Mon. Not. R. Astron. Soc. **340,** 1240–1248 (2003)

The mass function of primordial star clusters

Fernando Santoro and Peter A. Thomas★
Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QJ

Accepted 2002 December 19. Received 2002 December 13; in original form 2002 September 9

ABSTRACT

We use the block model to generate merger trees for the first star clusters in a Λ-cold dark matter cosmology. Using a simple collapse model and cooling criterion, we determine which haloes are able to form stars before being disrupted by mergers. We contrast the mass functions of all the resulting star clusters and those of primordial composition, i.e. star clusters that have not been contaminated by subclusters inside them. In confirmation of previous work, two generations of primordial star clusters are identified: low-temperature clusters that cool via molecular hydrogen, and high-temperature clusters that cool via electronic transitions. The former dominate by number, but the two populations contain a similar mass with the precise balance depending upon the details of the model. We speculate on the current-day distribution of Population III stars.

Key words: Galaxy: formation – Galaxy: halo – Galaxy: stellar content – dark matter.

1 INTRODUCTION

In a cold dark matter (CDM) cosmology, small structures are the first to collapse and these then cluster together in a hierarchical fashion, giving rise to the bottom–up picture of galaxy formation. In this paper, we use the term 'primordial star cluster' for the first objects that are able to cool and form, zero-metallicity Population III stars. They are of interest both in their own right and because they may be responsible for reionization of the intergalactic medium.

In a primordial gas, for which the main elements are hydrogen and helium and their derivatives, there are two main cooling mechanisms, dependent on the temperature: for haloes with virial temperatures below 8600 K the cooling is dominated by rotovibrational excitations of hydrogen molecules, while those with higher temperatures cool mainly via electronic transitions.

In a landmark paper entitled 'How small were the first cosmological objects?', Tegmark et al. (1997, hereafter T97) analytically tracked a top-hat collapse to the point of virialization, at which point the gas was cooled at constant density. They accepted an object as having cooled if it met the criterion $T(0.75z_{\rm vir}) \leqslant 0.75\,T_{\rm vir}$, where $T_{\rm vir}$ is the virial temperature and $z_{\rm vir}$ is the virialization redshift. They found that the first generation of objects that cooled in a standard CDM scenario virialized at a redshift of 27 and had a baryonic mass of about $10^5\,{\rm M}_\odot$. In a later paper, Abel et al. (1998) repeated the calculation with a different H_2 cooling function and estimated a very similar virialization redshift but a smaller baryonic mass, $7 \times 10^3\,{\rm M}_\odot$.

In a paper somewhat cheekily entitled 'How big were the first cosmological objects?' (Hutchings et al. 2002, hereafter HSTC02) extended the previous work to include haloes of higher mass. They found two distinct generations of haloes: Generation 1 haloes dominated by molecular cooling as in the previous work and higher-mass Generation 2 haloes dominated by electronic cooling. The properties of these haloes in three different versions of the CDM cosmology are listed in Table 1.

Note that the difference in cooling redshift between the two generations of haloes is small (partly because the CDM fluctuation spectrum is flat on small scales and partly because cooling is more efficient in Generation 2 haloes). This led HSTC02 to speculate that both generations of haloes may form Population III stars (whereas others have considered only the smaller, Generation 1 haloes to be important).

There were two main deficiencies in the model of HSTC02. First, they neglected substructure: the referee suggested that all Generation 2 haloes will have Generation 1 haloes inside them and so they will not be of primordial composition. We show below that primordial Generation 2 haloes can exist. Secondly, they considered only 3σ fluctuations. In reality there will be a Gaussian distribution of overdensities leading to a wide range of halo masses virializing at any given redshift.

The present paper, as a continuation of the previous work, addresses these two points by using the block model of Cole & Kaiser (1988) to generate a merger history of collapsed haloes. This allows us to follow a wide dynamic range of halo masses very efficiently. Within the merger tree, the properties of haloes are calculated using the same chemical model as in HSTC02.

Our work complements that of other authors who are investigating first object formation using numerical simulations (e.g. Abel et al. 1998; Bromm, Coppi & Larson 1999, 2002; Abel, Bryan & Norman 2000; Fuller & Couchman 2000; Nakamura & Umemura 2001, 2002). They are able to follow the dynamical evolution of

★E-mail: p.a.thomas@sussex.ac.uk

Table 1. Properties of the two generations of 3σ haloes to form in each of the cosmologies (from HSTC02): cosmological model; redshift at which the halo cools to 75 per cent of the virial temperature; virial temperature; total mass of the halo; baryonic mass of the halo.

Model	$z_{0.75}$	$T_{\rm vir}/{\rm K}$	$M_{\rm tot}/{\rm M}_\odot$	$M_{\rm bary}/{\rm M}_\odot$
Generation 1				
SCDM	19.5	3 600	1.6×10^6	1.2×10^5
τCDM	10.8	4 500	5.0×10^6	9.2×10^5
ΛCDM	21.9	3 400	8.6×10^5	3.3×10^4
Generation 2				
SCDM	18.5	10 800	9.9×10^6	7.5×10^5
τCDM	10.7	10 600	2.1×10^7	3.9×10^6
ΛCDM	20.6	10 400	5.7×10^6	2.2×10^5

single star clusters in great detail, whereas we learn instead about the properties of the cluster population.

We describe our numerical method in Section 2. The properties of individual haloes for a merger tree corresponding to an overdense region of the Universe are presented in Section 3 and the corresponding mass function is described in Section 4. Finally, Section 5 explores variations of the basic model and discusses the nature of the star clusters.

2 METHODOLOGY

In this paper, we use the popular ΛCDM cosmology, for which parameters are tabulated in Section 2.1. The generation of a merger tree of collapsed haloes is described in Section 2.2 and the criteria whereby we determine which of these form star clusters is outlined in Section 2.3.

2.1 Cosmology

HSTC02 investigated the cooling of haloes in three different CDM cosmologies. In this paper, we restrict our attention to the currently favoured ΛCDM cosmology, for which the parameters are listed in Table 2.

We have used the transfer function calculated by CMBFAST. There have been recent suggestions that the normalization of the power spectrum may be closer to 0.7 than 0.9 (e.g. Seljak 2002; Allen et al. 2002) this would have the effect of moving the formation epoch of the first star clusters to lower redshift and also lowering the amount of substructure.

2.2 Block model

We generate a halo merger tree using the block model of Cole & Kaiser (1988). This starts with a 'root' block of mass M_0 and density fluctuation δ_0. In this paper we fix $M_0 = 10^{11}\,{\rm M}_\odot$ and choose two different values of δ_0 corresponding to a 3σ fluctuation ($\delta_0 = 10.98$) and the mean density ($\delta_0 = 0$).

