of Schaye 2004, valid for $Z = 10^{-4}\text{--}10 Z_{\odot}$)

$$n_{\text{H}}^*(Z) = 10^{-1} \text{ cm}^{-3} \left(\frac{Z}{0.1 Z_{\odot}} \right)^{-0.64}, \quad (4)$$

where Z is the gas metallicity and we used $Z_{\odot} = 0.02$ for consistency with Schaye (2004). If the metallicity is zero, then we set $n_{\text{H}}^* = 10 \text{ cm}^{-3}$.

Fig. 9 compares models *SFTHRESZ* and *REF*. At very high redshift, the metallicity is low and the threshold density is higher than in the reference model. This results in a decrease in the SFR that drops rapidly from 0.3 dex at $z = 10$ to zero by $z = 6$. For

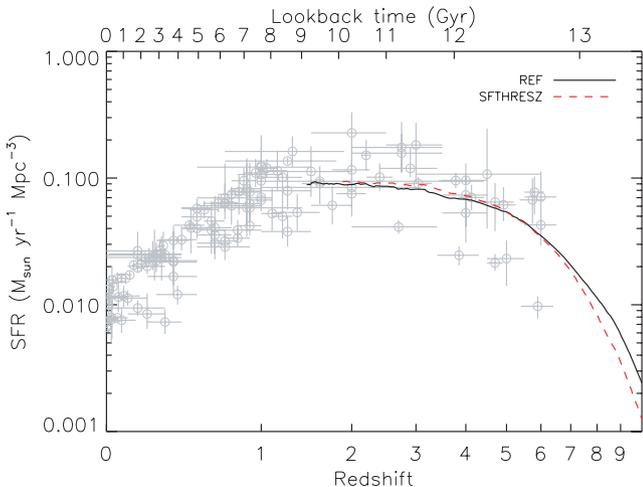


Figure 9. As Fig. 5, but comparing the SFHs for the reference model (black, solid), which uses a fixed threshold density for SF, to that of model *SFTHRESZ* (red, dashed), for which the SF threshold decreases with metallicity as predicted by Schaye (2004). The threshold densities in the two models agree for a metallicity of $0.1 Z_{\odot}$. Both simulations use a $25 h^{-1}$ Mpc box and 2×512^3 particles. At very high redshift, the metallicity is low and the total SFR is smaller for *SFTHRESZ* because it has a higher threshold density at this point. Below $z = 6$ the situation is reversed, but the difference between the SFHs is very small, suggesting that the cosmic SFR is dominated by galaxies that are able to regulate their SFRs.

$z < 6$ the SFR is slightly higher than in the *REF* model, which indicates that the metallicity of the star-forming gas is typically higher than $0.1 Z_{\odot}$, but the effect is marginal. Apparently, after a brief period in which the SFR is dominated by haloes that are just resolved and therefore just starting to form stars, the predicted SFRs become insensitive to the SF threshold. This suggests that the galaxies are able to regulate their SFRs. We will provide more evidence for this below.

4.5.2 The Kennicutt–Schmidt star formation law

As discussed in Section 3.1.3, gas on the effective EOS is allowed to form stars at a pressure-dependent rate that reproduces the observed Kennicutt–Schmidt law (Kennicutt 1998), $\dot{\Sigma}_* = A(\Sigma_g/1 M_{\odot} \text{pc}^{-2})^n$, with $A = 1.515 \times 10^{-4} M_{\odot} \text{yr}^{-1} \text{kpc}^{-2}$ and $n = 1.4$. The normalization (A) and slope (n) are constrained by observations, but remain controversial (e.g. Blanc et al. 2009). To develop an understanding of the physical role of the SF law, we have carried out one run with a different slope and two with different amplitudes.

Fig. 10 compares a run with $n = 1.75$ (model *SFSLOPE1p75*; red, dashed) with our reference model, which uses $n = 1.4$. The SF laws are in both cases normalized at $\Sigma_g = 1 M_{\odot} \text{pc}^{-2}$, which is below the threshold and hence implies that the SFR is higher for all densities in the run with the steeper slope SF law. For $z > 6$, the cosmic SFR is indeed higher in the run with $n = 1.75$. This is expected because at these high redshifts, the SFR is dominated by haloes that are just resolved and therefore just starting to form stars. These galaxies have not yet had time to become self-regulating, and their SFRs are inversely proportional to the gas consumption time-scales implied by the SF law.

Below $z = 6$, however, the SFRs in the two runs are nearly indistinguishable. This strongly suggests that the galaxies are regulating their SFRs such that they produce the same number of stars, and thus

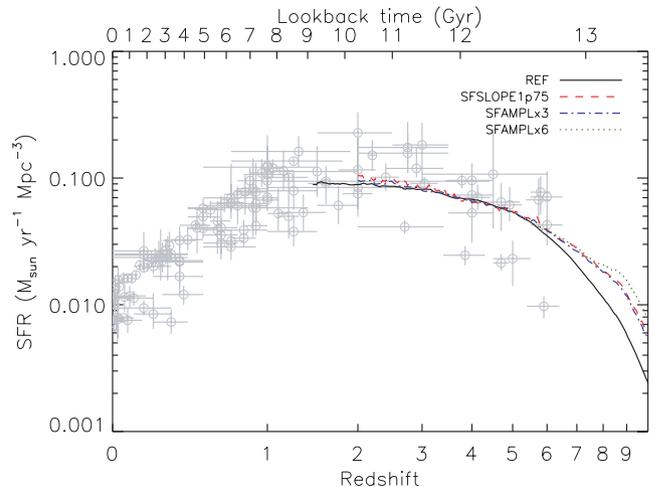


Figure 10. As Fig. 5, but comparing the SFHs for models with varying Kennicutt–Schmidt SF laws. Model *SFSLOPE1p75* (red, dashed) assumes a power-law slope $n = 1.75$ whereas the other models use our fiducial value $n = 1.4$. For models *SFAMPLx3* (blue, dot–dashed) and *SFAMPLx6* (olive, dotted) the amplitude of the SF law has been multiplied by factors of 3 and 6, respectively. All simulations use a $25 h^{-1}$ Mpc box and 2×512^3 particles. At very high redshift, when the SFR in the simulations is dominated by poorly resolved haloes, a more efficient SF law yields a higher SFR. After this initial phase the SFH is insensitive to the assumed SF law, which suggests that it is dominated by galaxies that are able to regulate their SFRs.

the same amount of SN energy, irrespective of the gas consumption time-scale. If a galaxy of a given halo mass, and hence with a fixed accretion rate, injects too little SN energy for a galactic outflow to balance the accretion rate, then the gas fraction, and hence the SFR, will increase. If, on the other hand, the SN rate is higher than required to balance the infall, then the gas fraction, and thus the SFR, will decrease. We thus expect that when the SF efficiency is changed, the galaxies will adjust their gas fractions so as to keep their SFR fixed. In Haas et al. (in preparation), we show that this is indeed what happens.

Finally, Fig. 10 shows that models in which the amplitude of the SF law is multiplied by factors of 3 (*SFAMPLx3*; blue, dot–dashed) and 6 (*SFAMPLx6*; olive, dotted), respectively, show the same behaviour. Initially, the SFR increases with A , but the SFR then quickly asymptotes to a fixed SFH. Observe that the two runs with higher amplitudes converge to a common evolution before the reference model joins them. This is because galaxies can regulate their SF more quickly if the SF efficiency is higher. Apparently, the cosmic SFR in the reference model only becomes dominated by self-regulated galaxies by $z = 6$. Note that higher resolution simulations may well find that self-regulation dominates already at higher redshifts because they can resolve SF in the progenitors of our lowest mass galaxies.

4.6 Intermediate-mass stars

Previous numerical studies of the cosmic SFH have mostly used the instantaneous recycling approximation (but see e.g. Oppenheimer & Davé 2008; Crain et al. 2009), which means that star particles eject all the products of stellar evolution immediately following their formation. Moreover, individual elements are typically not tracked. Instead, each gas element carries only a single metallicity variable and relative abundances are assumed to be solar. Furthermore, such simulations neglect mass loss, i.e. star particles change

the metallicity of their neighbours, but not their masses. As discussed in Section 3.1.4, we follow the timed release of 11 elements by intermediate-mass stars (SN Ia and AGB stars) and massive stars.

To check the importance of intermediate-mass stars, which eject much of their mass hundreds of millions to billions of years after their formation and which are responsible for most of the mass lost by stellar populations, we have run two *L100N512* simulations. In simulation *NOAGB_NOSNIa* we do not allow intermediate-mass stars to release mass, leaving massive stars, which evolve on time-scales of $\lesssim 10^7$ yr, as the only mechanism for releasing metals. To assess the impact of our choice of the distribution of SN Ia progenitor lifetimes, we ran a simulation (*SNIaGAUSS*) that uses a Gaussian rather than an e-folding time-delay function. This was motivated by the high redshift observations of Dahlen et al. (2004), which show a marked decline in the SN Ia rate beyond $z = 1$. The parameters of the delay model are $\sigma = 0.66$ Gyr and $\tau = 3.3$ Gyr (see Wiersma et al. 2009b).

Fig. 11 shows that the shape of the SN Ia delay function does not have a significant effect on the predicted SFH. Turning off both mass loss by SN Ia and AGB stars results, however, in a strong reduction of the SFR at late times. While the reduction factor is still very small at $z = 2$, it increases steadily thereafter to about 0.21 dex at $z = 0$, which corresponds to a 40 per cent reduction. Given that the SN Ia delay function does not matter, the difference must come mostly from mass loss by AGB stars. It is not important before $z = 2$ because there has not been sufficient time for a substantial fraction of the stars to reach the AGB phase. Note, however, that higher resolution simulations will predict higher SFRs at high redshift and may therefore find that AGB mass loss becomes important earlier.

Mass loss by AGB stars provides fresh fuel for SF and releases metals that were locked up in stars. This reduces the sharpness of the drop in the SFR with time, worsening the agreement with observations. Simulations that ignore this process will overestimate the steepness of the drop following the peak in the cosmic SFR.

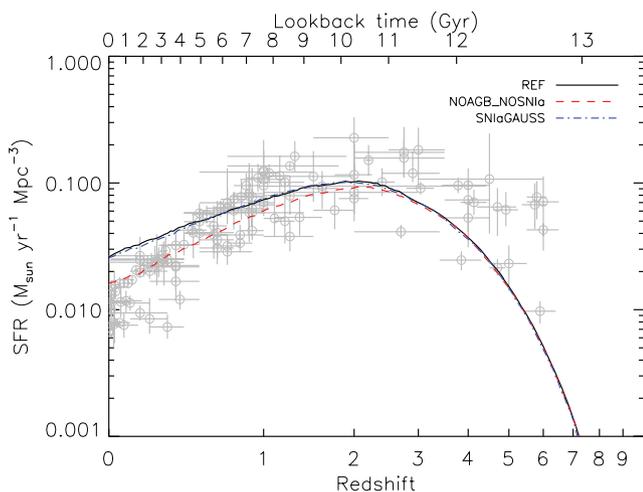


Figure 11. As Fig. 5, but comparing the SFHs for models that vary in their treatments of intermediate-mass stars. In model *NOAGB_NOSNIa* (red, dashed), mass loss by intermediate-mass stars and SN Ia has been turned off. While model *SNIaGAUSS* (blue, dot-dashed) includes these processes, it assumes a Gaussian time-delay function for SNe Ia instead of the e-folding model used in the other simulations. All simulations use a $100 h^{-1}$ Mpc box and 2×512^3 particles. While the SN Ia time-delay function is unimportant, mass loss by AGB stars provides fresh fuel for SF and releases metals back into the ISM, thereby boosting the SFR at late times.

