The first generation of star-forming haloes

Darren S. Reed,1⋆ Richard Bower,1 Carlos S. Frenk,1 Liang Gao,2 Adrian Jenkins,1 Tom Theuns1,3 and Simon D. M. White2

1Department of Physics, Institute for Computational Cosmology, University of Durham, South Road, Durham DH1 3LE
2Max Planck Institut für Astrophysik, Karl-Schwarzschild Strasse 1, Garching, Munich D-85740, Germany
3Department of Physics, University of Antwerp, Campus Drie Eiken, Universiteitsplein 1, Antwerp, Belgium

Accepted 2005 July 6. Received 2005 July 4; in original form 2005 April 5

ABSTRACT

We model gas cooling in high-resolution N-body simulations in order to investigate the formation of the first generation of stars. We follow a region of a Lambda cold dark matter (ΛCDM) universe especially selected to contain a rich cluster by the present day. The properties of the dark haloes that form in these subsolar mass-resolution simulations are presented in a companion paper by Gao et al. The first gas clouds able to cool by molecular hydrogen (H2)-line emission collapse at extremely high redshift, z ≈ 47, when the mass of the dark halo is 2.4 × 10^5 h^-1 M⊙. By z ≈ 30, a substantial population of haloes are capable of undergoing molecular hydrogen cooling although their ability to form stars is dependent on the efficiency of feedback processes such as dissociating Lyman–Werner radiation. The mass of the main halo grows extremely rapidly and, by z ≈ 36, its virial temperature has reached 10^4 K, at which point gas cooling becomes dominated by more effective atomic processes. By z ≈ 30, a small ‘group’ of such potential galaxies will have formed unless prevented from doing so by feedback processes. By this redshift, massive (>100 M⊙) Population III stars are able to ionize gas well beyond their own host halo and neighbouring H II regions can percolate to form an ionized superbubble. Such patches would be too widely separated to contribute significantly to reionization at this time. The large number density of early cooling haloes in the pre-reionized universe raises the exciting prospect that this ultra-early generation of stars may be observable as gamma-ray bursts or supernovae.

Key words: methods: N-body simulations – galaxies: formation – galaxies: haloes – cosmology: theory – dark matter.

1 INTRODUCTION

The first stars in the Universe are believed to have formed from primordial metal-free gas in haloes with virial temperatures less than the ~10^4-K threshold for atomic hydrogen-line cooling [see reviews by Bromm & Larson (2004), Ciardi & Ferrara (2005), and references therein]. Primordial gas in these small haloes is instead able to cool by molecular hydrogen (H2) transitions, a less-effective process (Saslaw & Zipoy 1967; Peebles & Dicke 1968). H2 is produced in a two-step reaction chain in which free electrons left over from recombination serve as a catalyst via H− production (when z ≲ 100). Metal-free (Population III) stars are the first potential producers of ultraviolet (UV) photons that can contribute to the reionization process, and are the first producers of the metals required for the formation of Population II stars.

Tegmark et al. (1997) developed analytic methods to model early baryonic collapse via H2 cooling. The early stages of primordial star formation were later directly modelled in high-resolution hydrodynamic numerical simulations by Bromm, Coppi & Larson (1999, 2002) and Abel, Bryan & Norman (2002). Yoshida et al. (2003b) further utilized simulations to develop a semi-analytic model based on the Tegmark et al. methods and included the effects of dynamical heating caused by the thermalization of kinetic energy of infall into a deepening halo potential. This heating effect varies from halo to halo, depending on mass-accretion rates, and has the potential to delay baryonic collapse.

Determining the first epoch of baryonic cooling in high-redshift haloes thus requires the use of numerical simulations to follow their growth. In this paper, we study the growth of structure in a region selected to be the site where a rich cluster forms by z = 0 in a large cosmological simulation. We repeatedly ‘resimulate’ this region at increasing resolution and redshift. In this way, we are able to model the site of an extremely early generation of star formation. The simulations employed here are those described in a companion paper (Gao et al. 2005; hereafter G05).

We estimate baryonic cooling rates using similar methods to those described by Yoshida et al. (2003b), who performed a detailed hydrodynamic simulation of a 1-h^-1 Mpc volume in order to
determine under what conditions gas in low-mass haloes was able to cool sufficiently to allow collapse. Our use of a base region almost 500 h⁻¹ Mpc on a side allows us to include the important effects of large-scale modes (White & Springel 2000; Barkana & Loeb 2004), and enables us to model a much earlier generation of baryonic cooling in rare haloes with low comoving number density. In addition, we model H₂ cooling by explicitly integrating the time-dependent H₂ production rate as described in Section 3.

Population III stars are generally expected to be massive (~100 M☉) due to the large Jeans mass (~1000 M☉) during initial baryonic collapse, and to the inability of zero-metallicity gas clouds to subfragment into small masses during collapse (e.g. Bromm et al. 1999, 2002; Nakamura & Umemura 2001; Abel et al. 2002; Omukai & Palla 2003; Bromm & Loeb 2004). We will assume that baryonic cooling leads to a single massive star per halo in order to make some predictions about their effects on their surroundings and their potential observability. By determining the mean separation and clustering properties of the first star-forming haloes, we can infer their contribution to reionization.

If the haloes hosting Population III stars are highly clustered, their associated H II regions could overlap, creating percolating bubbles of reionization and, later, metal enrichment. Alternatively, if they are weakly clustered, then reionization from the first stars would be confined to isolated H II regions. H II regions from Population III stars have been studied in the redshift range z = 10–30 using 1D codes by, for example, Kitayama et al. (2004) and Whalen, Abel & Normal (2004), who find that most UV photons produced by a massive (~100 M☉) star are able to escape the virial radius of small haloes such as we have simulated here. However, the extent to which high-redshift ionized regions may overlap with those associated with neighbouring haloes is not yet established. If Lyman–Werner (L–W) radiation or metal enrichment from the first generation of stars is effective at suppressing the formation of further Population III stars (e.g. Omukai & Nishi 1999; Glover & Brand 2001), overlap will not occur and the original H II region will simply collapse once the stars that power it reach the end of their lifetime.

As a first step in distinguishing between these scenarios, our simulations focus on the largest progenitors at high redshifts of a present day cluster region. The structure of this paper is as follows: the numerical techniques and assumptions of these simulations are discussed in Section 2. In Section 3, we describe the criteria required for baryonic collapse. In Section 4, we show that the first haloes to collapse do so at extremely high redshift, and show that after a small redshift interval, many other nearby haloes also satisfy the collapse criterion. In Section 5, we consider the implications of these results; we discuss potential sources of feedback, and we consider whether this early generation of star formation is open to observational study. Our conclusions are summarized in Section 6. Throughout, we assume a flat ΛCDM model with the following cosmological parameters, which are consistent with the combined Wilkinson Microwave Anisotropy Probe (WMAP)/2-degree Field Galaxy Redshift Survey (2dFGRS) results (Spergel et al. 2003): matter density, Ω_m = 0.3; dark energy density, Ω_Λ = 0.7; baryon density, Ω_baryon = 0.04; fluctuation amplitude, σ_8 = 0.9; Hubble constant H_0 = 0.7 (in units of 100 km s⁻¹ Mpc⁻¹); and no tilt (i.e. a primordial spectral index of 1).

## 2 NUMERICAL TECHNIQUES

### 2.1 The simulations

We use the parallel gravity solvers Gadget (Springel, Yoshida & White 2001) and Gadget-2 (Springel 2005) to resimulate the region surrounding a rich cluster identified in a cosmological simulation of a 479-h⁻¹ Mpc cube [the WLS run (very large simulation) of Jenkins et al. (2001), Yoshida et al. (2001)]. This cluster and its environment were part of the sample resimulated and analysed by Navarro et al. (2004) and Gao et al. (2004a,b,c). We identified the largest dark matter halo at z = 5 in this resimulation, and then resimulated it again with resolution improved by an additional factor of 200. We repeated this procedure three more times stepping progressively back in redshift at increasing resolution.