Table 2. Cosmological parameters: density parameter; cosmological constant in units of $\lambda_0 = \Lambda/3H_0^2$; current baryon density in units of the critical density; Hubble parameter in units of $h = H_0/100\ {\rm km\ s^{-1}\ Mpc^{-1}}$; power spectrum shape parameter; root-mean-square dispersion of the density within spheres of radius $8h^{-1}$ Mpc.

Ω_0	λ_0	$\Omega_{\rm b0}$	h	Γ	σ_8
0.35	0.65	0.038	0.7	0.21	0.90

Geometrically the block can be visualized as a cuboid with sides in the ratio $1 : 2^{1/3} : 2^{2/3}$, but the density fluctuations are calculated as for a spherical top-hat model of the same mass; given the uncertainty in the normalization of the power spectrum, the distinction is of no importance. The block can be bisected by a plane perpendicular to its longest axis, creating two similar blocks of half the mass, $M_1 = M_0/2$. To generate density fluctuations in these daughter blocks, we add power drawn at random from a Gaussian distribution with variance $\sigma^2(M_1) - \sigma^2(M_0)$; this approximation has been tested by Cole & Kaiser (1988), where they found a very good agreement between the mass fluctuations generated by the block model and those from Press–Schechter theory (see fig. 3 in their paper). A positive fluctuation is added to one block and an equal negative fluctuation to the other, so as to conserve the overall level of fluctuations in the root block.

The same procedure is then repeated, with each parent block being divided into two equal-mass daughters until the desired resolution is reached. In this paper, we use 21 levels, creating a total of $2^{21} - 1 \approx 2.1 \times 10^6$ blocks with a minimum block mass of $9.5 \times 10^4\,{\rm M}_\odot$.

We use a simple model in which the collapse of blocks to form bound objects is determined only by their overdensity. To be precise, we assume them to virialize once their linear overdensity reaches $\delta_{\rm c} = 1.69$ (see Eke, Cole & Frenk 1996 and Lokas & Hoffman 2001, where only a very weak dependence on Ω has been found).

The equation for the rate of growth of δ with redshift, z, is

$$\delta(z) = \frac{\delta(0)}{1+z}\frac{g(\Omega)}{g(\Omega_0)}, \tag{1}$$

where

$$g(\Omega) = 2.5\frac{\Omega}{\left(\frac{1}{70} + \frac{209}{140}\Omega - \frac{1}{140}\Omega^2 + \Omega^{4/7}\right)} \tag{2}$$

is the growth suppression factor (Viana & Liddle 1996) and

$$\Omega = \frac{\Omega_0}{\left[\Omega_0 + \left(\frac{1}{1+z}\right)^3(1-\Omega_0)\right]}. \tag{3}$$

In the binary tree generated by the block model, most of the blocks are contained in larger blocks of greater overdensity. Often this will be the immediate parent (one of the two daughters of each parent will have lesser density), but it could also be a block further up the tree. Under these circumstances, the larger block will collapse before the smaller one, and so the latter will never attain an independent existence as a virialized structure. We eliminate these underdense blocks from the tree and call the remaining blocks 'haloes.'

We have performed 100 realizations of the block model with different number seeds. We use these for both values of δ_0 as the set of haloes is the same in each case. Unless stated otherwise, the results presented below are averages over all realizations.

2.3 Halo evolution and the formation of star clusters

We wish to know whether a halo can form stars before it becomes incorporated into some larger structure. To do this, we construct an artificial model in which the halo has no substructure and cools at constant density. The actual structure of haloes will be highly complex but our model gives a reasonable estimate of what is going on, short of performing a prohibitively time-consuming simulation.

We begin with the smallest haloes and work our way up the merger tree. Each halo is treated as an isolated, isothermal sphere, as described in Section 3.3 of HSTC02. The mean baryon density within the virial radius, $r_{\rm vir}$, is taken to be equal to

$$\rho_{\rm vir} = \left(\frac{\Delta_{\rm c}}{\Omega}\right)\rho_{\rm b0}(1+z_{\rm vir})^3, \quad (4)$$

where $z_{\rm vir}$ is the virialization redshift, $\rho_{\rm b0}$ is the current mean density of baryons in the Universe and $\Delta_{\rm c}$ is the mean overdensity of the virialized halo in units of the critical density at that time, which we take to be $\Delta_{\rm c} \approx 18\pi^2\Omega^{0.45}$ (Eke, Navarro & Frenk 1998).

We define the dynamical time for each halo to be equal to the free-fall time which, in the top-hat model, is

$$t_{\rm dyn} = \frac{1}{4\sqrt{2}}t_{\rm vir}, \quad (5)$$

where $t_{\rm vir}$ is the age of the Universe at the time of virialization.

The virial temperature of the halo is

$$T_{\rm vir} = \frac{\mu m_{\rm H}}{k_{\rm B}}\frac{GM_{\rm tot}}{2r_{\rm vir}} \approx 40.8\frac{\mu}{1.225}(1+z_{\rm vir})\left(\frac{\Delta_{\rm c}h^2\Omega_0}{18\pi^2\Omega}\right)^{1/3}\left(\frac{M_{\rm tot}}{10^5\,{\rm M}_\odot}\right)^{2/3}{\rm K}, \quad (6)$$

where $M_{\rm tot}$ is the total mass (dark plus baryonic), $m_{\rm H}$ is the mass of a hydrogen atom, $k_{\rm B}$ is the Boltzmann constant, G is the gravitational constant and μ is the mean mass of particles in units of $m_{\rm H}$. Note that equation (6) differs slightly from the equivalent expression in HSTC02 as the latter contains a typographical error.

The initial fractional abundance of molecular hydrogen is taken to be 1.1×10^{-6} as calculated by Galli & Palla (1998), and the initial ionization fraction is taken to be the maximum of the equilibrium value at $T_{\rm vir}$ and the residual value from the early Universe, 1.33×10^{-4}.

Starting from these initial conditions, we determine the time that it would take the gas to cool isochorically to $T_{0.75} = 0.75T_{\rm vir}$, using the minimal model presented in Section 2 of HSTC02. This includes molecular hydrogen cooling, collisional excitation and ionization of hydrogen and helium, and inverse Compton cooling from cosmic microwave background photons.

If haloes are unable to cool to $T_{0.75}$ before being swallowed by a more massive halo, then we assume that they are heated to the new virial temperature and that any substructure (which would be minimal anyway because of the long cooling time) is erased. In contrast, haloes that can cool to $T_{0.75}$ are assumed to be able to (instantaneously) cool further to low temperatures and to form a star cluster.

