Mon. Not. R. Astron. Soc. **369,** 1055–1080 (2006) doi:10.1111/j.1365-2966.2006.10426.x

The epoch of reionization

Andrew J. Benson,[1⋆] Naoshi Sugiyama,[2⋆] Adi Nusser[2,3⋆] and Cedric G. Lacey[4⋆]

[1]*Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH*
[2]*NAOJ, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[3]*Technion – Israel Institute of Technology, Technion City, Haifa 32000, Israel*
[4]*Institute for Computational Cosmology, University of Durham, Durham DH1 3LE*

Accepted 2006 March 27. Received 2006 March 27; in original form 2005 December 14

ABSTRACT

We have modelled the process of reionization of the intergalactic medium (IGM) by photoionization by galaxies, in order to learn what galaxy formation in the framework of the cold dark matter (CDM) model predicts for the epoch of reionization. We use a sophisticated semi-analytic model of galaxy formation to track the formation of these galaxies, their influence on the IGM and the back reaction of the state of the IGM on further galaxy formation. Our study represents a much more complete and physically consistent modelling of reionization than has been conducted in the past. In particular, compared to previous work by ourselves and others, our new calculations contain significant improvements in the modelling of the effects of reionization of the IGM on the collapse of baryons into dark matter haloes (this is now computed self-consistently from the properties of model galaxies), and in the model for the cooling and condensation of gas within haloes (our new model includes photoheating from a self-consistently computed ionizing background and also includes cooling due to molecular hydrogen). We find that reionization can be achieved by $z \sim 10$–20 in a ΛCDM cosmological model with $\sigma_8 \approx 0.9$. However, a cosmological model with a running spectral index is only able to achieve reionization before $z \approx 9$, and thus be consistent with an optical depth of 0.1, if very extreme assumptions are made about the physics of feedback at high redshifts. We also consider the specific galaxy formation model recently discussed by Baugh et al., which includes a top-heavy initial mass function (IMF) in starbursts, and find that it is able to reionize the Universe by $z \approx 12$. The previous results assume that all of the ionizing photons produced by stars in galaxies are able to escape and ionize the IGM. If this is not the case, then the redshift of reionization could be substantially reduced. We find that extended periods of partial reionization and double reionizations can occur in models in which the first stars formed via cooling by H_2 molecules, are very massive, and in which the escape fraction of ionizing photons $\sim$10–30 per cent. Such models do not fully reionize until $z \approx 6$–7, but predict an electron scattering optical depth as large as 0.15. The recent *Wilkinson Microwave Anisotropy Probe* (*WMAP*) three-year data suggest lower values of $\sigma_8 = 0.7$–0.8. Models with lower σ_8 have reduced redshifts of reionization, consistent with the lower optical depth also suggested by the *WMAP* three-year data.

Key words: galaxies: high-redshift – intergalactic medium – cosmology: theory.

1 INTRODUCTION

While gas in the intergalactic medium (IGM) is highly ionized in the present-day Universe, it is known to have been mostly neutral in the past. Evidence for this comes from observations of the Gunn–Peterson effect in the spectra of high redshift quasars (Becker et al. 2001; Djorgovski et al. 2001; Fan et al. 2006) and from the value of the optical depth to scattering measured from the cosmic microwave background (CMB; Spergel et al. 2003). Thus, there was a period in the history of the Universe when it transitioned from being neutral to ionized – the epoch of reionization.

Measurements of the CMB polarization made during the first year of operation of the *Wilkinson Microwave Anisotropy Probe* (*WMAP*)

⋆E-mail: abenson@astro.ox.ac.uk (AJB); naoshi@yso.mtk.nao.ac.jp (NS); adi@physics.technion.ac.il (AN); Cedric.Lacey@durham.ac.uk (CGL)

© 2006 The Authors. Journal compilation © 2006 RAS

satellite[1] have been used to estimate a value $\tau = 0.17 \pm 0.04$ for the electron scattering optical depth to reionization (Kogut et al. 2003), which, in the simple case that the IGM reionized instantaneously, implies a reionization redshift $z_{\rm reion} = 17 \pm 5$ (Spergel et al. 2003). This reionization redshift is significantly higher than was previously thought on the basis of the observed Gunn–Peterson effect and on the basis of theoretical modelling (see for example Haiman & Loeb 1997; Gnedin & Ostriker 1997; Baugh et al. 1998; Valageas & Silk 1999; Bruscoli et al. 2000; Chiu & Ostriker 2000; Ciardi et al. 2000; Benson et al. 2001). Although the *WMAP* measurement could be explained in scenarios in which reionization happened twice, or happened only partially for some time (Cen 2003; Wyithe & Loeb 2003; Naselsky & Chaing 2004),[2] the simple interpretation (that the Universe was fully reionized at $z = 17$ and remained so ever since) implies the presence of a significant population of ionizing sources at high redshift. Coupled with the hints that the effective spectral index of the matter power spectrum is changing with scale (i.e. a 'running spectral index'), this scenario becomes hard to explain in the standard cold dark matter (CDM) scenario of structure and galaxy formation.[3]

This tension has been explored recently by several authors (Ciardi, Ferrara & White 2003; Haiman & Holder 2003; Somerville, Bullock & Livio 2003; Yoshida et al. 2003b; Onken & Miralda-Escudé 2004), who generally reach the conclusion that $z_{\rm reion} = 17$ is incompatible with the running spectral index model, and hard to achieve even in a power-law model. One possible escape from this apparent problem is to reionize the Universe in some other way, i.e. without relying on ionizing photons from stars in high redshift galaxies. One such mechanism which has received attention recently is a contribution to reionization from particle decays during the cosmic dark ages (Chen & Kamionkowski 2004; Hansen & Haiman 2004; Kasuya & Kawasaki 2004; Kasuya, Kawasaki & Sugiyama 2004), another is a contribution from 'mini-QSOs' (see for example Ricotti & Ostriker 2004).

In this paper, we undertake detailed calculations of the galaxy formation process at high redshift to investigate whether galaxies alone can ionize the Universe sufficiently early to explain the *WMAP* measurement of the optical depth. For this purpose, we employ the GALFORM semi-analytic model of galaxy formation (Cole et al. 2000; Benson et al. 2002a). The advantage of our model over previous attempts to address this question is that it incorporates much more of the actual physics of galaxy formation (e.g. it follows the cooling of gas within haloes and incorporates a model of star formation and feedback), even if this physics is treated in a simplified way. The model used in this paper is significantly improved over that employed in our earlier study of reionization in Benson et al. (2001). The effects of reionization on subsequent galaxy formation (due to the heating of the IGM and the subsequent rise in the Jeans mass, and due to photoheating of gas in haloes) are now self-consistently computed within our model. Furthermore, we now include a detailed treatment of cooling due to molecular hydrogen in our calculations. Thus, we are able to provide a more reliable answer as to what the CDM model of structure formation predicts for the epoch of reionization.

The remainder of this paper is arranged as follows. In Section 2, we outline the model used to follow the reionization of the Universe. In Section 3, we examine the reionization histories predicted by our models and assess the importance of the various physical ingredients of the models. In Section 4, we present results for the reionization epoch and the imprint of reionization on the CMB. Finally, in Section 5, we assess the implications of our results. Several appendices describe further details of the modelling, and present additional results.

2 MODELLING

2.1 Galaxy formation modelling

We employ the GALFORM semi-analytic model of galaxy formation throughout this work to predict the properties of the galaxy population as a function of redshift. Most importantly for the current work is the ability of this model to compute the star formation rates (and associated luminosities in ionizing photons) needed to predict the reionization history of the Universe. A full description of the GALFORM model is beyond the scope of this paper. We instead refer the reader to previous papers which document the model in more detail (Cole et al. 2000; Benson et al. 2002a). Here, we give a very brief description of key components of the model and provide somewhat more detail about two ingredients which are critical to the current work: the cooling of hot gas in haloes (see Section 2.2) and star formation (see Section 2.3). It should be noted that the major innovation with respect to Paper I is our improved model of gas cooling, which now includes cooling due to H_2 and which is affected by feedback from the reionization process. The implications of this new modelling are discussed in Section 2.2.

The basic structure of the GALFORM calculation can be summarized as follows.

(i) The hierarchical formation of dark matter haloes is followed using the extended Press–Schechter theory (Bond et al. 1991; Bower 1991). This allows the construction of merger trees which form the backbone of the galaxy formation calculation.

(ii) A galaxy can form in each branch of the merger tree if the hot gas in the halo is able to cool and condense to the halo centre (see Section 2.2) where it forms a rotationally supported disc.

(iii) This condensed gas is assumed to begin to form stars at a rate governed by simple, phenomenological rules (see Section 2.3). As a consequence of this star formation, supernovae inject energy into the galaxy causing some of the condensed gas to be re-ejected from the galaxy. This feedback loop is crucial to producing a population of galaxies which matches the observed low-redshift Universe (Benson et al. 2003a). Star formation also produces metals, enriching the gas in the galaxy. When this enriched gas is ejected from the galaxy by supernova explosions, the hot atmosphere of gas pervading the halo also becomes enriched in metals.

(iv) When two branches of the merger tree join, the most massive galaxy from the branches becomes the central galaxy of the new branch. Any other galaxies become satellite galaxies, free to orbit within the gravitational potential of the new dark matter halo. Satellite galaxies are no longer able to gain gas from the surrounding hot atmosphere.

[1] While this paper was being completed, results based upon the first three years of data from *WMAP* became available (Spergel et al. 2006). These show a much lower optical depth to reionization. We will comment upon these results in Section 5.

[2] The physical plausibility of these scenarios has been questioned by Furlanetto & Loeb (2005). We present evidence, in Section 3.3, that such double or partial reionization may in fact occur under plausible physical conditions.

[3] The running spectral index reduces the amplitude of fluctuations on small, galactic scales relative to that on large scales. As a result there are many fewer haloes present at high redshift with mass sufficient to form galaxies.

(v) Satellite galaxies experience dynamical friction as they orbit, which eventually forces them to merge into the galaxy residing at the halo centre. These mergers are assumed to transform stellar discs into stellar spheroids if the masses of the two galaxies are comparable. Thus, in this model, mergers are the drivers of morphological evolution. Mergers of this type may also trigger bursts of star formation (see Section 2.3).

2.2 Cooling of gas in haloes

When a dark matter halo reaches virial equilibrium, it is assumed to have accreted into itself some mass of gas from the surrounding IGM. For high-mass haloes, the mass of gas accreted is equal to $\Omega_b/(\Omega_0 - \Omega_b)$ times the dark matter mass of the halo (where Ω_0 and Ω_b are the mean densities of total and baryonic mass in the Universe in units of the critical density). For lower mass haloes, the mass of gas accreted is reduced, due to the finite gas pressure in the IGM, in accordance with the filtering mass formalism of Gnedin (2000). The filtering mass, M_F, is discussed in detail in Appendix A3. The mass of gas accreted by a halo of mass M_{tot} is then given by

$$M_{gas} = \frac{(\Omega_b/\Omega_0)M_{tot}}{\left[1 + (2^{1/3} - 1)M_F(z)/M_{tot}\right]^3}. \quad (1)$$

The gas is assumed to be heated by shocks as it accretes into the halo, reaching a uniform temperature equal to the halo virial temperature. The gas is then assumed to be distributed through the halo with a density profile

$$\rho_{gas}(r) \propto \frac{1}{r^2 + c^2}, \quad (2)$$

where c is a scale radius which we set equal to $r_s/3$, where r_s is the scale radius of the dark matter density profile [assumed to have the Navarro, Frenk & White (NFW) form].

To compute the rate of cooling of the gas, we construct a table of net radiative cooling rates as a function of temperature, density, metallicity and redshift. (The redshift dependence arises through the inclusion of photoheating from a time-varying ionizing background which is computed self-consistently from the emission from the model galaxies.) This net cooling/heating rate includes contributions from atomic processes (e.g. recombinations, collisional excitations etc), Compton cooling (Peebles 1968), cooling due to molecular hydrogen (see Section A1) and photoheating by the ionizing background (see Section 2.4). Given these cooling rates and the assumed gas density profile and temperature for each halo, we then calculate, at each time-step of our calculation, the cooling radius, defined as the radius in the halo at which the local cooling time equals the time since the halo formed. (A halo is defined to form anew whenever its mass is doubled by mergers.) Any gas within this cooling radius is assumed to accrete on to the galaxy forming at the centre of the halo, provided that the time for the gas to free fall to the halo centre is less than the age of the halo. Occasionally, for example when the ionizing background is rapidly rising, the cooling radius at a given time-step may be smaller than that at the previous time-step. In such cases no gas is allowed to cool until the cooling radius once again begins to increase.

The inclusion of cooling due to H_2 is crucial to the present work as, at high redshifts, this is the only cooling channel available to gas in the majority of haloes. Without H_2 cooling included, we risk seriously underestimating the epoch of reionization. The IGM pressure and the ionizing background act as sources of negative feedback on star formation (i.e. star formation leads to a rise in the IGM pressure and ionizing background, which in turn suppress later star formation). By including these effects self-consistently in our model, we are therefore able to check whether this negative feedback is sufficient to significantly quench star formation. It should be noted that if we neglect H_2 cooling and set the filtering mass and ionizing background to zero our revised cooling model produces results identical to those of Paper I (i.e. it simply follows the cooling prescription of Cole et al. (2000) with the addition of Compton cooling). We will examine the effects of including H_2 cooling in our calculations in Section 3.1.

The effects of including the filtering mass and ionizing background in our cooling model have been explored in detail by Benson et al. (2002a). Benson et al. (2002a) find that it is the filtering mass which has the greatest effect on galaxy formation at low redshifts, resulting in a suppression of star formation after the epochs of H I and He I reionization. For the Cole et al. (2000) model, Benson et al. (2002a) found this suppression to be quite small. However, for a model with weak feedback from supernovae the suppression effect of the filtering mass was significant (reducing the global star formation rate by a factor of 2). This occurs because, in models with weak SNe feedback, star formation can otherwise progress efficiently in low-mass haloes, and therefore such haloes make a significant contribution to the global star formation rate. Once the filtering mass rises star formation is shut off in these low-mass haloes, thereby greatly suppressing the global star formation rate. The filtering mass also acts to reduce the number of faint galaxies which can form, as discussed in detail by Benson et al. (2002b). Note that the filtering mass begins to rise as the process of reionization begins, but does not reach its maximum value until significantly after the epoch of reionization. As such, we may expect it to have only minimal impact on the epoch of reionization, a point which we will explore in Section 3.2.

2.3 Star formation and supernova feedback

Once gas has cooled inside a dark matter halo it is assumed to collapse, conserving its angular momentum, until it reaches a radius where it becomes rotationally supported against further collapse. At this point, the gas is assumed to form an exponential disc. The gas in this disc then begins to form stars at a rate ψ given by

$$\psi = M_{cold}/\tau_\star, \quad (3)$$

where M_{cold} is the total mass of gas available in the galaxy and $\tau_\star$ is a characteristic time-scale for star formation. Following Cole et al. (2000), we adopt the parametrization

$$\tau_\star = \epsilon_\star^{-1}\tau_{disc}\left(V_{disc}/200\,\mathrm{km\,s^{-1}}\right)^{\alpha_\star}, \quad (4)$$

where $\epsilon_\star$ and $\alpha_\star$ are dimensionless parameters, V_{disc} is the circular velocity of the disc at its half-mass radius and τ_{disc} is the dynamical time of the disc at that radius.

We assume that supernova explosions in the galaxy cause gas to be ejected from the galaxy at a rate $\dot{M}_{eject} = \beta\psi$, where

$$\beta = (V_{disc}/V_{hot})^{-\alpha_{hot}}, \quad (5)$$

where α_{hot} is a dimensionless parameter and V_{hot} is a parameter with dimensions of speed.

Star formation can also occur in a burst mode as the result of a violent merger between two galaxies. When such a merger occurs, all of the gas in the two merging galaxies undergoes star formation on a very short time-scale. We use the same rules as for quiescent star formation in discs, but using the dynamical properties of the newly formed spheroid instead.

© 2006 The Authors. Journal compilation © 2006 RAS, MNRAS **369,** 1055–1080

We compute the chemical enrichment of each galaxy using the instantaneous recycling approximation, with yield of heavy elements p and recycled fraction R (as defined by Cole et al. 2000 – we use the same values for these parameters as did Cole et al. 2000 unless otherwise noted), including the exchange of metals between the galaxy and the halo gas. Once the star formation history and associated chemical enrichment history of a galaxy have been determined, we compute the spectral energy distribution (SED) of the galaxy at each time-step using a stellar population synthesis model based on the Padova stellar evolution tracks (Granato et al. 2000). This allows us to compute the ionizing photon luminosity for each galaxy.

2.4 Evolution of the IGM and ionizing background

We model the process of reionization and its feedback on the IGM and galaxy formation using the methods of Benson et al. (2001, hereafter Paper I) and Benson et al. (2002a). Briefly, we compute the volume averaged emissivity in ionizing photons at each redshift by summing the ionizing luminosity of all galaxies in our calculation. This emissivity is used to evolve a distribution of gas elements drawn from an appropriate PDF (Paper I). We track the ionization (H and He) and thermal state of these gas elements as the Universe evolves. From the resulting thermal history of the IGM we are able to compute the filtering mass describing the effects of the IGM pressure on the collapse of baryons into haloes. This allows us to estimate the mass of baryons accreted into each dark matter halo forming at any redshift. We also have, at each redshift, the spectrum of the ionizing background, which allows us to compute the rate of photoionization heating experienced by gas in dark matter haloes. These two processes (reduced accretion due to the IGM pressure and photoheating of halo gas), which act to suppress galaxy formation, are incorporated into a second iteration of our galaxy formation calculation. This iterative process is repeated until convergence is reached (i.e. the input ionizing background and filtering mass are consistent with those resulting from the galaxy formation calculation) as in Paper I. Schematically, the iterative procedure works as follows.

(i) Initially assume that there is no ionizing background (and associated photoheating) and that the filtering mass is always zero.

(ii) Run the GALFORM galaxy formation calculation.

(iii) Compute the evolution of the ionizing background based on the GALFORM calculation.

(iv) Compute associated photoheating rates for gas of varying densities, temperatures and metallicities.

(v) Compute the evolution of the filtering mass based on the GALFORM calculation.

(vi) Using these photoheating rates and filtering mass, go back to step (ii) and repeat.

Iteration stops once the star formation history produced in GALFORM calculations on successive iterations is sufficiently well converged.

Benson et al. (2001, 2002a) considered cooling due to both atomic processes and Compton cooling due to the scattering of CMB photons from hot electrons. Here we employ an improved cooling model which is described in detail in Appendix A. To explore the high redshifts and low-mass scales of interest here we must also consider the effects of cooling due to molecular hydrogen (see Section A1). We must also reconsider Compton cooling in the non-equilibrium regime (i.e. before free electrons in the gas have had sufficient time to recombine to reach their abundance in ionization equilibrium at the density of the halo gas) which occurs in some of the lower mass haloes that we consider (see Section A2).

2.5 Escape fractions

A very important parameter in our models is the fraction, $f_{\rm esc}$, of ionizing photons produced in a galaxy which escape into the IGM. For most of our models (and unless stated otherwise), we assume $f_{\rm esc} = 1$. The only exceptions are the models H_2WeakFB and H_2WeakFB-VMS (to be defined in Section 2.7.2), for which we also present some results for lower escape fractions (for the purpose of producing models with 'double reionization' – see Section 3.3). The assumption that $f_{\rm esc} = 1$ is obviously an extreme one (although such an assumption has been used in other works studying reionization e.g. Haiman & Holder 2003), since it ignores the absorption of ionizing photons by gas and dust in the galaxy, but it should result in the earliest possible reionization if other parameters are held fixed. We note, however, that observational evidence (and simple physical models for $f_{\rm esc}$) imply a much lower escape fraction. Therefore, in Appendix B, we present a simple, physically motivated calculation of $f_{\rm esc}$ which is in reasonable agreement with observational data, and show the resulting reionization histories.