Table 1 lists all our simulations, giving their initial and final redshifts, the mass of an individual particle in the highest-resolution region, the mass of the main halo at the final time (the original cluster in the case of the WLS simulation) and the gravitational softening employed. The object that we picked at z = 5 did not end up as part of the original cluster selected in the WLS simulation, but rather as a nearby cluster of smaller mass. This accounts for the difference in masses quoted for the final ‘main halo’ in the R1 and VLS simulations. The highest-resolution resimulation ends at redshift of 49, has a particle mass of 0.55 h⁻¹ M☉ and includes a high-resolution subvolume of side 90 h⁻¹ kpc. We continue to follow the same halo in a resimulation ending at redshift of 29 with particle mass of 29 h⁻¹ M☉. Initial conditions were created using the CMBFAST transfer function (Seljak & Zaldarriaga 1996), extrapolated to small scales by imposing $P(k) \propto k^{-3}$.

### 2.2 Halo properties

Our virialized haloes are identified using the spherical overdensity (SO) algorithm (Lacey & Cole 1994), assuming the spherical tophat model (Eke, Cole & Frenk 1996) in which the ΛCDM virial overdensity, $\Delta_{vir}$, in units of the mean density is 178 at high redshifts when $\Omega_m \approx 1$. For a complex mass distribution such as that expected to arise from the flat fluctuation spectrum of our high-redshift simulations, so has an advantage over the simpler friends-of-friends halo identification algorithm (Davis et al. 1985) in that it is less likely spuriously to link together neighbouring haloes or to misclassify highly ellipsoidal but unvirialized structures as haloes.

The temperature of gas in virial equilibrium within a dark matter halo is given by

$$T_{vir} = \frac{\mu M_{vir}}{1.2} \left(\frac{\Omega_m}{3/2}\right)^{2/3} \left(\frac{\Omega(z)}{\Omega(z)_{LA}}\right)^{1/3} \left(\frac{1+z}{50}\right) K,$$

(1)

for example, Eke et al. (1996), where $\mu$ is the mean molecular weight (in units of proton mass) and $M_{vir}$ is the mass within the virial radius, $r_{vir}$. Simulations by Machacek, Bryan & Abel (2001) and Yoshida et al. (2003b) confirm that the local gas temperature is relatively...
close to the virial temperature of high-redshift haloes, even within the central density peaks.

Throughout this paper, processes occurring within the haloes are calculated assuming uniform density and temperature within $r_{200}$ because they depend in complex ways on mixing processes, which cannot be treated realistically without simulation. This includes the assumption that gas is shock-heated upon initial infall. We neglect the possibility that some infalling H$_2$ may be destroyed during shock-heating, but note that H$_2$ production is rapid enough that it would be quickly replenished to a level that allows baryonic cooling. Moreover, collisional H$_2$ dissociation rates are small below $10^4$ K, so any shock-heating strong enough to cause dissociation would be accompanied by ionization, providing free electrons that would increase the eventual H$_2$ fraction. Detailed hydrodynamic simulations are required to model the heating of infalling gas accurately. We assume that the baryonic mass fraction in the halo is equal to the universal mean value. Lower baryon fractions are possible due to the suppression of linear baryon fluctuations on small scales before the decoupling epoch (e.g. Yamamoto, Sugiyama & Sato 1998; Singh & Ma 2002; Yoshida, Sugiyama & Hernquist 2003a). However, even for haloes somewhat below the linear-theory Jeans mass, gravitational infall into non-linear dark matter halo potential wells allows the baryons to largely catch up to the dark matter (e.g. Tegmark et al. 1997). Yoshida et al. (2003b) were able to show that all these approximations are reasonably accurate for determining whether a halo undergoes sufficient H$_2$ cooling for baryonic collapse of primordial haloes to occur.

One inevitable uncertainty involves the masses and the number of the objects formed in the collapse. In order to follow the evolution of the baryonic component to the point at which individual stars are formed, fully 3D simulations with radiative transfer would be required. Such simulations are currently beyond the limit of numerical simulation techniques [Abel et al. (2002); Bromm et al. (2002); see Bromm & Larson (2004) for a review of recent progress]. We thus assume a fiducial collapse model in which the baryonic cooling of a halo results in a single $\sim 120$-M$_\odot$ star.

3 STAR FORMATION CRITERIA

We apply the methodology of Tegmark et al. (1997) and Yoshida et al. (2003b) to determine if the baryonic content of a halo is able to collapse and form stars. The two requirements are: (i) the cooling rate due to H$_2$ must be fast enough for the halo to cool within a Hubble time; and (ii) the dynamical heating rate due to mass accretion must be slower than the baryonic cooling rate. We assume that efficient baryon cooling leads to the formation of an increasingly dense gas core that eventually becomes Jeans unstable (e.g. Abel, Bryan & Norman 2000).

Fig. 1 shows that halo growth rates are remarkably fast, the haloes typically doubling in mass over a redshift interval $\Delta z \simeq 2$. Here, the growth of the 20 largest haloes in simulation R5 is illustrated by tracking the evolution of their largest progenitor. We show later that because of the rapid halo growth, the associated dynamical heating causes a delay in the baryonic collapse for most haloes. Gas cooling is assumed to depend purely on the H$_2$ abundance, which we calculate by integrating the density and temperature-dependent H$_2$-production rate over the lifetime of each halo, as we now discuss.

The amount of H$_2$ in each dark matter halo is computed according to the prescription of Tegmark et al. (1997). The primary H$_2$-production reaction, $H + H^- \rightarrow H_2 + e^-$, requires the presence of free electrons to form H$^-$. The H$_2$-production rate,

$$\frac{df_{H_2}}{dt} \propto T^{3.88} \left( \frac{n_{H_2}}{n_H} \right)$$

(2)

where $k_B$ is the production rate, which is limited mainly by the supply of free electrons for the reaction $H + e^- \rightarrow H^- + h\nu$. H$_2$ production depends on the density and temperature of the gas, which are taken to be the volume-averaged virial quantities, as well as on the free-electron fraction $x_e$. We use $x_e \equiv x_0 = 10^{-4}$ as obtained for a halo that virializes at $z \sim 50$ from the tophat collapse approximation described in Tegmark et al.

We calculate $f_{H_2}$ using a time-integrated production rate. Other authors have taken a different approach, assuming that the production of H$_2$ stalls at a temperature-dependent value when the remaining free electrons have largely recombined (e.g. Tegmark et al. 1997; Yoshida et al. 2003b) but, as we show in Appendix A, our haloes grow too quickly to reach this state. Since halo growth is so rapid, we assume that infalling gas mixes uniformly with existing baryons in the halo, and this has two primary effects. The first is that the H$_2$ within a halo is diluted by infalling material. The second is that electrons depleted through recombination are quickly replenished by infalling gas such that $x_e \simeq x_0$. The production of H$_2$ in a halo is then described by

$$\frac{dM_{H_2}}{dt} \simeq \frac{2\Omega_{H_2}}{\Omega_m} \left( \frac{dM}{dt} f_{H_2} + \left< M \frac{df_{H_2}}{dt} \right> \right).$$