4.7 The stellar initial mass function

4.7.1 A Salpeter IMF

Our reference model assumes a Chabrier IMF, but much of the literature uses a Salpeter IMF. The two IMFs have similar shapes above $1 M_{\odot}$, but while the Salpeter IMF is a pure power law, the Chabrier IMF includes a lognormal decrease at the low-mass end which results in a much lower stellar mass-to-light ratio. Because most of the ejected metal mass and all of the energy from core collapse SNe are produced by massive stars, the Salpeter IMF is less efficient in enriching the gas and driving outflows per unit stellar mass formed.

In order to assess the effect of changing the IMF we ran a simulation employing a Salpeter IMF (*IMFSALP*), using the same range of stellar masses as we used in the reference model (i.e. $0.1\text{--}100 M_{\odot}$). This simulation used Kennicutt’s original normalization for the amplitude of the SF law ($A = 2.5 \times 10^{-4} M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$; Kennicutt 1998), as he assumed the same IMF. Recall that this value is a factor of 1.65 greater than the amplitude assumed in *REF* (see Section 3.1.3).

Fig. 12 compares the SFHs predicted by the two IMFs. Initially, the Salpeter IMF gives a slightly higher SFR because of the higher SF efficiency implied by the change in the SF law (see also Section 4.5.2). However, after a short initial phase the SFR falls below that of the reference model. By $z = 0$, the difference has increased to 0.2 dex for *L100N512*. This behaviour can be explained by the fact that a Salpeter IMF produces less metals and returns less mass (and thus releases less metals that were locked up in stars) per unit

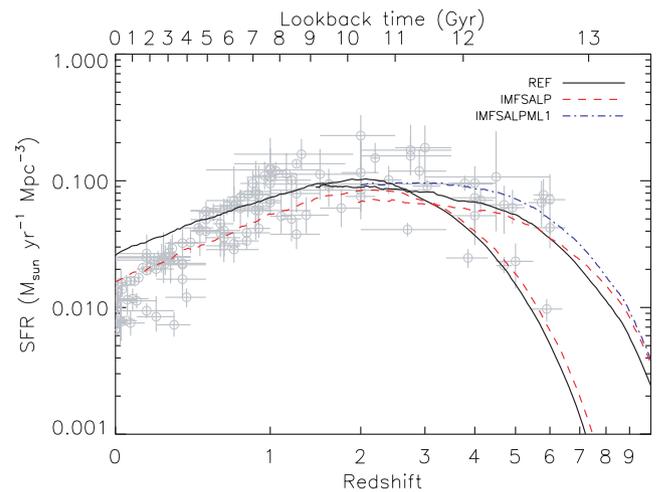


Figure 12. As Fig. 5, but comparing the SFHs for models assuming a Salpeter IMF (*IMFSALP*; red, dashed) to the reference model (black, solid), which assumes a Chabrier IMF. The amplitude of the SF law is taken from observations and has therefore been rescaled to the assumed IMF. It is a factor of 1.65 higher for the Salpeter IMF. Model *IMFSALPML1* assumes a Salpeter IMF and uses a wind mass loading factor that is a factor of 1.65 smaller (i.e. $\eta = 2/1.65$) than that used in the other models, which accounts for the change in the number of SNe per unit stellar mass. Note that the observed data points assume a Chabrier IMF. They need to be shifted upwards by a factor of 1.65 (0.22 dex) to compare with models assuming a Salpeter IMF. Initially, the SFR scales with the amplitude of the SF law and a Salpeter IMF produces a higher SFR. Later on, the SFR is smaller for a Salpeter IMF because of the decreased importance of metal-line cooling (because less metals are produced and a greater fraction are locked up in stars) and stellar mass loss. However, the smaller number of SNe per unit stellar mass more than compensates for this effect, at least for $z > 2$.

stellar mass formed. Hence, metal-line cooling is less efficient for a Salpeter IMF. Indeed, the predicted SFRs fall in between those for the reference model and the run without metal-line cooling (cf. Fig. 6).

However, a Salpeter IMF does not only produce less metal mass per unit stellar mass, it also produces fewer SNe. Assuming that the total energy in SNe scales as the total number of ionizing photons, the difference is a factor of 1.65 (Section 3.1.3). Thus, model *IMFSALP* uses 66 per cent of the SN energy to drive winds, whereas *REF* used only 40 per cent. For consistency, we therefore ran another *L025N512* simulation, model *IMFSALPML1*, that is identical to *IMFSALP*, except that the wind mass loading factor was reduced by a factor of 1.65 to $\eta = 1.2$. Fig. 12 shows that, as expected, this run yields a higher SFR than *IMFSALP*, although the two converge for $z > 9$ where there has not been sufficient time for the simulated galaxies to regulate their SF. In fact, with this change, a Salpeter IMF yields a higher SFR than a Chabrier IMF.

4.7.2 A top-heavy IMF at high pressures

Observational determinations of the IMF are extremely difficult. In particular, extragalactic observations are usually only sensitive to the light emitted by massive stars, either directly or indirectly via dust grains. While the IMF is usually assumed to be universal, it is expected to be top-heavy (or bottom-light) at very high redshift and low metallicity (e.g. Larson 1998) and both observations and theory suggest that it is top-heavy in extreme environments like the Galactic Centre and starburst galaxies (e.g. Padoan, Nordlund & Jones 1997; Baugh et al. 2005; Klessen, Spaans & Jappsen 2007; Maness et al. 2007; Bartko et al. 2009; Dabringhausen, Kroupa & Baumgardt 2009).

We have performed a series of runs to investigate the possible effects of an IMF that is top-heavy at high pressures, as may be the case in starbursts and in galactic nuclei. For simplicity, we assume

that the IMF switches suddenly from Chabrier to the top-heavy power law proposed by Baugh et al. (2005), $dN/dM \propto M^{-1}$ (as compared to $\propto M^{-2.3}$ for the high-mass tail of the Chabrier IMF). The transition is assumed to take place at the pressure $P/k = 2.0 \times 10^6 \text{ cm}^{-3} \text{ K}$ (evaluated at the resolution limit of the simulations), which was chosen because for this value $\sim 10^{-1}$ of the stellar mass in our simulations forms at higher pressures. This ensures that the top-heavy IMF is important, but not dominant. Of course, a discontinuous dependence on pressure is not physical, but it is simple and serves to illustrate the qualitative effects of a top-heavy IMF in starbursts.

Assuming that the SN energy scales with the emissivity in ionizing photons, our top-heavy IMF yields 7.3 times more SN energy per unit stellar mass formed. We have therefore increased the energy injected into the galactic wind by the same factor for star particles born out of high-pressure gas. In terms of our kinetic prescription for winds, we can either increase the mass loading or the wind velocity. We have tried both. Model *ML14* use a 7.3 times larger mass loading, while model *VI618* use a $\sqrt{7.3}$ times higher wind velocity.

If the actual SF law were continuous with pressure, then a sudden change in the IMF would imply a sudden change in the rate of formation of massive stars, which would manifest itself as a discontinuity in the apparent SF law inferred from observations under the assumption of a universal IMF. However, the observed SF law appears to be a continuous power law [though Krumholz, McKee & Tumlinson (2009) suggest that there may be a kink at $\Sigma_g \sim 10^2 \text{ M}_\odot \text{ pc}^{-2}$, which corresponds roughly to the pressure (see equations 20 and 21 of Schaye 2004) at which we switch IMFs]. We therefore tried two possibilities: models *DBLIMF* (*DBLIMFCONTSF*) assume a continuous (discontinuous) rate of formation of massive stars, but a discontinuous (continuous) SF law.

The left-hand panel of Fig. 13 compares the SFHs of all four models for each of the two box sizes. Note that models *DBLIMFML14_L100N512* and *DBLIMFCONTSFML14_L025N512* were stopped

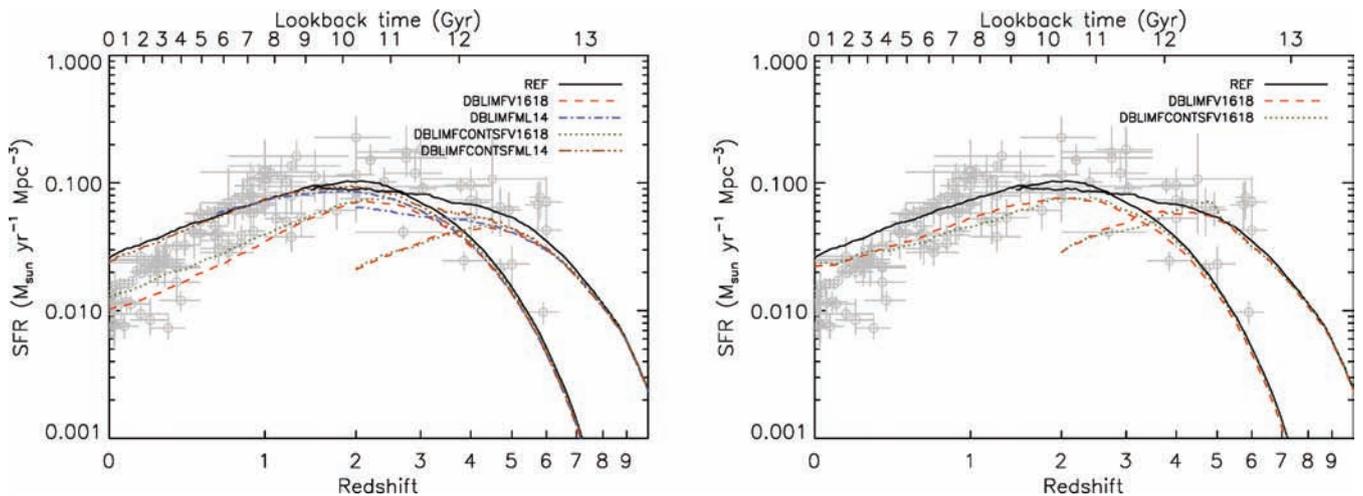


Figure 13. As Fig. 5, but comparing the SFHs for models with a top-heavy IMF in starbursts. The transition to a top-heavy IMF ($dN/dM \propto M^{-1}$ compared to $\propto M^{-2.3}$ for Chabrier) happens suddenly at a pressure $P/k = 2 \times 10^6 \text{ cm}^{-3} \text{ K}$. The top-heavy IMF produces 7.3 times more core collapse SNe per unit stellar mass formed. Models *ML14* and *VI618* use this extra energy to increase the wind mass loading and velocity, respectively. Model *DBLIMFCONTSF* assume a continuous SF law, whereas model *DBLIMF* assume the rate of formation of massive stars to be continuous. The left-hand panel shows actual SFRs, whereas the SFRs have been rescaled to the ones that would be inferred under the assumption of a Chabrier IMF in the right-hand panel. Comparisons to the observed data points, which assumed a Chabrier IMF, are only self-consistent for the right-hand panel. The models start to differ when some galaxies have become sufficiently massive to form a fraction of their stars with a top-heavy IMF. Models with a top-heavy IMF form less stars, which indicates that the relative increase in the SN rate is more important than the increase in the metal production rate. Whether the SF law is continuous or not is unimportant. Using the extra SN energy to increase the wind mass loading is less effective than increasing the wind velocity. A top-heavy IMF in starbursts reduces the SFR in massive galaxies, but the effect on the formation rate of massive stars is much less strong.