The choice of $T_{0.75}$ is rather arbitrary but it has been given some credence by the simulations of Fuller & Couchman (2000). They showed that collapsing low-mass haloes achieved a critical H_2 fraction of 5×10^{-4}, the same as that found by Tegmark et al. (1997) using the $T_{0.75}$ criterion. Note also that the assumption of isochoric cooling is not critical as the constant-pressure cooling time is typically shorter than the constant-density cooling time by a factor of the order of unity, which depends upon the cooling function and the structure of the halo.

For haloes for which cooling times are much shorter than their dynamical times, then the gas will probably never get heated to the virial temperature in the first place and the assumption of instantaneous star formation will be a good one. However, for haloes in which the cooling time exceeds the dynamical time then it seems likely that our model will underestimate the time taken to form stars. We consider in Section 5.2 the effect of adding a time-delay before star formation and show that it favours the second generation of haloes.

Star clusters may or may not survive subsequent halo mergers, but either way they are assumed to instantly contaminate their surroundings with metals. Thus the primordial star clusters are those that contain no smaller star clusters within them. We assume that the metals do not propagate into haloes on other branches of the tree. Thus metals may be ejected from star clusters but are confined within the next level of the merger hierarchy. This can be justified by a self-regulated model of star formation in which star formation is terminated once gas is expelled from the star cluster.

We also assume that there is no external radiation field, other than that provided by the cosmic microwave background. Primordial star clusters will be surrounded by neutral gas and the propagation of ionizing photons will be severely limited. Nevertheless, these first objects will be highly clustered and so they will at some stage begin to interact with each other. We hope to consider this in a future paper.

3 HALO PROPERTIES

In this section, we present results for the overdense, 3σ, root halo. A comparison with the mean-density root halo will be presented in Section 5.3.

3.1 Collapsed haloes

We start by considering the properties of all collapsed haloes, i.e. haloes for which the linear overdensity exceeds that of all the (more-massive) haloes within which they are contained. In Fig. 1 we plot the virialization redshift of such haloes (drawn from all 100 realizations) against their virial temperature.

The banding comes about because the haloes come in fixed masses. Thus the smallest haloes of mass $9.5 \times 10^4\,{\rm M}_\odot$ correspond to the leftmost band; they have a wide range of virialization redshifts and temperatures that vary between about 200 and 1200 K; the topmost point corresponds to a 6.2σ and the lowest one to a 0.9σ fluctuation. 20 other bands are then visible, one for each factor of 2 in mass until we reach the parent halo in the bottom right-hand corner, which has a mass of $10^{11}\,{\rm M}_\odot$, a virial temperature of $T_{\rm vir}$

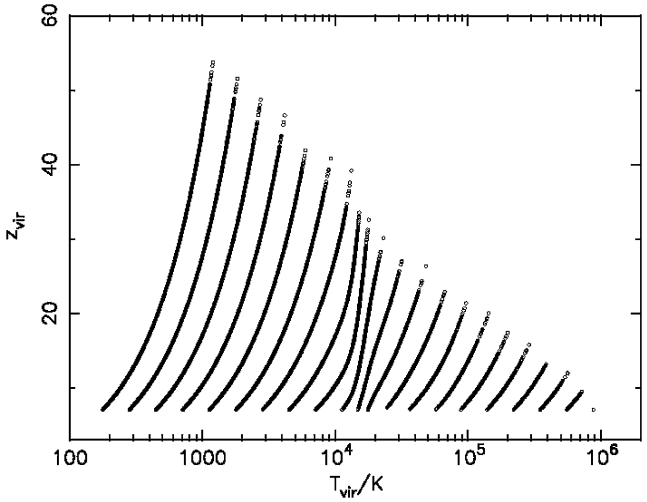

Figure 1. Virialization redshift, $z_{\rm vir}$, versus virial temperature, $T_{\rm vir}$, for all haloes in the 100 realizations.

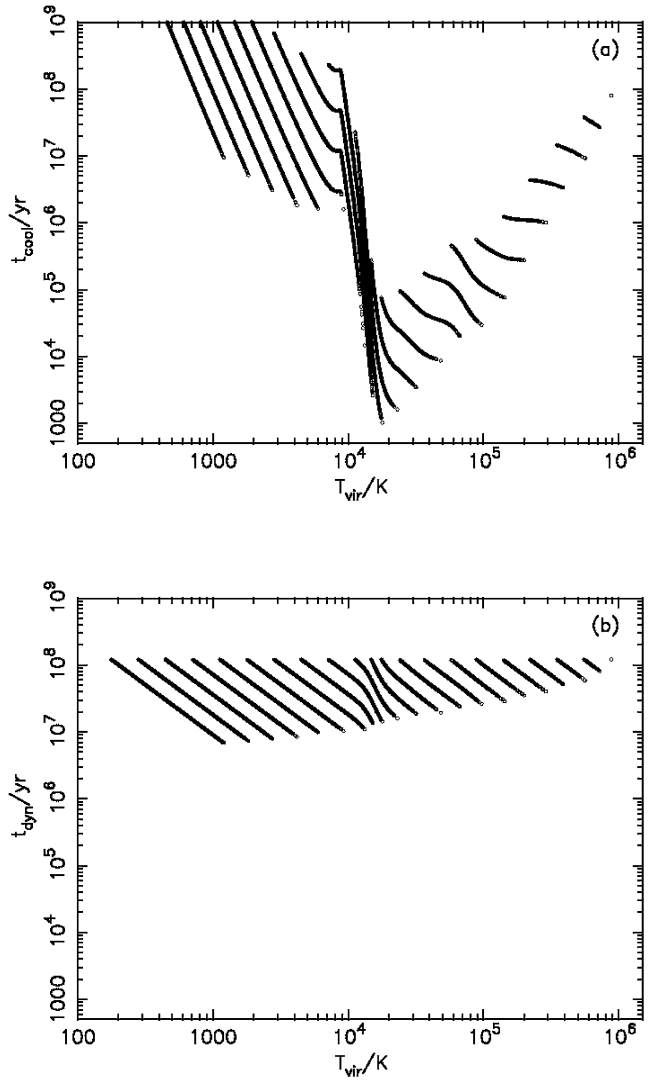

Figure 2. (a) The cooling time, $t_{\rm cool}$, and (b) the dynamical time, $t_{\rm dyn}$, for all collapsed haloes, plotted against virial temperature.