2.6 Calculation of filling factor

To compute the filling factor of H II we follow the growth of ionized regions around each halo in the galaxy formation model using the methodology of Paper I. Specifically, we compute the total emission rate of ionizing photons from stars in each dark matter halo, $S(t)$, and compute the number of hydrogen atoms which will be ionized by this radiation by solving

$$\frac{{\rm d}N_{\rm H}}{{\rm d}t} = S(t) - \alpha_{\rm H}^{(2)} a^{-3} f_{\rm clump} n_{\rm H} N_{\rm H}, \quad (6)$$

where $N_{\rm H}$ is the total number of hydrogen atoms ionized, $\alpha_{\rm H}^{(2)}$ is the case B recombination coefficient for hydrogen (for a temperature of 10^4 K), a is the cosmic expansion factor, $f_{\rm clump}$ is the clumping factor for ionized gas and $n_{\rm H}$ is the mean number density of hydrogen atoms in the Universe. We compute the clumping factor following the method described in Paper I for $f_{\rm clump}^{\rm (haloes)}$, but with the cosmological parameters appropriate to the current work. In this model, the main contribution to the clumping factor comes from gas residing in haloes with virial temperatures around 10^4 K, which have deep enough potential wells to retain gas, but which are cool enough to be not completely collisionally ionized. The specific values for the clumping factors used in this work will be discussed in Section 2.7.1.

Note that we do not assume any particular geometry for the ionized region around each source. Ionized regions from two or more sources may overlap (particularly just before the epoch of reionization or if the sources are strongly clustered). This does not affect our results, as we are simply counting the total number of hydrogen atoms which have been ionized.

The total number of hydrogen atoms, $N_{\rm H,tot}$, which will have been ionized in some representative volume of the Universe, V, is then found by simply summing over the contributions from galaxies in haloes of all masses

$$N_{\rm H,tot} = \sum_i N_{{\rm H},i} n_i V, \quad (7)$$

where $N_{{\rm H},i}$ is the number of hydrogen atoms which have been ionized by galaxies in halo i, of which there are n_i such haloes per unit volume. The sum is taken over all haloes. The mean ionized fraction in volume V is found by simply dividing by the total number of

hydrogen atoms in the volume, $n_{\rm H}V$, where $n_{\rm H}$ is the mean density of hydrogen atoms in the Universe. If we assume a uniform IGM, we can identify this mean ionized fraction with the filling factor of ionized regions,

$$F_{\rm fill} = \frac{1}{n_{\rm H}} \sum_i N_{{\rm H},i} n_i . \quad (8)$$

We consider the Universe to have been fully reionized once $F_{\rm fill} = 1$. Note that with these definitions $F_{\rm fill}$ can exceed unity. In such cases, the Universe is fully ionized and ionizing photons will begin to build up an ionizing background rather than be absorbed by neutral hydrogen atoms.

The IGM evolution model of Benson et al. (2002a) gives an alternative means of following the reionization of the Universe. In that model, the mean ionization state of the IGM is determined by computing the ionization and recombination rates in the presence of an evolving ionizing background, allowing for a distribution of IGM densities. While these two methods for following reionization are quite different, we find that they predict very similar redshifts of reionization, with the method of Benson et al. (2002a) typically predicting $z_{\rm reion}$ to occur roughly $\Delta z \approx 1$ earlier. Throughout this work we will show results based on the approach described in this section.

2.7 Cosmological and galaxy formation models

We will explore three different cosmological models and several different models of galaxy formation in this work. The details of these models are outlined below.

2.7.1 Cosmological parameters

For the cosmological parameters we adopt the values determined from the first-year *WMAP* data set (Spergel et al. 2003). In particular, we consider cosmological models with the best-fitting parameters for a power-law ΛCDM model with constraints from *WMAP* alone (table 7 of Spergel et al. 2003) and a running spectral index ΛCDM model with constraints from *WMAP*, the 2dF Galaxy Redshift Survey and Lyman α forest observations (table 8 of Spergel et al. 2003). We refer to these cosmological models as *WMAP* PL and *WMAP* RSI, respectively, and their parameters are given in Table 1. We also compute reionization histories in the cosmological model corresponding to the specific galaxy formation model of Baugh et al. (2005), discussed further below.

We include these three cosmologies because they have complementary strengths and weaknesses. The cosmologies based on *WMAP* results give good fits to the available CMB data (by definition). However, we do not have a set of galaxy formation model parameters in these cosmologies which result in galaxy populations with properties which match local observations. The Baugh et al. (2005) model, on the other hand, was not designed to match the CMB data (although its cosmological parameters are not too far from those measured by *WMAP*), but does reproduce the properties of galaxies in the local and high-redshift Universe.

We show in Fig. 1, for all three cosmological models, the fractional fluctuation of the density measured in spheres, $\sigma(M)$, for the linear density field extrapolated to $z = 0$. The quantity $\sigma(M)$ determines the abundances of dark matter haloes at different redshifts. It is clearly seen that the *WMAP* RSI model has much smaller fluctuations on small scales.

The IGM clumping factor which we use depends only on the cosmological model, and is shown as a function of redshift in Fig. 2. For the *WMAP* PL cosmology it is very similar to that found in Paper I ($f_{\rm clump} \approx 18$ at $z = 10$, dropping to approximately 2.5 at $z = 20$ for example), while for the *WMAP* RSI cosmology it is much lower ($f_{\rm clump} \approx 1.5$ at $z = 10$, dropping to close to 1 at $z = 20$) due to the reduction of small-scale fluctuations in the latter cosmology. Inclusion of the clumping factor makes almost no difference to our results for the *WMAP* RSI cosmology, as the clumping factor is still small by the time reionization occurs. However, in the *WMAP* PL cosmology, clumping can delay reionization by $\Delta z = 2$–4 (the earlier reionization occurs, the less effect clumping has). In the Baugh et al. (2005) cosmology, the clumping factor is comparable to that of the *WMAP* PL cosmology.

2.7.2 Galaxy formation parameters

To model the formation of galaxies in the Universe, we employ the semi-analytic model of galaxy formation GALFORM. A full description of this model can be found in Cole et al. (2000) and Benson et al. (2002a). For our 'Standard model' in the *WMAP* RSI and *WMAP* PL cosmologies, we adopt parameters which are very similar to those of Cole et al. (2000). (We ignore the cosmological parameters specified by Cole et al. (2000) and instead use those described in Section 2.7.1.) However, we use the improved merging calculation of Benson et al. (2002a), which results in somewhat fewer mergers on average than the simpler model used by Cole et al. (2000). Because of the change in cosmological parameters and in the merger rates, we adjust the values of a few of the parameters in the galaxy formation model, increasing the star formation efficiency $\epsilon_\star$ from 0.005 to 0.0067 and reducing the mass-to-light ratio parameter Υ from 1.38 to 1.00.

We regard the resulting model as being a plausible model for galaxy formation in the high-redshift Universe, since it has been shown to reproduce well the properties of galaxies in the low-redshift universe, and we simply assume that the same physical processes apply to galaxy formation at high redshift as at low redshift. Our model is therefore motivated by what we currently know about the local Universe – the extrapolation to high redshifts is a simple assumption as our current understanding of the physics that governs the high-redshift Universe does not allow us to devise more accurate physical prescriptions for early times. Note, however, that the Cole et al. (2000) model assumed $\Omega_{\rm b} = 0.02$. The higher baryon densities in the cosmologies considered here will result in this model producing too many galaxies of all luminosities at $z = 0$. We do

Table 1. Cosmological parameter sets used in this work. Values are taken from Spergel et al. (2003), *WMAP* and MacTavish et al. (2005), Boomerang. $n_{\rm s}$ is the scalar spectral index at $K = 0.05\,{\rm Mpc}^{-1}$, while ${\rm d}\,n_{\rm s}/{\rm d}\ln k$ is the running of the spectral index as defined by equation (5) of Spergel et al. (2003).

Name	Ω_0	Λ_0	$\Omega_{\rm b}$	$H_0/{\rm km\,s^{-1}\,Mpc^{-1}}$	σ_8	$n_{\rm s}$ (at $K = 0.05\,{\rm Mpc}^{-1}$)	${\rm d}\,n_{\rm s}/{\rm d}\ln k$
WMAP PL	0.270	0.730	0.0463	72.0	0.90	0.99	0.000
WMAP RSI	0.268	0.732	0.0444	71.0	0.83	0.93	−0.031
Baugh et al. (2005)	0.300	0.700	0.0400	70.0	0.93	1.00	0.000
Boomerang (MacTavish et al. 2005)	0.300	0.700	0.0469	69.6	0.85	0.95	0.000

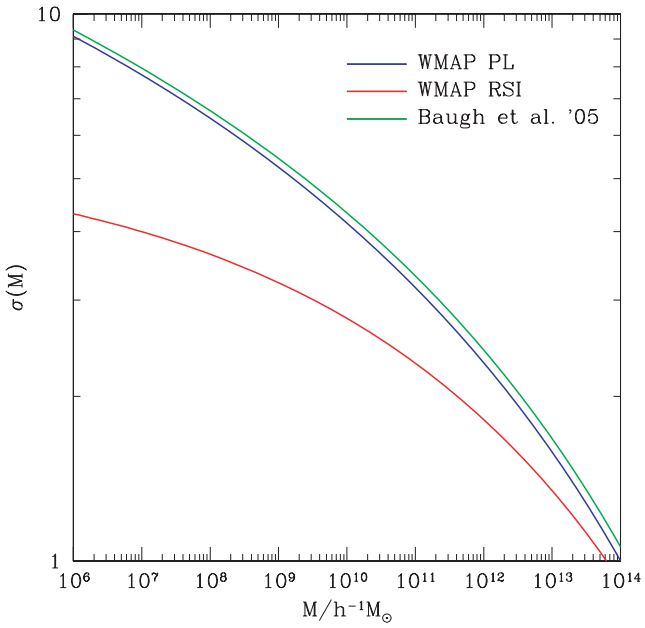

Figure 1. The fractional root mean square fluctuation of the density averaged in spheres, $\sigma(M)$, as a function of mass, M, for each of our cosmological models: *WMAP* RSI (red line), *WMAP* PL (blue line) and Baugh et al. (2005) (green line).

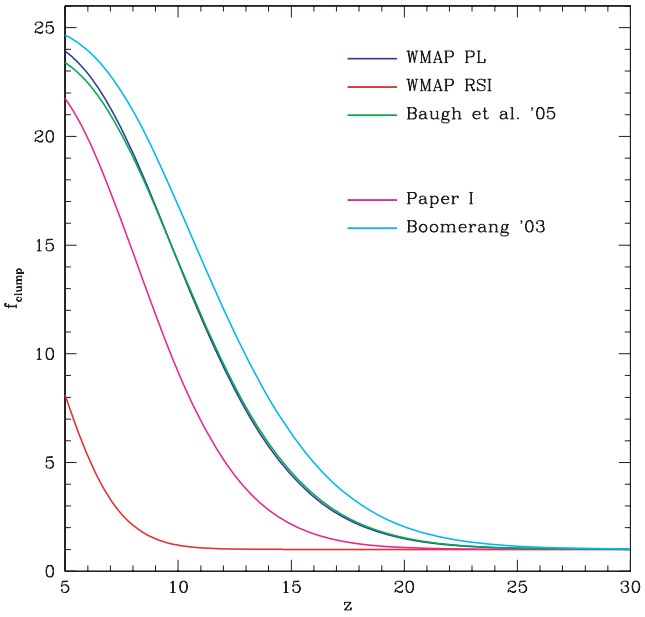

Figure 2. The clumping factor, $f_{\rm clump}$, as a function of redshift for the three cosmological models considered in this work. Also shown, for reference, are the clumping factors corresponding to the cosmological parameters of Paper I (see Appendix C1) and the Boomerang (MacTavish et al. 2005) cosmology discussed in Section 5.

not attempt to correct for this deficiency for two reasons: (1) we are interested here in determining how early reionization can occur – overproduction of galaxies will push reionization to higher redshifts and so give us a conservative answer to the question of whether a particular model can reionize sufficiently early to match the *WMAP* results; and (2) while the Cole et al. (2000) model works well at $z = 0$ it does not match observational data for high-redshift galaxies and, in any case, we are here extrapolating our model to $z > 20$ where observational constraints on galaxy formation are practically non-existent. It is for these reasons that we consider our model to be plausible but not definitive.

While there are numerous variations of the model parameters which it would be interesting to explore, we limit ourselves here to examining a handful which might be expected to have a particularly large impact on the epoch of reionization. These various models are summarized in Table 2.

(i) *Standard:* For the *WMAP* RSI and *WMAP* PL cosmologies, this model is very similar to that described in Cole et al. (2000) – the differences from Cole et al. (2000) are the parameter changes described in Section 2.7.2 and the inclusion of Compton cooling, which is necessary for studying the high redshifts of interest here. It assumes that all stars form with a Kennicutt (1983) initial mass function (IMF), with slope $x = 0.4$ for $0.1 < m < 1$ $M_\odot$ and $x = 1.5$ for $1 < m < 120$ $M_\odot$. (A Salpeter IMF has $x = 1.35$.) For this IMF, the total number of ionizing photons (with $h\nu > 13.6\,\mathrm{eV}$) produced per $M_\odot$ of stars formed is $N_{\rm Lyc}/M_* = 3.2 \times 10^{60}$, for a metallicity $Z = 0.1$ $Z_\odot$. Note that, for this model (and all others unless stated otherwise) we adopt a value of $V_{\rm hot} = 200$ km s^{-1} for the feedback parameter. This value, chosen to reproduce the low abundance of faint galaxies in the $z = 0$ Universe, results in quite strong feedback. For the Baugh et al. (2005) cosmology, the Standard model corresponds to the galaxy formation parameters given by Baugh et al. (2005). Note that this includes a higher value for $V_{\rm hot} = 300$ km s^{-1}. Fig. 3 shows the halo mass with virial velocity equal to $V_{\rm hot} = 200$ km s^{-1} as a function of redshift. At high redshifts, the vast majority of haloes able to form will be strongly affected by feedback for $V_{\rm hot} \geqslant 200$ km s^{-1}.

(ii) *MetFree:* In this model, we compute stellar luminosities and spectra assuming that all stars are metal free (while retaining the chemical enrichment of the interstellar and intergalactic gas, as this affects gas cooling rates), thereby enhancing the emission of ionizing photons from galaxies. In this case, we have $N_{\rm Lyc}/M_* = 4.4 \times 10^{60}$.

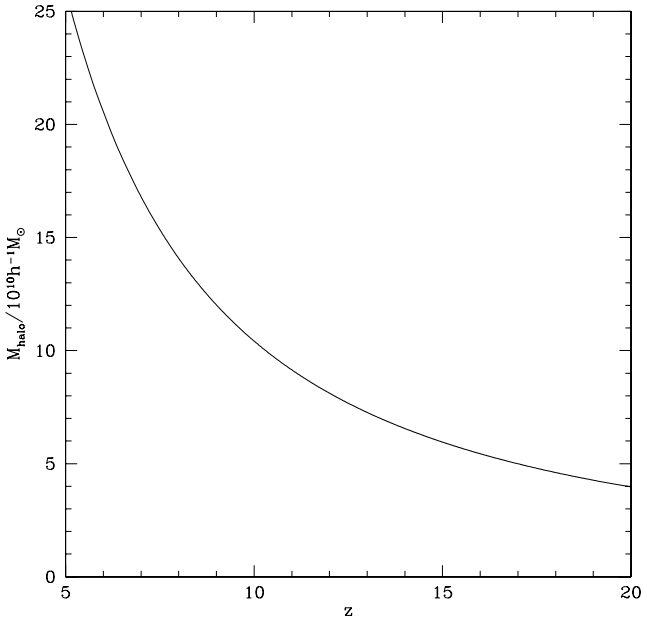

Figure 3. The mass of a halo with virial velocity corresponding to the feedback parameter, $V_{\rm hot} = 200$ km s^{-1}, is shown as a function of redshift for the *WMAP* PL cosmology (results for the *WMAP* RSI and Baugh et al. 2005 cosmologies are practically identical).

Table 2. Names and details of the galaxy formation models used in this work. Column 1 describes the model while column 2 gives the name by which we refer to the model throughout this work. Column 3 lists the changes in model parameters made relative to the standard model.

Model	Name	Variation from standard model
Standard model	Standard	–
Metal free stellar spectra	MetFree	$Z_\star = 0$
Top-heavy IMF	TopHeavy	IMF Salpeter above 10 $M_\odot$, zero below; $R = 0.91$; $p = 0.14$
Weak feedback	WeakFB	$V_{\rm hot} = 1$ km s^{-1}
Compton + H_2 cooling and dynamical heating	H_2	H_2 cooling on
Compton + H_2 cooling + weak feedback	H_2WeakFB	H_2 cooling on; $V_{\rm hot} = 1$ km s^{-1}
Compton + H_2 cooling + weak feedback + VMS for $T_{\rm vir} < 10^4$ K	H_2WeakFB-VMS	H_2 cooling on; $V_{\rm hot} = 1$ km s^{-1}; UV flux × 20 for stars formed via H_2 cooling

(iii) *TopHeavy:* We assume an extreme IMF which has the Salpeter slope $x = 1.35$, but only covers the mass range $10 < m < 100$ $M_\odot$, i.e. it is truncated for masses below 10 $M_\odot$. As massive stars produce the majority of ionizing photons, this model is expected to produce earlier reionization. For this IMF, we have $N_{\rm Lyc}/M_* = 2.5 \times 10^{61}$.

(iv) *WeakFB:* This model has very weak feedback from supernova explosions, such that even the lowest mass galaxies which form in the model will not experience significant mass loss due to feedback. In GALFORM, the strength of feedback is controlled by two parameters $V_{\rm hot}$ and $\alpha_{\rm hot}$, as defined in equation (5). We fix $\alpha_{\rm hot} = 2$ throughout this work, and use $V_{\rm hot} = 200(300)$ km s^{-1} for the *WMAP* (Baugh et al. 2005) cosmologies except for models with weak feedback, for which we assume $V_{\rm hot} = 1$ km s^{-1} (see Table 2). The lower value of $V_{\rm hot} = 1$ km s^{-1} corresponds to an efficiency of supernova energy injection and of mass ejection which is lower by a factor 4×10^4 than for the standard case $V_{\rm hot} = 200$ km s^{-1}, so it is fairly extreme. This weak feedback will greatly enhance the efficiency of star formation in low-mass haloes, and hence will also increase the production of ionizing photons at high redshift.

(v) *H_2*: This model is identical to the Standard model except that it incorporates cooling due to molecular hydrogen – this should allow galaxies to form earlier, thereby shifting the epoch of reionization to higher redshifts.

(vi) *H_2WeakFB:* This model combines a weak feedback with molecular hydrogen cooling.

(vii) *H_2WeakFB-VMS:* This model investigates the possibility that stars forming via H_2 cooling ('Population III') may be very massive stars (VMS), with $m \sim$ 100–1000 $M_\odot$ (Abel, Bryan & Norman 2000, 2002; Bromm, Coppi & Larson 2002). If all of the stars are very massive, then the total ionizing luminosity will be greatly enhanced relative to stars formed with a normal IMF. Therefore, in this model, we increase the ultraviolet (UV) luminosity of any stars forming in a halo with virial temperature below 10^4 K by a factor of 20 relative to the Kennicutt IMF at all metallicities.[4] We therefore have $N_{\rm Lyc}/M_* = 2.0 \times 10^{54}$.