(3)

where $\left< M \frac{df_{H_2}}{dt} \right>$ is the average over adjacent simulation outputs $t_i$ and $t_{i-1}$, $\Omega_{H_2}$ is the hydrogen density in units of the critical density and $M$ is the halo virial mass. The primordial H$_2$ fraction, $f_{H_2,0}$, is quite small ($\sim 10^{-6}$) due to H$_2$ photodissociation by cosmic microwave background (CMB) photons at high redshift (Anninos & Norman 1996; Galli & Palla 1998; Nishi & Susa 1999), and so can be ignored. The H$_2$ abundance at time $t_i$ is thus:

$$f_{H_2,0} \simeq \frac{f_{H_2,0}^i M_{H_2} + \left< M \frac{df_{H_2}}{dt} \right> \Delta t}{M_{H_2}}.$$

(4)
We consider a halo to be a star-forming candidate when the H$_2$ fraction becomes large enough that its baryons can cool within a Hubble time:

$$\tau_{\text{cool,H}_2} = \left(\frac{1}{\gamma - 1} \right) \frac{k_B T}{\rho \Lambda(T)} < \tau_{\text{Hubble}},$$

where $k_B$ is Boltzmann’s constant and $\gamma$ is the adiabatic index of the gas. The cooling time, $\tau_{\text{cool,H}_2}$, is calculated using the cooling function, $\Lambda(T)$, from Galli & Palla (1998) for H$_2$ rotational line transitions. Fig. 2 compares the evolution of $f_{\text{H}_2}$ with the critical abundance, $f_{\text{H}_2,crit}$, required for cooling in the Hubble time. Sufficient H$_2$ for baryonic cooling is produced at $z = 50$, if we assume dilution by newly added gas, at which point the halo has a mass of $10^9 M_{\odot}$. However, we will now demonstrate that the dynamical heating due to increasing halo mass and the corresponding increasing virial temperature delays collapse.

Our second criterion for collapse is that the rate of dynamical heating must be lower than the rate of H$_2$ cooling. If the dynamical heating rate is larger than the H$_2$ cooling rate, then nett cooling will not occur and baryonic collapse will be prevented. Following Yoshida et al. (2003b), we require that:

$$\frac{d Q_{\text{H}_2,\text{cool}}}{d t} > \frac{d Q_{\text{dyn,heat}}}{d t},$$

where $Q$ is the thermal energy per hydrogen atom. The dynamical heating rate is given by

$$\frac{d Q_{\text{dyn,heat}}}{d t} = \frac{d Q}{d T} \frac{dT}{d t} = \frac{k_B}{\gamma - 1} \frac{dT}{d t},$$

and the H$_2$ cooling rate is

$$\frac{d Q_{\text{H}_2,\text{cool}}}{d t} \simeq \Lambda(T) f_{\text{H}_2} n_{\text{H}}.$$
4.2 Smaller haloes

In this section, we contrast the cooling of gas in the largest halo with subsequent cooling in other large haloes in the surrounding region. Fig. 4 compares the properties of the 100 most-massive haloes in the simulation at \( z = 45 \) with the two criteria required for gas to cool and collapse. Although five of the largest haloes have sufficient \( H_2 \) for rapid cooling at redshift of 45, their dynamical heating rates are still so high that only two haloes are able to cool by this time. The mass of the smaller of these two haloes is \( \approx 10^5 h^{-1} \)\( M_\odot \). All of these and eight other haloes meet both baryonic cooling criteria by redshift of 40, and are thus able to cool finally, as reflected in Fig. 5.

At redshift of 40, the characteristic collapse mass for cooling haloes, which we define as the mass above which half of the haloes meet both baryonic cooling criteria, is approximately \( 2 \times 10^5 h^{-1} M_\odot \). This characteristic mass increases slowly to \( \approx 3 \times 10^5 h^{-1} M_\odot \) by \( z = 29 \), at which time 80 haloes in the simulation are capable of cooling (Fig. 6). Most haloes in Figs 4–6 lie along a rough path that follows a general sequence of increasing mass from lower-left to upper-right covering \( 1.5 \times 10^4 - 5 \times 10^5 h^{-1} M_\odot \) in Fig. 4 and \( 6 \times 10^4 - 6 \times 10^5 h^{-1} M_\odot \) in Fig. 6. Collapse is delayed by dynamical heating for the majority of our haloes, as evidenced by the fact that the path of the mass sequence crosses the \( \tau_{\text{Hubble}}/\tau_{\text{H}_2,\text{cool}} = 1 \) line before reaching the \( Q_{\text{H}_2,\text{cool}}/Q_{\text{dyn},\text{cooling}} = 1 \) line.

4.3 Atomic hydrogen cooling

Cooling by atomic hydrogen is possible as early as redshift of 36 when the most-massive halo reaches a virial temperature of \( 10^4 \) K, \( 1.2 \times 10^7 \) \( \text{yr} \) after it formed its first star. At this point, if gas has not all been expelled by prior energy injection, for example from a supernova, the halo should experience much more rapid baryonic cooling. If the high efficiency of metal-free atomic cooling leads to efficient star formation, this could lead to the formation of an extremely early galaxy-like system. The star formation efficiency and stellar initial mass function (IMF) will depend on the details of the molecular cooling (or dust and metals if the gas is previously enriched) required to cool the gas below 8000 K, although isothermal contraction could also lead to star formation (Omukai 2001). By redshift of 29, 13 haloes are hot enough for atomic cooling to occur, and a small ‘group’ of galaxies may already have formed. Note that this would occur well before the first stars form in the simulations by Abel et al. (2002), Bromm et al. (1999, 2002) and Yoshida et al. C⃝ 2005 RAS, MNRAS 363, 393–404
4.4 Abundance and spatial distribution of the first star-forming haloes

We cannot use the simulation to determine the abundance of the first star-forming haloes directly because the high-resolution subvolume encompassed only 90 $h^{-1}$ kpc (comoving) in diameter. The Press & Schechter (1974; hereafter P–S) prediction for the number density of objects more massive than $2.4 \times 10^5 h^{-1} M_{\odot}$ at $z = 47$ is $n_{p-s} (>m) \sim 6 \times 10^{-4} h^3$ Mpc$^{-3}$, while the Sheth & Tormen (1999; hereafter S–T) formula give a much larger value of $5.4 \times 10^{-2} h^3$ Mpc$^{-3}$. We do not know a priori which of these should be preferred since neither formula has been tested against $N$-body simulations in the regime that is relevant here. Reed et al. (2003) found the S–T mass function to be reasonably accurate even for rare massive objects at $z = 10$, and Springel et al. (2005) showed that the S–T formula is much more accurate than the P–S formula out to similar redshift. However, these simulations are still orders of magnitude away from probing the mass range relevant to this study. Nevertheless, in what follows, we will adopt the S–T predictions, but the reader has been warned that this extrapolation is uncertain. Somewhat surprisingly, G05 found that the extended Press–Schechter theory (Bond et al. 1991; Bower 1991) in the form worked out by Mo & White (1996) does accurately reproduce the abundance of haloes in the rare high-density regions tracked in our R4 and R5 simulations over the range $30 < z < 50$.

It is interesting to consider how rare is the initial overdensity that leads to the halo where baryons were first able to cool in our simulation. To estimate this, we identify the Lagrangian volume $V$ of the halo by tracing the particles that make-up the halo at $z = 45$ (which would correspond to the life-span of a $\sim 100 M_{\odot}$ star formed at $z = 47$) back to the starting redshift, $z_{\text{start}}$, of the simulation. The Lagrangian volume $V$ of such a rare halo is nearly spherical because it corresponds to a high peak in the initial density field (e.g. Bardeen et al. 1986). Therefore, we can characterize $V$ by its mean tophat linear overdensity, $\delta$, and radius, $L$. We compute the expected rms density fluctuations on scale $L$ and $\delta(L)$, from our initial power spectrum. At $z_{\text{start}} = 599$, when the fluctuation amplitude is very nearly Gaussian-distributed, $\delta/\sigma(L) = 6.5$, that is, the halo forms from a 6.5$\sigma$ fluctuation.