earlier than the other runs. The SFHs agree at early times, when SF is confined to low-mass haloes for which the gas pressure remains low. The models in which the extra SN energy was used to increase the wind mass loading predict SFHs that are similar to that of the reference model. This could mean that increasing the wind mass loading does not strongly boost the efficiency of SN feedback at high gas pressures. This agrees well with the results of Dalla Vecchia & Schaye (2008), who found that at high pressures the kinetic feedback becomes inefficient due to gas drag and that the pressure above which this occurs increases with the wind velocity. Indeed, as can be seen from Fig. 13, the models in which the wind velocity is increased for the top-heavy IMF do show a strong reduction in the SFR.

However, a top-heavy IMF not only yields more SN energy, but also more metal mass per unit stellar mass formed. The associated increase in the metal-line cooling rates will boost the SFRs (see Section 4.2). The relatively small difference between the *ML14* and *REF* models could therefore also mean that the increased wind mass loading compensates for the higher cooling rates.

The differences between *DBLIMF* and the corresponding *DBLIMFCNTSF* runs are small. This was to be expected, as we already showed in Section 4.5.2 that the SFHs are insensitive to the SF law because galaxies regulate their gas fractions so as to produce the same amount of SN energy, irrespective of the assumed SF efficiency.

To compare with the observed data points, which were derived from observations of massive stars under the assumption of a Chabrier IMF, we have to multiply the rate of SF in the top-heavy mode by a factor of 7.3. The right-hand panel of Fig. 13 shows that doing so reduces the drop at late times, i.e. the SFR inferred under the assumption of a universal IMF falls off less steeply than the actual SFR. Observe that the differences between models *DBLIMF* and *DBLIMFCNTSF* are also reduced, particularly for $z < 1$. This supports our proposal that because galaxies regulate their SF, they inject a fixed amount of SN energy for a given halo mass.

Finally, we note that the agreement between the 25 and $100 h^{-1}$ Mpc boxes is much poorer for the models with a top-heavy IMF in starbursts than it was for the other models. This reflects our choice to make the IMF a function of the pressure, which is one-to-one related to the gas density in our simulations since star-forming gas follows a polytropic EOS. Increasing the resolution decreases the mass above which haloes contain enough particles to sample the high-density tail of the gas distribution. Hence, lower halo masses will be able to form some fraction of their stars with a top-heavy IMF and thus suppress subsequent SF. Clearly, using prescriptions for feedback that are functions of density or pressure will make the results more prone to resolution effects.

We conclude that although our toy models are too simple and sensitive to resolution, it is clear that a top-heavy IMF in starbursts can serve to suppress SF in high-mass haloes. This can result in a steeper drop in the SFR at late times, as suggested by observations (although the effect is less strong when only the rate of formation of massive stars is considered). The stronger suppression in massive haloes can also shift the peak in the SFH to higher redshifts.

4.8 SN-driven winds

It is well known that simulations without galactic winds suffer from a severe overcooling problem: the SFR greatly exceeds the observational constraints and is usually only limited by numerical resolution (e.g. Balogh et al. 2001). As Fig. 6 shows, except at very high redshift when the SFR in the real Universe is expected

to be dominated by haloes that are below the resolution limit of our simulations, the runs without SN feedback do indeed produce far too many stars. Moreover, comparison of the SFHs for models *REF* and *WML4* in Fig. 5 shows that doubling the SN energy that is injected, in this case by doubling the wind mass loading, further reduces the SFR.

In Sections 4.5.2 and 4.7.2, we showed that the SFH is insensitive to the assumed SF law. We concluded from this that SF in galaxies is self-regulating: the SFR adjusts such that outflows driven by feedback from massive stars balance the infall driven by gas accretion on to haloes and radiative cooling. If the SF law is changed, then galaxies simply adjust their gas fractions in order to inject the same amount of SN energy into haloes of a given mass. If galaxies do indeed regulate their SFRs in this manner, then we would expect the rate of energy injection into the winds to remain constant if the fraction of the SN energy that is injected is varied. In other words, the SFR should be inversely proportional to this fraction. We can test this by comparing model *REF* to simulation *WML4*, which injects twice as much SN energy per unit stellar mass. Fig. 5 shows that while the SFR is indeed lower for *WML4*, the difference is always smaller than 0.25 dex, whereas we would have expected 0.30 dex (i.e. a factor of 2) at late times, when the SFR is dominated by galaxies that have had enough time to become self-regulating.

There are, however, good reasons why we would not expect the SFR to be exactly inversely proportional to the efficiency of the SN feedback. First, a change in the SFR does not only change the rate of energy injection into winds, but it also changes the rate of metal injection and the rate at which mass loss from AGB stars supplies the galaxy with fresh fuel for SF. Both of these effects would, however, tend to lower the amount of SN energy that is needed for self-regulation, which means that we would expect the SFR to decrease faster than linear with the SN efficiency. This is opposite to what is actually happening. Secondly, a reduction in the SFR implies a reduction in the stellar mass and hence in the gravitational force. Again, this would imply that less SN energy is required for self-regulation, which would lead to an even larger drop in the SFR, contrary to what the simulations predict.

The reason why the SFR varies more slowly with the SN efficiency than the inverse proportionality we would naively expect is likely that SN winds are not effective in high-mass galaxies, at least when kinetic feedback is used with a velocity of 600 km s^{-1} . If the feedback is inefficient, then we would not expect the galaxies to be able to self-regulate. Indeed, we will show in Haas et al. (in preparation) that, at $z = 2$, the SFR in *WML4* is in fact half that of *REF* for haloes with total mass less than $10^{11} M_{\odot}$ and that the SN feedback becomes inefficient for higher halo masses.

4.8.1 Varying the parameters at constant wind energy

As discussed in Section 3.1.5, we inject the energy from SNe in kinetic form. Newly formed star particles kick their gaseous neighbours with a constant velocity v_w in a random direction. On average, the mass kicked is η times the mass of the star particle. While the product ηv_w^2 determines the energy of the winds and is therefore constrained by the energy available from SNe, it is not clear a priori what values should be chosen for the individual parameters. Note that they cannot be taken directly from observations because the parameter values refer to the properties of the wind at the inter-particle distance (neighbours of new stars) which varies and will typically not agree with the scales relevant for the observations. The observational constraints are usually inferred from the velocity offset and

column densities of blueshifted absorption lines, but it is unclear at what distance from the source the absorption occurs (e.g. Veilleux et al. 2005). Moreover, the absorption lines probe only the cold part of the outflow. The constraints on the velocity and mass loading of the hot wind are also very poor.

Given the lack of observational constraints, one would hope that the results are insensitive to the amount of mass that a fixed amount of SN energy is distributed over. Dalla Vecchia & Schaye (2008) showed that this is indeed the case for the SFRs provided the galaxies are well resolved and the wind velocity exceeds a critical value that increases with the pressure of the ISM and thus also with halo mass. If, however, the wind velocity is too low, then the wind particles are immediately stopped by drag forces and never leave the ISM. If the discs are unresolved, then the hydrodynamic drag is underestimated and for sufficiently low resolutions all particles that are kicked are able to escape the ISM.

The runs with a top-heavy IMF in starbursts (Fig. 13) show that SF is much more efficiently suppressed if the extra SN energy (relative to a Chabrier IMF) is used to increase the wind velocity v_w than if it is used to increase the mass loading factor η . Since in these models the feedback energy is only boosted in high-pressure gas, this suggests that the wind velocity of 600 km s^{-1} that was used in the reference model is insufficient at the high pressures that we required for the IMF to become top-heavy. This implies that our default prescription for SN feedback is inefficient in high-mass galaxies, which could account for the fact that the SFR drops off less rapidly at late times than is observed.

To further investigate the dependence on the two individual wind parameters, we have run a series of simulations that all inject the same amount of SN energy per unit stellar mass as the reference model, but assume mass loading factors that differ by factors of 2, ranging from 1 to 8 in the $25 h^{-1} \text{ Mpc}$ box ($v_w \propto \eta^{-1/2}$ varies from 848 to 300 km s^{-1}) and from 1 to 2 in the $100 h^{-1} \text{ Mpc}$ box (v_w varies from 848 to 600 km s^{-1}). Fig. 14 compares the SFHs of these runs. Clearly, the results are not just determined by the total energy injected into the wind, which is identical for all the runs. At high redshifts the SFHs are similar, although the feedback is slightly more efficient for higher values of η . However, at late times the different SFHs start to diverge, with higher wind velocities suppressing the SF more strongly.

These results are consistent with the conclusions of Dalla Vecchia & Schaye (2008). As the universe evolves, stars typically form in more massive haloes and the minimum wind velocity for which the feedback remains effective thus increases. Hence, the redshift for which the feedback becomes inefficient decreases with increasing wind velocity. At early times, when many stars form in haloes that are poorly resolved, higher mass loading factors are more efficient because all particles that are kicked from poorly resolved haloes are able to escape the ISM.

The SFH, including the redshift at which it peaks, is clearly sensitive to the poorly constrained parameters η and v_w . The same is likely to be true for other types of subgrid prescriptions than kinetic feedback. Thus, unless one varies the parameters of the wind model, which is unfortunately not always done in the literature, one risks overinterpreting the results.

If the main goal were to reproduce the observed SFH and if one were willing to accept the lack of ‘ab initio predictive power’ displayed by Fig. 14, then one could choose to take an approach similar to that of semi-analytic models and take advantage of the sensitivity of the results to the wind parameters. By varying the parameters with halo mass or with the physical properties of the star-forming gas, one could match a wide range of SFHs.