Baugh et al. (2005) recently presented a GALFORM model employing a significantly different set of galaxy formation parameters from those used by Cole et al. (2000). We will use the Baugh et al. (2005) parameters to describe our 'Standard model' in the Baugh et al. (2005) cosmology. The cosmological parameters of this model are listed in Table 1. The Baugh et al. (2005) model was constrained to produce a good match to observations of sub-mm galaxies and Lyman-break galaxies at high redshifts, as well as the properties of galaxies at low redshifts. In order to match the numbers of sub-mm galaxies, Baugh et al. (2005) assumed that stars which form in merger-induced bursts have a top-heavy IMF (with slope $x = 0$ covering the mass range $0.1 < m < 120$ $M_\odot$), while stars which form quiescently in discs have a Kennicutt (1983) IMF (as in our Standard model). In the Baugh et al. (2005) model, the fraction of stars formed in the burst mode increases with redshift, with the burst mode dominating the total star formation density at $z \gtrsim 3$ (see fig. 1 in Baugh et al. 2005). In subsequent papers, Le Delliou et al. (2005, 2006) have investigated the properties of Lyman α emitting galaxies at high redshift, and Nagashima et al. (2005a,b) have investigated the chemical enrichment of elliptical galaxies and intracluster gas in the same model. Given the success of the Baugh et al. (2005) model in reproducing observations of various types of star-forming galaxies at high redshift, as well as observations of chemical abundances at low redshift, it is of interest to see what it predicts for the reionization of the IGM. We have therefore used the galaxy formation and cosmological parameters of the Baugh et al. (2005) model to compute the reionization history of the Universe.[5]

For stars formed in bursts in the Baugh et al. (2005) model, the number of ionizing photons produced per $M_\odot$ of stars formed $N_{\rm Lyc}/M_* = 3.7 \times 10^{61}$ for a metallicity $Z = 0.1$ $Z_\odot$, i.e. nearly 20 times larger than for a Kennicutt IMF. The only model variants we consider for the Baugh et al. (2005) cosmology at Standard, H_2 and H_2WeakFB.[6]

[4] Our choice of a UV enhancement factor of 20 is motivated by the results of Bromm, Kudritzki & Loeb (2001), who find an enhancement factor of 10–20. (Their enhancement is relative to a Salpeter IMF. We use the Kennicutt (1983) IMF here, but the total ionizing photon production per unit mass of star formation for this IMF differs from that of the Salpeter IMF by approximately 10 per cent for subsolar metallicities.) We have taken the value at the upper end of their range to maximize the effects. We also note that theoretical considerations suggest that VMS should form when the metallicity of the interstellar medium (ISM) is below a critical value, and not merely when gas cooling occurs through the H_2 channel. If we model VMS formation in this way (i.e. we enhance the UV emissivity only for stars with a metallicity below, for example, 10^{-4} $Z_\odot$ (Schneider et al. 2002; Bromm & Loeb 2004) we find little enhancement in the net UV emission. This occurs because the duration of the subcritical metallicity phase of star formation is very short. However, our current model does not account for the different heavy element yield of these Population III stars – a low yield could dramatically extend the duration of the subcritical metallicity phase of star formation. We defer a detailed examination of this possibility to a future work.

[5] Note that Baugh et al. (2005) included a simple model of photoionization feedback in their calculation, in which they suppressed the cooling of gas in dark matter haloes with virial velocities below 60 km s^{-1} for $z < 6$. As we are here modelling reionization and photoionization feedback in much greater detail, we remove this simple suppression from our implementation of the Baugh et al. (2005) model.

[6] Model H_2WeakFB-VMS is not considered, as the Baugh et al. (2005) model already includes VMS which are assumed to form in bursts.

3 RESULTS

Below, we present results from our calculations of reionization. We begin with a summary of the key results, before concentrating on the importance of H_2 cooling and photoionization feedback in setting the epoch of reionization. A comparison of the predicted epochs of reionization to the observational determinations and an exploration of the signal imprinted in the CMB by the patchy reionization process are given in Section 4.

Some further results, including comparisons to Paper I and tests of the convergence of our models are given in Appendix C. Appendix B briefly examines the effects on the reionization history of using physically motivated calculations of the escape fraction of ionizing photons.

Fig. 4 shows key results from all of our calculations. Each row corresponds to a different model (as indicated along the right-hand edge of the figure) and within each panel different colours correspond to different cosmologies: *WMAP* RSI (red), *WMAP* PL (blue) and Baugh et al. (2005) (green). The three columns, respectively, show the average star formation rate per unit volume, the ionizing luminosity density and the H II filling factor (as indicated by labels at the top of the figure) all as functions of redshift. It is clear from this figure that the suppression of small mass haloes in the *WMAP* RSI cosmology leads to a severe reduction in the star

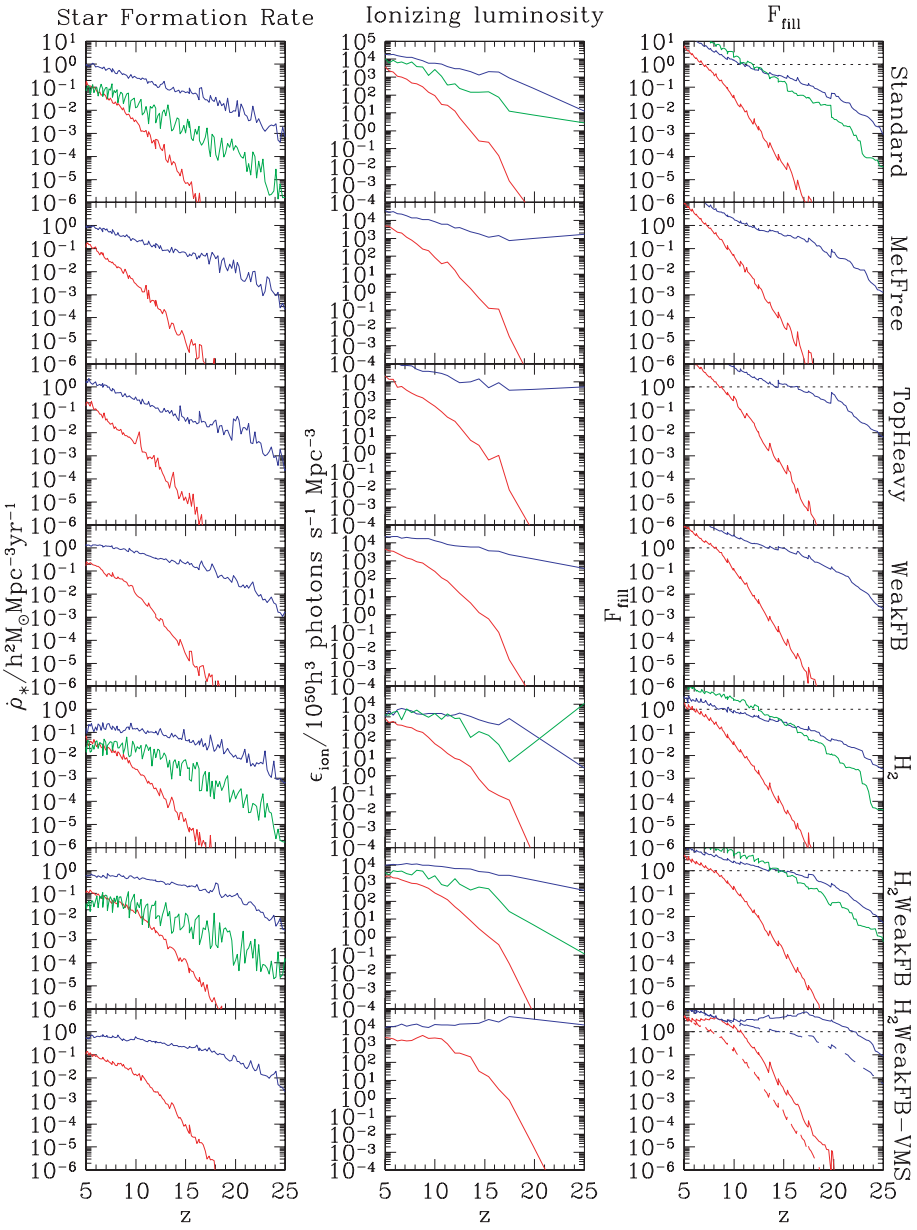

Figure 4. The evolution of key physical parameters in our models. The three columns show mean star formation rate density, ionizing luminosity and H II filling factor as indicated at the top of each panel. Each row corresponds to a different galaxy formation model, as indicated on the right-hand side of the figure. The three lines in each panel correspond to three different cosmological models: *WMAP* RSI (red), *WMAP* PL (blue) and Baugh et al. (2005) (green – not shown for models MetFree, TopHeavy, WeakFB and H_2WeakFB-VMS). In the H II filling factor columns, the horizontal dashed line indicates the point of complete reionization, while dashed coloured lines for the H_2WeakFB-VMS model show the contribution of Pop I and II stars to the total filling factor.

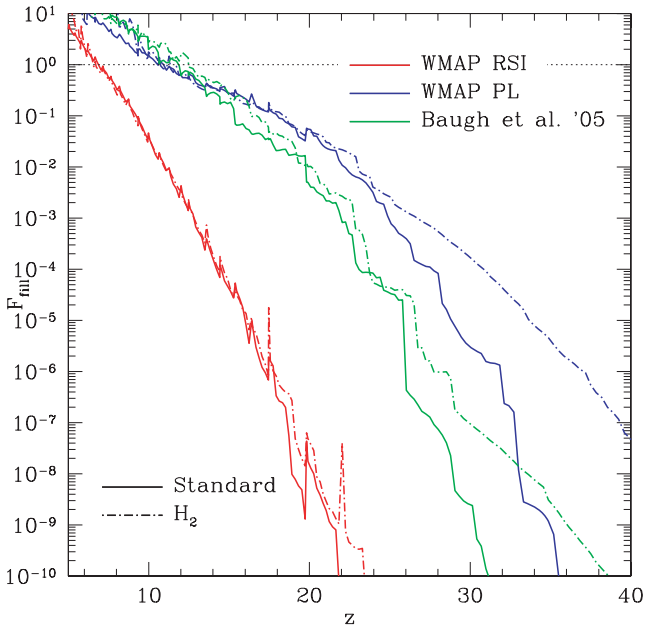

Figure 5. The H II filling factor as a function of redshift for models with (dot–short dashed lines) and without (solid lines) H_2 cooling. Colours indicate the cosmological model as indicated in the figure. Results are shown to higher redshifts and lower filling factors than in Fig. 4.

formation rate, ionizing luminosity and H II filling factor compared to our other cosmological models. This effect is larger than that of any of the other variations we consider in this work. As such, it may be possible to use a measured reionization epoch to place strong constraints on the cosmological model. It is also clear that reionization redshifts high enough to explain the *WMAP* measurement of the optical depth *can* be attained in the *WMAP* PL and Baugh et al. (2005) cosmologies, albeit only in rather extreme models. These points will be discussed further in the remainder of this section and in Section 4.

3.1 Does H_2 cooling help?

We can examine whether the inclusion of H_2 cooling channels in our calculations actually aids in pushing the epoch of reionization to higher redshifts. Fig. 5 shows the H II filling factor as a function of redshift for models with and without cooling due to H_2. It can clearly be seen that the addition of H_2 cooling channels has almost no effect on the epoch of reionization. At earlier redshifts, H_2 cooling *does* result in an enhancement in the filling factor (e.g. by a factor of 10^4 at $z = 35$ in the *WMAP* PL cosmology). However, by the time $F_{\rm fill}$ is approaching unity, the ionizing emissivity is dominated by emission from haloes which are able to cool through atomic processes, and so the addition of H_2 cooling channels does not help to increase $z_{\rm reion}$ significantly.

To understand the effects of cooling through molecular hydrogen, we can explore the cooled baryonic fraction in haloes as a function of their mass. The upper panels of Fig. 6 show the condensed baryonic fractions of haloes at redshifts 5 (left-hand panel) and 10 (right-hand panel) for models without the effects of photoionization and photodissociation included. Blue circles indicate the cooled fraction, $f_{\rm cond}$ (i.e. the mass of material which has been able to cool and condense into the galactic phase), for a model with no molecular hydrogen cooling, while red circles show results for a model with molecular hydrogen cooling. We indicate, with a vertical dashed magenta line, the halo mass for which the virial temperature corresponds to the supernova feedback parameter $V_{\rm hot}$ in our semi-analytic model. Vertical, dashed cyan lines indicate the filtering masses for the two models. In the lower panels of Fig. 6 are curves which show the cooling measure $t_{\rm age}/(1+\beta)t_{\rm cool}$, the ratio of the age of the Universe at the specified redshift $t_{\rm age}$ to the cooling time $t_{\rm cool}$ (calculated at 200 times the mean density of the Universe, which is approximately the mean density of haloes), and divided by $1+\beta$, where β (given by equation 5) is the ratio of the mass ejection rate due to supernova feedback to the star formation rate in model galaxies. Solid blue lines are for the model without molecular hydrogen cooling and without Compton cooling, while dashed red lines include molecular hydrogen cooling (and Compton cooling). We show curves for primordial gas and for gas of half-solar metallicity (the upper curves at large halo masses).

Our cooling measure, the quantity $t_{\rm age}/(1+\beta)t_{\rm cool}$, is interesting since it roughly determines whether a halo has had time to turn most of its supply of gas into stars. Roughly speaking, a galaxy will accrete (from its surrounding hot atmosphere) all available gas in a time $t_{\rm cool}$. Assuming haloes to have survived for a period roughly equal to $t_{\rm age}$, the ratio of these two time-scales then determines whether or not a halo has been able to accrete all available gas. However, much of this gas will be ejected once again due to feedback. This gas may be subsequently re-accreted. Consequently, a galaxy must accrete its hot halo $\sim(1+\beta)$ times before it can turn the majority of the available gas into stars.[7]

When $t_{\rm age}/(1+\beta)t_{\rm cool}$ is greater than 1 (the point $t_{\rm age}/(1+\beta)t_{\rm cool} = 1$ is indicate by green horizontal lines in Fig. 6), the effective cooling time is short compared to the age of the Universe and we do not expect the cooling function to have strongly determined the shape of the $f_{\rm cond}$ curve. When it is less than unity, the effective cooling time is long and so the shape of the $f_{\rm cond}$ curve should have been determined strongly by the shape of the cooling curve.

From Fig. 6, we can see that the inclusion of H_2 cooling makes no difference to the fraction of condensed mass in haloes with virial temperatures above about 2×10^4 K. H_2 cooling does, however, result in a tail of low-mass haloes (with virial temperatures below 10^4 K) which are able to condense some of their gas, unlike in the case of cooling through purely atomic processes. However, the amount of gas condensed is minimal, with $\lesssim$10 per cent of the total available gas being condensed (a number which falls rapidly to zero in lower mass haloes). This inefficient condensation occurs because $t_{\rm age}/t_{\rm cool}(1+\beta) \ll 1$ (typically 10^{-3}–0.1) for these haloes due mostly to the fact that $\beta \gg 1$ (typically, $\beta \gtrsim 100$ for these haloes) – this makes star formation a very inefficient process in these galaxies, with the majority of the gas which was able to condense being re-ejected due to feedback.

Further plots of the condensed fraction for all cosmologies and models can be found in Appendix D.

3.2 Effects of including photoionization and photodissociation feedback

We have run all models both with and without the effects of photoionization feedback (i.e. the suppression of accretion due to the

[7] Assuming the cooling/reheating loop to occur n times (and assuming cooling and reheating to happen instantaneously), the fraction of available gas eventually bound into stars will be $1-(\beta/(1+\beta))^n$. For $n = 1+\beta$ this converges, for large β to $1-1/e \approx 0.63$.

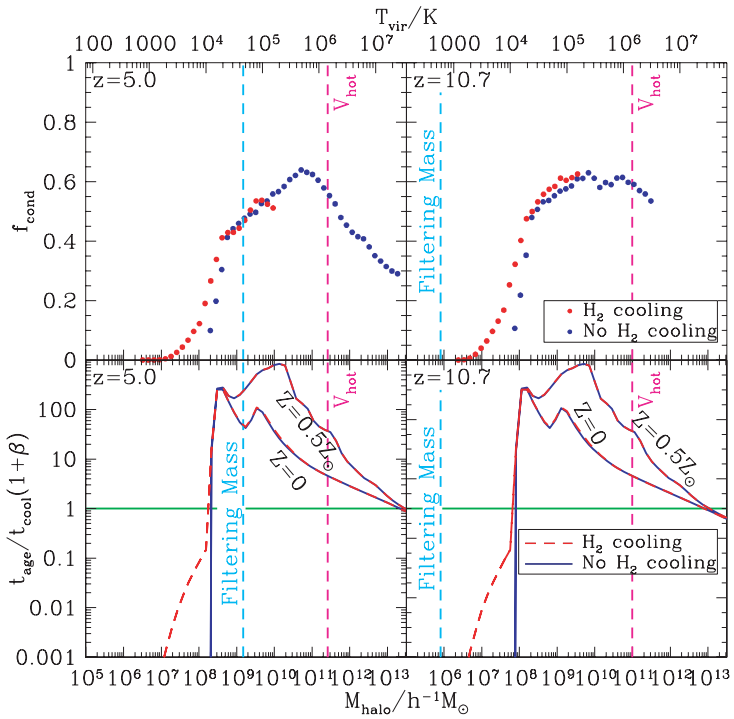

Figure 6. Measures of the ability of baryonic material to cool into the galactic phase are shown for $z = 5$ (left-hand panels) and $z = 10$ (right-hand panels) for models *without* the effects of photoionization and photodissociation feedback. Vertical dashed magenta lines indicate the halo mass for which the virial temperature corresponds to the feedback parameter $V_{\rm hot}$. Vertical, dashed cyan lines indicate the filtering masses for the two models. Upper panels: the average fraction of baryonic material in haloes which has been able to cool and condense into galaxies, as a function of halo mass in the Standard model in the *WMAP* RSI cosmology. Blue circles indicate the cooled fraction for a model with no molecular hydrogen cooling, while red circles show results for a model with molecular hydrogen cooling. Lower panels: cooling measures for metallicities $Z = 0$ and $Z = 0.5\,{\rm Z}_\odot$ (the upper curves at large halo masses): solid blue lines are for the model without molecular hydrogen cooling, while dashed red lines include molecular hydrogen cooling (and Compton cooling). The cooling measure is the ratio $t_{\rm age}/(1+\beta)t_{\rm cool}$. Here β is the ratio of rate of mass ejection due to supernova feedback to the star formation rate in model galaxies. The horizontal green lines indicate the point where the cooling measure equals 1 (i.e. the point above which cooling is efficient in the sense described in the text.

IGM pressure and photoheating of gas in haloes by an ionizing background) and photodissociation feedback (i.e. the destruction of H_2 molecules by the dissociating background). In Fig. 7, we show the resulting star formation rates and H II filling factors as a function of redshift. Photoionization and photodissociation feedback reduce the redshift of reionization in general due to the suppression of galaxy and star formation. Fig. 7 also illustrates once more that the inclusion of H_2 cooling does little to increase the redshift of reionization. In fact, it appears that including H_2 cooling actually reduces the redshift of reionization in the curves plotted here. This is a numerical artifact, arising from the fact that the most massive haloes simulated in our calculations are of lower mass in models which include H_2 cooling (see Section C.2.2). Once this is corrected for, reionization redshifts are almost identical whether or not H_2 cooling is included (see Section 4.1).

Fig. 8 shows the same information as Fig. 6 but for models which now include the effects of photoionization and photodissociation feedback. By comparing this figure with Fig. 6, it can be seen from this figure that the main causes of the suppression of galaxy formation are photoheating of hot gas in haloes and photodissociation of H_2. The filtering mass is approximately the same in models with and without photoionization and photodissociation feedback included[8] (it is not until after reionization that the filtering mass has the chance to respond to the increase in the temperature of the gas). At $z = 5$, we can clearly see that the cooling times in haloes with virial temperatures below around 10^5 K are greatly increased due to this photoheating and photodissociation. At $z = 10$, the effect is much weaker, due both to the lower background at that redshift and the fact that gas in haloes is denser at higher redshifts and so less affected by photoheating.

It can be seen that the inclusion of photoionization and photodissociation feedback has only a small effect on the redshift of reionization. This is not surprising, as the mechanisms through which photoionization feedback works (photoheating of gas in haloes by the ionizing background and an increase in the filtering mass) become important only *after* reionization. Fig. 9 shows the evolution of the H II filling factor,[9] filtering mass and ionizing background for

[8] Note that, even when photoionization feedback is not included the filtering mass still rises due to the heating of the IGM by ionizing photons emitted by galaxies. All that is left out of the calculation is the effect of the filtering mass on subsequent galaxy formation.