The clustering properties of the first metal-free stars will determine their role in the reionization and metal enrichment of the surrounding gas. If these stars are highly clustered, then their $H$ II regions could overlap and percolate leading to growing patches of reionization. If, on the other hand, they are more uniformly distributed, percolation will be delayed and the eventual reionization will be more uniform. G05 used the Mo & White (1996) formalism to estimate spatial correlation functions at $z = 47$ for haloes at least as massive as our largest halo. They found a bias factor of 24, which results in a comoving correlation length of $2.5 h^{-1}$ Mpc, only slightly smaller than that of star-forming galaxies today. The comoving abundances estimated above are lower than those of such galaxies, however, and the luminosities predicted for our first star systems are very much lower. As a result, the first light sources are relatively weakly clustered despite their high bias. Neighbouring haloes that lie close enough to allow the possibility of overlapping $H$ II regions are extremely rare at redshift $z = 47$. As we discuss next, the $H$ II regions of haloes in the region we have followed may, nevertheless, percolate.

5 DISCUSSION

We have shown that gas is able to cool and collapse in haloes of mass $2.4 \times 10^5 h^{-1} M_{\odot}$ already at $z = 47$, and that this is followed quickly by the collapse of the gas in less-massive haloes in the surrounding region. In just another 22 Myr, at $z = 36$, the first haloes able to cool by atomic cooling have formed. In this section, we consider the impact of the collapse of the first haloes on their surroundings, focusing in particular on the size of the ionization patches that are likely to be formed. We consider the observable consequences of the formation of the first stars in these haloes at the end of this section.

5.1 Ionizing radiation from the first stars

We begin by calculating the role of the first stars in ionizing the surrounding intergalactic medium (IGM). In the next subsection, we will consider the effect of the first stars on the star formation rate in surrounding haloes. In order to estimate limits on the impact of this first baryonic collapse, we use the flux and life-span calculations for Population III stars by Schaerer (2002). We first establish some fiducial values. The efficiency of production of ionizing photons peaks at a stellar mass of $120 M_{\odot}$. We parametrize the total ionizing photon flux resulting from the collapse of the gas in a halo as

$$r_{\text{ionize}} = 1.1 \times 10^{58} m_{120} f_{1200},$$

(9)

where $m_{120}$ is the mass, in units of $120 M_{\odot}$ of ionizing stars formed per halo and $f_{120}$ is the total number of ionizing photons emitted per stellar mass relative to the specific ionizing flux from a $120 M_{\odot}$ star. Since $f_{120}$ is a weak function of stellar mass, $r_{\text{ionize}}$ is a weak function of the actual mass of individual star(s) formed and is mainly dependent on the total mass turned into stars. Later, we will assume that the $2.4 \times 10^5 h^{-1} M_{\odot}$ dark matter halo (which contains $3.2 \times 10^5 M_{\odot}$ baryons) forms a single $120 M_{\odot}$ star (i.e. $m_{120} f_{1200} = 1$). The maximum possible size of the $H$ II region around such a star-forming object is $76(m_{120} f_{1200})^{1/3}$ kpc (comoving), if we assume no recombination (i.e. the final number of ionized atoms is equal to the total number of ionizing photons produced by the star) and that the sphere expands into the mean baryon density. In practice, the high densities at such high redshift give rise to high recombination rates that are likely to limit the $H$ II region significantly.

The size of the ionized region is uncertain for several reasons. First, the recombination rate depends sensitively on the local gas density. For simplicity, we assume that gas traces dark matter, which is likely to be a good approximation outside the virial radius in the absence of feedback effects, although there is little reason to expect this to be correct within the virial radius. Secondly, a further complication arises from the effect of Compton cooling on the gas. The Compton cooling time-scale is shorter than the recombination time-scale in low-density regions at high redshifts. This could serve to quench the growth of ionized regions because the recombination rate increases as gas cools. Thirdly, the expected short lifetimes of these massive stars could result in only a fraction of haloes hosting their first star at any one time. However, after the death of the first star, the enhanced electron fraction within its $H$ II region may promote subsequent star formation, albeit of stars possibly of lower mass (O’Shea et al. 2005). Fourthly, SNe from the first stars may create bubbles larger than their $H$ II regions. Finally, as we discuss in Section 5.2, L–W radiation may destroy $H_2$ in neighbouring haloes, thus limiting the number of stars that form in $H_2$-cooled hosts, although stars in atomic-cooled haloes may soon dominate the stellar population. In practice, of course, the ionized region will not be spherical but will depend on the morphology of the surrounding gas distribution.

© 2005 RAS, MNRAS 363, 393–404
We have estimated the size of an H II region using a steady-state 1D code. The recombination rate is governed by the rate coefficient,

\[ k_{\text{recomb}} = 1.88 \times 10^{10} T^{-0.64} \text{(cm}^3 \text{s}^{-1}) \] (Hutchins 1976), and the temperature of the ionized gas is set to \( T = 10^4 \text{ K} \). The size of the H II region is taken simply to be the radius at which the recombination rate within the sphere is equal to the ionizing photon production rate. We assume that the gas follows an NFW profile (Navarro, Frenk & White 1996, 1997) with a concentration of 1.75 at \( z = 47 \) and 1.87 at \( z = 29 \), which provides a good fit to the dark matter profile in the simulation (G05). To estimate the effects of the shock front, we smoothen the gas in the central portion of the H II region to have a uniform density over the sound-crossing distance for a star with a 2.5-Myr life-span (assuming ionized gas at \( 10^4 \text{ K} \)). This smoothing also makes the estimated H II region size less sensitive to the assumed profile shape. The mass density surrounding the largest halo in the simulation approximately follows an NFW profile extrapolation until it reaches \( \sim 10 \) times the mean density at \( \sim 3r_{\text{vir}} \), beyond which we assume a uniform density.

For the fiducial \( m_{120} f_{120} = 1 \) case, at redshift of 47, the H II region extends just beyond the virial radius to \( 1.2r_{\text{vir}} \) in our calculation. However, by redshift of 29, physical densities are low enough that over half the ionizing photons from a single 120-\( \text{M}_\odot \) star will escape a halo of the characteristic mass for \( \text{H}_2 \)-cooled haloes at this redshift \((\sim 3 \times 10^5 \text{ h}^{-1} \text{ M}_\odot)\). Additionally, a much larger number of haloes are now susceptible to \( \text{H}_2 \) cooling, raising the possibility that percolation of H II regions could create a large H II bubble. At \( z = 29 \), the H II region around a single 120-\( \text{M}_\odot \) star would extend to \( 9r_{\text{vir}} \) \((21 \text{ h}^{-1} \text{ kpc, comoving})\). This is large enough for the entire \( \sim 75-\text{h}^{-1} \text{ kpc} \) (comoving) radius region that we have simulated to have an H II filling factor of \( \sim 1 \), provided that each halo maintains a similar sized H II region. Of course, such regions are rare and the contribution to the ionized volume of the Universe as a whole is small.

