Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **416,** 232–241 (2011) doi:10.1111/j.1365-2966.2011.19025.x

The non-linear evolution of baryonic overdensities in the early universe: initial conditions of numerical simulations

Smadar Naoz,[1,2★] Naoki Yoshida[3] and Rennan Barkana[2]

[1]*CIERA, Northwestern University, Evanston, IL 60208, USA*

[2]*Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel*

[3]*IPMU, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8583, Japan*

Accepted 2011 May 7. Received 2011 April 11; in original form 2010 September 4

ABSTRACT

We run very large cosmological *N*-body hydrodynamical simulations in order to study statistically the baryon fractions in early dark matter haloes. We critically examine how differences in the initial conditions affect the gas fraction in the redshift range $z = 11$–21. We test three different linear power spectra for the initial conditions. (1) A complete heating model, which is our fiducial model; this model follows the evolution of overdensities correctly, according to Naoz & Barkana (2005), in particular including the spatial variation of the speed of sound of the gas due to Compton heating from the CMB. (2) An equal-δ model, which assumes that the initial baryon fluctuations are equal to those of the dark matter, while conserving σ_8 of the total matter. (3) A mean $c_{\rm s}$ model, which assumes a uniform speed of sound of the gas. The latter two models are often used in the literature. We calculate the baryon fractions for a large sample of haloes in our simulations. Our fiducial model implies that before reionization and significant stellar heating took place, the minimum mass needed for a minihalo to keep most of its baryons throughout its formation was $\sim 3 \times 10^4$ M$_{\odot}$. However, the alternative models yield a wrong (higher by about 50 per cent) minimum mass, since the system retains a memory of the initial conditions. We also demonstrate this using the 'filtering mass' from linear theory, which accurately describes the evolution of the baryon fraction throughout the simulated redshift range.

Key words: galaxies: haloes – galaxies: high-redshift – dark ages, reionization, first stars.

1 INTRODUCTION

Recent measurements of anisotropies of the cosmic microwave background (CMB) radiation have revealed the detailed distribution of matter in the Universe a few hundred thousand years after the big bang (Spergel et al. 2003, 2007; Komatsu et al. 2009, 2011). Observations utilizing large ground-based telescopes and space telescopes have discovered galaxies and black holes that were in place when the age of the Universe was less than a billion years. Moreover, many galaxies have been found at $z > 7$ (Bouwens et al. 2010; McLure et al. 2010) in the *Hubble Ultra Deep Field*, whereas already a few γ-ray bursts at $z > 6$ have been detected by the *Swift* satellite (Salvaterra et al. 2009; Tanvir et al. 2009; Lin, Liang & Zhang 2010). These first objects are probably the building blocks of the present day galaxies, thus, solving the puzzle behind their formation will have a profound implication on our understanding of the Universe (see for recent reviews Bromm et al. 2009; Yoshida 2010, and references therein).

The formation of the first generation of galaxies in the Universe has been studied for many years. High-resolution cosmological simulations can follow complex astrophysical processes, while analytical calculations can provide an overall understanding, and can be used to decouple different physical effects seen in simulations. Analytic models are also useful for estimating the limitations of numerical simulations such as insufficient resolution and small boxsizes (Yoshida et al. 2003; Barkana & Loeb 2004; Naoz & Barkana 2005). Combining the two approaches may offer many of the advantages of both.

The initial conditions (hereafter ICs) in a cosmological simulation can have a large effect on the formation of the first galaxies in simulations, i.e. both on the formation time (or on the halo abundance at a given time) and the halo properties at formation time (such as the average gas fraction). Yoshida et al. (2003b) studied high-redshift structure formation and reionization while testing two different models for power spectra as their ICs. They found that different models have a profound effect on the abundance of primordial star-forming gas clouds and thus on when the reionization was initiated and its progress. In the analytical point of view, Naoz, Noter & Barkana (2006) and Naoz & Barkana (2007) showed that

★E-mail: snaoz@northwestern.edu

the ICs at high redshift have a significant effect on the halo abundance and the gas fraction at virialization. While these effects are largest at the highest redshift, e.g. $z \sim 65$ for the first star in the Universe (Naoz et al. 2006), they are still significant for haloes forming at $z \sim 10$–30. The first gas rich haloes at these redshifts are expected to host the first stars ($z \sim 65$–30; Naoz, Noter & Barkana 2006; Yoshida 2006; Gao et al. 2007; Trenti, Stiavelli & Michael Shull 2009a) and even the first γ-ray bursts (e.g. Bromm & Loeb 2006; Naoz & Bromberg 2007). Thus, investigating the formation properties of these haloes is of prime importance.