[9] The 'spikes' seen in the H II filling factor are an artifact of the semi-analytic model. As in Paper I, we compute the reionization process by drawing a set of halo masses from the halo mass function at some redshift z and constructing merger trees for each halo back to very high redshifts. The galaxy formation calculation then proceeds in each tree. Our merger tree algorithm is not completely accurate in the sense that, when used to produce merger trees that span large redshift ranges, it does not exactly reproduce the expected distribution of progenitor halo masses. This problem becomes more severe as the redshift range spanned increases. Consequently, we repeat our calculations for several starting redshifts z_i (ensuring that at each z_i we have

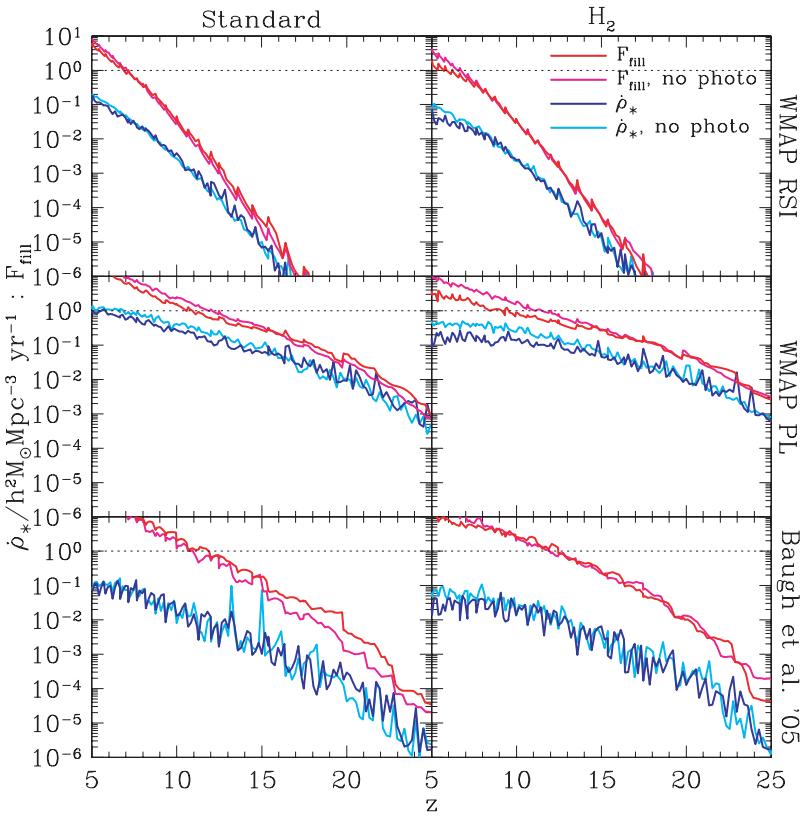

Figure 7. Star formation histories (blue and cyan lines) and H II filling factors calculated using ionized spheres (red and magenta lines). Cyan and magenta lines show results for models where photoionization and photodissociation feedback are neglected, while blue and red lines show results for models including these feedback mechanisms. Results are shown for Standard and H_2 models in *WMAP* PL, *WMAP* RSI and Baugh et al. (2005) cosmologies as indicated for each column and row.

the standard model in the *WMAP* RSI cosmology with and without the effects of photoionization feedback included. The filtering mass begins to rise shortly before reionization occurs, when photoheating first dominates over cooling of gas due to the expansion of the Universe, but does not achieve any significant value until after the Universe is already reionized. Similarly, the ionizing background cannot build up to any significant level (and thereby efficiently photoheat gas in haloes) until the Universe has already been ionized and thereby become transparent to ionizing radiation. Benson et al. (2002a) found that it was the filtering mass which caused the greatest suppression of galaxy formation, seemingly in contradiction with the above result. Benson et al. (2002a) were, however, considering the effects of photoionization feedback on the galaxy population at $z = 0$, by which time the filtering mass had had time to fully respond to the change in the IGM temperature during reionization. When considering the effects of photoionization feedback on the epoch of reionization itself, as we do here, the filtering mass is much less important as it does not have time to fully respond to the heating of the IGM until well after reionization is over. The ionizing background, however, rises very rapidly as reionization occurs, allowing photoheating to have some (albeit small) effect on the redshift of reionization.

the correct distribution of halo masses). The reionization histories shown are constructed by splicing together results from each calculation, using quantities between z_i and z_{i+1} from each calculation. At each z_i, we find discontinuities in the predicted properties due to this splicing process.

3.2.1 Effects of H_2 dissociation?

In general, we find that stars formed via H_2 cooling contribute little to the reionization of the Universe, if they form with a normal IMF. Therefore, it is not surprising that the inclusion or otherwise of H_2 dissociation in our model makes no significant difference to the end results on the ionized filling factor, as shown in Fig. 10.

3.3 Enhanced emission from early generations of stars

In model 'H_2WeakFB-VMS', we enhance the UV emission from stellar populations which form via H_2 cooling (i.e. those which form in dark matter haloes with virial temperatures below 10^4 K) by a factor of 20 relative to a standard IMF, to account for the possibility that these stars might be supermassive 'Population III' stars. The left-hand panel of Fig. 11 shows the resulting H II filling factors versus redshift for this model and for model H_2WeakFB (which is identical apart from having no enhancement of UV emission due to supermassive stars) for the *WMAP* cosmologies. The results are plotted for an escape fraction chosen to give a double reionization: 25 per cent (*WMAP* RSI) and 15 per cent (*WMAP* PL). This results in all models producing full reionization at $z \approx 6$. Note, however, that model H_2WeakFB-VMS displays an extended period of partial reionization back to higher redshifts (from $z \approx 8$ for the *WMAP* RSI cosmology to $z \approx 17$ for the *WMAP* PL cosmology). As a consequence, model H_2WeakFB-VMS has an increased electron scattering optical depth as shown in the right-hand panel of Fig. 11. For example, in the *WMAP* PL cosmology we find $\tau = 0.15$, compared to 0.08 for

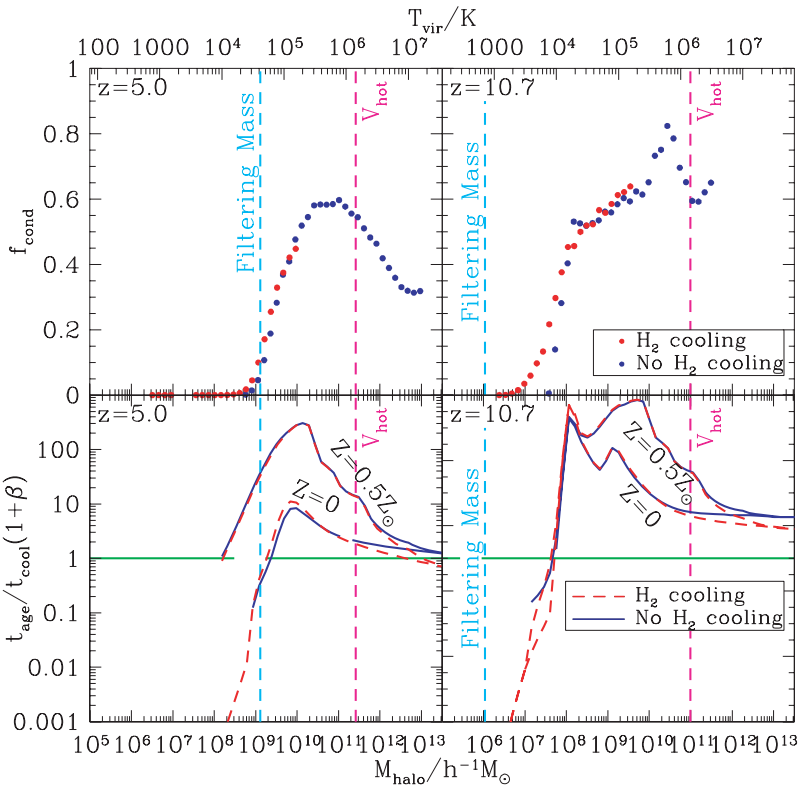

Figure 8. As Fig. 6 but for models *including* the effects of photoionization and photodissociation feedback.

model H_2WeakFB. Values of τ are lower for the other cosmologies, but still come significantly closer to the *WMAP* measurement of $\tau = 0.17 \pm 0.04$. As such, model H_2WeakFB-VMS is in much better agreement with the data obtained from the *WMAP* satellite (Kogut et al. 2003). Enhanced UV emission from stellar populations formed via H_2 cooling is therefore able to reconcile the apparent disparity between a large optical depth and evidence for reionization at $z \approx 6$ in the *WMAP* PL cosmology, by inducing a period of partial reionization.

The dotted line in Fig 11(a) shows the contribution to the H II filling factor from stars forming via H_2 cooling in the H_2WeakFB-VMS model for the *WMAP* PL cosmology. Before $z \approx 20$ these stars dominated the reionization of the Universe, as the steepness of the dark matter halo mass function results in there being very few haloes hot enough to cool via atomic processes at these redshifts. At lower redshifts the abundance of haloes with virial temperature above 10^4 K increases, while galaxy formation in lower temperature haloes becomes inefficient due to increased cooling times and the rising filtering mass.[10] Therefore, haloes hotter than 10^4 K rapidly become the dominant contributors to the H II filling factor.

Obtaining a period of extended partial reionization, or a double reionization, requires an interplay of several key factors. First, the cosmological model must produce enough haloes with virial temperatures below 10^4 K to permit a significant H II filling factor to be obtained from the early generation of stars (assumed to be supermassive). The enhancement factor (taken to be 20 in this work) is also important: if it is too high then reionization may happen once only, if it is too low the early generation of stars will not be able to produce a significant H II filling factor. The decline in the contribution from early generation stars as time progresses must also be sufficiently rapid (otherwise the filling factor will continue to rise, leading to a single reionization event) – requiring either that the filtering mass grows rapidly or that cooling times for haloes cooler than 10^4 K increase rapidly shortly after the first reionization. Despite this requirement that many factors combine to produce a double reionization, we find, for quite reasonable values of the UV enhancement factor, that double reionization is a natural outcome of our H_2WeakFB-VMS model in all cosmologies considered. An extensive survey of parameter space would be required to determine how common such double reionizations are, but this is beyond the scope of this work.

While our modelling of this enhanced UV emission from early stars is very simplified (e.g. we assume that supernova feedback is the same for VMS and for normal stellar populations), and based on uncertain assumptions (e.g. that all haloes which cool by H_2 produce VMS), it is still relatively sophisticated compared to previous studies of double reionization (Cen 2003; Wyithe & Loeb 2003) to the extent that ours is the only model which follows the full hierarchical merging history of each forming galaxy and, furthermore, computes the fraction of baryons able to cool and the fraction able to turn into stars using physically motivated prescriptions rather than simply making these fractions parameters of the model.

4 OBSERVABLE QUANTITIES

4.1 The epoch of reionization

Table 3 lists the redshifts of reionization in all of the models considered (all results shown include the effects of photoionization and photodissociation feedback), along with the corresponding

[10] The filtering mass in the H_2WeakFB-VMS model in the *WMAP* PL cosmology rises above the mass corresponding to a 10^4 K halo at $z = 15$.

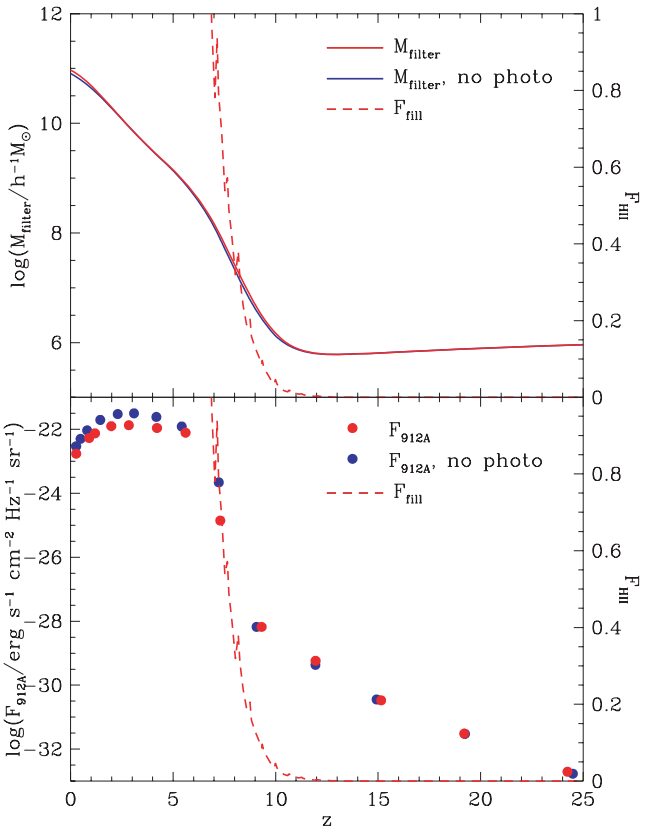

Figure 9. Indicators of the strength of photoionization feedback mechanisms as a function of redshift for the Standard model in the *WMAP* RSI cosmology. Results are shown for this single model and cosmology only as they are typical of all of the models we consider. The solid red (blue) line in the upper panel shows the filtering mass (scale on left-hand axis) when photoionization feedback is (is not) included. In the lower panel red (blue) circles show the ionizing background intensity (scale on left-hand axis) when photoionization feedback is (is not) included. The dashed red line in both panels indicates the H II filling factor (scale on right-hand axis) for a model including photoionization feedback.

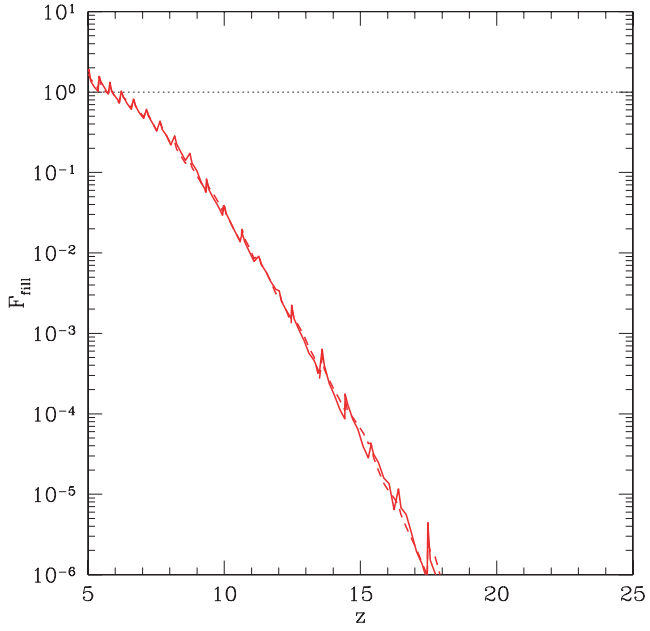

Figure 10. The H II filling factor as a function of redshift for model H_2 with (solid line) and without (dashed line) dissociation of H_2 molecules included in our calculations. Results are shown for the *WMAP* RSI cosmology only, but are typical of all other cosmologies and models.

electron-scattering optical depths. Observations of the CMB suggest a large value for $z_{\rm reion}$. Kogut et al. (2003) used *WMAP* measurements of the temperature-polarization (TE) cross-power spectrum to estimate the optical depth τ. They found a best-fitting value $\tau = 0.17$, with a 95 per cent confidence range $0.09 \leqslant \tau \leqslant 0.28$. For the *WMAP* cosmology, this corresponds to a best-fitting value of the reionization redshift $z_{\rm reion} = 17$, and a 95 per cent confidence range $11 < z_{\rm reion} < 24$, if one assumes instantaneous reionization, i.e. if the ionized fraction increased from 0 to 1 at $z = z_{\rm reion}$. Spergel et al. (2003) found very similar results from their multiparameter fit to the whole *WMAP* data set combined with data on galaxy clustering.

Since our models predict a gradual rather than an instantaneous reionization, we compare our models with the *WMAP* data using the value of τ rather than the value of $z_{\rm reion}$. From Table 3, we see that all of our models based on the *WMAP* PL cosmology and assuming $f_{\rm esc} = 1$ are consistent with the measured τ (at 95 per cent confidence), apart from the H_2WeakFB-VMS model, which predicts $\tau = 0.30$, marginally too high. These models predict $z_{\rm reion}$ in the range 11–22. However, when $f_{\rm esc}$ is reduced to 15 per cent, the H_2WeakFB-VMS model in this cosmology predicts $\tau = 0.15$ (and $z_{\rm reion} = 17$), in good agreement with the *WMAP* measurement. In contrast, all but one of our models based on the *WMAP* RSI cosmology, and assuming $f_{\rm esc} = 1$, are excluded (at 95 per cent confidence) by the *WMAP* measurement of τ, due to the predicted optical depths being too low, the one exception being the H_2WeakFB-VMS model, which (with $\tau = 0.10$) is marginally compatible. The models in the *WMAP* RSI cosmology predict $z_{\rm reion}$ in the range 6–11. Thus, while the *WMAP* RSI cosmology cannot be completely excluded based on the *WMAP* measurement of τ, it does appear very unlikely. If the escape fraction $f_{\rm esc}$ is in fact significantly less than unity, then the predicted reionization epochs will be shifted to considerably lower redshifts in all of the models. For the Baugh et al. (2005) cosmology, we find that the Standard and H_2 models are marginally consistent with the *WMAP* constraint, while models with weak feedback are able to match that constraint within current errors.

Observations of the spectra of the highest redshift quasars at $z \sim 6$ show evidence of Gunn–Peterson absorption by neutral hydrogen in the IGM (Becker et al. 2001; Djorgovski et al. 2001; Fan et al. 2003, 2006; White et al. 2003), which has been interpreted as evidence that the Universe reionized at redshifts only slightly above 6, since simple theoretical models predict that the transition from the IGM being almost completely neutral to almost completely ionized should have happened over a rather narrow redshift range. A low $z_{\rm reion} \sim 7$ is in apparent contradiction with the results from *WMAP*. This interpretation of the observed quasar spectra has however been questioned by some authors (e.g. Songaila 2004). Even if current measurements of the Gunn–Peterson optical depth $\tau_{\rm GP}$ at $z \sim 6$ are correct, they can be explained by very small neutral hydrogen fractions in the IGM, e.g. White et al. (2003) measure $\tau_{\rm GP} \sim 5$–20 for two quasars at $z \approx 6.3$, which can be produced with neutral fractions of only $f_{\rm HI} \sim 1.5$–6×10^{-5}. Therefore, in principle the IGM might have ceased to be mainly neutral at much higher redshifts, and still produce detectable Gunn–Peterson absorption at $z \sim 6$. Note that the only way in which we are able to produce a 'double reionization' is by assuming that stars forming via H_2 cooling (PopIII stars) are very massive objects producing around 20 times more ionizing

Table 3. The epochs of reionization for all models considered in this work. Columns 1–3 specify input parameters of the models: column 1 specifies the cosmological model while column 2 specifies the galaxy formation model used and column 3 specifies the value of $f_{\rm esc}$ used. The remaining columns list output quantities. Column 4 lists the redshift of reionization – defined as the redshift at which the filling factor in the ionized sphere reionization calculation reaches unity. The first value gives $z_{\rm reion}$ as found directly from our calculations, while the second value accounts for the finite upper mass limit of the haloes which we simulate. The true $z_{\rm reion}$ should lie between these two values. The next column gives the electron scattering optical depth (assuming that helium is singly ionized at the same time as hydrogen is ionized, and that helium is doubly ionized below $z = 3$). The two values again refer to the calculation without and with the correction for the finite upper mass limit of simulated haloes. The next two columns list the redshifts at which the neutral hydrogen fraction in the IGM drops to 10^{-2} and 10^{-5}, respectively, as determined using the IGM evolution model of Benson et al. (2002a). The final column lists the redshift at which the mean Gunn–Peterson optical depth reaches a value of 3.