**Figure 7.** In this schematic image, an H II region powered by a 120-\( \text{M}_\odot \) Population III star is centred on each \( z = 29 \) R4 halo that meets the \( \text{H}_2 \)-cooling criteria of Section 3 at redshifts of 45, 40 and 29. In the left-hand panel, the 21-kpc (comoving) radius H II regions (see text) are assumed to be maintained after the end of the first star in each halo by subsequent ionizing sources. In the right-hand panel, H II regions are assumed to survive for only one recombination time \((\sim 1 \text{ Myr})\) beyond the life of the first star in each host. Significant local percolation of H II regions is seen in both cases. In Section 5.2, we discuss feedback mechanisms that may greatly alter the number of star-forming haloes in regions such as this.
(see Section 5.3). The percolation of these H II bubbles is illustrated in Fig. 7, left-hand panel. This scenario is probably unrealistic due to short recombination times of $\sim 1 \text{ Myr}$ in the high density ($> 10^6$ times the mean density) vicinity of these haloes, and due to potential removal of gas by feedback. The short recombination times imply that percolation of ionized bubbles will be significantly reduced if a new ionizing source(s) do(es) not form before the end of the lifetime of the first star. The right-hand panel in Fig. 7 illustrates how percolation will be reduced if each H II region survives only one recombination time after the lifetime of the primary star. Finally, we have considered whether Compton cooling could affect the size of individual H II regions by enhancing the recombination rate, and find that this effect is likely to be unimportant over the lifetime of a massive star. At redshift of 29, the 1.4-Myr Compton cooling timescale is of similar order, but generally longer than the typical local recombination time.

Thus, while the size of an H II region is quite uncertain, it seems possible that a large bubble around the first star-forming halo could become ionized well before the rest of the Universe. One might then expect similarly large and widely separated H II bubbles to be scattered throughout the Universe prior to reionization.

5.2 Feedback from the first generation of stars

In Section 4, we neglected the impact of the collapse of the first stars on neighbouring haloes. The picture we have presented so far only takes into account the ability of a halo to cool and form stars according to local criteria without taking into account external influences. We now briefly discuss how negative feedback from the first stars may impact star formation in the surrounding region.

There are three (possibly more) feedback mechanisms that could mean that the ability of a halo to cool is not a local process, and does not depend solely on the mass of the dark matter halo and its formation history.

(i) L–W radiation. Radiation in the L–W bands can dissociate H$_2$. If the destruction rate exceeds the rate of production, the halo will not build up the amount of H$_2$ needed to cool. As a result, the formation of the first star may ‘sterilise’ the surrounding region and prevent other stars from forming (e.g. Haiman, Rees & Loeb 1997; Ciardi, Ferrara & Abel 2000; Haiman, Abel & Rees 2000; Ricotti, Gnedin & Shull 2002; Mackey, Bromm & Hernquist 2003; Yoshida et al. 2003b).

The photodissociation time-scale for H$_2$ is given by

$$t_{\text{diss}} = k_{\text{diss}}^{-1} F_{\text{trans}}^{-1}$$

(10)

$$k_{\text{diss}} = 1.1 \times 10^{-6} j_{\text{LW}} \text{ (s}^{-1}\text{)}$$

(11)

(Abel et al. 1997), where $j_{\text{LW}}$ is the L–W flux at $h\nu = 12.87$ eV in units of $10^{-20}$ erg s$^{-1}$ cm$^{-2}$ Hz$^{-1}$. In the absence of self-shielding, the transmission factor $F_{\text{trans}}$ is equal to 1.

It is convenient to parameterize the dissociation time-scale in terms of the photodissociating luminosity (12.24–13.51 eV) for the fiducial 120–M$_{\odot}$ star $L_{\odot}$:

$$t_{\text{diss}} = 1.3 \frac{1}{L_{\odot}} \left(\frac{d}{1 \text{ Mpc}}\right)^2 \left(\frac{50}{1 + z}\right)^2 \frac{1}{F_{\text{trans}}} \text{ (Myr)}$$

(12)

where $d$ is the distance from the star in comoving Mpc.

In the absence of self-shielding (see below), a single 120-M$_{\odot}$ star produces enough L–W flux to dissociate H$_2$ within a volume of $\sim 1$ comoving Mpc$^3$. Here, the ‘sterilisation by dissociation’ radius is determined by estimating the distance at which the dissociation time-scale is comparable to the H$_2$-formation time-scale, which is typically $\geq 1$ Myr.

This calculation ignores the fact that the surrounding haloes are likely to contain dense cores. If the optical depth in these regions is sufficiently high, the inner mass will escape dissociation because of self-shielding in the outer parts (e.g. Glover & Brand 2001; Kitayama et al. 2001; Machacek et al. 2001). Unfortunately, the effectiveness of self-shielding depends on the internal velocity structure of the halo gas caused by infall and turbulent motions. The properties of these flows are uncertain and they will be difficult to model adequately even in the highest-resolution simulations.

If the gas velocities are negligible, the transmission factor is given approximately by

$$F_{\text{trans}} = \min\left[1, \left(\frac{N_{\text{HI}}}{10^{14} \text{ cm}^{-2}}\right)^{-3/4}\right]$$

(Draine & Bertoldi 1996), where $N_{\text{HI}}$ is the H$_2$ column density, integrated from outside-in. Because it neglects gas motions, this calculation gives the maximum possible self-shielding (Machacek et al. 2001). For an NFW halo profile with a concentration of 1.75, and assuming that the H$_2$ density traces the mass density, equation (13) implies that 90 per cent of the H$_2$ mass ($r > 0.2r_{\text{vir}}$) will be prone to dissociation by a single 120-M$_{\odot}$ star located at a distance of 65 kpc at redshift of 45, for a typical H$_2$ fraction of $10^{-4}$. This implies that dissociation is still important at large distances. For the typical halo separation in our simulations ($\sim 10 h^{-1}$ kpc), a transmission factor of less than $1.5 \times 10^{-4}$ would be required to ensure nett H$_2$ production ($t_{\text{diss}} \gtrsim 1$ Myr) at redshift of 45. However, at this distance, the central baryonic mass that is adequately shielded is much less than that required to form a massive star unless either the baryon density or the H$_2$ fraction is significantly enhanced. On the other hand, baryonic cooling is likely to concentrate baryons strongly towards the halo centre (e.g. Yoshida et al. 2003b), increasing the L–W optical depth and decreasing the effect of L–W feedback. These issues can only be addressed with detailed numerical hydrodynamic simulations.

We estimate that L–W shielding, either by intergalactic H$_2$ or by intergalactic neutral hydrogen, will not be important. Using the transmission factor of equation (13), and assuming a primordial H$_2$ fraction of $10^{-6}$ at redshift of 50, L–W shielding by intergalactic H$_2$ will be negligible except over very large distances ($F_{\text{trans}} = 0.1$ at 1 Mpc, comoving). This estimate likely overpredicts the intergalactic shielding effect because doppler shifts due to Hubble expansion and intrinsic gas motions will decrease the optical depth. More importantly, any shielding by intergalactic H$_2$ will be shortlived because it will quickly be dissociated. Shielding by intergalactic neutral hydrogen is also ineffective because L–W absorption will be largely confined to the closely spaced Lyman series lines in the upper energy L–W bands. Even at a distance of 1 Mpc (comoving), less than half of total L–W radiation will be absorbed by neutral hydrogen at redshift of 50.

We have rerun our H$_2$–production model including the dissociating effects of L–W radiation under the assumption that self-shielding is unimportant. The L–W flux from a single 120-M$_{\odot}$ star, which is enough to quickly sterilize the entire ($75 h^{-1}$ kpc comoving radius) high-resolution subvolume of the R4 simulation. However, if we then switch-off the L–W flux after a 2.5–Myr stellar lifetime, a second halo is able to produce enough H$_2$ to meet the baryonic cooling criteria within 1–2 Myr. Thus, the sterilizing effects of a single massive star are temporary, but may nevertheless limit the star formation density to $\sim 1$ star per comoving Mpc per Myr in H$_2$-cooled hosts. However, if self-shielding within dense baryonic...
cores can cause a significant time delay between the onset of baryonic collapse and subsequent sterilization of neighbouring haloes, then there may exist a brief window during which multiple haloes will have time to produce enough H₂ for baryonic cooling. A sufficiently long window could allow a rapid ‘mini-burst’ of H₂-induced Population III star formation in a previously sterilized region. Thus, depending on the detailed structure of collapsing clouds, the radius out to which L–W feedback is effective could be as large as ∼1 Mpc, which would substantially limit the number of haloes that undergo collapse via H₂ cooling. Self-shielding is likely to mediate this feedback to some degree, but it is probably unimportant except for dense baryonic cores. However, we stress that even if all the H₂ in a halo were destroyed by external L–W radiation, atomic cooling alone should still be sufficient to trigger baryonic collapse in our haloes once they reach a sufficiently high virial temperature (e.g. Omukai 2001).