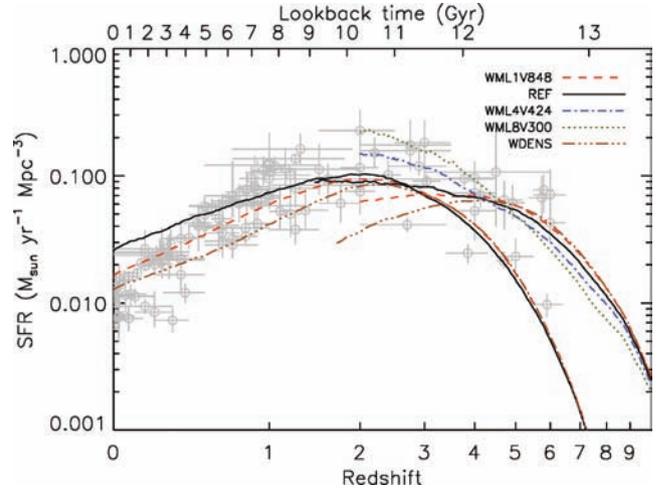


Figure 14. As Fig. 5, but comparing the SFHs for models that all inject the same amount of SN energy per unit stellar mass (i.e. ηv_w^2 is constant), but that assume different wind velocities. Models $WMLxVyyy$ assume a wind mass loading $\eta = x$ and velocity $v_w = yyy \text{ km s}^{-1}$. Model $WDENS$ assumes that the wind velocity scales with the local sound speed, which implies $v_w \propto \rho^{1/6}$ for our EOS, normalized to the value used in the reference model at the SF threshold. At high redshift, when the SFR is dominated by poorly resolved, low-mass haloes, the SN feedback is more efficient if it is distributed over more mass. However, at low redshift the situation is reversed. This can be explained if the feedback becomes inefficient when the velocity falls below a critical value that increases with galaxy mass. Scaling the wind parameters with local properties, such as the density for model $WDENS$, can help to keep the feedback efficient, but it also makes the results sensitive to the resolution.

While we have not tried to tune the SFH, we have investigated a toy model that uses the same amount of SN energy as the reference model, but in which the wind velocity scales with the local effective sound speed, $c_{s,\text{eff}}$, as might be the case for thermally driven winds. If it is indeed hydrodynamic drag that stalls low-velocity winds, then this scaling could keep the winds efficient at all pressures. We implement this model, which we term $WDENS$, by making the wind parameters functions of the density of the gas from which the star particle formed:

$$v_w = v_w^* \left(\frac{n_H}{n_H^*} \right)^{1/6}, \quad (5)$$

$$\eta = \eta_* \left(\frac{v_w}{v_w^*} \right)^{-2} = \eta_* \left(\frac{n_H}{n_H^*} \right)^{-1/3}, \quad (6)$$

which implies $v_w \propto c_{s,\text{eff}}$ since star-forming gas follows the effective EOS $P = \rho_g c_{s,\text{eff}}^2 \propto \rho_g^{4/3}$. We set $v_w^* = 600 \text{ km s}^{-1}$ and $\eta_* = 2$, so that the values of the wind parameters agree with those of the reference model for stars formed at the density threshold n_H^* , while the wind velocity is greater (and the mass loading smaller) at higher pressures.

Comparing the $25 h^{-1} \text{ Mpc}$ runs shown in Fig. 14, we see that $WDENS$ predicts a nearly identical SFH as $WML1V848$ and REF down to $z = 4$, but that it generates much more efficient winds at later times. The $100 h^{-1} \text{ Mpc}$ shows qualitatively the same behaviour, with $WDENS$ predicting significantly lower SFRs below $z = 2$. Because the winds in $WDENS$ remain effective for higher galaxy masses, the drop in the SFR below redshift 2 is steeper than for the reference model, although it is still less steep than observed.

Comparing the two $WDENS$ runs, we see that the agreement between the different box sizes is much worse than for the reference

model. Clearly, making SN feedback a function of the local gas density increases the sensitivity to numerical resolution as we already concluded from the models that used a top-heavy IMF at high densities (see Section 4.7.2). This is probably because the high-density tail of the PDF can only be sampled if the galaxy contains a sufficient number of particles. One would therefore expect somewhat better convergence if the wind parameters were a function of the properties of the dark matter halo rather than the local gas properties. We will investigate such models in Section 4.9.

4.8.2 Hydrodynamically decoupled winds

In recent years, a large fraction of the results from cosmological, SPH simulations discussed in the literature were obtained from simulations run with *GADGET2* (Springel 2005) and employing the Springel & Hernquist (2003a) prescription for kinetic SN feedback. This prescription for galactic winds differs in two respects from ours. First, the wind particles are selected stochastically from all the star-forming (i.e. dense) particles in the simulation and are therefore not local to the star particles as is the case for us. Secondly, the wind particles are subsequently decoupled from the hydrodynamics for 50 Myr (i.e. 31 kpc if travelling at 600 km s^{-1}) or until their density has fallen below 10 per cent of the threshold for SF, which ensures that they escape the ISM.

Dalla Vecchia & Schaye (2008) investigated the effects of this decoupling in detail and found them to be dramatic. While decoupled winds remove fuel for SF, they cannot blow bubbles in the disc, drive turbulence or create channels in gas with densities typical of the ISM. Decoupled winds are less efficient at suppressing SF in low-mass galaxies, because they cannot drag gas along. They are, however, much more efficient for high-mass galaxies, because they do not suffer the large energy losses due to drag in the high-pressure ISM. As the numerical resolution is decreased, the disc responsible for the drag disappears and the two prescriptions tend to converge. Hence, the decoupled winds are less sensitive to resolution, essentially because they only need to resolve the Jeans scales at the relatively low density for which the wind particles are recoupled.

Fig. 15 shows the SFH predicted by a model in which the wind particles were decoupled from the hydrodynamics in the manner described in Springel & Hernquist (2003a) (*WHYDRODEC*; note that the wind particles were, however, still local to the newly formed star particles). Compared with the reference model, the decoupled winds are less efficient at high redshift, because the wind particles cannot drag other gas particles out of low-mass galaxies. We expect this difference to increase for higher resolution simulations, because they can resolve the discs of such galaxies better. Decoupled winds are, however, much more effective at lower redshifts when the galaxies dominating the SFR are better resolved and more massive, leading to large energy losses due to drag in the ISM for the reference model. Consequently, the peak in the SFR shifts from $z \sim 2$ for the reference run to $z \approx 4$ for the decoupled winds.

4.8.3 Thermal SN feedback

In the previous sections we found that the results of simulations employing kinetic SN feedback are sensitive to the parameters of the subgrid model, even if the total energy in the wind is kept constant. It is therefore of interest to consider alternatives to kinetic feedback. In particular, thermal feedback would seem a natural choice as it would allow the simulation itself to determine the properties of the wind such as the mass loading, velocity and geometry, based on

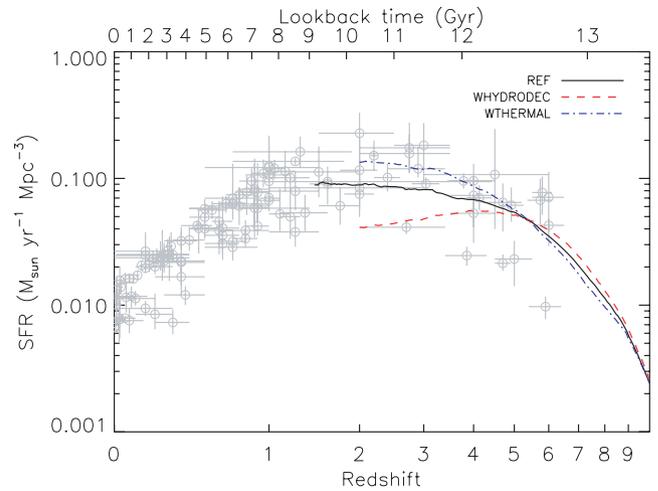


Figure 15. As Fig. 5, but comparing the SFHs for models that use different implementations of SN feedback. Model *WHYDRODEC* is identical to the reference model, except that the wind particles are temporarily decoupled from the hydrodynamics as in Springel & Hernquist (2003a). Model *WTHERMAL*, on the other hand, injects the SN energy in thermal form following the prescription of Dalla Vecchia & Schaye (2009). Hydrodynamically decoupled wind particles can freely escape the ISM, but are unable to drag other particles along. They are therefore less efficient at high redshift, when low-mass galaxies dominate the SFR, but they are much more efficient at low redshift, when hydrodynamical drag within the high-pressure ISM of massive galaxies is important. Injecting the same amount of SN energy in thermal form increases the efficiency of the feedback in poorly resolved, low-mass galaxies, but the winds are somewhat less effective at low redshift, i.e. for higher mass galaxies.

the properties of the starburst. Unfortunately, current simulations of cosmological volumes cannot resolve the energy-conserving phase in the evolution of SN remnants, causing any thermal energy input to be mostly radiated away before it can be converted into kinetic form. This is why alternatives such as kinetic feedback have been developed in the first place.

As discussed in Dalla Vecchia & Schaye (2009), this overcooling problem is mainly caused by the fact that in naive implementations of thermal feedback the ratio of the heated mass to that of the star particle is too large, which means that the temperature jump is too small and hence that the radiative cooling times are too short. The problem can thus be overcome if the SN energy produced by a star particle is injected in a sufficiently small amount of mass, such that its cooling time becomes long compared to the time-scale on which the local gas density can change in response to the energy injection. Dalla Vecchia & Schaye (2009) propose a stochastic method, which generalizes a prescription introduced by Kay, Thomas & Theuns (2003), that uses the temperature increase of the heated gas, ΔT , and the fraction of the SN energy that is injected, f_{th} , as parameters. These parameters then determine the probability for a gas element neighbouring a newly formed star particle to be heated. Dalla Vecchia & Schaye (2009) demonstrated that as long as the temperature increase is greater than a value that depends weakly on the resolution and the gas density, the factor by which the feedback suppresses SF is insensitive to the value of the temperature increase.

We have performed a run, *WTHERMAL*, in the $25 h^{-1} \text{ Mpc}$ box using the thermal feedback prescription of Dalla Vecchia & Schaye (2009), setting $\Delta T = 10^{7.5} \text{ K}$ and $f_{\text{th}} = 0.4$. The latter value agrees with the one implied by the product ηv_w^2 of the parameters of the kinetic feedback used in the reference run. Substituting these

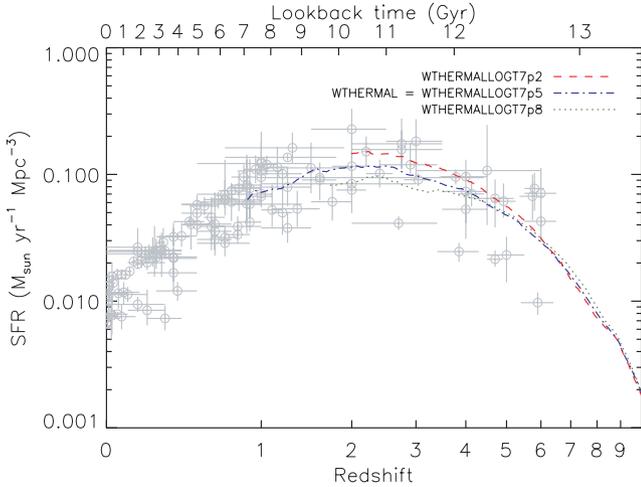


Figure 16. As Fig. 5, but comparing the SFHs for models that all employ thermal SN feedback using the method of Dalla Vecchia & Schaye (2009). All models inject 40 per cent of the available SN energy per unit stellar mass, but they use different temperature jumps for the heated gas. Model *WThermalLogT_{xpy}* uses temperature jumps $\log_{10} \Delta T = x \cdot y$ (e.g. *LOGT7p5* implies $\Delta T = 10^{7.5}$ K). Model *WThermalLogT7p5* assumed identical parameters as model *WThermal* shown in Fig. 15. All simulations use a $12.5 h^{-1}$ Mpc box and 2×256^3 particles. Enforcing a greater temperature jump, which implies a smaller heating probability, makes the feedback less efficient at high redshift, but more efficient at low redshift. However, the feedback efficiency is less sensitive to the temperature jump than it is to the wind velocity for the case of kinetic SN feedback.

parameter values into the equations presented in Dalla Vecchia & Schaye (2009), we estimate that for our resolution the thermal feedback will be efficient up to at least the density $n_H \sim 2 \times 10^2 \text{ cm}^{-3}$, which exceeds our SF threshold by more than three orders of magnitude. Per star particle, the average number of gas particles that receive SN energy is about 0.54.