Cosmology	Model	$f_{\rm esc}$	$z_{\rm reion}$	τ	$z_{f_{\rm HI}=10^{-2}}$	$z_{f_{\rm HI}=10^{-5}}$	$z_{\tau_{\rm GP}=3}$
WMAP RSI	Standard	1.00	6.9–6.9	0.061–0.061	7.36	7.25	7.10
WMAP RSI	MetFree	1.00	7.3–7.4	0.064–0.068	7.60	7.50	7.38
WMAP RSI	TopHeavy	1.00	8.7–8.7	0.079–0.080	9.02	8.94	8.79
WMAP RSI	WeakFB	1.00	8.1–8.1	0.073–0.079	8.50	8.45	8.33
WMAP RSI	H_2	1.00	6.1–6.6	0.055–0.063	6.91	6.69	6.51
WMAP RSI	H_2WeakFB	1.00	7.8–8.1	0.074–0.075	8.44	8.38	8.22
WMAP RSI	H_2WeakFB	0.25	4.4–4.4	0.048–0.055	7.60	7.34	7.17
WMAP RSI	H_2WeakFB-VMS	1.00	10.6–10.6	0.102–0.102	8.46	8.39	8.21
WMAP RSI	H_2WeakFB-VMS	0.25	4.4–4.4	0.076–0.080	7.55	7.36	7.13
WMAP PL	Standard	1.00	10.8–10.8	0.141–0.141	11.38	10.74	10.54
WMAP PL	MetFree	1.00	11.4–11.4	0.157–0.157	11.92	11.20	10.93
WMAP PL	TopHeavy	1.00	15.3–15.3	0.213–0.213	14.54	14.46	14.33
WMAP PL	WeakFB	1.00	14.1–14.1	0.191–0.191	13.73	12.77	12.42
WMAP PL	H_2	1.00	9.2–10.8	0.131–0.141	7.87	7.79	7.73
WMAP PL	H_2WeakFB	1.00	14.1–14.1	0.206–0.206	13.87	12.56	12.31
WMAP PL	H_2WeakFB	0.15	6.1–8.3	0.095–0.107	7.99	7.85	7.84
WMAP PL	H_2WeakFB-VMS	1.00	22.2–22.2	0.299–0.299	13.77	12.71	12.40
WMAP PL	H_2WeakFB-VMS	0.15	6.9–8.8	0.183–0.196	7.98	7.85	7.84
Baugh et al. (2005)	Standard	1.00	11.9–11.9	0.108–0.108	7.97	7.44	7.39
Baugh et al. (2005)	H_2	1.00	12.5–12.5	0.118–0.118	7.55	7.23	6.87
Baugh et al. (2005)	H_2WeakFB	1.00	14.4–14.4	0.144–0.144	8.62	8.19	7.76

photons per unit mass than for a standard IMF. Nevertheless, under such an assumption, double reionization is a natural consequence of our model.

To allow a rough comparison of our models with observations of Gunn–Peterson absorption, we have calculated the average neutral hydrogen fraction in the IGM, $f_{\rm HI}$ (defined as the ratio of volume averaged values of the abundances of neutral hydrogen $n_{\rm HI}$ and total hydrogen $n_{\rm H}$, which is the quantity which is most directly comparable to measurements derived from the Gunn–Peterson effect,[11]) and used this to predict the mean Gunn–Peterson optical depth $\langle\tau_{\rm GP}\rangle$, both as functions of redshift. Results for these quantities for several of our models (using $f_{\rm esc} = 1$) are plotted in Fig. 12. We also give in Table 3 for each model the (lowest) redshift at which the mean neutral fraction drops to $f_{\rm HI} = 10^{-2}$ and 10^{-5}, respectively, and the redshift for which $\langle\tau_{\rm GP}\rangle = 3$.

The results for $f_{\rm HI}$ and $\langle\tau_{\rm GP}\rangle$ for the different models favour the models in the *WMAP* RSI cosmology if $f_{\rm esc} = 1$, or alternatively $f_{\rm esc}$ appreciably smaller than unity if the *WMAP* PL cosmology is assumed. Note, however, that none of our models produce a neutral hydrogen fraction at $z = 6$ which is sufficiently high to explain the observed Gunn–Peterson optical depths (e.g. White et al. 2003; Fan et al. 2006) when $f_{\rm esc} = 1$. (The H_2 model in the *WMAP* RSI cosmology comes closest, achieving $\tau_{\rm GP} = 3$ by $z = 6.5$, but by $z = 6$ the optical depth has already dropped to 0.35.) The neutral fraction falls very rapidly during reionization as can be seen from the fifth and sixth columns of Table 3, which list the redshifts at which neutral fractions of 10^{-2} and 10^{-5} are reached. Neutral fractions comparable to those required by the observational data could be obtained if $f_{\rm esc}$ were reduced, at the expense of a slightly lower reionization redshift.

We find that the Baugh et al. (2005) model achieves reionization at $z \approx 12$ for an escape fraction of 100 per cent, while adding H_2 cooling to this model raises the reionization redshift to $z \approx 12.5$. The corresponding values of the electron scattering optical depth are $\tau = 0.108$ and 0.118, respectively. This model is therefore consistent with the *WMAP* measurement of the optical depth, given the rather large errors on that datum. However, the neutral fraction in this model is too low at $z \approx 6$ to explain observations of the Gunn–Peterson trough (the optical depth at $z = 6$ for the Baugh et al. (2005) model lies in the range 0.01–0.1 for the various models considered here). While this could be rectified by lowering the escape fraction in order to reduce the reionization redshift to be closer to $z \approx 6$, this would also reduce the optical depth.

4.2 CMB anisotropies

We have computed the temperature and polarization power spectra predicted by our model reionization histories. Fig. 13 shows results for the *WMAP* PL and *WMAP* RSI cosmologies.[12] To

[11] That is, $f_{\rm HI} = \int n_{\rm HI}\,{\rm d}V / \int n_{\rm H}\,{\rm d}V$. This is equivalent to a *mass weighted* average of the local neutral fraction $x_{\rm HI}$.

[12] The Baugh et al. (2005) cosmology was not designed to accurately match available CMB data and so is not considered here.

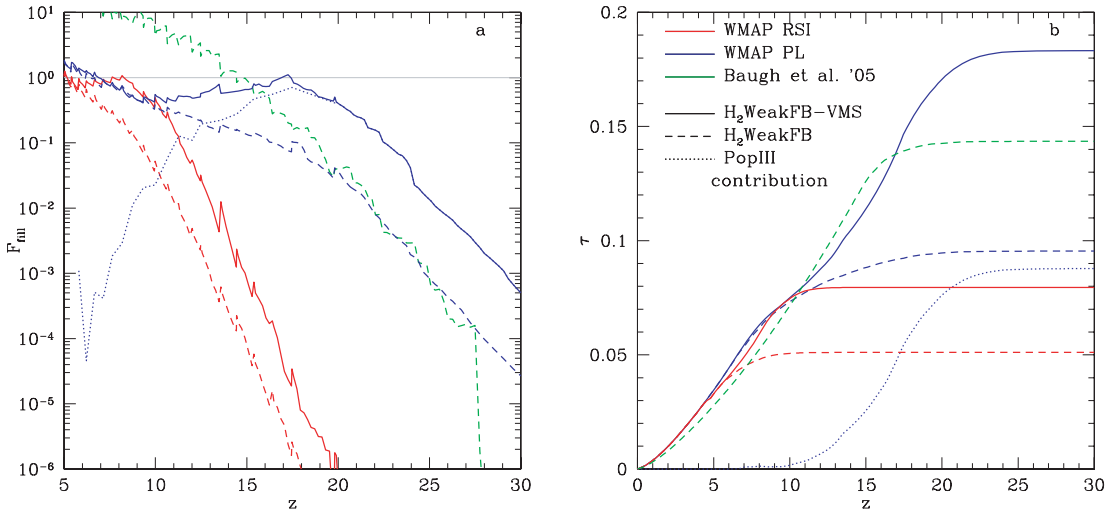

Figure 11. Left-hand panel: the H II filling factor as a function of redshift for models H_2WeakFB-VMS (solid lines) and H_2WeakFB (dashed lines) in our three cosmologies (see key in right-hand panel). The dotted line shows, for the *WMAP* PL cosmology only, the contribution to the H II filling factor from stars forming in haloes with virial temperatures below 10^4 K. Escape fractions have been chosen so as to produce double reionizations in each model, apart from the Baugh et al. (2005) model (for which we do not compute the H_2WeakFB-VMS model) in which case we simply choose $f_{\rm esc} = 1$. The escape fractions required are $f_{\rm esc} = 0.25$, 0.15 and 1.0 for the *WMAP* RSI, *WMAP* PL and Baugh et al. (2005) cosmologies, respectively. Right-hand panel: the electron scattering optical depth as a function of redshift for the same models.

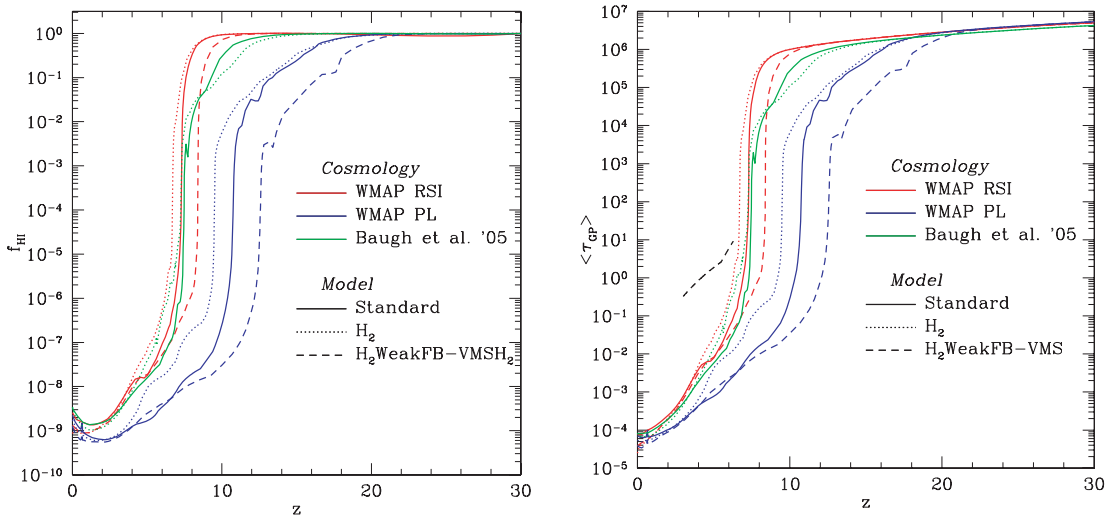

Figure 12. The mean mass-weighted H I fraction (left-hand panel) and Gunn–Peterson optical depth (right-hand panel) as functions of redshift for various models in our three cosmologies (see key in figure for colour and line-type coding). The model results plotted here all assume $f_{\rm esc} = 1$. The dashed black line in the right-hand panel indicates the fit to the mean optical depth reported by Fan et al. (2006).

calculate CMB temperature spectra, we employed the Boltzmann code of Sugiyama (1995), which has been used to calibrate the accuracy of publicly available codes, e.g. CMBFast (Seljak et al. 2003). The models shown in the figure have $\sigma_8 = 0.9$ for the *WMAP* PL cosmology and $\sigma_8 = 0.83$ for the *WMAP* RSI cosmology. Furthermore, all models are shown for $f_{\rm esc} = 1$, except for the model labelled 'grad: $f_{\rm esc} = 0.15$' for which we plot the H_2WeakFB-VMS model with $f_{\rm esc} = 0.15$ (which exhibits a double reionization). For comparison, we also show the CMB anisotropies predicted for a toy model of instant reionization, in which the IGM instantaneously becomes fully reionized at $z_{\rm reion} = 17$, chosen to give the optical depth $\tau = 0.17$ that best fits the *WMAP* data.

For the *WMAP* PL cosmology, it is clear that the H_2WeakFB-VMS model with $f_{\rm esc} = 1.0$ is ruled out from the TT spectrum. The reason for the failure of this model is the combination of the large optical depth (which suppresses CMB temperature anisotropies on small scales) and $\sigma_8 = 0.9$. One could try to make this model agree with the TT data by increasing σ_8 by ∼20 per cent. However, this increase leads to galaxy formation at higher redshifts and hence a larger optical depth. Thus, while increasing σ_8 helps to reconcile this model with the high ℓ part of the TT spectrum, it produces a very large polarization signal at small ℓ, which is inconsistent with that measured by *WMAP*.

The H_2 and Standard models in the *WMAP* PL cosmology also do not agree well with the *WMAP* TT spectrum, due to their optical depths being too low.

It is very interesting that the gradual reionization model (the H_2WeakFB-VMS model with $f_{\rm esc} = 0.15$; light blue line in Fig. 13)

© 2006 The Authors. Journal compilation © 2006 RAS, MNRAS **369,** 1055–1080

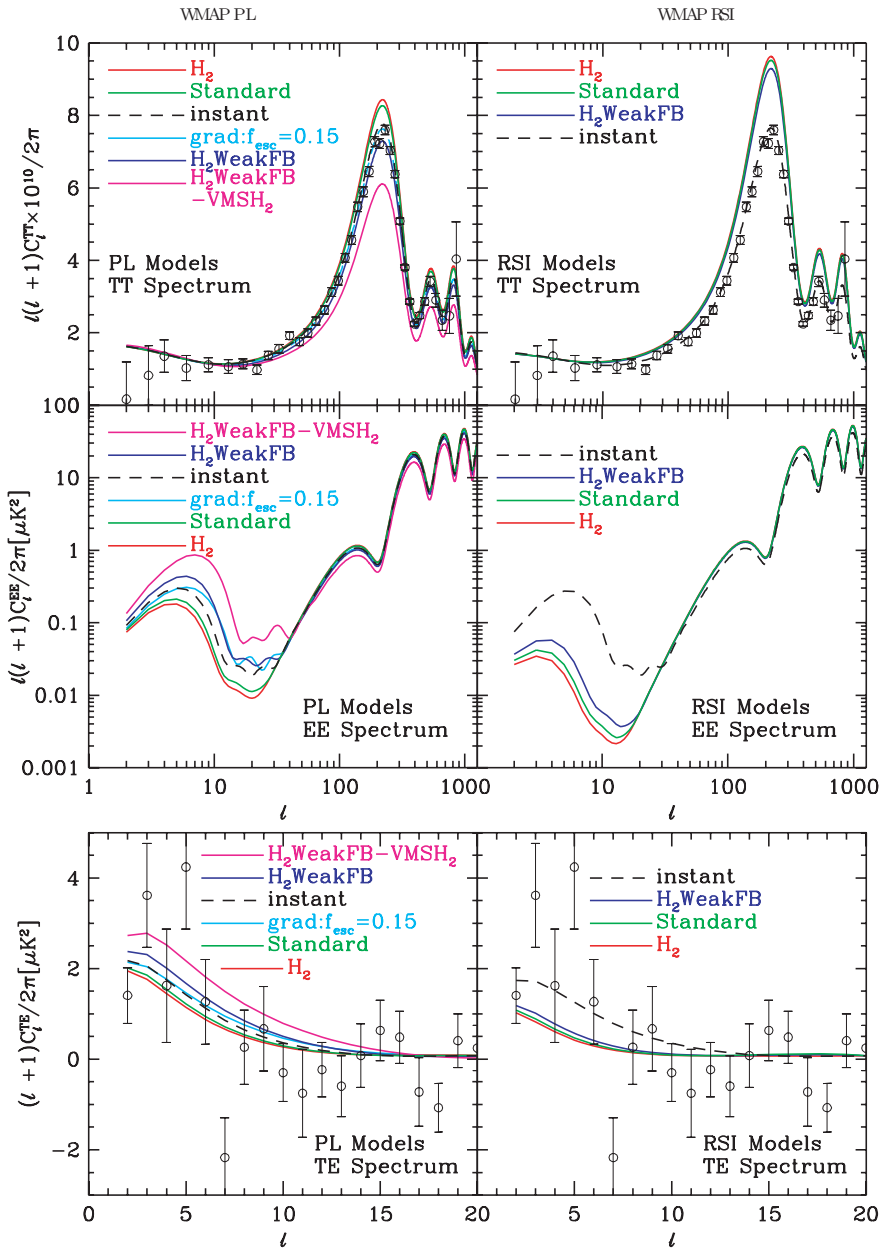

Figure 13. CMB anisotropy power spectra for the *WMAP* PL (left-hand panels) and *WMAP* RSI (right-hand panels) cosmologies. The power spectra of temperature fluctuations (TT) are shown in the upper panels, while the middle and lower panels show polarization power spectra (EE in the middle panels and TE in the lower panels). The circles with error bars show observational data from the *WMAP* satellite (Hinshaw et al. 2003; Kogut et al. 2003). The dashed black lines (labelled instant) show results for a simple toy model in which reionization is instantaneous, and occurs at a redshift selected to best fit the *WMAP* data. The coloured lines show the predictions for various models from this work, assuming $f_{\rm esc} = 100$ per cent unless indicated otherwise. The cyan lines (labelled grad) show results for the H_2WeakFB-VMS model with an escape fraction of 15 per cent, which exhibits gradual reionization.

shows a unique behaviour in the EE spectrum compared with the H_2WeakFB and instant reionization models, while the optical depths of these three models are similar. In particular, this model shows a feature (a bump) at $\ell \approx 20$. The location of the $\ell \approx 8$ peak and the width of this peak are slightly different from other models of comparable optical depth (i.e. the width is slightly greater and the peak location is shifted to the right). These differences are due to the fact that the reionization process occurs over a longer period in this model. On the other hand, there are only small differences in TT and TE spectra. We expect the EE spectrum to be the best probe of the details of reionization such as the duration and the epoch (Holder et al. 2003).

None of the models in the *WMAP* RSI cosmology fits the *WMAP* CMB data, since we employ the σ_8 normalization here (i.e. $\sigma_8 = 0.83$ in all of the RSI models plotted), and the optical depths of all of the models are very small.

The calculations here do not include the effects of patchiness in the reionization process. Patchy reionization can produce secondary anisotropies in the CMB on small angular scales, $\ell \gtrsim 3000$ (Paper I). These anisotropies must be studied in detail using numerical

simulations which provide the spatial distribution and peculiar motions of the ionized gas. We defer this work to a future paper.

5 DISCUSSION

We have described the most detailed and physically realistic semi-analytic calculation of the reionization of the Universe that has been carried out to date. These calculations rely on a physical model of galaxy formation which includes cooling through Compton and molecular hydrogen channels, a physically motivated model of star formation and feedback, and the feedback exerted on galaxy formation by the process of reionization itself.

We find that, in the *WMAP* PL cosmological model, the Universe can be reionized by $z \approx 14$–15 either if early generations of stars are preferentially very massive or if feedback in the high-redshift Universe is much weaker than it is today. A direct measurement of the reionization redshift is not yet possible. Instead, we can compute, from our reionization histories, optical depths for electron scattering which can be compared directly with the measurement of this quantity from *WMAP*. Reionization at these redshifts produces optical depths in the range $\tau = 0.19$–0.21. This is consistent with the *WMAP* result of $\tau = 0.17 \pm 0.04$. In the *WMAP* RSI cosmology, however, reionization cannot occur before $z \approx 9$, due to the much smaller number of massive haloes which are able to form at high redshifts in this model. As such, the *WMAP* RSI cosmological model is barely consistent with the *WMAP* result for τ, achieving at most $\tau \approx 0.1$. This suggests one of three possibilities:

(i) the *WMAP* estimate of τ is too high (either due to systematic effects or a simple statistical fluctuation);

(ii) the RSI cosmology is incorrect (perhaps due to problems in determining the normalization of the Lyman α forest power spectrum; see, for example, Viel, Weller & Haehnelt 2004);

(iii) our understanding of galaxy formation, or the CDM hierarchical model itself, are incorrect at high redshifts.