Gas rich haloes in the early Universe may very well be a nurturing ground for dwarf galaxies, which at high redshift can form stars (e.g. Bromm, Coppi & Larson 2002, 1999; Abel, Bryan & Norman 2002; Yoshida et al. 2006; Yoshida, Omukai & Hernquist 2008, and references therein) perhaps even at a high star formation rate (Ricotti, Gnedin & Shull 2002; Clark et al. 2011; Greif et al. 2010). Their properties are very important as they are responsible for metal pollution and the ionizing radiation at these early times (e.g. Shapiro, Iliev & Raga 2004; Ciardi et al. 2006; Gnedin, Kravtsov & Chen 2008; Trenti & Stiavelli 2009). Moreover, haloes that are too small for efficient cooling via atomic hydrogen, i.e. minihaloes, are most susceptible to the effect of ICs. While they may not normally host astrophysical sources, minihaloes may produce a 21-cm signature (Kuhlen, Madau & Montgomery 2006; Shapiro et al. 2006; Naoz & Barkana 2008 but see Furlanetto & Oh 2006), and they can block ionizing radiation and produce an overall delay in the initial progress of reionization (e.g. Barkana & Loeb 2002; Iliev et al. 2003; Iliev, Scannapieco & Shapiro 2005; McQuinn et al. 2007). The evolution of the halo gas fraction at various epochs of the Universe is of prime importance, particularly in the early Universe.

In this paper, we examine the effect of using different initial conditions in simulations on the resulting minimum gas-rich halo mass in the redshift regime $z = 11$–21. We perform GADGET-2 (Springel, Yoshida & White 2001; Springel 2005) simulations using a total of $768^3 \times 2$ particles. We compare the ICs presented in Naoz & Barkana (2005), which describe the linear evolution of overdensities in a fully consistent way, to two other alternative ICs, often used in the literature. We also compare to the prediction of the gas-rich mass from linear theory. We describe our different ICs and simulations in Sections 2 and 3, respectively. Our simulation results are presented in Section 4 where we divide our discussion to the evolution of the non-linear power spectra (Section 4.1) and to the minimum gas-rich halo mass resulting from either linear theory or from the simulations (Section 4.2). Finally, we discuss our conclusions (Section 5).

Throughout this paper, we adopt the following cosmological parameters: $(\Omega_\Lambda, \Omega_M, \Omega_b, n, \sigma_8, H_0)$ = (0.72, 0.28, 0.046, 1, 0.82, $70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$) (Komatsu et al. 2009).

2 DIFFERENT INITIAL CONDITION MODELS – BASIC EQUATIONS

2.1 The fiducial ICs – 'fid'

We follow Naoz & Barkana (2005), who studied the linear evolution of both dark matter (DM) and baryon overdensities. The fluctuations of the temperature of the baryons (δ_T) cannot be described as a simple function of a spatially uniform baryonic sound speed $c_s(t)$, as was previously assumed (e.g. Ma & Bertschinger 1995). Furthermore, at high redshifts, the baryon density fluctuations (δ_b) are not equal to those of DM (δ_{dm}) (contrary to a common assumption in simulations; four redshift examples are shown in fig. 1 of Naoz & Barkana 2005). We label the power spectrum model as the 'fid'

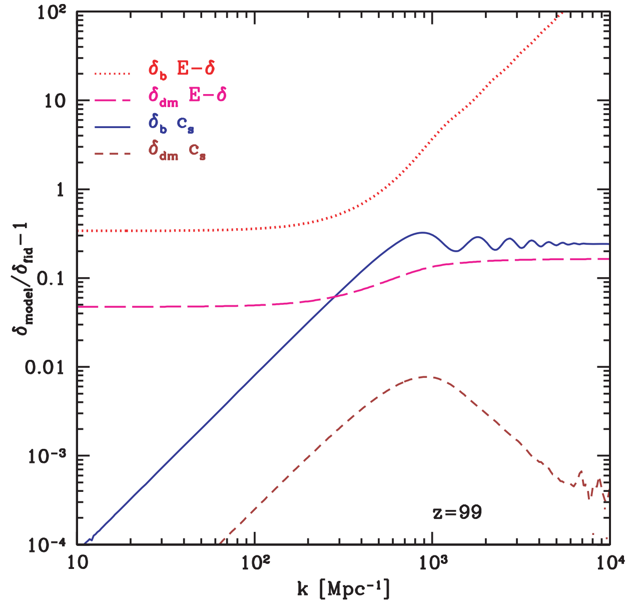

Figure 1. The relative difference (specifically, $\delta_{model}/\delta_{ch} - 1$) between the fiducial linear initial conditions and the alternative models at $z = 99$. We consider the relative difference between the fid ICs and the mean c_s ICs for both the baryons and dark matter (solid and short-dashed curves, respectively), and the relative difference between the fid ICs and the E-δ ICs for both the baryons and dark matter (dotted and long-dashed curves, respectively). Note that we have plotted here the absolute value; the mean c_s model gave a negative value (i.e. an underestimate compared to the fid model) while the E-δ model gave a positive value (i.e. an overestimate).