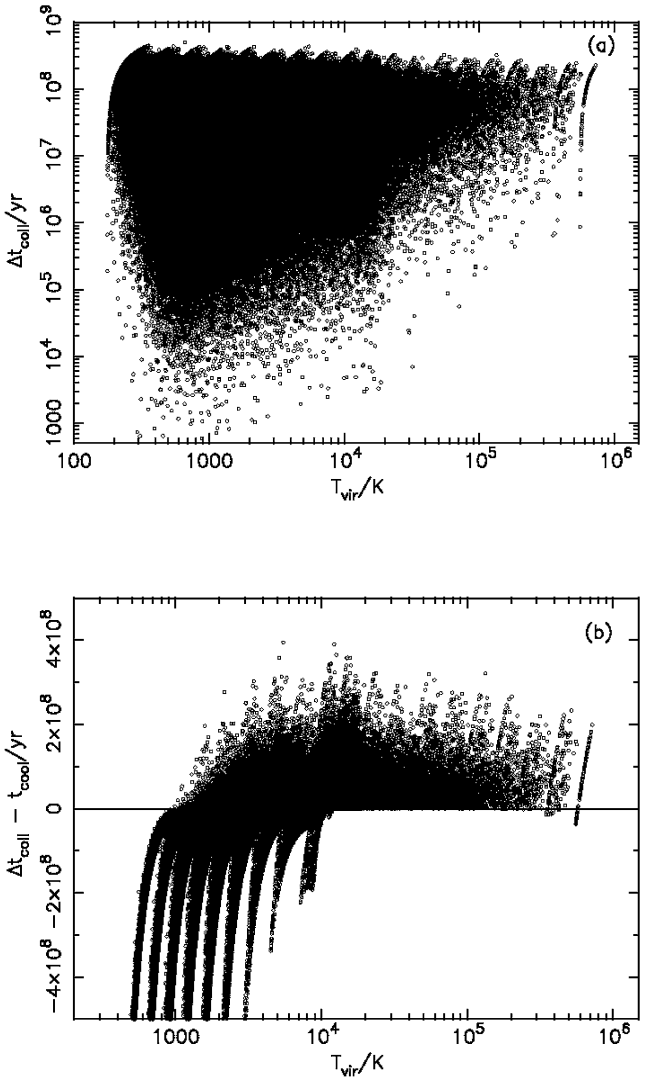

Figure 3. (a) The difference in collapse time of a halo and its parent, $\Delta t_{\rm coll}$, and (b) the difference between the collapse time of a halo and its parent minus the cooling time of the halo, $\Delta t_{\rm coll} - t_{\rm cool}$, versus virial temperature. For clarity, the lower vertical axis in (b) has been truncated at -5×10^8 yr.

$= 8.81 \times 10^5$ K and a virialization redshift of $z_{\rm vir} = 7.1$ (corresponding to a 3σ fluctuation on this scale). The bands are mostly parallel, except for the temperature range $T_{\rm vir} \approx 10\,000$–$20\,000$ K within which the ionization level is changing.

The top panel of Fig. 2 shows the cooling time, $t_{\rm cool}$, versus the virial temperature for each of the haloes shown in Fig. 1. There is a sharp decline in the cooling time at $T_{\rm vir} \approx 10\,000$ K corresponding to the ionization temperature of hydrogen. Haloes with higher virial temperatures than this (Generation 2 haloes) are able to cool rapidly via electronic processes and so have relatively short cooling times. Those with lower virial temperatures (Generation 1) have to rely on cooling via molecular hydrogen, which is formed only in very small quantities. Although our definition of $t_{\rm cool}$ only follows cooling down to $0.75T_{\rm vir}$, we note that Generation 2 haloes have a high residual ionization at lower temperatures that acts as a catalyst to form molecular hydrogen: thus their cooling rates at low temperature are faster than for Generation 1 haloes.

The lower panel of Fig. 2 shows the dynamical time for the collapsed haloes, as defined in equation (5). As a rule of thumb, haloes with virial temperatures above about 10 000 K are able to cool in less than a dynamical time; cooler haloes take longer.

3.2 Star clusters

The time difference, $\Delta t_{\rm coll}$, between the collapse time of each halo and that of its parent is shown in the top panel of Fig. 3. For a random location in space, one might expect that the time difference would be largest for massive, high-temperature haloes. However, that is not the case for these realizations in which the top level itself is constrained to collapse at $t = 0.69$ Gyr (it is true for the haloes considered in Section 5.3 for which the overdensity of the top-level halo is zero).

Those haloes for which $t_{\rm cool}$ is less than $\Delta t_{\rm coll}$ can cool to a fraction 0.75 of their virial temperature before being swallowed up by their parent halo in the merger hierarchy. To begin with, we assume that this is a sufficient criterion to allow them to form stars and we identify them with star clusters. In reality the time-delay before star formation will be larger as the gas has to cool to low temperatures and to congregate into regions of high density. We will consider the effect of allowing a longer time-delay in Section 5.2.

The lower panel in Fig. 3 plots $\Delta t_{\rm coll} - t_{\rm cool}$ against the virial temperature. Those haloes that lie above the line are those that form

star clusters. Just 2 per cent of all collapsed haloes satisfy this condition. However, these are not distributed evenly over mass. For example, in the first three levels of the merger tree (i.e. the three levels with the lowest mass) only a fraction 1.2×10^{-7}, 8.4×10^{-5} and 3.2×10^{-3} of collapsed haloes are able to form stars, whereas a successively higher fraction do so at higher mass. Only for the most massive haloes with virial temperatures in excess of 10^5 K does the cooling time again begin to exceed the lifetime of haloes.

3.3 Primordial star clusters

We assume that metal enrichment from star formation is instantaneous but that it does not extend beyond the immediate environment of a star cluster and its parent halo. Then star clusters of primordial composition are simply those which do not have any smaller star clusters contained within them.

Approximately half of these clusters (56.8 per cent) in our 3σ realizations satisfy this condition with the bias swinging back towards low masses. The most massive primordial star cluster has a mass of $9.76 \times 10^7\,\mathrm{M}_\odot$ and a virial temperature of 1.54×10^4 K.

4 HALO MASS FUNCTIONS

The number of star clusters as a function of virial temperature, averaged over all 100 realizations, is shown in the upper panel of Fig. 4. The upper histogram shows all star clusters, whereas the lower one is restricted to primordial star clusters. There is a clear minimum at about 8600 K corresponding to the division between Generation 1 haloes on the left and Generation 2 haloes on the right. Note that the star clusters that make up the upper histogram are not all independent, i.e. many of the low-temperature clusters are subcomponents of the higher-temperature ones. However, the primordial star clusters are all distinct objects. The y-scale in panel (a) could be multiplied by $3.94h^3$ Mpc^{-3} to convert to a number density but we have not done this as the 3σ region that we consider is not representative of all space.

The lower panel of Fig. 4 shows the same distribution but weighted by mass. From this it is apparent that, whereas the majority of the primordial star clusters in this region are Generation 1 objects, the division of mass between the two generations is much more even, with only about twice as much mass being contained in Generation 1 as compared with Generation 2 objects.

Note the sharp cut-off in the mass density of primordial star clusters at virial temperatures greater than about 15 000 K. This is because all higher-temperature clusters contain a Generation 2 subcluster for which the cooling time is very short and which can itself form stars on a short time-scale. This situation changes when we introduce a time-delay for star formation in Section 5.2.