(ii) X-ray and far UV radiation. The first SNe and miniquasars will bathe the surrounding region with ionizing photons (e.g. Madau et al. 2004). Those of sufficiently high energy have small ionization cross-sections and will escape from the H II region. However, even if they only ionize a small fraction of the neutral hydrogen in a nearby halo, the increase in the free-electron density will catalyze the production of H₂. If this process is significant in the centres of these haloes, it would promote the collapse of their gas, acting as a positive feedback mechanism (e.g. Haiman, Rees & Loeb 1996; Haiman et al. 2000; see however, Machacek, Bryan & Abel 2003, who find that a uniform X-ray background has little effect on H₂-induced cooling and collapse of haloes in simulations).

(iii) Metal enrichment and SNe. The energy released by the explosion of the first SNe sends a blast wave into the surrounding IGM. If cooling haloes are fragile, the associated mechanical energy can remove portions of their outer envelope and prevent further cooling (e.g. Rees 1985; Bromm, Yoshida & Hernquist 2003; Kitayama & Yoshida 2005).

Population III stars with masses in a narrow range around 200 M⊙ are thought to end their life as a pair production supernova (e.g. Heger & Woosley 2002). In this case, the supernova ejecta is very metal-rich and, if this material mixes in with the gas in nearby haloes, the overall gas cooling efficiency will be enhanced. Metal enrichment is, in any case, required in order to enhance baryon cooling at low temperatures (<10⁴ K) so as to promote the fragmentation that results in the low-mass IMF typical of Population II stars (Schneider et al. 2002, 2003; Mackey et al. 2003). At redshift of ∼30, energetic (∼10⁷⁵) SNe blastwaves may expand up to ∼20 kpc (comoving) before the initial fast expansion is ‘stalled’ [estimated according to Yoshida et al. (2004)], suggesting that metal enrichment from neighbouring haloes is possible.

Thus, SNe blast waves can produce both positive and negative feedback. However, the region affected by these processes is likely to be small compared to the region potentially affected by the L–W radiation.

(iv) H⁻ photodetachment. Another possible mechanism suppressing H₂ formation is the photodetachment of H⁻ ions before they can combine with a hydrogen atom to produce H₂. We estimate the radius over which photodetachment suppresses H₂ formation by equating the photodetachment rate to the rate at which H⁻ reacts to form H₂ [H + H⁻ → H₂ + e⁻; rate taken from Hirasaki (1969)], assuming virial overdensity. We assume the effective temperature and luminosity of a 120-M⊙ Population III star from Schaerer (2002), and employ the Tegmark et al. (1997) fitting function to photodetachment cross-sections from Wishart (1979). This yields a photodetachment distance of 2.5 kpc (comoving) at z = 29, roughly equal to the virial radius of a typical halo on the verge of baryonic cooling at this time, and much smaller than the H II region expected from such a massive star. Thus, we do not expect H⁻ photodetachment by ∼100-M⊙ Population III stars to impact baryon cooling in neighbouring haloes.

In summary, our calculations show that L–W radiation could ‘sterilize’ a region as large as ∼1 Mpc (comoving) around a primordial massive (120-M⊙) star. However, the effects of self-shielding are so uncertain that we cannot rule out the possibility that dissociation feedback may be ineffective at suppressing further star formation. The uncertainties in other feedback mechanisms are also considerable. In the preceding section, we discussed the possibility that the H II regions around the first stars might percolate giving rise to large multi-halo H II bubbles. Whether or not this happens depends primarily on how effective the L–W sterilization is at suppressing star formation in the vicinity of a star, and on the impact of the other feedback schemes discussed above. However, even if sterilization does occur, this can only delay star formation in the region, since atomic cooling, already possible in 13 haloes in our simulation by z = 29, will ultimately become the dominant path for star formation. For the case of maximally effective sterilization, the S–T number density of haloes capable of atomic cooling surpasses that of H₂-cooled haloes (∼1 h⁻³ Mpc⁻³) by z ∼ 25. Gas in atomic-cooled haloes could dominate the total stellar mass and UV radiation considerably earlier than this if, as seems likely, this gas turns into stars more efficiently than gas within H₂-cooled haloes due to faster rate and greater temperature sensitivity of atomic cooling.

Unfortunately, it is not yet possible to rule out any particular scenario. We proceed by exploring the extreme case where dissociative feedback is unimportant, and percolation leads to the formation of a large ionized region.

5.3 Contribution to reionization

The sequence of events following the formation of the first generation of Population III stars (z ≥ 45) is uncertain because of our poor understanding of the interplay between negative feedback from L–W radiation and positive feedback from the X-ray/UV background. If these processes are unimportant or if positive feedback dominates, our calculations suggest that large (∼100 h⁻¹ kpc comoving) patches of the Universe will be reionized at z ∼ 30. The first regions of the Universe to collapse may thus become reionized much earlier than the z ∼ 20 inferred from the WMAP microwave background radiation data (Kogut et al. 2003). This could have dramatic effects on low-mass galaxies forming in these regions, whose gas could be removed suppressing subsequent star formation.

Observational constraints derived from observations of high-redshift QSOs (Fan et al. 2001) and from the temperature of the high-redshift IGM (Theuns et al. 2002) imply that complete reionization occurs at z = 6–9. In contrast, the high optical depth for Thomson scattering inferred from the WMAP data suggests a much earlier epoch for reionization, at least over significant fractions of the Universe. It is interesting to consider whether large patches of early reionization could reconcile these contradictory constraints (see e.g. Bruscoli, Ferrara & Scannapieco 2002; Cen 2003; Ciardi, Ferrara & White 2003; Haiman & Holder 2003; Sokasian et al. 2004). However, according to the number densities inferred from the S–T formula, regions such as the one we have simulated are quite rare and so they can only make a small contribution to global reionization and to the Thompson optical depth. Moreover, stars in H₂-cooled haloes could even delay reionization if their sterilizing
effects prevent H$_2$-cooled star formation over large distances. Alternatively, strong L–W feedback from the first Population III stars might actually expedite reionization by preventing the expulsion of gas by feedback associated with star formation in early low-mass haloes, leaving plentiful gas supplies when haloes grow large enough for efficient atomic cooling.

5.4 Observability

Finally, we consider whether the first generation ($z \gtrsim 45$) of Population III stars that is predicted by our simulations could actually be observed. We focus first on the case where H$_2$ is unaffected by feedback processes, which are highly uncertain. The SNe resulting from the collapse of these stars are potentially observable, particularly if they can be detected as gamma-ray bursts (GRBs). The time dilation of bursts occurring at higher redshifts implies that the source can be observed closer to maximum light, thus compensating for the increasing luminosity distance. In a recent paper, Gou et al. (2004) showed that GRBs can be detected in X-rays out to $z = 30$ and beyond in exposure times as short as 300 s.