Option (iii) is almost certainly true to some degree – there are clearly large gaps remaining in our understanding of galaxy formation. We include within option (iii) the possibility that some neglected galaxy formation physics, for example a contribution to the ionizing emissivity from supermassive black holes forming in high redshift galaxies (Ricotti & Ostriker 2004), affects our results. However, in the *WMAP* RSI cosmology, the basic problem is that there are simply too few haloes existing at high redshifts to allow enough galaxies to form and so cause reionization. Only if 100 per cent of the baryons which have the potential to cool in high redshift haloes are allowed to turn instantly into stars can a sufficiently large τ be attained in the *WMAP* RSI cosmology (Haiman & Holder 2003; Somerville et al. 2003). Within our models, even with rather extreme assumptions about the strength of feedback, the IMF of early generations of stars, and the escape fraction of ionizing photons, we are unable to achieve reionization prior to $z = 9$ for the *WMAP* RSI cosmology. Thus, the *WMAP* RSI cosmology seems hard to reconcile with our understanding of galaxy formation in the CDM hierarchical model.

The recent three-year data release from the *WMAP* satellite confirms (i) above. Spergel et al. (2006) report $\tau = 0.093 \pm 0.029$. This optical depth is well within the reach of our models with $f_{\rm esc} = 1$, although physically motivated escape fractions (see Appendix B) would result in models still struggling to reach such large optical depths.

Using the model of Baugh et al. (2005), which was designed to match the properties of low- and high-redshift galaxies, we find that reionization can occur by $z \approx 12$ for $f_{\rm esc} = 1$, producing an optical depth of $\tau = 0.11$ which is marginally consistent with the *WMAP* data. This optical depth can be increased to $\tau = 0.12$ by including H_2 cooling (which does not affect the ability of this model to match local galaxy data), and can be increased further if the strength of feedback is reduced (which *would* affect the ability of the model to match local galaxy data).

Recently, the Boomerang experiment has provided an improved measurement of the cosmological parameters (MacTavish et al. 2005). We have computed reionization models using these revised cosmological parameters (specifically, we use the 'CMBall+B03+LSS' parameter set from MacTavish et al. 2005, as listed in Table 1). We find that at high redshifts ($\gtrsim$20) the Boomerang cosmology produces a lower H II filling factor than the *WMAP* PL cosmology (for all of our different galaxy formation model parameter sets). However, this difference becomes less towards lower redshifts, and we find that the epoch of reionization in the Boomerang cosmology is usually very close to that for the corresponding model in the *WMAP* PL cosmology. The one exception is for the H_2WeakFB-VMS model, in which reionization occurs significantly later in the Boomerang cosmology (since reionization happens early in the H_2WeakFB-VMS model, the differences between the Boomerang and *WMAP* cosmologies are more pronounced). None the less, double reionization can still occur in our model, even in the Boomerang cosmology. We conclude, therefore, that the conclusions we present for the *WMAP* cosmology can also be applied to the revised cosmological parameter set suggested by the Boomerang data.

Throughout this work we have used the values of σ_8 which best fit the *WMAP* data ($\sigma_8 = 0.90$ and 0.83 for the PL and RSI models, respectively, Spergel et al. 2003). However, the value of σ_8 remains somewhat uncertain, with plausible values lying between about 0.7 and 1.1. In particular, Spergel et al. (2006) report $\sigma_8 = 0.76 \pm 0.05$. We have rerun our *WMAP* PL models using these two extreme values of σ_8. Results for the ionized filling factor are shown in Fig. 14. Not surprisingly, lowering σ_8 reduces the redshift of reionization, while increasing σ_8 increases it. The effect is greater for models in which reionization happens at higher redshifts, since here those haloes which are able to form galaxies and emit ionizing photons correspond to very rare fluctuations in the initial Gaussian density field. As such, their abundance is very sensitive to the normalization of that density field. Typically, models which reionize at $z \approx 10$ have their epoch of reionization reduced/increased by $\Delta z \approx 1$ for a reduction/increase of 0.2 in σ_8, while models with reionization at $z \approx 20$ experience a Δz of 4–5 for the same change in σ_8 (these values of Δz do depend also on the parameters of the galaxy formation calculation). Thus, σ_8 can have a large effect on predictions of the epoch of reionization.

These results show that, even for the low σ_8 reported by Spergel et al. (2006), sufficiently early reionization can be achieved (i.e. to match the optical depth also reported by Spergel et al. 2006) providing $f_{\rm esc} = 100$ per cent.

Most of the results we present here are based on the assumption that the fraction of ionizing photons which escape from galaxies into the IGM is $f_{\rm esc} = 1$. In the local Universe, and at $z \sim 3$, observational estimates give much smaller values. If $f_{\rm esc}$ were low at higher redshifts also, then our estimates of $z_{\rm reion}$ would be substantially too high. A value of $f_{\rm esc} = 0.1$ would reduce the reionization redshift in even our best models to $z \lesssim 10$. In Appendix B, we present reionization histories computed using a simple physically motivated model for the escape fractions. Clearly, a better theoretical understanding of $f_{\rm esc}$ is crucial if we are to use measures of the

© 2006 The Authors. Journal compilation © 2006 RAS, MNRAS **369,** 1055–1080

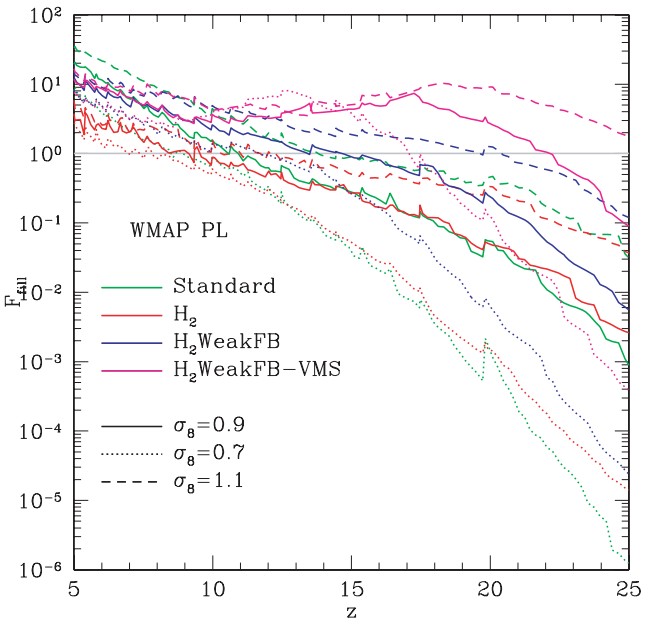

Figure 14. The H II filling factor as a function of redshift in various models (distinguished by line colour – see key in figure) for different values of σ_8 (distinguished by line type – see key in figure). We show results for the *WMAP* PL cosmology only – results for other cosmologies are qualitatively similar. All results are shown for an escape fraction of 100 per cent.

reionization epoch to put interesting constraints on the process of galaxy formation.

We have found that extended periods of partial reionization and double reionization can occur in models in which early generations of stars formed via H_2 cooling are very massive (with enhanced ionizing luminosities relative to a standard IMF). Such models are able to delay full reionization until $z \approx 6$–7, while achieving an electron scattering optical depth large enough to be consistent with the results from *WMAP*. While a more extensive study of such models would be required to determine whether or not they can match all of the current observational constraints, they appear to be a promising route to reconciling the apparently discrepant estimates of the reionization redshift from the CMB polarization measured by *WMAP* and from the Gunn–Peterson absorption detected in the spectra of $z \sim 6$ quasars. It should be noted, however, that these models rely on the assumption of very weak feedback from supernovae. Within CDM cosmogonies, it has long been argued that feedback from supernovae must be very effective in order to reduce the number of low-luminosity galaxies at low redshifts which are otherwise produced (see Benson et al. 2003a and references therein). Thus, if double reionization (or an extended period of partial reionization) is required to meet observational constraints, then it appears that the strength of supernova feedback must vary with redshift, or else some other feedback mechanism must act at low redshifts to reduce the number of faint galaxies formed. This underlies the urgent need for the inclusion of physical recipes for feedback into semi-analytic models of galaxy formation, rather than the phenomenological feedback rules currently employed.

In conclusion, the *WMAP* RSI cosmology is very difficult to reconcile with our current understanding of galaxy formation (although a wide parameter space of models remains to be searched) if the optical depth were as large as suggested by the first year *WMAP* data. The *WMAP* PL model, in contrast, is able to achieve a sufficiently high optical depth with reasonable assumptions about the nature of galaxy formation at high redshifts. The three-year *WMAP* data imply an optical depth significantly lower than that found in the first year data (i.e. $\tau = 0.093 \pm 0.029$; Spergel et al. 2006). Reionization can then be achieved sufficiently early in both the *WMAP* PL and *WMAP* RSI cosmologies, providing $f_{\rm esc}$ is close to 1 and σ_8 is not too low. We will present a more detailed analysis of the implications of the three-year *WMAP* data in a follow-up paper.

In principle, the value of the optical depth to reionization is one of the few strong constraints on galaxy formation at high redshifts. The uncertainty mentioned in option (i) has already been reduced by the *WMAP* three-year data. Thus, the ability to reionize the Universe sufficiently early will now become one of the key tests that any model of galaxy formation must pass before it can be considered to be viable.

ACKNOWLEDGMENTS

AJB acknowledges support from a Royal Society University Research Fellowship. NS is supported by a Grant-in-Aid for Scientific Research from JSPS (No. 17540276). This work was also supported by the PPARC rolling grant for extragalactic astronomy and cosmology at Durham. We wish to thank Carlton Baugh, Shaun Cole and Carlos Frenk for permitting us to use the GALFORM model of galaxy formation in this work. We also acknowledge helpful suggestions from an anonymous referee.

REFERENCES

Abel T., Bryan G. L., Norman M. L., 2000, ApJ, 540, 39
Abel T., Bryan G. L., Norman M. L., 2002, Sci, 295, 93
Anninos P., Norman M. L., 1996, ApJ, 460, 556
Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 1998, ApJ, 498, 504
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
Becker R. H. et al., 2001, AJ, 122, 2850
Benson A. J., Cole S., Frenk C. S., Baugh C. M., Lacey C. G., 2000, MNRAS, 311, 793
Benson A. J., Nusser A., Sugiyama N., Lacey C. G., 2001, MNRAS, 320, 153 (Paper I)
Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002a, MNRAS, 333, 156
Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002b, MNRAS, 333, 177
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Cole S., 2003a, ApJ, 599, 38
Benson A. J., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2003b, MNRAS, 343, 679
Benson A. J., Lacey C. G., Frenk C. S., Baugh C. M., Cole S., 2004, MNRAS, 351, 1215
Bergvall N., Zackrisson E., Andersson B.-G., Arnberg D., Masegosa J., Östlin G., 2006, A&A, 448, 513
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bower R. G., 1991, MNRAS, 248, 332
Bromm V., Loeb A., 2004, Nat, 425, 812
Bromm V., Kudritzki R. P., Loeb A., 2001, ApJ, 552, 464
Bromm V., Coppi P. S., Larson R. B., 2002, ApJ, 564, 23
Bruscoli M., Ferrara A., Fabbri R., Ciardi B., 2000, MNRAS, 318, 1068
Bruzual A. G., Charlot S., 1993, ApJ, 405, 538
Cazaux S., Spaans M., 2004, ApJ, 611, 40
Cen R., 2003, ApJ, 591, L5
Chen X., Kamionkowski M., 2004, Phys. Rev. D, 70, 3502
Chiu W. A., Ostriker J. P., 2000, ApJ, 534, 507
Ciardi B., Ferrara A., Governato F., Jenkins A., 2000, MNRAS, 314, 611
Ciardi B., Ferrara A., White S. D. M., 2003, MNRAS, 344, 7
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Djorgovski S. G., Castro S. M., Stern D., Mahabal A. A., 2001, ApJ, 560, 5

Dove J. B., Shull M. J., 1994, ApJ, 423, 196
Dove J. B., Shull M. J., Ferrara A., 2000, ApJ, 531, 486
Fan X. et al., 2003, AJ, 128, 515
Fan X. et al., 2006, preprint (astro-ph/0512082)
Ferrara A., Bianchi S., Cimatti A., Giovanardi C., 1999, ApJS, 123, 437
Fujita A., Martin C. L., Mac Low M.-M., Abel T., 2003, ApJ, 599, 50
Furlanetto S. R., Loeb A., 2005, ApJ, 634, 1
Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, ApJ, 613, 16
Galli D., Palla F., 1998, A&A, 335, 403
Gnedin N. Y., 2000, ApJ, 542, 535
Gnedin N. Y., Ostriker J. P., 1997, ApJ, 486, 581
Granato G. L., Lacey C. G., Silva L., Bressan A., Baugh C. M., Cole S., Frenk C. S., 2000, ApJ, 542, 710
Haiman Z., Holder G., 2003, ApJ, 595, 1
Haiman Z., Loeb A., 1997, ApJ, 483, 21
Hansen S. H., Haiman Z., 2004, ApJ, 600, 26
Hinshaw G. et al., 2003, ApJS, 148, 135
Holder G. P., Haiman Z., Kaplinghat M., Knox L., 2003, ApJ, 595, 13
Hutchins J. B., 1976, ApJ, 205, 103
Jenkins A. et al., 1998, ApJ, 499, 20
Kasuya S., Kawasaki M., 2004, Phys. Rev. D, 70, 103 519
Kasuya S., Kawasaki M., Sugiyama N., 2004, Phys. Rev. D, 69, 3512
Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Kennicutt R. C., 1983, ApJ, 272, 54
Kogut A. et al., 2003, ApJS, 148, 161
Le Delliou M., Lacey C., Baugh C. M., Guiderdoni B., Bacon R., Courtois H., Sousbie T., Morris S. L., 2005a, MNRAS, 357, L11
Le Delliou M., Lacey C., Baugh C. M., Morris S. L., 2006, MNRAS, 365, 712
Leitherer C., Ferguson H., Heckman T. M., Lowenthal J. D., 1995, ApJ, 454, 19
Machacek M. E., Bryan G. L., Abel T., 2001, ApJ, 548, 509
MacTavish et al., 2005, ApJ submitted (astro-ph/0507503)
Nagashima M., Lacey C. G., Baugh C. M., Frenk C. S., Cole S., 2005a, MNRAS, 358, 1247
Nagashima M., Lacey C. G., Okamoto T., Baugh C. M., Frenk C. S., Cole S., 2005b, MNRAS, 363, L31
Naselsky P., Chiang L.-Y., 2004, MNRAS, 347, 795
Onken C. A., Miralda-Escudé J., 2004, ApJ, 610, 1
Peebles P. J. E., 1968, ApJ, 153, 1
Ricotti M., Ostriker J. P., 2004, MNRAS, 352, 547
Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, ApJ, 571, 30
Seager S., Sasselov D. D., Scott D., 2000, ApJS, 128, 407
Seljak U., Sugiyama N., White M., Zaldarriaga M., 2003, Phys. Rev. D, 68, 083 507
Somerville R. S., Bullock J. S., Livio M., 2003, ApJ, 593, 616
Songaila A., 2004, AJ, 127, 2598
Spergel D. N. et al., 2003, ApJS, 148, 175
Spergel D. N. et al., 2006, ApJ, submitted (astro-ph/0603449)
Steidel C. C., Pettini M., Adelberger K. L., 2001, ApJ, 546, 665
Sugiyama N., 1995, ApJS, 100, 281
Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, ApJ, 474, 1
Tumlinson J., Giroux M. L., Shull M. J., Stocke J. T., 1999, AJ, 118, 2148
Valageas P., Silk J., 1999, A&A, 347, 1
Viel M., Weller J., Haehnelt M. G., 2004, MNRAS, 355, L23
White R. L., Becker R. H., Fan X., Strauss M., 2003, AJ, 126, 1
Wood K., Loeb A., 2000, ApJ, 545, 86
Wyithe J. S. B., Loeb A., 2003, ApJ, 586, 693
Yoshida N., Abel T., Hernquist L., Sugiyama N., 2003a, ApJ, 592, 645
Yoshida N., Sokasian A., Hernquist L., Springel V., 2003b, ApJ, 598, 73

APPENDIX A: REVISED COOLING MODEL

In this appendix, we discuss how we incorporate cooling by molecular hydrogen and Compton cooling into our model and how we apply the effects of the filtering mass.

A1 Cooling by molecular hydrogen

We incorporate cooling due to molecular hydrogen in our model following the algorithms of Yoshida et al. (2003a). Below, we give a brief account of our calculation of the H_2 contribution to the cooling rate. The reader is referred to Yoshida et al. (2003a) and references therein for a full description.

We begin by estimating the fraction, f_{H_2}, of molecular hydrogen which would be present if there were no background of H_2-dissociating radiation from stars. As shown by Tegmark et al. (1997), at the redshifts of interest here, H_2 is formed mostly via the formation of H^- through the reaction $H + e^- \rightarrow H^-$, with rate coefficient

$$k_{H^-} = 1.83 \times 10^{-18} (T/\mathrm{K})^{0.88}\,\mathrm{cm^3\,s^{-1}} \quad \text{(A1)}$$

(Hutchins 1976). The rate of formation of H^- depends on the abundance of free electrons, which depends on recombinations $H^+ + e^- \rightarrow H + \gamma$ with rate coefficient

$$k_1 = 1.88 \times 10^{-10} (T/\mathrm{K})^{-0.64}\,\mathrm{cm^3\,s^{-1}} \quad \text{(A2)}$$

(Hutchins 1976). (We ignore photoionization and collisional ionization of hydrogen in this part of the calculation, since cooling by H_2 is only important in low-temperature haloes at high redshifts.) Tegmark et al. (1997) solved the rate equations for electron recombination and H_2 formation to obtain a characteristic H_2 abundance $f_{H_2,c}$ for gas with hydrogen number density n_H and temperature T at redshift z:

$$f_{H_2,c} = 3.5 \times 10^{-4} T_3^{1.52} \left[1 + \left(7.4 \times 10^8 / n_{H,1}\right) \times (1+z)^{2.13} \exp(-3173/[1+z])\right]^{-1}, \quad \text{(A3)}$$

where $n_{H,1}$ is the hydrogen number density in units of $\mathrm{cm^{-3}}$ and T_3 is temperature in units of 1000 K. (The above result depends on the values of various rate coefficients, and is consistent with the rate coefficients given in equations A1, A2 and A8.) The factor in square brackets corresponds to photodissociation of H^- by CMB photons, but is important only for $z \gtrsim 100$. For our own estimate of the H_2 abundance in the absence of a photodissociating background from stars, we then use

$$f_{H_2,0} = \max\left[f_{H_2,\mathrm{prim}},\, f_{H_2,c}\right], \quad \text{(A4)}$$

where $f_{H_2,\mathrm{prim}}$ is the primordial H_2 fraction, which we take to be 3.5×10^{-7} (Anninos & Norman 1996) (although our results are highly insensitive to this choice of value).

Next, we account for the effects of a background of H_2-dissociating radiation from stars. If we assume equilibrium between H_2 production via H^- and H_2 destruction by photodissociation, we obtain an H_2 abundance,

$$f_{H_2,\mathrm{eq}} = k_{H^-} x_e n_H / k_{\mathrm{diss}}. \quad \text{(A5)}$$

We calculate the electron fraction, x_e, appearing in equation (A5) from

$$x_e = \frac{x_{e,0}}{1 + x_{e,0} t_{1/2}/\tau_{\mathrm{rec}}}, \quad \text{(A6)}$$

based on equation (6) from Tegmark et al. (1997), where

$$\tau_{\mathrm{rec}} = 1/(k_1 n_H). \quad \text{(A7)}$$

This estimate of x_e assumes that recombinations have been proceeding in gas at its present density and temperature for a time $t_{1/2}$ equal to half the age of the Universe at redshift z, starting from the primordial post-recombination electron fraction $x_{e,0}$. We calculate $x_{e,0}$ using RECFAST (Seager, Sasselov & Scott 2000). Equation (A6)

accounts crudely for the age of the halo, and therefore the time available for recombinations. The electron fraction is never allowed to fall below the value corresponding to collisional ionization equilibrium at the specified temperature. Note that, post-reionization, the electron fraction should in fact tend towards the equilibrium value for photoionization equilibrium (as opposed to collisional ionization equilibrium). Our current model does not account for this fact, but this is unimportant here as the determination of electron fractions in the post-reionization Universe does not affect our determination of the epoch of reionization.