Fig. 5 is similar to Fig. 4 except that the ordinate is now virial mass rather than virial temperature (to convert to baryonic mass, the x-scale should be multiplied by 0.12). In panel (b), the dotted line shows the contribution to the total mass of Generation 1 haloes (for which the virial temperature is less than 8600 K), while the dot-dashed line is for Generation 2 haloes. The fractional mass contained in the two generations is 0.109 and 0.049, respectively. Thus about 16 per cent of all baryons in this 3σ region of space will have at one time been part of a primordial star cluster.

Fig. 6 contrasts the collapse and star formation redshifts of both generations of halo. The spiky features visible in the distributions are caused by the factor of 2 mass resolution of our haloes and would be smoothed out in a more general merger tree. Because of the long cooling times of Generation 1 haloes, the difference in the peaks of

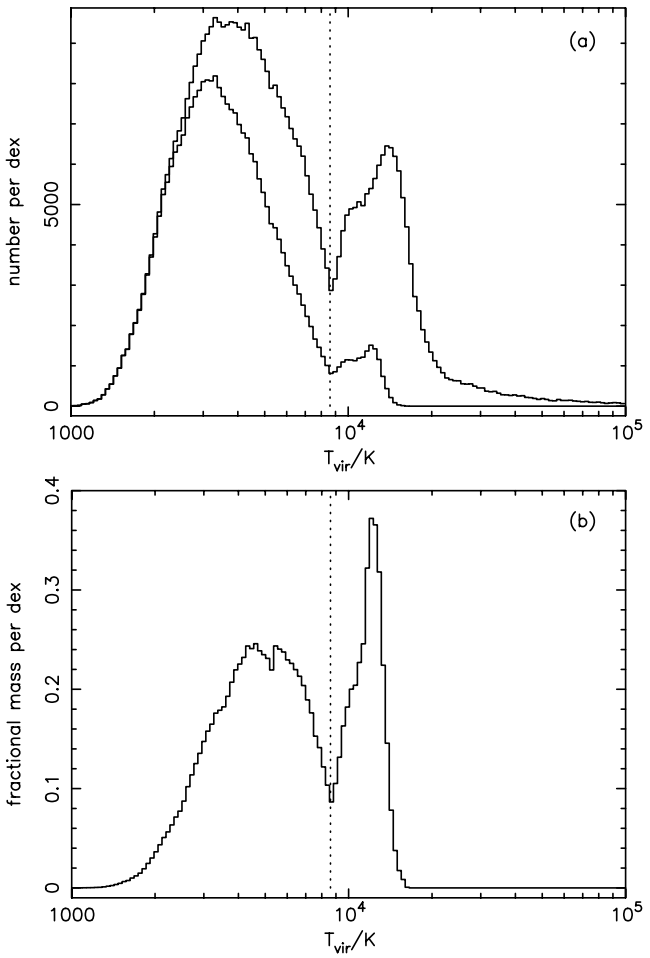

Figure 4. Histograms showing (a) the number and (b) the fractional mass contained in, star clusters as a function of their virial temperature. In (a) the upper line shows all star clusters, whereas the low line is for clusters of primordial composition. Panel (b) shows only primordial haloes. The minimum at $T_{\rm vir} \approx 8600$ K is used to demarcate between the two generations of haloes.

the distributions of star formation redshifts of the two generations is not so great as for their collapse redshifts. Nevertheless, it is clear from the figure that a significant fraction of Generation 1 haloes both collapse and form stars before Generation 2 haloes begin to form in numbers. This highlights the need for a more sophisticated model of feedback than we attempt in this paper.

5 DISCUSSION

5.1 Numerical considerations

It is legitimate to ask to what extent our results are limited by the factor-of-2 mass resolution inherent in the block model. If we could have subhalos with a wider range of masses would that lead to a greater probability of contamination by star formation and a reduction in the fraction of primordial haloes? The dramatic reduction in the cooling time of Generation 2 haloes compared with their low-mass, Generation 1 subhalos, as illustrated by the upper panel in Fig. 2, suggests that this is unlikely to be the case. We expect to move to a more realistic merger tree in future work.

Meanwhile, we have tested the sensitivity of our results to the precise choice of halo masses by performing a second 100 realizations

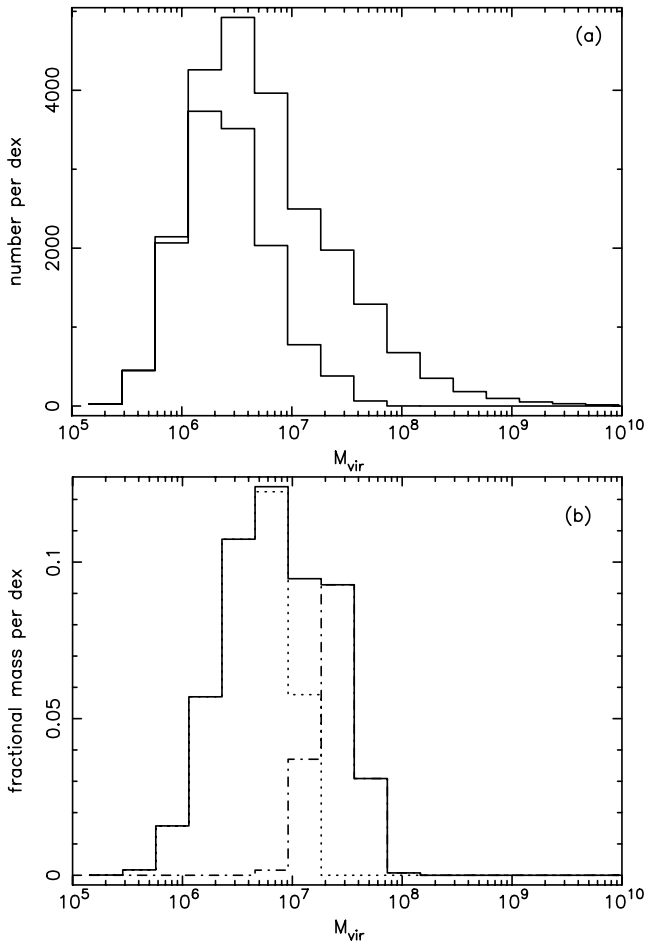

Figure 5. Histograms showing (a) the number and (b) the mass fraction, of star clusters as a function of their virial mass. In (a) the upper line shows all star clusters, whereas the low line is for clusters of primordial composition. Panel (b) shows only primordial clusters. The dotted line corresponds to Generation 1 haloes ($T_{\rm vir} < 8600$ K) and the dotted-dashed line correspond to Generation 2 haloes.