The rest-frame UV emission from the burst will be strongly absorbed shortwards of Lyman $\alpha$ (Ly$\alpha$), so that detection will need to be made in the mid-infrared. Gou et al show that the James Webb Space Telescope will have the required sensitivity to detect GRBs out to $z = 37$ in the 4.8-µm band in only 1 h. Detection of GRBs at higher redshift should also be possible in longer wavelength bands and larger exposure times. Radio afterglows may also be possible at $z \sim 30$ (Ioka & Meszaros 2005).

The number of haloes massive enough to support star formation lying in our past lightcone at $z \gtrsim 45$ is of the order of $10^4$, according to the S–T prediction. If we assume that each of these haloes forms a single massive star with a 2.5-Myr life-span whose endpoint is a GRB, then the event rate over the entire sky would be of the order of one per month. Our estimated abundances (which, as we discussed in Section 4.4, are uncertain) are higher than, but in rough agreement with those of Miralda-Escudé (2003) who, on the basis of the P–S formula, predicts the highest-redshift star in our past lightcone to be at redshift of 48, assuming that such stars form in haloes with a virial temperature of 2000 K. Our $z \sim 50$ SNe rates are of similar order to those predicted by Wise & Abel (2004) on the basis that SNe are hosted by haloes of mass $\sim 10^5 M_{\odot}$, whose abundances are given by the S–T model. Note that we have ignored GRB beaming in this calculation; the number of observable GRBs that are beamed towards us is likely to be lower by a factor of $\sim 100$–1000 (e.g. Frail et al. 2001).

The rate of potentially observable GRBs or SNe forming in haloes, where gas has cooled via H$_2$, increases rapidly at lower redshifts. At $z = 29$ in our simulation, we expect $\sim 10$ SNe per day over the entire sky or 1 per $\sim 10$ square degrees per year (again, ignoring beaming effects). If sterilisation by L–W radiation strongly suppresses star formation, the number of observable GRBs at $z \gtrsim 29$ could be reduced by a significant factor, depending on the uncertain effects of self-shielding. However, even in that case, the number density of haloes capable of sustaining atomic cooling would soon surpass that of H$_2$-cooled haloes, leaving open the possibility of abundant pre-reionization observational targets. The increase in the expected event rate at lower redshifts suggests that GRBs or SNe should indeed be observable in the pre-reionized universe.

We conclude that if the highest-redshift Population III stars end their life as GRBs, the prospects for their detection are promising. Even if these stars do not have a GRB phase or if the beaming angle is small, observational evidence of the first stars may still be available, for example, through direct detection of SNe (Panagia 2003) or of emission from an early generation of accreting black holes (e.g. Madau et al. 2004). The difficulty, however, will lie in identifying such rare and faint objects and distinguishing them from foreground contaminants.

6 CONCLUSIONS

We have modelled the cooling of gas by H$_2$-line emission in a series of N-body simulations in order to identify the first generation of dark matter haloes capable of forming stars. We followed the formation of structure in a region selected to host a rich cluster today, repeatedly simulating, at increasing resolution, the largest progenitors of a massive halo at the final time. To determine whether the gas is able to cool, we estimated the production rate of H$_2$ and considered the balance between radiative cooling and dynamical heating. The dramatic growth of the halo at early times has the nett effect of delaying, but not preventing the cooling of gas. Our main results are as follows.

(i) The first star-forming haloes. Since we have followed evolution in a special region of the Universe (selected to have a rich cluster today), we find that formation of the first stars occurs significantly earlier than commonly found in simulations of volumes too small to sample such rare overdensities. At a redshift of 47, there is already a population of haloes in the simulation that are undergoing baryonic collapse via H$_2$ cooling and can host the first stars.

The mass of these haloes is typically $2.5 \times 10^7 h^{-1} M_{\odot}$ and their comoving space density is comparable to that of dwarf galaxies today. By redshift of 29, 80 haloes in the high-resolution simulation subvolume of radius $75 h^{-1}$ kpc are capable of forming stars. The largest halo reaches a virial temperature of $10^5$ K, at which gas in equilibrium can begin to cool by more efficient atomic processes, at $z = 36$, potentially giving rise to a first-generation galaxy. By $z = 29$, 13 haloes in the simulation have grown large enough to sustain atomic cooling.

(ii) Ionising effects. The formation of the first stars will ionize the gas around them. At $z = 47$, the hydrogen recombination rates are high enough that these H II regions are likely to be contained almost entirely within their own host halo. However, by $z = 29$, gas densities have dropped sufficiently that massive ($\sim 100 M_{\odot}$) Population III stars are able to ionize regions significantly beyond the halo. Percolation of neighbouring regions could lead to the formation of an H II superbubble. However, such regions would be too widely separated to contribute significantly to reionization at this time.

(iii) Feedback. Feedback effects that could alter the cooling properties of the gas are likely to occur but their nature (and, in some cases, even their sign) is difficult to calculate. L–W radiation from just a single massive star could sterilize a large surrounding volume by dissociating H$_2$, severely limiting star formation throughout the volume. However, the effectiveness of this feedback channel would be limited if the H$_2$ within dense halo cores is self-shielded by the outer halo layers. Supernova feedback could hinder star formation by removing gas, or could enhance star formation by blast wave compression, by metal enrichment, or through high-energy radiation, which might produce electrons that catalyze H$_2$ formation.

Even if feedback is effective enough to suppress H$_2$ cooling within a region of size $\sim 1$ Mpc (comoving) around the first star, further star formation is still expected in haloes where gas can cool by atomic processes. This happens as early as $z = 36$ in our simulated region.

(iv) Observability, The prospects for observing the very earliest generation of stars are encouraging. The James Webb Space...
Telescope may be able to view directly pre-reionization SNe and gamma-ray bursts from Population III objects in the rest-frame mid-infrared (Panagia 2003). X-ray emission or radio afterglow from these early gamma-ray bursts may also be detectable (Gou et al. 2004). The event rate of SNe at z = 29 and higher could be as large as ~10 per day. However, if feedback due to L–W radiation is important, the event rate could be substantially lower.

ACKNOWLEDGMENTS

DR is supported by PPARC. RGB is a PPARC Senior Fellow. TT thanks PPARC for the award of an Advanced Fellowship. The simulation DR is supported by PPARC. RGB is a PPARC Senior Fellow. TT

REFERENCES

Abel T., Anninos P., Zhang Y., Norman M., 1997, NewA, 2, 181
Abel T., Bryan G., Norman M. L., 2000, ApJ, 540, 39
Abel T., Bryan G., Norman M. L., 2002, Sci, 295, 93
Anninos P., Norman M. L., 1996, ApJ, 460, 556
Bardeen J., Bond R., Kaiser N., Szalay L., 1986, ApJ, 305, 15
Barkana R., Loeb A., 2004, ApJ, 609, 474
Bond R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bower R., 1991, MNRAS, 248, 332
Bromm V., Larson R., 2004, ARA&A, 42, 79
Bromm V., Loeb A., 2004, NewA, 9, 353
Bromm V., Coppi P. S., Larson R. R., 1999, ApJ, 527, L5
Bromm V., Coppi P. S., Larson R. R., 2002, ApJ, 564, 23
Bromm V., Yoshida N., Hernquist L., 2003, ApJ, 596, L135
Bruscoli M., Ferrara A., Scannapieco E., 2002, MNRAS, 330, L43
Cen R., 2003, ApJ, 591, 12
Ciardi B., Ferrara A., 2005, Space Sci. Rev., 116, 625
Ciardi B., Ferrara A., Abel T., 2000, ApJ, 533, 594
Ciardi B., Ferrara A., White S., 2003, MNRAS, 344, L7
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 381
Draine B., Bertoldi F., 1996, ApJ, 468, 269
Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
Fan X. et al., 2001, AJ, 122, 2833
Fraui D. et al., 2001, ApJ, 562, 55
Galli D., Palla F., 1998, A&A, 335, 403
Gao L., Loeb A., Peebles P., White S., Jenkins A., 2004a, ApJ, 614, 17
Gao L., White S., Jenkins A., Stoehr F., Springel V., 2004b, ApJ, 355, 819
Gao L., De Lucia G., White S., Jenkins A., 2004c, ApJ, 352, L1
Gao L., White S. D. M., Jenkins A., Frenk C. S., Springel V., 2005, MNRAS, in press (doi: 10.1111/j.1365-2966.2005.09509.x) (astro-ph/0503003)
(G05)
Glover S., Brand P., 2001, MNRAS, 321, 385
Gou L., Meszaros P., Abel T., Zhang B., 2004, ApJ, 604, 508
Haiman Z., Holder G., 2003, ApJ, 595, 1
Haiman Z., Rees M., Loeb A., 1996, ApJ, 467, 522
Haiman Z., Rees M., Loeb A., 1997, ApJ, 476, 458
Haiman Z., Abel T., Rees M., 2000, ApJ, 534, 11
Heger A., Woosley S., 2002, ApJ, 567, 532
Hirasawa T., 1969, Prog. Theor. Phys., 42, 523
Hutchings J. B., 1976, ApJ, 205, 103
Ioka K., Meszaros P., 2005, ApJ, 619, 684
Jenkins A., Frenk C. S., White S. D. M., Colberg J., Cole S., Evrard A., Couchman H., Yoshida N., 2001, MNRAS, 321, 372