Fig. 15 shows that the thermal feedback, when implemented in this manner, is indeed effective at suppressing the SFR. At high redshift it is more efficient than the kinetic feedback used in the reference model, which suggests that it results in higher mass loading factors for poorly resolved galaxies. For $z < 5$ the SFR is higher than in the reference run, but the difference is always less than 0.2 dex and remains constant below redshift 3.5.

Fig. 16 shows the SFHs for three different models that all employ thermal feedback using identical amounts of SN energy ($f_{\text{th}} = 0.4$), but different temperature jumps. From top to bottom at redshift 2, the models use $\Delta T = 10^{7.2}$, $10^{7.5}$ and $10^{7.8}$ K, respectively. Hence they differ by factors of 2, which matches the factors of $\sqrt{2}$ difference between the wind velocities used in Fig. 14. These three thermal feedback models use 2×256^3 particles in a $12.5 h^{-1}$ Mpc box, which means that the resolution is identical to that used in the *L025N512* runs shown in Fig. 14.

Comparing the three models, we see that a higher temperature increase makes the feedback slightly less efficient at high redshift ($z > 6$), although the effect is marginal. Interestingly, the small difference at $z > 6$ appears to be caused by the varying strength of the response to the reheating associated with reionization (which happens at $z_r = 9$ in our simulations). This agrees with the finding of Pawlik & Schaye (2009) that SN feedback and photoheating strengthen each other. Near the end of the simulations, the feedback is more efficient for greater ΔT . The redshift at which $\Delta T = 10^{7.8}$ K becomes more efficient than $\Delta T = 10^{7.5}$ K is lower than the

redshift at which the latter model becomes more efficient than $\Delta T = 10^{7.2}$ K. All of this can be explained if the thermal feedback becomes inefficient for haloes more massive than some value which increases with ΔT . This situation parallels that of the kinetic feedback with ΔT playing the role of v_w .

Interestingly, the SFH is much less sensitive to the value of ΔT than it is to v_w for the case of kinetic SN feedback. The difference between the models with $\Delta T = 10^{7.2}$ and $10^{7.8}$ K is always less than 0.24 dex, whereas the difference between the models with $v_w = 424$ and 848 km s^{-1} is 0.38 dex at $z = 2$ (see Fig. 14). Moreover, while the SFHs in the different kinetic feedback models diverge rapidly, the differences between the thermal models is nearly constant below $z = 3$.

Our findings that the thermal feedback is efficient and that it is less sensitive to the parameters of the model than is the case for kinetic feedback are very encouraging. We only used 40 per cent of the SN energy because we wanted to match the fraction used for the kinetic feedback in the reference model. Higher values can, however, easily be justified for thermal feedback since we are now simulating radiative losses.

4.9 ‘Momentum-driven’ winds

We have so far only considered galactic winds driven by SN feedback. It is, however, also possible that outflows are driven by the momentum that is deposited by photons from massive stars as they are absorbed by dust grains (Murray, Quataert & Thompson 2005). In such a momentum-driven wind one would expect the mass loading to be inversely proportional to the wind velocity and, according to the model of Murray et al. (2005), the terminal wind velocity would be similar to the velocity dispersion of the host galaxy. Martin (2005) argued that such scalings are in agreement with estimates of the mass and velocity in cold clouds as inferred from Na I absorption lines. Oppenheimer & Davé (2006) implemented several versions of momentum-driven winds into cosmological simulations of the chemical enrichment of the IGM and found good agreement with observations of C IV absorption.

Direct simulation of radiation pressure requires radiative transfer, which is too costly for cosmological simulations. Oppenheimer & Davé (2006) therefore chose to use the Springel & Hernquist (2003a) recipe for kinetic feedback, including the hydrodynamic decoupling discussed in Section 4.8.2, but to make the velocity kick, v_w , and the mass loading factor, η , functions of the properties of the galaxy so as to mimic the scalings expected for momentum-driven winds. Specifically, they assumed

$$v_w = \left(a_1 \sqrt{a_2 f_L(Z) - 1} + a_3 \right) \sigma, \quad (7)$$

$$\eta = a_4 / \sigma, \quad (8)$$

where a_1 – a_4 are free parameters, σ is the galaxy velocity dispersion and $f_L(Z)$ is a function that accounts for the dependence of the stellar luminosity on metallicity and that varies from 1.7 at $10^{-3} Z_{\odot}$ to unity for solar abundances. Parameter a_1 was set to 3 which is the value suggested by Murray et al. (2005). Parameter a_2 was either set to 2 or varied randomly between 1.05 and 2. Parameter a_3 was introduced after noting that for radiatively driven winds the outflow velocity increases out to large distances, whereas in the simulations the gas is no longer given momentum after it is ‘kicked’ out of the ISM. They tried both $a_3 = 0$ and $a_3 = 2$. Finally, parameter a_4 was set to 300 km s^{-1} in order to roughly match the observed cosmic SFR at high redshift. The velocity dispersion of the host galaxy was

estimated by taking the local gravitational potential Φ and estimate σ using the virial theorem ($\sigma = \sqrt{-\frac{1}{2}\Phi}$).

Later papers by the same authors compared to other types of observations, but used different parameter values and methods. Oppenheimer & Davé (2008) realized that the gravitational potential is dominated by large-scale structure rather than by the mass of individual haloes. For this reason, they moved away from using the gravitational potential to estimate wind properties and instead used friends-of-friends halo catalogues, generated on-the-fly throughout the simulation and set $\sigma = \sqrt{2}v_c$, where $v_c = \sqrt{GM/R_{\text{vir}}}$ is the halo circular velocity and R_{vir} is the virial radius. Oppenheimer & Davé (2008) also increased a_1 from 3 to 4.3, halved the value of a_4 to 150 km s^{-1} and imposed an upper limit on the wind velocity corresponding to twice the total SN energy. Later papers used the same values as Oppenheimer & Davé (2008) although Oppenheimer & Davé (2009) no longer imposed a limit on the total wind energy.

We first implemented the Oppenheimer & Davé (2006) method, which uses the local potential to estimate the velocity dispersion, but without decoupling the wind particles from the hydrodynamics. We used the parameter values advocated by Oppenheimer & Davé (2008), although we neglected the metallicity-dependent square root term, which simplifies equation (8) to $v_w = (a_1 + a_3)\sigma$. We ran two versions. While both assumed $a_1 = 3$ and $a_4 = 150 \text{ km s}^{-1}$, model *WPOTNOKICK* used $a_3 = 0$ whereas model *WPOT* used $a_3 = 2$. Fig. 17 shows the SFHs predicted by the models for both box sizes. *WPOT* gives lower SFRs than *WPOTNOKICK* which is not surprising since it uses higher wind velocities. Model *WPOTNOKICK*

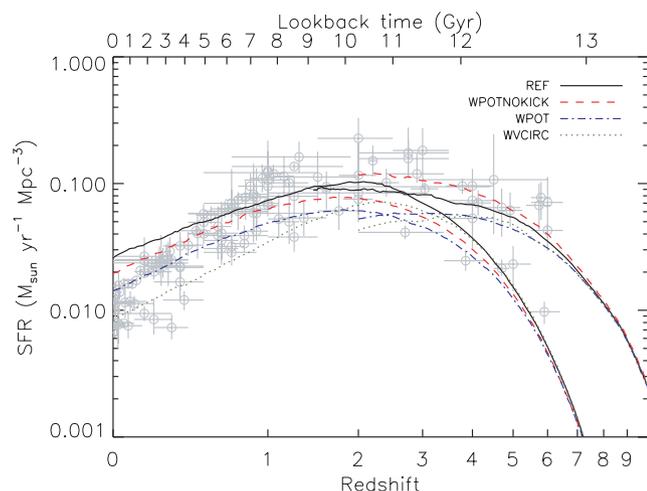


Figure 17. As Fig. 5, but comparing the SFHs for different implementations of ‘momentum-driven’ winds, following Oppenheimer & Davé (2006, 2008). While these models were motivated by the claim that galactic outflows may be driven by radiation pressure on dust grains (Murray et al. 2005), they do not actually include radiative transfer. Instead, the parameters of the kinetic feedback prescription are made functions of either the local potential (*WPOT* and *WPOTNOKICK*) or the mass of the dark matter halo (*WVCIRC*). The wind parameters scale such that the wind velocity increases with halo mass, while the mass loading is inversely proportional to the wind velocity. For massive galaxies, the winds use more energy than is available from SNe. See the text for additional details. Because the wind velocity increases with halo mass, the feedback is more efficient than for the reference model at late times. At high redshift it is also somewhat more efficient, thanks to the higher mass loading factors.

predicts higher SFRs than the reference model for the $25 h^{-1} \text{ Mpc}$ box, but the order is reversed for the $100 h^{-1} \text{ Mpc}$ box. This reflects the fact that there is more large-scale structure in the larger box and that the potential is dominated by the largest structures in the box. Clearly, this situation is not desirable.

We therefore also ran model *WVCIRC* which estimates σ using an on-the-fly halo friends-of-friends halo finder as in Oppenheimer & Davé (2008).¹⁰ The halo finder was run at times spaced evenly in $\log a$ with $\Delta a = 0.02a$, where a is the expansion factor. Haloes were found based on the distribution of dark matter particles using a linking length of 0.2 and requiring a minimum of 25 dark matter particles per halo.¹¹ Baryonic particles were attached to the nearest dark matter particle.

The predicted SFHs are also shown in Fig. 17. Except at very high redshift, the SFR is strongly reduced compared to the reference model. The difference increases with time so that the SFR falls off more rapidly below $z = 2$ than for *REF*, as required by the observations. As was the case for *WDENS* (see Section 4.8 and Fig. 14), the increased efficiency of the feedback at late times arises because the wind velocity increases with the halo mass, whereas it is constant for the reference model. Contrary to *WDENS*, the feedback is also more efficient than that of the reference model at high redshift. This is a consequence of the high-mass loading used for low-mass haloes (recall that for *WDENS*, the mass loading is never higher than for *REF*).