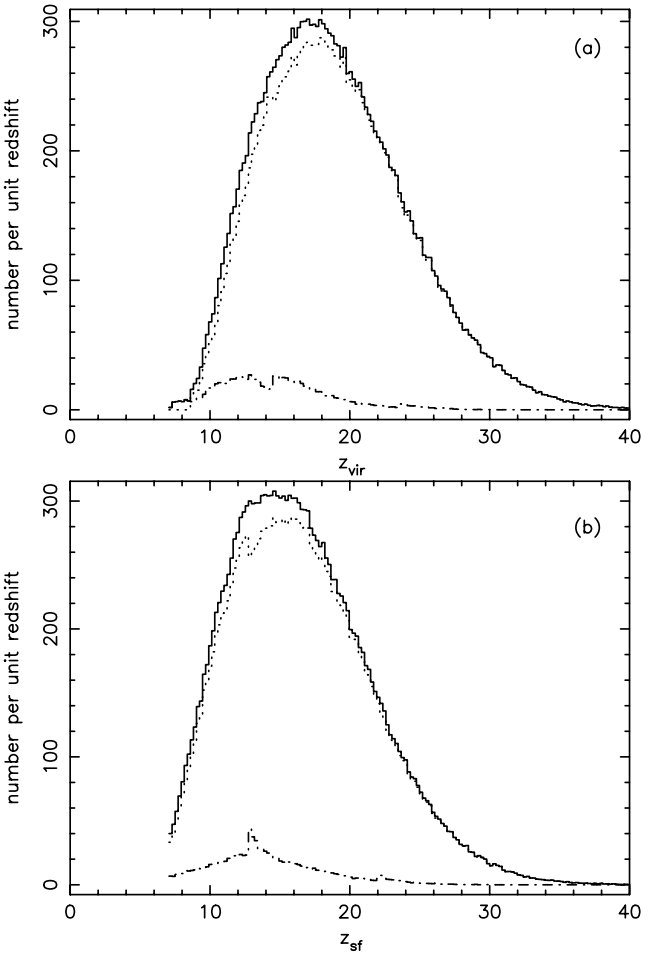

Figure 6. A histogram showing the number of primordial star clusters as a function of (a) their virialization redshift and (b) their star formation redshift. The dash-dotted and dotted lines correspond to Generation 1 and Generation 2 haloes, respectively.

of the merger tree with the root halo mass (and hence each level of the merger hierarchy) increased by a factor of $\sqrt{2}$, to $1.4 \times 10^{11}\,{\rm M}_\odot$. Fig. 7 shows the mass function of primordial haloes for these simulations contrasted with our original simulation (dotted line). There is no significant difference between the two.

The results that we have presented so far are an average over a large number of realizations. Fig. 8 shows the dispersion around the average, of five of the 100 realizations carried out in this paper. It can be seen that the scatter is significant but not enough to seriously affect the divide between the two generations of star clusters within each realization.

5.2 Time-delayed star formation

So far we have assumed that after the cooling of the gas to low temperatures (following the T97 criterion) stars form instantaneously. In reality, there will be a lapse of time until the gas reaches the high-density regime in which nuclear reactions can take place and the stars are born. In an attempt to include in our code a time delay between initial cooling and star formation, we consider in this section the effect of adding the dynamical time to the cooling time. The justification for this is simply that, following virialization, one would

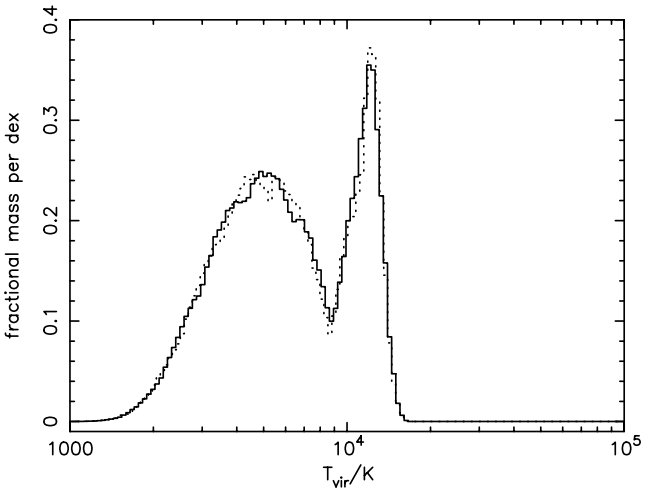

Figure 7. Histogram showing the mass fraction of primordial haloes for two different choices of halo mass. The dotted line is the same as in Fig. 4(b) while the solid line shows the mass fraction for a merger tree in which the mass of the root halo has been increased a factor of $\sqrt{2}$.

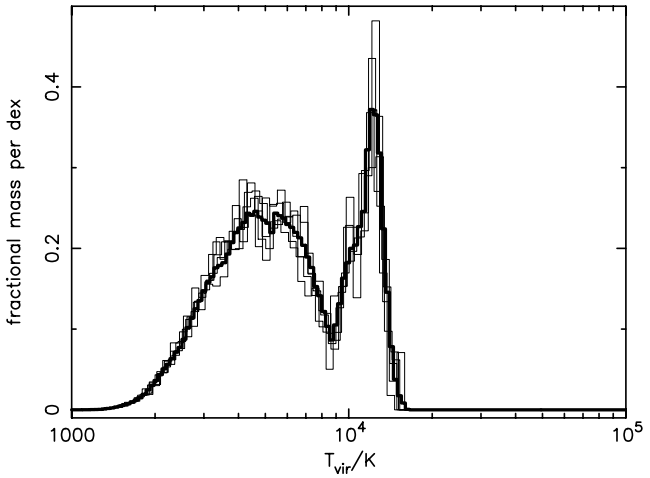

Figure 8. Histogram showing the effect of different realizations on the mass fraction of primordial objects as a function of virial temperature. The thick line corresponds to an average over 100 realizations.

expect the gas to take at least a dynamical time to contract within the potential well of the halo (this argument has less force for high-mass haloes for which the cooling time is very much shorter than their dynamical time and which may therefore never attain virial equilibrium in the first place). Some evidence for this delayed star formation comes from Abel et al. (2002) who describe the formation of a primordial star using a three-dimensional hydrodynamical simulation. Their results show that a cooled ($\sim$200 K) high-redshift molecular cloud is formed at $z = 24$ and then a protostar is formed at $z = 18.2$. The time that took to form this protostar is of the order of the dynamical time of the cloud.

Fig. 9(a) shows a histogram of the fractional mass of primordial star clusters as a function of their virial temperatures. The dotted line corresponds to the original case in which no time delay has been added, while the full line shows the distribution when a dynamical time has been added to their cooling time. The fractional mass has changed in such a way that now we have three times more mass in Generation 2 haloes than in Generation 1 haloes, and a greater total mass fraction than before. Note also that there is no longer a sharp cut-off at virial temperatures above 15 000 K because it is possible for subhalos to have short cooling times in this model and yet not to form stars.