For the photodissociation rate in equation (A5), we assume

$$k_{\rm diss} = 1.1 \times 10^{-13} J_{21} F_{\rm shield}\ {\rm s}^{-1} \tag{A8}$$

(Machacek, Bryan & Abel 2001), where J_{21} is average flux in units of 10^{-21} erg cm^{-2} s^{-1} Hz^{-1} in the Lyman–Werner bands (with photon energy around 12.87 eV), and is computed from the radiation background tracked by our model of IGM evolution (Benson et al. 2002a). We account for the effects of self-shielding using the algorithm of Yoshida et al. (2003a). Specifically, we estimate the shielding factor

$$F_{\rm shield} = \min\left[1, \left(\frac{N_{\rm H_2}}{10^{14}\,{\rm cm}^{-2}}\right)^{-3/4}\right], \tag{A9}$$

which depends on the H_2 column density $N_{\rm H_2}$. In order to avoid the computational cost and complexity of solving iteratively for the radial abundance profile of H_2 within each halo, we estimate this shielding factor using the abundance $f_{\rm H_2,0}$ computed in the absence of a photodissociating background, using

$$N_{\rm H_2} = C f_{\rm H_2,0}\, \overline{n}_{\rm H,vir}\, R_{\rm vir} \tag{A10}$$

with $C = 0.2$, where $R_{\rm vir}$ is the virial radius of the dark matter halo and $\overline{n}_{\rm H,vir}$ is the mean hydrogen density within $R_{\rm vir}$. (For a more detailed justification of this approach, see Yoshida et al. 2003a.)

We then calculate the final H_2 abundance used in our cooling calculation from

$$f_{\rm H_2} = \min\left[f_{\rm H_2,0}, f_{\rm H_2,eq}\right] \exp\left(-\frac{T}{51\,920\,{\rm K}}\right). \tag{A11}$$

We have introduced an exponential cut-off in the molecular hydrogen fraction at temperatures corresponding to the dissociation energy of H_2, to account for collisional dissociation of H_2. (Tegmark et al. 1997 ignored this because they were interested only in the relatively low temperature regime. The exact form and position of the cut-off are unimportant, as the cooling function is dominated by atomic processes at the temperatures for which collisional dissociation is important.)

Finally, we compute the cooling time-scale due to molecular hydrogen using[13] (Tegmark et al. 1997)

$$\tau_{\rm H_2} = 48\,200\,{\rm yr}\left(1 + \frac{10 T_3^{7/2}}{60 + T_3^4}\right)^{-1} \times \exp(512\,{\rm K}/T)(f_{\rm H_2} n_{\rm H,1})^{-1}. \tag{A12}$$

The net radiative cooling time-scale for gas, $\tau_{\rm cool}$, is then given by

$$\tau_{\rm cool}^{-1} = \tau_{\rm H_2}^{-1} + \tau_{\rm other}^{-1}, \tag{A13}$$

where $\tau_{\rm other}$ is the cooling time-scale due to all other cooling mechanisms (e.g. atomic processes, Compton cooling).

[13] Using the more recent estimate of the H_2 cooling function from Galli & Palla (1998) makes only small differences to our results.

Yoshida et al. (2003a) recommend the inclusion of a 'dynamical heating' term which accounts for the accretion of new material on to a halo. This accretion is assumed to heat the gas already in the halo (as the kinetic energy of the infalling material is converted into thermal energy), reducing the rate at which it can cool. Yoshida et al. (2003a) find that the inclusion of this term in their analytic model of gas cooling in haloes produces good agreement with the results of their gas-dynamical calculations. We have experimented with including such a term in our calculations. We note, however, that our semi-analytic model already effectively includes such dynamical heating. Whenever a halo's mass doubles we consider the halo to have reformed (Cole et al. 2000). The temperature of the gas in the halo is then recomputed, and the cooling clock is reset to zero. It is simple to show that this results in an effective heating rate for the gas which is comparable to that supplied by the dynamical heating term of Yoshida et al. (2003a). For example, following the approach of Yoshida et al. (2003a) and taking the virial temperature of a halo to be given by

$$T = \gamma M^{2/3}(1+z), \tag{A14}$$

where γ is a normalization constant, we find by differentiation that

$$\dot{T} = \frac{2}{3}\gamma \frac{(1+z)}{M^{1/3}}\left(\dot{M} + \frac{3}{2} M \frac{\dot{z}}{1+z}\right). \tag{A15}$$

(Note that Yoshida et al. 2003a include only the first of the two terms on the right-hand side of this equation. However, we expect them to be of similar magnitude.) In the semi-analytic calculation, the virial temperature is also given by equation (A14) and is updated after each halo reformation event. Consequently, when averaged over a halo lifetime, the amount of dynamical heating in the semi-analytic model must approximately equal that proposed by Yoshida et al. (2003a).

Not surprisingly, therefore, we find that the explicit inclusion or otherwise of a dynamical heating term (of the form proposed by Yoshida et al. 2003a) makes no significant difference to the results of our calculations.

While in primordial plasma H_2 must form in the gas phase via the H^- channel described above, in enriched gas there is the possibility of H_2 formation on the surfaces of dust grains. While we do not include this formation mechanism in our cooling model we have checked that it makes no significant difference to the evolution of the filling factor. To do this, we used the methodology of Cazaux & Spaans (2004) to compute the ratio of H_2 formation rates on dust grains and in the gas phase. We then scaled the abundance of H_2 in our cooling model appropriately and computed a selection of our models using this revised cooling model. We found no significant change in the evolution of the H_2 filling factors.

A2 Compton cooling

The Compton cooling time-scale does not explicitly depend on the gas density, but does depend on the electron abundance, and therefore on the ionization state of the gas. The electron abundance, in turn, *does* depend on the gas density, due to the density dependence of the electron recombination time-scale.

Prior to halo formation, the electron abundance in the IGM gas is computed from our model of the ionization state of the IGM. Prior to IGM reionization, the electron abundance is low, corresponding to that left over from recombination. Post-reionization, the electron fraction becomes large as the Universe is almost entirely ionized. During and after halo formation, the electron fraction will move towards the appropriate equilibrium value as

described in Section A1. The Compton cooling time-scale is then computed using the recombined electron abundance from equation (A6), or the equilibrium electron abundance, whichever is the larger.

A3 Filtering mass

The filtering mass is similar to the Jeans mass, but is the relevant quantity for describing the effect of pressure on the gravitational collapse of gas in an expanding Universe with a time-varying temperature. It is defined by

$$M_{\rm F} = (4\pi/3)\rho_0(2\pi a/k_{\rm F})^3, \quad \text{(A16)}$$

where ρ_0 is the mean density of the Universe, a is the expansion factor and

$$\frac{1}{k_{\rm F}^2(t)} = \frac{1}{D(t)}\int_0^t {\rm d}t' a^2(t')\frac{\ddot{D}(t')+2H(t')\dot{D}(t')}{k_{\rm J}^2(t')}\int_{t'}^t \frac{{\rm d}t''}{a^2(t'')}, \quad \text{(A17)}$$

where $D(t)$ is the linear theory growth factor, $H(t)$ is the Hubble variable and $K_{\rm J}(t)$ is the Jeans wavenumber, defined in the usual way. The collapse of gas into haloes with masses below $M_{\rm F}$ is greatly suppressed relative to the case of zero pressure.

APPENDIX B: ESCAPE FRACTIONS

Here, we apply simple, physical models to predict the escape fraction for each galaxy in our model. We follow Paper I and use the model of Dove & Shull (1994) (and the development thereof carried out in Paper I) to compute the amount of ionizing radiation absorbed by the gas in each galaxy. The resulting escape fractions are shown in Table B1 and in Fig. B1.

The mean escape fraction for H I ionizing photons for galaxies due to absorption by neutral gas are found to be a few percent. These numbers agree reasonably well with the limited observational determinations of escape fractions, e.g. Leitherer et al. (1995) find that less than 3 per cent of ionizing photons escape from low-redshift starburst galaxies, and similar results have been obtained by Tumlinson et al. (1999), while Steidel, Pettini & Adelberger (2001) estimate escape fractions $\sim$10 per cent for Lyman-break galaxies at $z = 3.4$. More recently, Bergvall et al. (2006) found an escape fraction of 4–10 per cent for a local starbursting dwarf galaxy.

It should be noted that these escape fraction calculations are extremely simplified – they model the ISM as a smooth distribution of gas (a clumpy distribution could enhance the escape fraction as noted by Wood & Loeb 2000) and ignore dynamical effects (such as expanding superbubbles as considered by Dove, Shull & Ferrara 2000 which significantly lower the escaping fraction). More detailed calculations, employing hydrodynamical simulations, suggest that high-redshift dwarf galaxies might have much higher escape fractions (Fujita et al. 2003). A better theoretical understanding of escape fractions is a crucial missing component of accurate calculations of reionization.

We estimate the effects of dust extinction on the ionizing radiation escape fractions by following the methods of Paper I. Briefly, we extrapolate the Milky Way extinction curve to the wavelengths corresponding to the ionization threshold of each species, H I, He I and He II, and compute an extinction averaged over galaxy inclinations using the calculations of Ferrara et al. (1999). The results of including dust extinction on escape fractions are shown in Table B1 and in Fig. B1. It is seen that dust can cause severe (frequently an order of magnitude) attenuation in the escaping fraction, at least in this highly simplified calculation of a smooth ISM. Clearly, a more detailed analysis of the effects of dust on the escaping fraction is urgently needed if we are to gain a better understanding of how the IGM was reionized.

The effects on the ionized filling factor of using different assumptions for the escape fraction are shown in Fig. B2, for different models and different redshifts. It is interesting to note that, in the *WMAP* cosmologies, only the TopHeavy model in the *WMAP* PL cosmology is able to achieve reionization by $z = 6$, and then only if the effects of dust on the escape fraction are ignored. In the Baugh et al. (2005), the majority of ionizing photons are produced during bursts of star formation. As such, dust (assumed to be located in galaxy discs in our simple model) has little effect on the escape fraction. It should be noted that our model for the escape fraction is likely to be very unreliable in cases where most of the ionizing photons are produced in bursts rather than in quiescent discs.

Table B1. Ionizing luminosity-weighted escape fractions for galaxies at two different redshifts ($z = 5$ and 11) are given for all cosmologies and models. Redshift is specified in the first column, with cosmology and model specified in columns 2 and 3. Column 4 specifies the escape fractions for H I ionizing radiation when dust is ignored, while column 5 specifies the same escape fractions when dust extinction is included.

z	Cosmology	Model	$f_{\rm esc}$	
			No dust	With dust
11	*WMAP* RSI	Standard	0.06	0.019
5	*WMAP* RSI	Standard	0.03	0.005
11	*WMAP* RSI	MetFree	0.05	0.012
5	*WMAP* RSI	MetFree	0.03	0.005
11	*WMAP* RSI	TopHeavy	0.05	0.006
5	*WMAP* RSI	TopHeavy	0.02	0.002
11	*WMAP* RSI	WeakFB	0.02	0.002
5	*WMAP* RSI	WeakFB	0.02	0.002
11	*WMAP* RSI	H_2	0.06	0.016
5	*WMAP* RSI	H_2	0.05	0.011
11	*WMAP* RSI	H_2WeakFB	0.06	0.007
5	*WMAP* RSI	H_2WeakFB	0.02	0.002
11	*WMAP* RSI	H_2WeakFB-VMS	0.04	0.001
5	*WMAP* RSI	H_2WeakFB-VMS	0.03	0.004
11	*WMAP* PL	Standard	0.03	0.005
5	*WMAP* PL	Standard	0.03	0.004
11	*WMAP* PL	MetFree	0.03	0.004
5	*WMAP* PL	MetFree	0.04	0.008
11	*WMAP* PL	TopHeavy	0.03	0.001
5	*WMAP* PL	TopHeavy	0.02	0.001
11	*WMAP* PL	WeakFB	0.02	0.001
5	*WMAP* PL	WeakFB	0.03	0.001
11	*WMAP* PL	H_2	0.05	0.014
5	*WMAP* PL	H_2	0.07	0.022
11	*WMAP* PL	H_2WeakFB	0.02	0.001
5	*WMAP* PL	H_2WeakFB	0.03	0.002
11	*WMAP* PL	H_2WeakFB-VMS	0.02	0.001
5	*WMAP* PL	H_2WeakFB-VMS	0.03	0.002
11	Baugh et al. (2005)	Standard	0.05	0.036
5	Baugh et al. (2005)	Standard	0.08	0.076
11	Baugh et al. (2005)	H_2	0.08	0.066
5	Baugh et al. (2005)	H_2	0.13	0.130
11	Baugh et al. (2005)	H_2WeakFB	0.06	0.062
5	Baugh et al. (2005)	H_2WeakFB	0.12	0.117
11	Baugh et al. (2005)	H_2WeakFB-VMS	0.07	0.069
5	Baugh et al. (2005)	H_2WeakFB-VMS	0.12	0.120

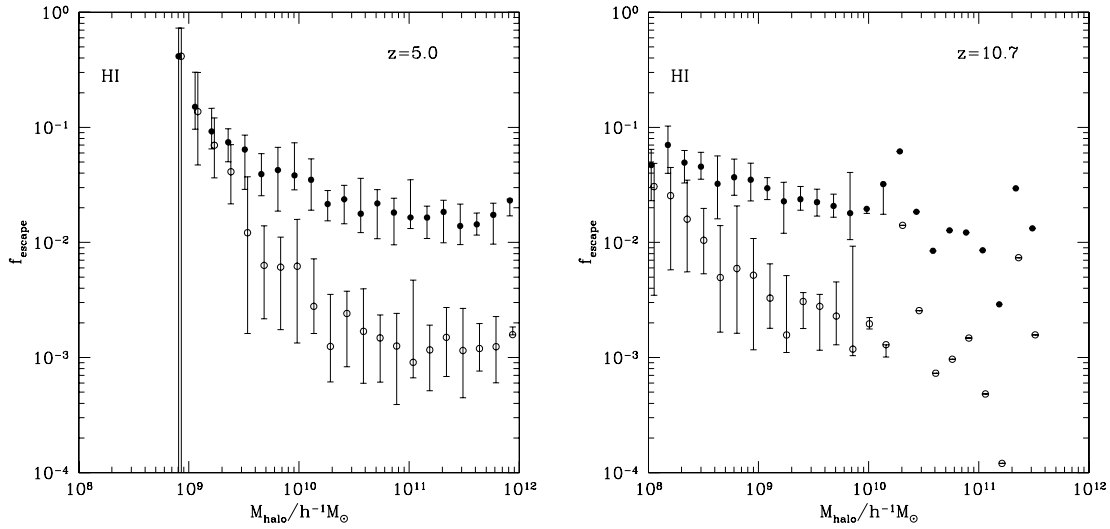

Figure B1. Ionizing luminosity-weighted mean H I escape fractions for haloes as a function of halo mass at $z = 5$ (left-hand panel) and $z = 10.6$ (right-hand panel) for the Standard model in the *WMAP* RSI cosmology. The points show the median escape fraction for all haloes of this mass, while the error bars show the upper and lower quartiles of the distribution. Filled points include only the effects of absorption by neutral gas, while open points also include the effects of dust extinction.

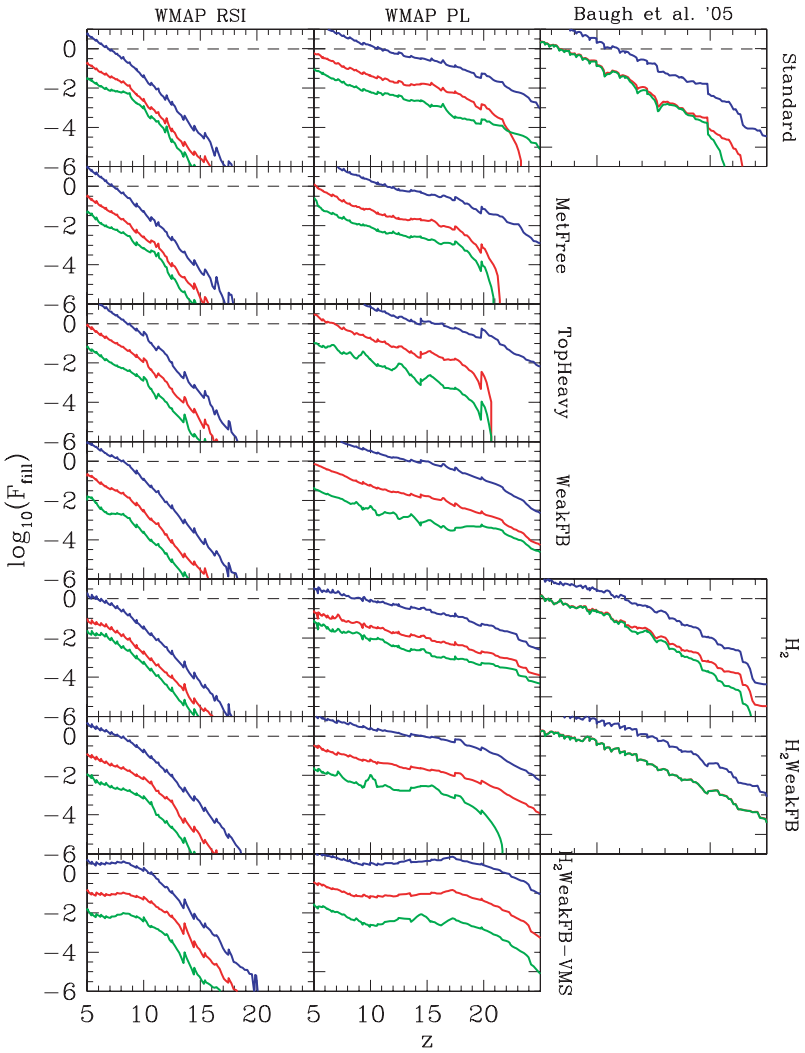

Figure B2. The H II filling factor for all cosmologies and models for a fixed escape fraction of $f_{esc} = 1$ (blue lines) and f_{esc} computed using the model described in Appendix B both for absorption by neutral gas only (red lines), and for absorption by gas and dust (green lines).

APPENDIX C: TESTS OF THE MODEL

In this appendix, we present results of various tests of our model. We begin by comparing to the results of Paper I after which we check that our methodology leads to model calculations that are suitably converged.

C1 Comparison with Paper I results

We begin by comparing our current results with those of Paper I. Our galaxy formation model, GALFORM, has undergone significant development over that used in Paper I. As such, we expect some differences in the results obtained. To assess the degree of change in our results we have performed the same calculations as in Paper I. Specifically, we adopted the same model parameters as in that paper and computed the H II filling factor as a function of redshift. The models of Paper I use cosmological parameters $(\Omega_0, \Lambda_0) = 0.3, 0.7$ and $\sigma_8 = 0.9$, while using similar galaxy formation prescriptions and parameters as used in this work. The Paper I models are similar to our Standard model, with major differences being a weaker feedback, $V_{\rm hot} = 150$ km s^{-1}, as opposed to 200 km s^{-1} as used here, and a simpler model for galaxy merging than used here (Paper I uses the model of Cole et al. 2000 whereas the current work uses that of Benson et al. 2004 which gives somewhat longer merger timescales). Furthermore, in Paper I a fraction of all star formation was assumed to result in the formation of brown dwarfs, such that the effective ionizing luminosity per unit mass of stars formed was reduced by a factor of 0.65.

Calculations were performed for both the low and high baryon fraction models used in Paper I ($\Omega_{\rm b} = 0.02$ and 0.04, respectively).

As can be seen in Fig. C1, differences do exist between the calculations of this work and those of Paper I. The revised modelling tends to produce slightly earlier reionization (by $\Delta z \approx 1$). These changes are due to several improvements and corrections made to the semi-analytic model and our calculation of ionized sphere growth made since Paper I.