The implementation of momentum-driven winds is rather crude. For example, the wind reaches its maximum velocity when it leaves the ISM rather than in the outer halo as expected for radiatively driven winds. Moreover, the parametrization leaves a lot of freedom, even more than for SN feedback, partly because the total amount of energy is no longer limited.¹² More to the point, Haas et al. (in preparation) show that the models inject much more momentum than is actually available in the form of star light. It would therefore be dangerous to use comparisons between observations and models such as these to discriminate between outflows driven by SNe and radiation pressure. Moreover, we note that it is not clear that observations of high-mass galaxies should agree with the ‘momentum-driven’ models, given that AGN feedback is thought to be crucial for such objects (e.g. Bower et al. 2006; Croton et al. 2006), but was not included. Conversely, if simulations such as the ones presented here were to disagree with observations, it would not necessarily mean that winds are not radiatively driven.

It is interesting that the scalings of the ‘momentum-driven’ prescription, i.e. a wind velocity that increases and a mass loading that decreases with galaxy mass, agree qualitatively with the results obtained for high-resolution simulations of energy-driven SN feedback (Dalla Vecchia & Schaye 2008, 2009). Even when hydrodynamical interactions are not temporarily ignored, the mass loading will be underestimated for galaxies that are poorly resolved (Dalla Vecchia & Schaye 2008). It may therefore be that the ‘momentum-driven’ wind scaling partly compensates for some of the resolution effects that plague cosmological simulations, particularly at high redshift.

¹⁰Note that Oppenheimer & Davé (2008) ran a friends-of-friends halo finder on the baryons using a linking length of 0.04 and assuming a fixed baryon to dark matter ratio equal to the universal value.

¹¹For star particles forming outside of any halo, we assume the minimum halo mass corresponding to 25 dark matter particles.

¹²Depending on redshift, *WVCIRC* injects more energy in the winds per unit stellar mass formed than model *REF* for halo masses that exceed $10^{11} - 10^{12} M_{\odot}$.

4.10 AGN feedback

The centres of galaxies are thought to harbour supermassive BHs. Matter accreting on to these BHs emits large amounts of high energy radiation. Even if only a small fraction of this energy gets coupled to the ISM, it could have a dramatic effect on the galaxies. Moreover, the magnetic fields carried by the accreting matter could lead to the formation of jets which can displace and heat gas in and around galaxies. Feedback from AGN has, for example, been invoked to explain the low SFRs of high-mass galaxies (e.g. Di Matteo, Springel & Hernquist 2005; Bower et al. 2006; Croton et al. 2006; Booth & Schaye 2009a) and the suppression of cooling flows in clusters of galaxies (e.g. Churazov et al. 2001; Dalla Vecchia et al. 2004; Ruszkowski, Brüggel & Begelman 2004; Cattaneo & Teyssier 2007).

To investigate the effect of AGN feedback, we have run a series of simulations employing the subgrid prescription for the growth of BHs and feedback from AGN described in Booth & Schaye (2009a) which is a substantially modified version of the model of Springel, Di Matteo & Hernquist (2005a). Below we will briefly summarize the main features of this model, but we refer the reader to Booth & Schaye (2009a) for further details and tests.

Seed BHs of mass m_{seed} are placed into every dark matter halo whose mass exceeds $m_{\text{halo, min}}$. Our fiducial AGN model uses $m_{\text{seed}} = 9 \times 10^4 M_{\odot}$, which corresponds to $10^{-3} m_{\text{g}} (6.4 \times 10^{-2} m_{\text{g}})$ in the $100 h^{-1} \text{Mpc}$ ($25 h^{-1} \text{Mpc}$) box. For the minimum dark halo mass we use $m_{\text{halo, min}} = 4 \times 10^{10} M_{\odot}$, which corresponds to $10^2 (6.4 \times 10^3)$ dark matter particles in the $100 h^{-1} \text{Mpc}$ ($25 h^{-1} \text{Mpc}$) box. Haloes are identified by running a friends-of-friends group finder on-the-fly as described in Section 4.9. BHs can grow via Eddington-limited accretion of the surrounding gas and through mergers with other BHs.

Booth & Schaye (2009a) show that any simulation that resolves the Jeans scales will also resolve the Bondi–Hoyle accretion radius for BHs whose mass exceeds the simulation’s mass resolution. At densities below the SF threshold n_{H}^* (10^{-1}cm^{-3} in our model), the gas is kept warm by the ionizing background (Schaye 2004) and we marginally resolve the Jeans scales in our highest resolution runs (see Section 3.2). For higher densities, however, a cold phase is expected to be present and naive application of the Bondi–Hoyle formula would lead us to strongly underestimate the accretion rate. We therefore assume that the accretion rate is given by the minimum of the Eddington rate and

$$\dot{m}_{\text{accr}} = \alpha \frac{4\pi G^2 m_{\text{BH}}^2 \rho}{(c_s^2 + v^2)^{3/2}}, \quad (9)$$

where m_{BH} is the mass of the BH, c_s and ρ are the sound speed and density of the local medium, respectively, v is the velocity of the BH relative to the ambient medium and α is a dimensionless efficiency parameter given by¹³

$$\alpha = \begin{cases} 1 & \text{if } n_{\text{H}} < n_{\text{H}}^* \\ \left(\frac{n_{\text{H}}}{n_{\text{H}}^*}\right)^{\beta} & \text{otherwise.} \end{cases} \quad (10)$$

Observe that for $\alpha = 1$ (i.e. $n_{\text{H}} < n_{\text{H}}^*$ or $\beta = 0$), equation (9) reduces to the Bondi–Hoyle–Lyttleton rate (Hoyle & Lyttleton 1939; Bondi & Hoyle 1944). Our fiducial AGN runs use $\beta = 2$, which results in

¹³Note that Springel et al. (2005a) used a fixed value $\alpha = 100$. Consequently, massive BHs need to suppress the ambient gas density to values far below n_{H}^* in order to reduce the accretion rate to sub-Eddington values (see Booth & Schaye 2009a).

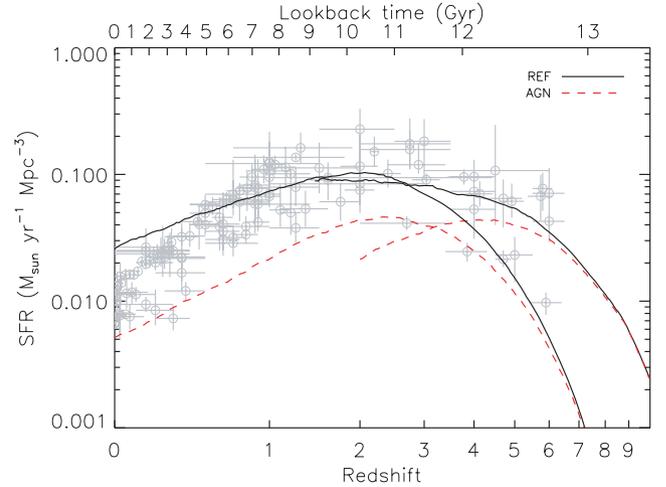


Figure 18. As Fig. 5, but comparing the SFHs for models with and without supermassive BHs. Feedback from AGN strongly suppresses the SFR in massive galaxies, which becomes more important at late times. When AGN are included the SFR appears to be too low compared with observations, but that could be changed by decreasing the fraction of the SN energy that is injected, which was in this case chosen to roughly match the peak of the SFH in the absence of AGN feedback.

efficient BH growth in haloes with stellar masses of $\gtrsim 10^{10.5} M_{\odot}$ in the $L100$ runs.

The amount of accreted mass is related to the rate of growth of the BH by $\dot{m}_{\text{BH}} = \dot{m}_{\text{accr}}(1 - \epsilon_r)$, where ϵ_r is the radiative efficiency of a BH, which we always assume to be 10 per cent, the mean value for the radiatively efficient Shakura & Sunyaev (1973) accretion on to a Schwarzschild BH.

We assume that a fraction ϵ_f of the radiated energy couples to the ISM. The amount of energy returned by a BH to its surrounding medium is thus given by

$$\dot{E}_{\text{feed}} = \epsilon_f \epsilon_r \dot{m}_{\text{accr}} c^2, \quad (11)$$

where c is the speed of light. We set $\epsilon_f = 0.15$ in order to match the observed cosmic mass density in BHs as well as the relation between BH and galaxy mass, both at redshift zero. BH particles store feedback energy until it suffices to heat n_{heat} of their neighbours by ΔT_{min} . The two parameters ΔT_{min} and n_{heat} are chosen such that AGN-heated gas obtains a long cooling time and so that the time taken to perform a feedback event is shorter than the Salpeter time for Eddington-limited accretion. The parameter choices $\Delta T_{\text{min}} = 10^8 \text{K}$ and $n_{\text{heat}} = 1$ are found to provide a good balance between these two constraints.

Fig. 18 shows that the addition of AGN feedback strongly suppresses the SFR. The difference with the reference model increases with time, which implies that AGN feedback is more important for higher mass galaxies. The drop in the SFR below $z = 2$ is much closer to the observed slope when AGN feedback is included, but the overall amplitude of the SFR is probably too low. It is, however, important to note that the observed SFRs are subject to large systematic uncertainties. Indeed, McCarthy et al. (in preparation) find that the predicted stellar masses are in fact in good agreement with observations of groups of galaxies.

Even if the simulation with AGN feedback really did form too few stars, it would not be a concern here. As discussed in Section 3.1.5, the fraction of the SN energy that is injected was chosen to roughly match the peak in the observed SFH. It is therefore unavoidable that including an extra form of efficient feedback, without making

any other changes, reduces the SFR to values that are lower than observed. It would have been possible to adjust the parameters of the prescription for SN feedback to obtain a better match to the observed SFH, but this is not the objective here.

Booth & Schaye (2009a) have carried out an extensive parameter study of the AGN model at the resolution of our $100 h^{-1}$ Mpc box, comparing the cosmic SFH as well as other observables. They found that the results are sensitive to the accretion model, i.e. the value of β , which sets the halo mass above which the BHs can grow on to the scaling relations. Remarkably, they found that the SFH is nearly completely independent of the feedback efficiency, but, in agreement with Di Matteo et al. (2005), the BH masses are inversely proportional to ϵ_f . They explained this in terms of self-regulation: the BHs grow until they have injected sufficient energy to balance the infall driven by gas accretion on to haloes and radiative cooling (see also Di Matteo et al. 2005; Robertson et al. 2006; Booth & Schaye 2009b). If half as much energy is injected per unit accreted mass, then the BHs need to grow twice as massive in order to inject the same amount of energy. Because the factor by which SF is suppressed depends on the amount of energy that is injected by the BHs, the SFH is insensitive to variations in the assumed efficiency of AGN feedback.