© 2005 RAS, MNRAS 363, 393–404

Kitayama T., Yoshida N., 2005, ApJ, in press (astro-ph/0505368)
Kitayama T., Susa H., Umemura M., Ikeuchi S., 2001, MNRAS, 326, 1353
Kitayama T., Yoshida N., Susa H., Umemura M., 2004, ApJ, 613, 631
Kogut A. et al., 2003, ApJS, 148, 161
Lacey C., Cole S., 1994, MNRAS, 271, 676
Machacek M., Bryan G. L., Abel T., 2001, ApJ, 548, 509
Machacek M., Bryan G. L., Abel T., 2003, MNRAS, 338, 273
Mackey J., Bromm V., Hernquist L., 2003, ApJ, 586, 1
Madau P., Rees M., Volotemi M., Haardt F., Oh S., 2004, ApJ, 604, 484
Miralda-Escudé J., 2003, Sci, 300, 1904
Mo H., White S. D. M., 1996, MNRAS, 282, 347
Nakamura F., Umemura M., 2001, ApJ, 548, 19
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Navarro J. F. et al., 2004, MNRAS, 349, 1039
Nishi R., Susa H., 1999, ApJ, 523, L103
Omont K. J., 2001, ApJ, 546, 635
Omont K. J., Nishi R., 1999, ApJ, 518, 64
Omont K., Palli F., 2003, ApJ, 589, 677
O’Shea B., Abel T., Whalen D., Norman M., 2005, ApJ, 628, L5
Panagia N., 2003, CHJA&A, 3, 115
Peekies P., Dicke R., 1968, ApJ, 567, 515
Press W. H., Schechter P., 1974, ApJ, 187, 425 (P-S)
Reed D., Gardner J., Quinn T., Stadel J., Fardal M., Lake G., Governato F., 2003, MNRAS, 346, 565
Rees M. J., 1985, MNRAS, 213, 75r
Ricotti M., Gnedin N., Shull M., 2002, ApJ, 575, 49
Saslaw W., Zipoy D., 1967, Nat, 216, 976
Schaerer D., 2002, A&A, 382, 28
Schneider R., Ferrara A., Natarajan P., Omukai K., 2002, ApJ, 571, 30
Schneider R., Ferrara A., Salvaterra R., Omukai K., Bromm V., 2003, Nat, 422, 869
Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437
Sheth R., Tormen G., 1999, MNRAS, 308, 119 (S-T)
Singh S., Ma C., 2002, ApJ, 569, 1
Sokasian A., Yoshida N., Abel T., Hernquist L., Springel V., 2004, MNRAS, 350, 47
Spergel D. et al., 2003, ApJS, 148, 175
Springel V., 2005, MNRAS, submitted (astro-ph/0505010)
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Springel V. et al., 2005, Nat, 435, 629
Tegmark M., Silk J., Rees M., Blanchard A., Abel T., & Palla F., 1997, ApJ, 474, 1
Theuns T., Schaye J., Zaroubi S., Kim T., Tzanavaris P., Carswell B., 2002,

APPENDIX A

In this appendix, we discuss the production of H2, comparing our method for calculating f_H2 with that employed by Tegmark et al.
Figure A1. Comparison of the estimated production rate of $H_2$, $df_{H_2}/dt$, in a halo (solid curve) versus $[df_{H_2}/dt]_{\text{asym}}$ (dashed curve), the rate of production needed to maintain the 'asymptotic' $H_2$ fraction. $f_{H_2,\text{asym}}$ is the resulting $H_2$ abundance after $H_2$ production slows dramatically due to the depletion of catalyzing free electrons. The growth of the halo is too rapid to maintain $[df_{H_2}/dt]_{\text{asym}}$ below $z \gtrsim 62$, and so at $z < 62$, the $H_2$ abundance of the halo will be below $f_{H_2,\text{asym}}$. In this regime, the $H_2$ abundance is thus limited primarily by the rapid growth of the halo rather than by electron depletion. We thus determine the $H_2$ abundance from a time integration of the $H_2$-production rate (see text). The curves that stop at redshift of 49 follow simulation ‘R5’ until its completion, while the curves that continue to redshift of 29 are for the same halo, which is modelled at lower resolution in simulation ‘R4’.

As the electrons left over from recombination are depleted, the production of $H_2$ slows down. This leads to the ‘asymptotic’ $H_2$ fraction, $f_{H_2,\text{asym}} \propto T^{1.52}$, where $T$ is assumed to be equal to the virial temperature of the halo [see equation (17) of Tegmark et al. (1997)].

Since $T_{\text{vir}} \propto M^{2/3}$, where $M$ is the halo mass, it follows that $df_{H_2,\text{asym}}/dt \propto dM/dt$ as the halo grows in mass, where $d f_{H_2,\text{asym}}/dt$ is the $H_2$-production rate needed to sustain $f_{H_2,\text{asym}}$ which increases as the halo grows. The asymptotic $H_2$ abundance is maintained only if the halo growth is sufficiently slow so that the actual production rate

$$\frac{df_{H_2}}{dt} > \frac{df_{H_2,\text{asym}}}{dt},$$

(A1)

where $df_{H_2}/dt$ is the halo $H_2$-production rate. If the halo grows too rapidly, $H_2$ production will not be able to keep up (even with the primordial abundance of free electrons) and the cooling efficiency of the halo will be reduced. In Fig. A1, we show that the $H_2$-production rate for our largest halo drops below the rate required to maintain the asymptotic $H_2$ abundance at $z = 62$. This means that for $z < 62$, $f_{H_2}$ for the halo increases more slowly than $f_{H_2,\text{asym}}$ because of the high halo growth rate. Similar curves are readily produced for other high-redshift haloes. It is thus necessary to estimate $H_2$ abundances by integrating $df_{H_2}/dt$ over time, rather than by simply adopting ‘asymptotic’ $H_2$ abundances (we also include dilution effects due to mergers as described in Section 3). Some haloes have periods where $df_{H_2}/dt$ [calculated from equation (2)] is larger than $df_{H_2,\text{asym}}/dt$, in which case we assume that electron depletion limits $H_2$ production to $df_{H_2,\text{asym}}/dt$.

This paper has been typeset from a TeX/LaTeX file prepared by the author.