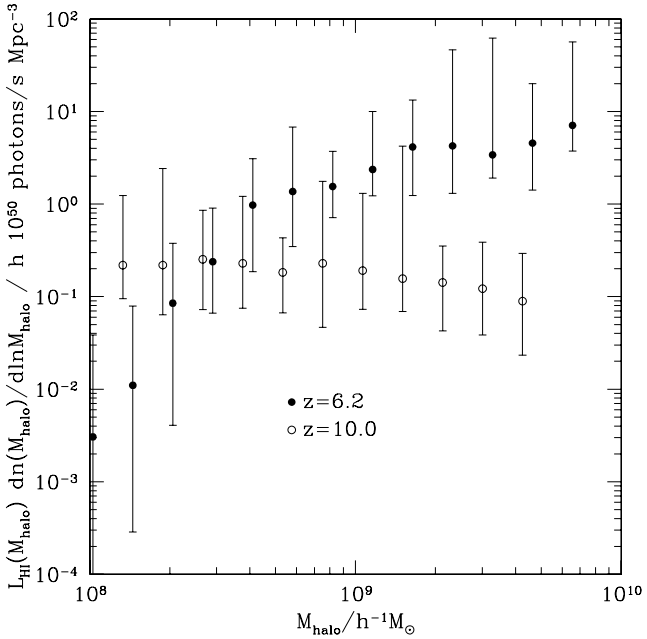

Figure C2. The H I ionizing luminosity of haloes of given mass weighted by the halo mass function, d n/d ln $M_{\rm halo}$. The points show the median value at fixed mass while the error bars enclose 80 per cent of the distribution. Results are shown for model H_2 in the *WMAP* RSI cosmology. Filled points correspond to $z = 6.2$ while open points correspond to $z = 10$. The highest halo masses shown correspond to the most massive haloes included in our galaxy formation calculations at the specified redshift.

C2 Tests of model convergence

We have examined two possible sources of numerical error in our computational method. We outline both below.

C2.1 Convergence of ionizing emissivity

As noted in Section 2, we use an iterative procedure to produce a self-consistent model of galaxy formation and IGM evolution.

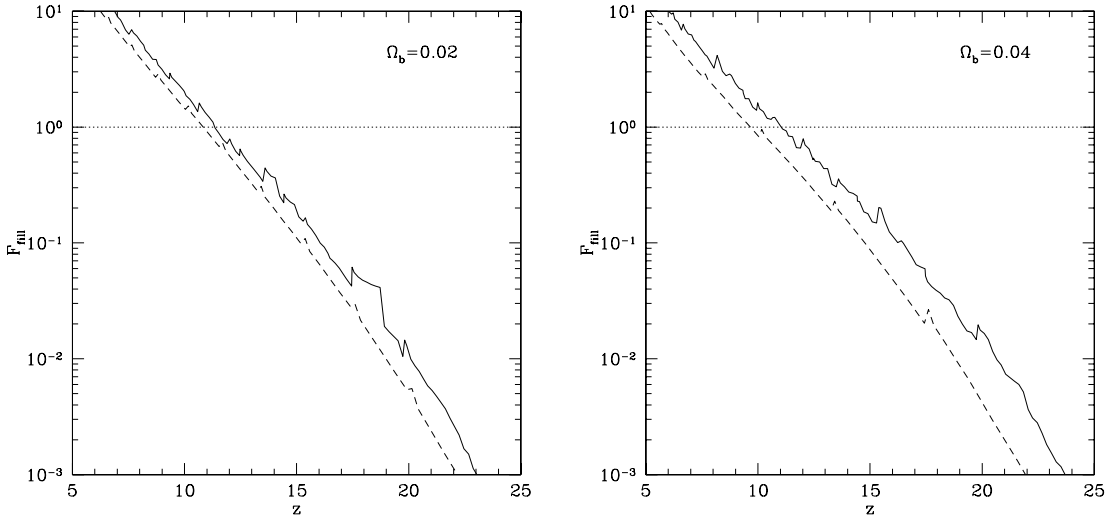

Figure C1. Comparison of H II filling factors using the parameters of Paper I. Dashed lines show the results from Paper I (for an escape fraction of 100 per cent and no dust absorption). Solid black lines show the results from this work using ionized spheres. Left panel: model with $\Omega_{\rm b} = 0.02$. Right panel: model with $\Omega_{\rm b} = 0.04$.

© 2006 The Authors. Journal compilation © 2006 RAS, MNRAS **369,** 1055–1080

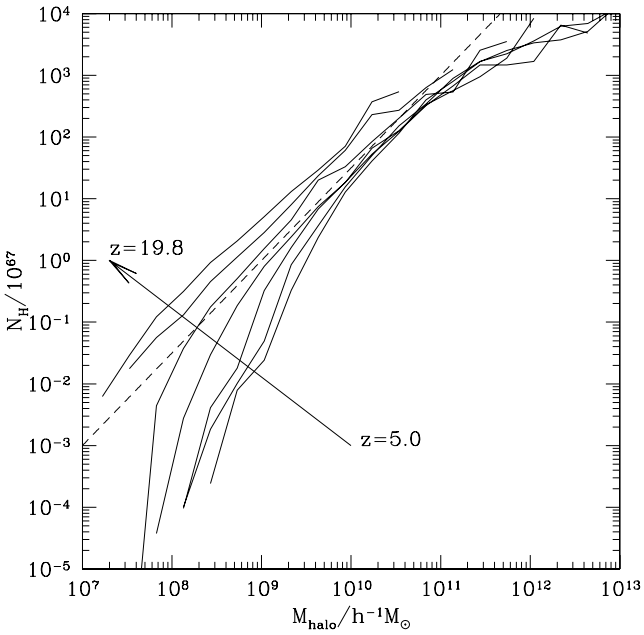

Figure C3. The median number of hydrogen atoms ionized by sources in a dark matter halo as a function of halo mass. Results are shown for the Standard model in the *WMAP* PL cosmology. Solid lines indicate the median number for redshifts 5.0, 6.2, 8.2, 10.7, 12.5, 15.4 and 19.8 (increasing from right to left as indicated by the arrow). The dashed line shows the relation $N_{\rm H} \propto M^{1.5}$ for comparison.

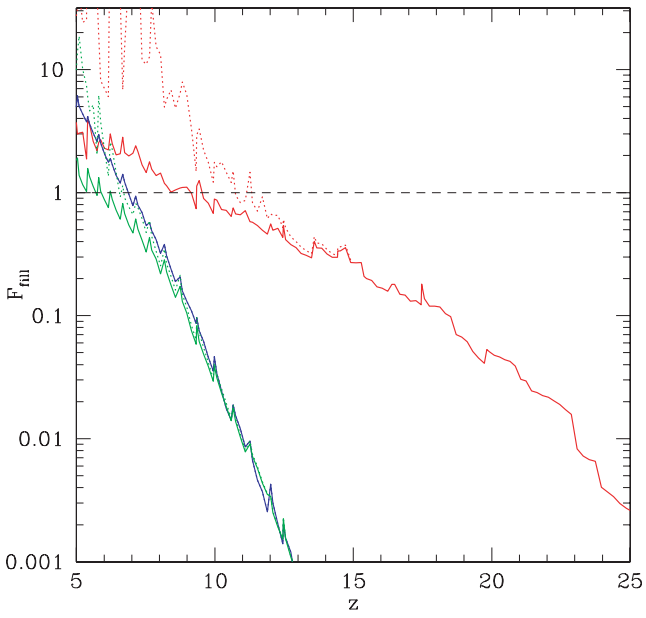

Figure C4. The H II filling factor as a function of redshift for the 'H_2' model in the *WMAP* PL cosmology is shown by red lines. The solid line shows the result direct from our model calculation while the dotted line shows the effects of correcting for the lack of the highest mass haloes in our calculation. The horizontal dashed line indicates a filling factor of unity. Note that this model has the largest fractional correction to the H II filling factor out of those listed in Table 2. For comparison, we show results from the Standard model in the *WMAP* RSI cosmology (blue lines) which is least affected by the lack of high-mass haloes and from the H_2 model in the *WMAP* RSI cosmology (green lines) in which the effect of missing the highest mass haloes is more typical.

To be self-consistent, the ionizing emissivity as a function of redshift produced by the galaxy formation model should be identical to that used to evolve the IGM in the previous iteration. We have checked, for all models used in this work, that such agreement has in fact been met to within the accuracy of the model calculation (which is limited by the number of merger trees we are able to simulate).

C2.2 Upper limit to halo masses

With finite computing resources, we can only simulate haloes with masses up to some maximum value, if we always resolve the formation histories of haloes down to the same minimum progenitor mass. Our approach is to ensure that the resolution of the merger trees used in the semi-analytic model is sufficient to resolve all haloes in which gas could cool (i.e. haloes with virial velocities above 13 km s^{-1} for models with no H_2 cooling or 1.3 km s^{-1} for models with H_2 cooling). In our standard runs, the most massive haloes we consider are a factor of 1.3×10^5 more massive than the lowest mass haloes. The contribution to the ionizing emissivity from galaxies in more massive haloes is therefore lost from our calculation. Fig. C2 shows the median H I ionizing luminosity from haloes as a function of mass weighted by the halo mass function, $\mathrm{d}n/\mathrm{d}\ln M_{\rm halo}$. This, therefore, gives an indication of the relative contribution to the total ionizing luminosity from haloes of different masses. At high redshifts, the contribution to the total luminosity declines with increasing halo mass (as the halo mass function is steeply declining), so we may expect the contribution from higher mass haloes to be negligible. At lower redshifts the contribution is an increasing function of halo mass. However, at the highest masses shown the relation is quite flat and is expected to turn over at still higher masses once the exponential break in the mass function is reached.

We can estimate the effects of the non-inclusion of higher mass haloes as follows. We find that, for all models, the dependence of the mean ionizing luminosity per halo on halo mass can be well represented by a power law.[14] We can use this relation, combined with the known mass function of dark matter haloes to determine the filling factor we would have obtained had we been able to model haloes of arbitrarily large mass. We have checked the accuracy of this correction by repeating some of our calculations including haloes up to a factor of 10 more massive than in our standard calculations.[15] We find that our correction does indeed overestimate the filling factor (as described above), but is typically accurate to better than 10 per cent. The correction is smaller for higher redshifts (due to the ever more rapidly declining halo mass function found at those redshifts), and is therefore less important prior to reionization (which is the period of interest in this work).

The relation between halo mass and ionizing luminosity which arises from our semi-analytic modelling of galaxy formation is non-trivial as can be seen in Fig. C3. While it is well represented by a power law for the highest mass haloes which we consider ($N_{\rm H} \propto M_{\rm halo}^{0.5-0.7}$ for haloes in the mass range 10^{12}–$10^{13}\,h^{-1}\,{\rm M}_\odot$) it steepens

[14] For even larger masses this relation flattens (see, for example, fig. 8 of Benson et al. (2000) which shows that, for high mass haloes, galaxy formation becomes gradually less efficient as halo mass increases), but this will merely make our estimate of the effects of ignoring high-mass haloes a conservative one.

[15] While this is possible, the computational time involved makes it impractical to do this for all of our models.

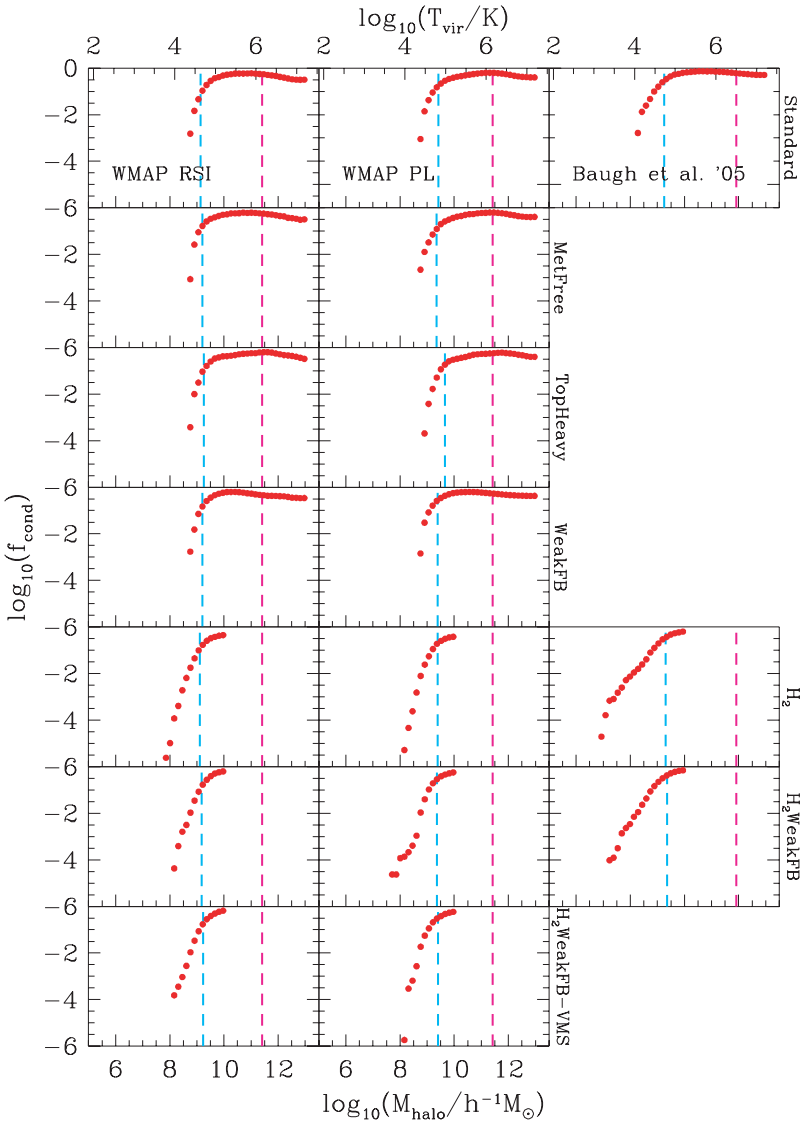

Figure D1. The condensed fraction, $f_{\rm cond}$, defined as the mass of material in a halo which has condensed into the galaxy phase, divided by $\Omega_{\rm b}M_{\rm halo}/\Omega_0$ is shown by red circles. Results are shown for $z = 5$ and for all three cosmological models (one per column as indicated by the labels in the upper row of panels), and for all models computed (one model per row as indicated at the right-hand edge of each row). The upper x-axis indicates the virial temperature corresponding to halo mass. Vertical dashed magenta lines indicate the virial temperature corresponding to the feedback parameter $V_{\rm hot}$ in each model, while vertical cyan lines indicate the filtering mass in each model.

for lower mass haloes, reaching $N_{\rm H} \propto M_{\rm halo}^{1.5-3.0}$ depending on redshift. The form of this relation arises from the various physical ingredients (e.g. cooling times, supernovae feedback etc) which go into our calculations. Other authors (e.g. Haiman & Holder 2003; Wyithe & Loeb 2003; Furlanetto, Zaldarriaga & Hernquist 2004) have used highly simplified assumptions (such as that the number of hydrogen atoms ionized is simply proportional to the halo mass; Furlanetto et al. 2004) which may be expected to give significantly different results.

We find that the loss of the highest mass haloes affects our determination of the epoch of reionization significantly only for models including H_2 cooling in the *WMAP* PL cosmology. Fig. C4 illustrates the effects of missing the highest mass haloes in such a model (red lines). For models without H_2 cooling our calculations span a sufficient range of halo masses to capture nearly all of the ionizing emissivity, while for the *WMAP* RSI cosmology the space density of high-mass haloes is so small that even models including H_2 cooling probe a sufficient range of halo masses. The typical correction to the H II filling factor in our models is only 3–5 per cent for $z > 10$ in models including H_2 cooling. For $z \approx 6$ the correction reaches a factor of 2–3 for models including H_2 cooling. In models which ignore H_2 cooling the correction is less than 1 per cent for $z > 6$. Fig. C4 also shows results for the model which is minimally affected by the lack of high-mass haloes (Standard model in the *WMAP* RSI cosmology; blue lines) and for a model in which the effect of missing high-mass haloes is more typical (model 'H_2' in the *WMAP* RSI cosmology; green lines).

When quoting results for the epoch of reionization we list the value computed directly from our calculations along with the value

© 2006 The Authors. Journal compilation © 2006 RAS, MNRAS **369,** 1055–1080

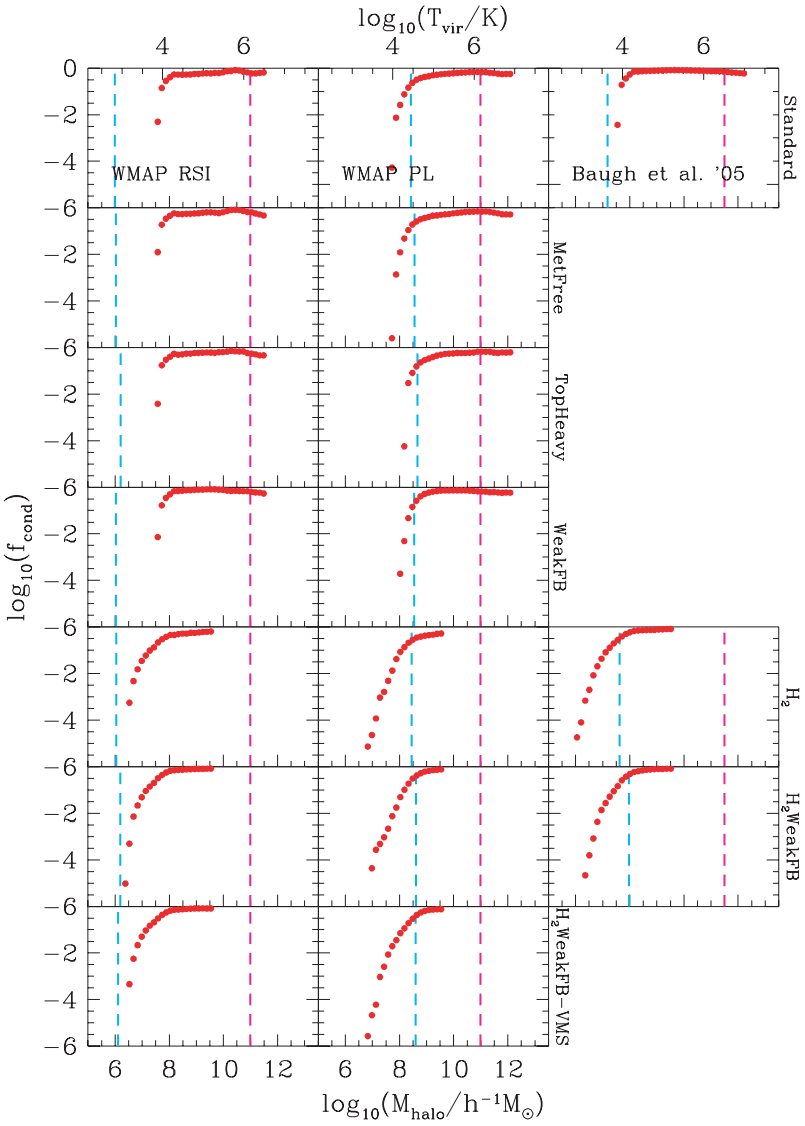

Figure D2. As Fig. D1 but for $z = 10$.

obtained after the above correction has been applied. The true epoch of reionization should then lie between these two values.

APPENDIX D: CONDENSED FRACTIONS

Figs D1 and D2 show condensed fraction (i.e. the mass of material condensed into the galaxy phase divided by $\Omega_b M_{\rm halo}/\Omega_0$) as a function of halo mass $M_{\rm halo}$ for redshifts $z = 5$ and 10, respectively. Columns indicate the cosmology (see labels in upper row) while rows indicate the model (see labels at right-hand edge of each row). In each panel, the vertical dashed magenta line indicates the virial temperature corresponding to the feedback parameter $V_{\rm hot}$ while the vertical dashed cyan line indicates the filtering mass. Red points indicate the condensed fraction (note that the vertical scale is logarithmic).

The effects of including H_2 cooling can be clearly seen as the presence of a tail of low $f_{\rm cond}$ for low-mass haloes. The effects of weak feedback are also apparent, resulting in a small increase in $f_{\rm cond}$ in low-mass haloes.

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