5 SUMMARY AND DISCUSSION

The cosmic SFH is perhaps the most fundamental observable in astrophysical cosmology. It is difficult to model, because of the large range of galaxy masses that contribute and because of the many feedback processes that may be important. Cosmological hydrodynamical simulations need to resort to subgrid prescriptions for the physics that remains unresolved, which makes them resemble semi-analytic models in some respects. It is, however, much more difficult to explore parameter space using fully numerical simulations because of the high computational cost.

Here, we have introduced the OWLS project, which consists of more than 50 large, cosmological, hydrodynamical simulations. The simulations were all run with a modified version of the SPH code GADGET3 (last described in Springel 2005), using new modules for radiative cooling (Wiersma et al. 2009a), SF (Schaye & Dalla Vecchia 2008), chemodynamics (Wiersma et al. 2009b), kinetic (Dalla Vecchia & Schaye 2008) or thermal (Dalla Vecchia & Schaye 2009) SN feedback, and accretion on to and feedback from supermassive BHs (Springel et al. 2005a; Booth & Schaye 2009a). With 2×512^3 particles, the OWLS runs are among the largest dissipative simulations ever performed. The simulations are repeated many times, each time changing a single aspect of the input physics or a single numerical parameter with respect to the reference model described in Section 3. We stress that this model merely functions as a reference point for our systematic exploration of parameter space and should therefore not be regarded as our ‘best’ model.

Generically, we find that SF is limited by the build-up of dark matter haloes at high redshift, reaches a broad maximum at intermediate redshift and then decreases as it is quenched by lower cooling rates in hotter and lower density gas, gas exhaustion and self-regulated feedback from stars and BHs, in broad agreement with previous work (e.g. White & Frenk 1991; Hernquist & Springel 2003). While we have limited ourselves to comparisons of the cosmic SFHs, future papers will investigate the behaviour of the models in other contexts.

The models are run down to redshift $z = 2$ in boxes of 25 comoving h^{-1} Mpc on a side and/or down to $z = 0$ in boxes of $100 h^{-1}$ Mpc. The mass (spatial) resolution is a factor of 64 (4) better in the smaller

box. Detailed convergence tests showed that a $25 h^{-1}$ Mpc box is sufficiently large to model the cosmic SFH down to $z = 0$. For the $25 h^{-1}$ Mpc ($100 h^{-1}$ Mpc) simulation of the reference model, the SFH is very close to converged for $z < 7$ ($z < 3$). Interestingly, we found that while the SFR typically increases if the resolution is improved, the situation reverses at low redshift before convergence has been attained. This calls into question the strategy to combine different box sizes to obtain a converged SFH all the way from high redshift to $z = 0$.

The cosmic SFR density can be decomposed into a halo mass function and the (distribution of the) SFR as a function of halo mass. The baryonic physics mainly affects the latter function, which we shall explore in more detail in Haas et al. (in preparation), while the mass function is determined mostly by the assumed cosmology. We showed that the SFH is sensitive to even relatively small changes in the cosmological parameters, such as the difference between the values inferred from the *WMAP* 3-year data used here and the earlier values assumed in the Millennium simulation. Clearly, the parameters of semi-analytic models build on to the Millennium simulation will have to be modified if they are used on a simulation assuming the current concordance cosmology. Comparisons to observations of rare objects and the high-redshift Universe will be particularly strongly affected due to the relatively large difference in the value of σ_8 .

Our systematic tests of the subgrid physics revealed that SF in intermediate-mass galaxies is highly self-regulated by feedback from massive stars. This explains our remarkable finding that the predicted SFH is nearly completely independent of the treatment of the unresolved ISM, including the assumed SF law. If the SF efficiency is increased, then the galaxies simply reduce their gas fractions so as to keep the SFR, and thus the rate at which stars inject energy into the ISM, constant. Similarly, if the efficiency of SN feedback is changed by injecting a different fraction of the SN energy, then, to first order, the galaxies simply adjust their SFRs so as to keep the rate of energy injection constant. The critical rate of energy injection that results from self-regulation is presumably the rate required to balance gas infall resulting from accretion on to haloes and radiative cooling and will therefore depend on the halo mass and redshift.

Note, however, that our findings apply to SFRs that have effectively been averaged over entire galaxies and over very long time-scales. Self-regulation may also occur on smaller length and time-scales. Indeed, we implicitly assume this to be the case through our use of empirical SF laws that have been averaged over spatial scales that are large compared to individual star-forming regions. Processes other than SN feedback may be important for the self-regulation that occurs on these smaller scales and perhaps even on large scales if SN feedback is inefficient (e.g. Robertson & Kravtsov 2008; Gnedin, Tassis & Kravtsov 2009).

For massive galaxies feedback from massive stars becomes inefficient and it is the self-regulated growth of supermassive BHs that ensures that a fixed amount of energy is injected into the ISM (e.g. Di Matteo et al. 2005; Booth & Schaye 2009a,b). As we showed in Booth & Schaye (2009a), the SFH is therefore independent of the assumed efficiency of AGN feedback. If the BHs inject twice as much energy per unit accreted mass, then they grow only half as massive so that they inject the same amount of energy and because they inject the same amount of energy, the SFR is suppressed by the same factor.

At very high redshift, the SFH is not yet controlled by self-regulation and will thus be limited by the gas consumption time-scale implied by the assumed SF law. In our simulations the

dependence on the SF law at high redshift may, however, mostly result from our limited resolution. Stars can only form in haloes that exceed the resolution limit and the formation of the first generation of stars within a halo is not hampered by winds. Because the minimum halo mass is set by the resolution, so is the SFH until it is dominated by haloes that are significantly more massive.

Radiative feedback from non-local SF is, however, very important at high redshift. In particular, reheating associated with hydrogen reionization quenches SF in haloes with virial temperatures of $\lesssim 10^4$ K and thus strongly affects the SFH when it would otherwise be dominated by such low-mass haloes. Our simulations underestimate this effect due to their limited resolution and because we assume that a photodissociating background is present at all redshifts. Contrary to hydrogen reionization, the milder reheating associated with helium reionization does not have a significant impact on the SFH because it is limited to the low-density IGM.

While cosmology and self-regulation via outflows are the principal ingredients controlling the SFH, there are other processes that are important. Metal-line cooling becomes very important at late times, as the metal content of the gas builds up and less gas falls in cold. The inclusion of cooling by heavy elements has two important effects. First, it allows more virialized gas to cool, an effect that has been widely discussed in the literature. However, we find that the dominant effect of metal-line cooling may be that it reduces the efficiency of galactic winds, as it increases thermal losses in the gas that has been shock-heated by the winds. The inclusion of metal-line cooling makes it much more difficult to reproduce the steep decline in the observed SFH below $z = 2$.

Another process that becomes important at late times is mass loss by intermediate-mass stars. On time-scales of hundreds of millions to billions of years, winds from AGB stars boost the SFR by providing fresh fuel for SF and by releasing metals that were locked up in stars into the ISM. As was the case for metal-line cooling, including mass loss by AGB stars makes it more difficult to match the sharp drop in the cosmic SFR that is observed at low redshift. The shape of the time-delay function for SNe Ia, on the other hand, turned out to be unimportant, presumably because its impact is limited to a shift in the timing of the release of a fraction of the iron.

Without AGN feedback, it is challenging to match the steep decline in the cosmic SFR below $z = 2$. It is, however, not yet totally clear that feedback from massive stars cannot solve the problem. If injected in kinetic form, SN feedback becomes inefficient once the wind velocity falls below some critical value that increases with galaxy mass and thus with the pressure in the ISM (Dalla Vecchia & Schaye 2008). By increasing the input wind velocity with galaxy mass, or with some correlated local property (e.g. gas density, pressure or velocity dispersion), while decreasing the input mass loading so as to keep the wind energy constant, we can keep SN feedback efficient in higher mass galaxies.

Using a sufficiently high constant wind velocity also results in efficient suppression of SF in relatively massive galaxies, but not for poorly resolved low-mass galaxies. If a galaxy is poorly resolved, then there is no disc from which wind particles can drag gas along, limiting the effective mass loading factor to the input value (Dalla Vecchia & Schaye 2008). Decreasing the input wind velocity (and thus increasing the input mass loading) with decreasing galaxy mass counteracts the resulting underestimate of the mass loading factor. If we allow ourselves to vary the parameters of the kinetic feedback with local properties, then we can even ‘design’ SFHs, but this freedom comes at the expense of introducing additional free parameters and an increased sensitivity to the numerical resolution.

Recently, the possibility that galactic outflows are driven by radiation pressure on dust grains has generated considerable interest (Murray et al. 2005). For such winds the mass loading is expected to be inversely proportional to the wind velocity, which, in the outer halo, is expected to be similar to the galaxy velocity dispersion. Oppenheimer & Davé (2006, 2008) implemented such ‘momentum-driven’ winds in a simplified fashion by kicking particles out of the ISM with several times the velocity that the wind is expected to reach in the outer halo and by tuning the normalization of the initial wind mass loading to match the observed SFR. We ran several simulations that employed such methods and found that the feedback is more efficient than our standard SN feedback, particularly at low redshift when relatively massive galaxies dominate the SFR. This is partly because the winds are allowed to carry more energy than is available from SNe, but it may be mostly due to the fact that the ‘momentum-driven’ scalings happen to overcome some of the numerical effects discussed above. Given the simplified nature and the limitations of the kinetic feedback models, it would be dangerous to use them to discriminate between winds driven by SNe and radiation pressure. Finally, we note that Haas et al. (in preparation) demonstrate that the amount of momentum that is injected in the models of Oppenheimer & Davé (2006, 2008), and thus also in our versions of these models, exceeds the amount of momentum that is actually available in the form of star light by up to an order of magnitude.

Besides the values of the wind parameters, the implementation of kinetic SN feedback is also important. For example, most studies in the literature employing GADGET temporarily decouple wind particles from the hydrodynamics, allowing them to freely escape the ISM without blowing bubbles or generating turbulence. We found that this decoupling (which we included in one of our runs) reduces the efficiency of SN feedback at high redshift, because of the inability of the winds to drag gas along, but strongly increases it at low redshift when hydrodynamic drag would otherwise quench the winds in massive galaxies.

Outflows driven by feedback from SF, be it SN or radiation pressure, are very important. It is therefore unfortunate that the predictions for the SFH of simulations that implement outflows in the form of kinetic feedback, i.e. by kicking particles, are not robust at the currently attainable resolution. The predicted SFH is sensitive to the values of poorly constrained wind parameters, even for a constant wind energy, and to the details of the implementation. Note that the same may well be true for other types of subgrid prescriptions for feedback from SF. Clearly, it is also crucial to vary the parameters of such models. On the other hand, we will show in future papers that many observables are much less sensitive to the implementation of feedback from SF than is the case for the cosmic SFH.