Fig. 10(a) shows the equivalent mass function of these clusters from which it can be seen that haloes as massive as $10^{10}\,M_\odot$ can contain primordial star clusters. While this does not seen very likely, the general conclusion that we draw is that delayed, rather than instantaneous, star formation will favour Generation 2 haloes over Generation 1.

We also tried a model with instantaneous star formation but with a time delay before energy and metallicity feedback. The idea is that if the time difference between the collapse of a parent halo and the cooling of its child is less than the time for the formation of supernovae, $\sim 10^7$ yr, then the parent halo will be of primordial composition and has to be added into the set of primordial objects. However, this makes only a minor difference to our results and we will not discuss it further.

5.3 Mean-density regions

The first star clusters will form in overdense regions of the Universe, hence our use of a 3σ root halo to this point. However, it

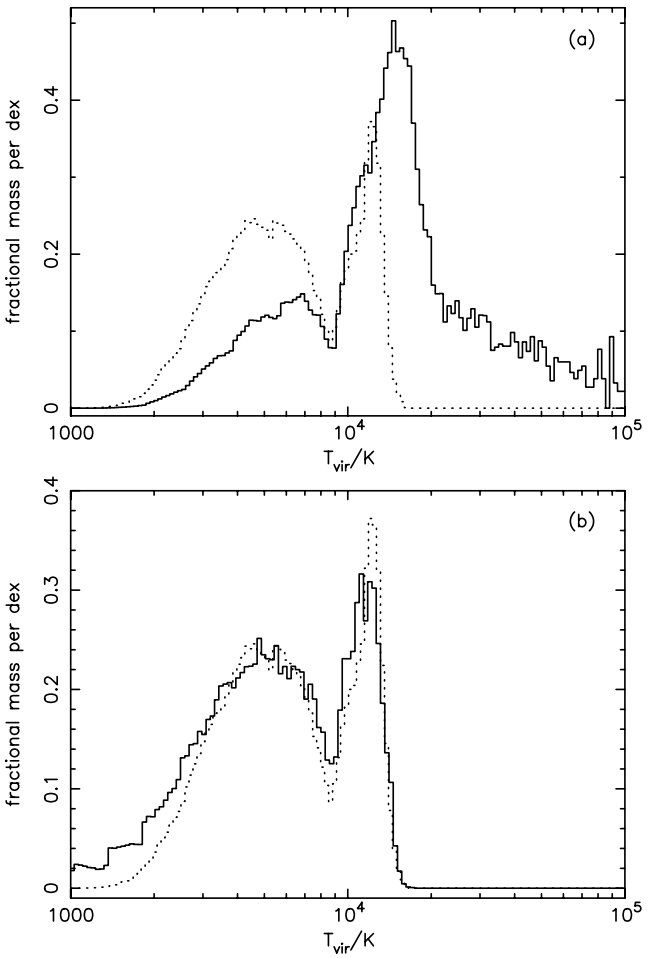

Figure 9. Histograms showing the mass fraction of primordial star clusters as a function of virial temperature. In (a) the solid line shows the mass fraction of haloes for which a dynamical time has been added to their cooling time. In (b) the solid line shows the mass fraction for a zero-overdensity top-level merger tree. In both panels the dotted line is the same as in Fig. 4(b).

is interesting to contrast these results with those expected for a more typical part of the Universe, with density equal to the cosmic mean.

Fig. 11 contrasts the collapse redshifts of primordial star clusters for the 3σ and mean-density regions. It can be seen that the haloes collapse at much lower redshifts in the mean overdensity case. Because the mean-density regions are likely to be far removed from the regions of the first star formation, they are unlikely to be affected by photoionizing photons at high redshift. However, observations (e.g. (Fan et al. 2002) and simulations (e.g. (Razoumov et al. 2002) both suggest that the Universe became re-ionized at a redshift of about 6 and our model will be invalid after this time. This will mostly affect the evolution of Generation 2 haloes.

The effect of the lower formation redshift on the virial temperatures and masses of the star clusters is shown in the lower panels of Figs 9 and 10. The fractional mass distribution over virial temperature is almost unchanged, but with a slight bias towards lower temperatures compared with the 3σ case. A greater effect is a shift in the mass function towards higher masses: because the haloes collapse at lower redshift and hence have lower densities, they have higher masses for a given virial temperature.

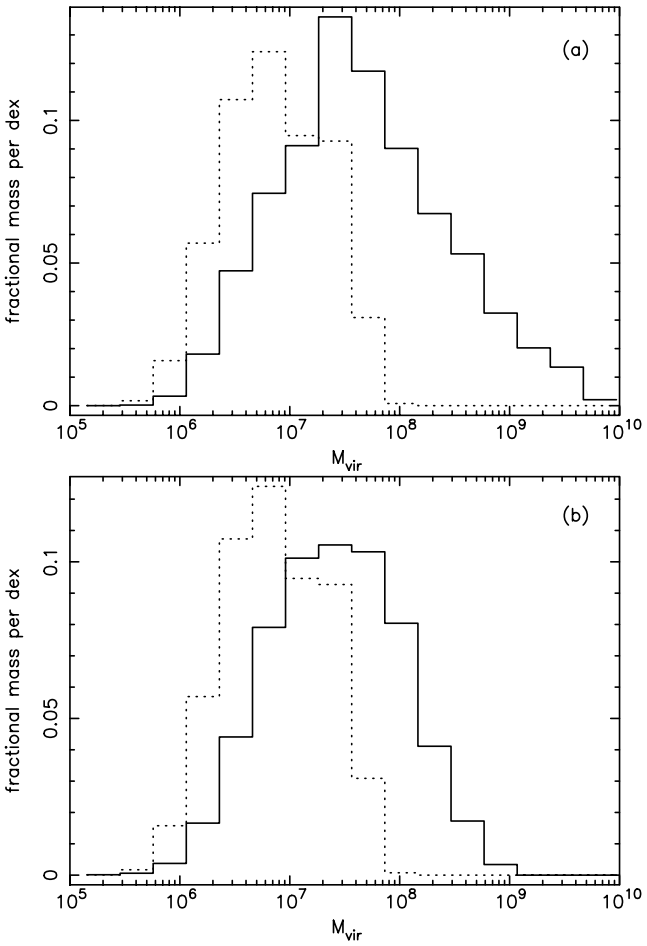

Figure 10. Histograms showing the mass fraction of primordial haloes as a function of their virial mass. The dotted line in both panels is the same as in Fig. 5(b). In (a) the solid line shows the mass fraction of haloes for which a dynamical time have been added to their cooling time. In (b) the solid line shows the mass fraction for a zero-overdensity top-level merger tree.