We investigated implementing SN feedback in thermal form, using the method of Dalla Vecchia & Schaye (2009) to overcome the overcooling problem that is commonly encountered when thermal SN feedback is used. Encouragingly, we found thermal feedback to be efficient and the predictions to be less sensitive to the values of the free parameters than is the case for kinetic feedback. Relative to kinetic feedback, thermal feedback is particularly efficient at high redshift. Because heated particles push all their neighbours, the winds remain highly mass loaded even in poorly resolved galaxies.

Another large uncertainty is the stellar IMF. The consequences of a change in the assumed IMF are difficult to predict because the IMF affects many things. The IMF determines the effective nucleosynthetic yields, the fraction of the stellar mass that is recycled by intermediate-mass stars and the amount of SN energy (and radiation) produced by massive stars. Moreover, a change in the IMF

also implies a change in the observed SF law and history, as SFRs are inferred from observations of the light emitted by massive stars. These two latter effects are often ignored. Interestingly, we found that an IMF that becomes top-heavy at high gas pressures would improve agreement with the observations because it preferentially suppresses SF in massive galaxies.

How do we proceed from here? Based on our findings, a better understanding of the mechanisms responsible for the generation of galactic winds, as well as more robust numerical implementations, is crucial. Paradoxically, a better understanding of SF may not directly improve predictions for the cosmic SFH if, as is the case for the models studied here, galaxies regulate their SFRs by adjusting their gas fractions via large-scale outflows. At very high redshift, photoheating is key and simulations including radiative transfer would constitute a clear improvement. At lower redshifts, it is very important to include mass loss from AGB stars and metal-line cooling. Better treatments of metal mixing and the inclusion of non-equilibrium cooling would certainly be helpful here. The steep drop in the observed SFR below $z = 2$ is difficult to reproduce without AGN feedback. Fortunately, the self-regulatory nature of the growth of supermassive BHs makes predictions for the SFH insensitive to the assumed efficiency of AGN feedback. Finally, as always, higher resolution would be very helpful as it would, for example, enable us to probe further down the halo mass function and to include more realistic treatments of galactic winds.

ACKNOWLEDGMENTS

It is a great pleasure to thank Volker Springel without whose generous help this project would not have been possible. We are also grateful to Anthony Aguirre, Richard Battye, Stefano Borgani, Richard Bower, Kathy Cooksey, Rob Crain, Scott Kay, Julio Navarro, Benjamin Oppenheimer, Andreas Pawlik, Xavier Prochaska, Olivera Rakic, Philipp Richter, Laura Sales, Debora Sijacki, Thorsten Tepper-García, Luca Tornatore and all the members of the Virgo consortium for discussions and help. We would also like to thank the referee, Brant Robertson, for a constructive report. The simulations presented here were run on Stella, the LOFAR Blue-Gene/L system in Groningen, on the Cosmology Machine at the Institute for Computational Cosmology in Durham as part of the Virgo Consortium research programme and on Darwin in Cambridge. This work was sponsored by National Computing Facilities Foundation (NCF) for the use of supercomputer facilities, with financial support from the Netherlands Organization for Scientific Research (NWO). This work was supported by Marie Curie Excellence Grant MEXT-CT-2004-014112 and by an NWO VIDI grant. IGM acknowledges support from a Kavli Institute Fellowship at the University of Cambridge.

REFERENCES

Abel T., Haehnelt M. G., 1999, *ApJ*, 520, L13
 Arnouts S. et al., 2005, *ApJ*, 619, L43
 Ascasibar Y., Yepes G., Gottlöber S., Müller V., 2002, *A&A*, 387, 396
 Balogh M. L., Pearce F. R., Bower R. G., Kay S. T., 2001, *MNRAS*, 326, 1228
 Bartko H. et al., 2009, *ApJ*, 697, 1741
 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
 Birnboim Y., Dekel A., 2003, *MNRAS*, 345, 349
 Blanc G. A., Heiderman A., Gebhardt K., Evans N. J. II, Adams J., 2009, *ApJ*, 704, 842
 Bondi H., Hoyle F., 1944, *MNRAS*, 104, 273

Booth C. M., Schaye J., 2009a, *MNRAS*, 398, 53
 Booth C. M., Schaye J., 2009b, *MNRAS*, submitted (arXiv:9011.0935)
 Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, *ApJ*, 670, 928
 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
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Cattaneo A., Teyssier R., 2007, *MNRAS*, 376, 1547
 Chabrier G., 2003, *PASP*, 115, 763
 Chiosi C., Bertelli G., Bressan A., 1992, *ARA&A*, 30, 235
 Choi J.-H., Nagamine K., 2009, *MNRAS*, 393, 1595
 Churazov E., Brügggen M., Kaiser C. R., Böhringer H., Forman W., 2001, *ApJ*, 554, 261
 Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
 Couchman H. M. P., Rees M. J., 1986, *MNRAS*, 221, 53
 Cowie L. L., Barger A. J., 2008, *ApJ*, 686, 72
 Crain R. A. et al., 2009, *MNRAS*, 399, 1773
 Croton D. J. et al., 2006, *MNRAS*, 365, 11
 Dabringhausen J., Kroupa P., Baumgardt H., 2009, *MNRAS*, 394, 1529
 Dahlen T. et al., 2004, *ApJ*, 613, 189
 Dalla Vecchia C., Schaye J., 2008, *MNRAS*, 387, 1431
 Dalla Vecchia C., Schaye J., 2009, *MNRAS*, submitted
 Dalla Vecchia C., Bower R. G., Theuns T., Balogh M. L., Mazzotta P., Frenk C. S., 2004, *MNRAS*, 355, 995
 Di Matteo T., Springel V., Hernquist L., 2005, *Nat*, 433, 604
 Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, *ApJ*, 676, 33
 Efstathiou G., 1992, *MNRAS*, 256, 43P
 Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, *PASP*, 110, 761
 Fontanot F., De Lucia G., Monaco P., Somerville R. S., Santini P., 2009, *MNRAS*, 397, 1776
 Gnedin N. Y., Tassis K., Kravtsov A. V., 2009, *ApJ*, 697, 55
 Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*, Proc. XXXVI Rencontres de Moriond. EDP Sciences, Les Ulis
 Hernquist L., Springel V., 2003, *MNRAS*, 341, 1253
 Hopkins A. M., Beacom J. F., 2006, *ApJ*, 651, 142
 Hoyle F., Lyttleton R. A., 1939, *Proc. Cambridge Philosophical Soc.*, 34, 405
 Kay S. T., Thomas P. A., Theuns T., 2003, *MNRAS*, 343, 608
 Kennicutt R. C. Jr., 1998, *ApJ*, 498, 541
 Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, 363, 2
 Klessen R. S., Spaans M., Jappsen A.-K., 2007, *MNRAS*, 374, L29
 Komatsu E. et al., 2009, *ApJS*, 180, 330
 Krumholz M. R., McKee C. F., Tumlinson J., 2009, *ApJ*, 699, 850
 Larson R. B., 1998, *MNRAS*, 301, 569
 Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, *ApJ*, 460, L1
 Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, *MNRAS*, 283, 1388
 Maness H. et al., 2007, *ApJ*, 669, 1024
 Marigo P., 2001, *A&A*, 370, 194
 Martin C. L., 2005, *ApJ*, 621, 227
 McCarthy I. G. et al., 2009, *MNRAS*, submitted (arXiv:0911.2641)
 Murali C., Katz N., Hernquist L., Weinberg D. H., Davé R., 2002, *ApJ*, 571, 1
 Murray N., Quataert E., Thompson T. A., 2005, *ApJ*, 618, 569
 Nagamine K., Cen R., Ostriker J. P., 2000, *ApJ*, 541, 25
 Nagamine K., Ostriker J. P., Fukugita M., Cen R., 2006, *ApJ*, 653, 881
 Okamoto T., Gao L., Theuns T., 2008, *MNRAS*, 390, 920
 Oppenheimer B. D., Davé R., 2006, *MNRAS*, 373, 1265
 Oppenheimer B. D., Davé R., 2008, *MNRAS*, 387, 577
 Oppenheimer B. D., Davé R., 2009, *MNRAS*, 395, 1875
 Ouchi M. et al., 2004, *ApJ*, 611, 660
 Padoan P., Nordlund A., Jones B. J. T., 1997, *MNRAS*, 288, 145
 Pawlik A. H., Schaye J., 2009, *MNRAS*, 396, L46
 Pawlik A. H., Schaye J., van Scherpenzeel E., 2009, *MNRAS*, 394, 1812

- Pearce F. R., Jenkins A., Frenk C. S., White S. D. M., Thomas P. A., Couchman H. M. P., Peacock J. A., Efstathiou G., 2001, *MNRAS*, 326, 649
- Portinari L., Chiosi C., Bressan A., 1998, *A&A*, 334, 505
- Reddy N. A., Steidel C. C., 2009, *ApJ*, 692, 778
- Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, *ApJS*, 175, 48
- Robertson B. E., Kravtsov A. V., 2008, *ApJ*, 680, 1083
- Robertson B., Hernquist L., Cox T. J., Di Matteo T., Hopkins P. F., Martini P., Springel V., 2006, *ApJ*, 641, 90
- Romeo A. D., Sommer-Larsen J., Portinari L., Antonuccio-Delogu V., 2006, *MNRAS*, 371, 548
- Ruszkowski M., Brüggem M., Begelman M. C., 2004, *ApJ*, 611, 158
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2005, *MNRAS*, 364, 552
- Schaye J., 2004, *ApJ*, 609, 667
- Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210
- Schaye J., Theuns T., Rauch M., Efstathiou G., Sargent W. L. W., 2000, *MNRAS*, 318, 817
- Schimminovich D. et al., 2005, *ApJ*, 619, L47
- Seljak U., Zaldarriaga M., 1996, *ApJ*, 469, 437
- Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337
- Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, *MNRAS*, 391, 481
- Sommer-Larsen J., Götz M., Portinari L., 2003, *ApJ*, 596, 47
- Spergel D. N. et al., 2007, *ApJS*, 170, 377
- Springel V., 2005, *MNRAS*, 364, 1105
- Springel V., Hernquist L., 2003a, *MNRAS*, 339, 289
- Springel V., Hernquist L., 2003b, *MNRAS*, 339, 312
- Springel V., Di Matteo T., Hernquist L., 2005a, *MNRAS*, 361, 776
- Springel V. et al., 2005b, *Nat*, 435, 629
- Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, *ApJ*, 519, 1
- Sutherland R. S., Dopita M. A., 1993, *ApJS*, 88, 253
- Thielemann F.-K. et al., 2003, in Hillebrandt W., Leibundgut B., eds, *From Twilight to Highlight: The Physics of Supernovae*. Springer, Berlin, p. 331
- Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, *MNRAS*, 382, 1050
- Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, 43, 769
- White S. D. M., 1996, in Schaeffer R., Silk J., Spiro M., Zinn-Justin J., eds, *Cosmology and Large Scale Structure*. Elsevier, Amsterdam, p. 349
- 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
- Wilkins S. M., Trentham N., Hopkins A. M., 2008, *MNRAS*, 385, 687
- Zel'Dovich Y. B., 1970, *A&A*, 5, 84

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.