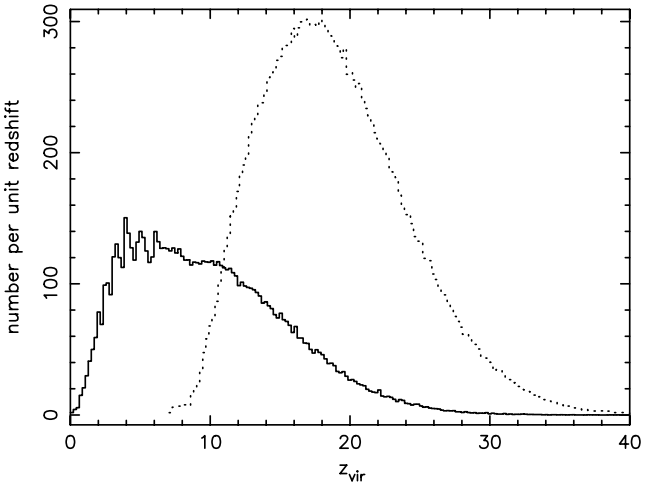

Figure 11. Histogram showing the distribution of collapse redshifts for primordial star clusters. The solid line is for a mean-density and the dotted line for the previously investigated 3σ regions.

5.4 What and where are they now?

Our model predicts the masses, virial temperatures and formation redshifts of primordial star clusters, but says nothing concerning their internal structure. Hydrodynamical simulations (see the references in the introduction) have made a start in this direction but are as yet still in their infancy. There has been some theoretical speculation concerning the masses of the first stars but no consensus has emerged. In this section, we use our results to discuss two possible fates of primordial star clusters, but note that the physics is sufficiently uncertain that we may even have got the roles of the two generations mixed up.

The baryonic mass of our root haloes, $1.2 \times 10^{10}\,\mathrm{M}_\odot$, is similar to that of a normal galaxy of mass approximately one-tenth that of an L_* galaxy. The space density for the 3σ fluctuations on this scale is $3.0 \times 10^3 h^3\ \mathrm{Mpc}^{-3}$, similar to that of groups of galaxies, so that we would perhaps expect one such galaxy in a typical group. The other galaxies will form slightly later and so the star clusters will be biased to higher masses, although the total mass contained in primordial star clusters will be similar (see Section 5.3). Our model therefore suggests that approximately one-tenth of the baryons in a typical galaxy will have passed through a primordial star cluster. The majority of these are probably enriched with processed material without themselves forming a zero-metallicity star.

The majority of Population III stars will be born in regions that are destined to end up in normal galaxies. However, our model does not preclude the formation of some zero-metallicity stars in low-density regions of the Universe at relatively low redshift (but before re-ionization). The resulting star clusters would be of low density and therefore very hard to detect.

The bulk of Generation 1 star clusters at high redshift have masses in the range 10^6–$10^7\,\mathrm{M}_\odot$; the baryonic mass is lower, 10^5–$10^6\,\mathrm{M}_\odot$. It is natural, therefore, to identify these objects with the low-metallicity globular clusters found in the bulges and haloes of normal galaxies. The relatively long cooling times of Generation 1 haloes compared with their dynamical times would have aided dissipative collapse within the dark matter halo and survivability of the star cluster. One objection to this is that zero-metallicity stars have not been discovered in globulars, but of course the first stars may have been of high mass and may have burnt out long ago. A more serious objection is that we know that the amount of material in globular clusters is much less than one-tenth of all the baryons in a galaxy. It is therefore probable that feedback of energy from the first supernovae (or hypernovae) will disrupt the star clusters and that the majority of zero-metallicity stars, should they still exist, will be spread throughout the bulges of normal galaxies.

The main differences with Generation 2 clusters is that they are more massive by about a factor of 10 and that their cooling times are much shorter. It is interesting to speculate that electronic cooling when it did turn on would lead to catastrophic accumulation of cold gas at the centre of the collapsing halo and perhaps to the formation of a massive black hole. Observations (e.g. (Ferrarese & Merritt 2000; Gebhardt et al. 2000) give a galactic black hole to bulge mass ratio of about 0.001. To be consistent with this would require an accretion efficiency of just 2 per cent, creating seed holes of mass 2×10^4–$2 \times 10^5\,\mathrm{M}_\odot$. Subsequent merging of these seed holes could lead to the formation of supermassive black holes in the centres of normal galaxies today.

ACKNOWLEDGMENTS

FS thanks his parents for their support; PAT is a PPARC Lecturer Fellow.

REFERENCES

Abel T., Anninos P., Norman M.L., Zhang Y., 1998, ApJ, 508, 518
Abel T., Bryan G.L., Norman M.L., 2000, ApJ, 540, 39
Abel T., Bryan G.L., Norman M.L., 2002, Sci, 295, 93
Allen S.W., Schmidt R.W., Fabian A., Ebeling H., 2002, Cosmological constraints from the local X-ray luminosity function of the most X-ray luminous galaxy clusters, MNRAS, submitted (astro-ph/0208394)
Bromm V., Coppi P.S., Larson R.B., 1999, ApJ, 527, L5
Bromm V., Coppi P.S., Larson R.B., 2002, ApJ, 564, 23
Cole S., Kaiser N., 1988, MNRAS, 233, 637
Eke V.R., Cole S., Frenk C.S., 1996, MNRAS, 282, 263
Eke V.R., Navarro J.F., Frenk C.S., 1998, ApJ, 503, 569
Fan X., Narayanan V.K., Strauss M.A., White R.L., Becker R.H., Pentericci L., Rix H., 2002, AJ, 123, 1247
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Fuller T.M., Couchman H.M.P., 2000, ApJ, 544, 6
Galli D., Palla F., 1998, A&A, 335, 403
Gebhardt K. et al., 2000, ApJ, 539, L13
Hutchings R.M., Santoro F., Thomas P.A., Couchman H.M.P., 2002, MNRAS, 330, 927 (HSTC02)
Lokas E.L., Hoffman Y., 2001, in Identification of Dark Matter, The spherical collapse model in a universe with cosmological constant. World Scientific, Singapore, p. 121
Nakamura F., Umemura M., 2001, ApJ, 548, 19
Nakamura F., Umemura M., 2002, ApJ, 569, 549
Razoumov A.O., Norman M.L., Abel T., Scott D., 2002, ApJ, 572, 695
Seljak U., 2002, 337, 769
Tegmark M., Silk J., Rees M.J., Blanchard A., Abel T., Palla F., 1997, APJ, 474, 1 (T97)
Viana P.T.P., Liddle A.R., 1996, MNRAS, 281, 323

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