(fiducial) ICs since it follows the evolution of linear overdensities in a complete and consistent way.

Following Naoz & Barkana (2005) we write the basic equations that describe the evolution of the DM, baryon density and temperature fluctuations:

$$\ddot{\delta}_{dm} + 2H\dot{\delta}_{dm} = \frac{3}{2}H_0^2\frac{\Omega_m}{a^3}\left(f_b\delta_b + f_{dm}\delta_{dm}\right), \tag{1}$$

where f_{dm} and f_b are the mean cosmic DM and baryonic fraction, respectively. Here we follow the standard notations for cosmological parameters such as Ω_m, H_0. The baryons are also subject to a pressure term:

$$\ddot{\delta}_b + 2H\dot{\delta}_b = \frac{3}{2}H_0^2\frac{\Omega_m}{a^3}\left(f_b\delta_b + f_{dm}\delta_{dm}\right) - \frac{k^2}{a^2}\frac{k_B\bar{T}}{\mu}\left(\delta_b + \delta_T\right), \tag{2}$$

where μ is the mean molecular weight, k_B is the Boltzmann constant and k is the wavenumber. Using the first law of thermodynamics, Naoz & Barkana (2005) derived the equations for the evolution of the baryon average temperature and temperature fluctuations:

$$\frac{d\bar{T}}{dt} = -2H\bar{T} + \frac{x_e(t)}{t_\gamma}\left(\bar{T}_\gamma - \bar{T}\right)a^{-4}, \tag{3}$$

where $\bar{T}_\gamma = [2.725\ \mathrm{K}]/a$ is the mean CMB temperature, and the first-order equation for the perturbation:

$$\frac{d\delta_T}{dt} = \frac{2}{3}\frac{d\delta_b}{dt} + \frac{x_e(t)}{t_\gamma}a^{-4}\left\{\delta_\gamma\left(\frac{\bar{T}_\gamma}{\bar{T}} - 1\right) + \frac{\bar{T}_\gamma}{\bar{T}}\left(\delta_{T_\gamma} - \delta_T\right)\right\}, \tag{4}$$

with the second term on the right-hand side accounting for the Compton scattering of the CMB photons on the residual electrons from recombination, where $x_e(t)$ is the electron fraction out of the total number density of gas particles at time t, and

$$t_\gamma^{-1} \equiv \frac{8}{3}\bar{\rho}_\gamma^0 \frac{\sigma_{\rm T}\, c}{m_{\rm e}} = 8.55 \times 10^{-13}\ {\rm yr}^{-1}, \tag{5}$$

where $\sigma_{\rm T}$ is the Thomson scattering cross-section and ρ_γ is the photon energy density. The first term on the right-hand side of each of these two equations (3) and (4) accounts for adiabatic expansion of the gas, and the remaining terms capture the effect of the thermal exchange with the CMB. Following Naoz & Barkana (2005) we have numerically calculated the evolution of the perturbations by modifying the CMBFAST code (Seljak & Zaldarriaga 1996) according to these equations. Note that similar physics was also explored by Yamamoto, Sugiyama & Sato (1997, 1998).

We solve the complete set of equations to obtain the power spectrum at different redshifts which can be used as initial conditions for our simulations. Fig. 1 shows the ratio between this initial condition to the two alternative models tested in this paper.

2.2 Alternative model I – equal δ – 'E-δ'

In many cosmological ICs for N-body simulations and semi-analytical calculations, the fluctuations of the baryons are assumed to be equal to the fluctuations of the DM. We construct a model that includes this incorrect assumption while maintaining the correct overall $\delta_{\rm tot}$ (i.e. conserving σ_8 at $z = 0$, see Appendix A for more details). Thus, in our 'E-δ' model we calculate the correct $\delta_{\rm tot}$ as a combination of $\delta_{\rm b}$ and $\delta_{\rm dm}$ from the fiducial calculation in Section (2.1), but then take the baryon perturbation to be the same as for the DM, namely

$$\delta_{\rm b}^{E\delta} = \delta_{\rm dm}^{E\delta} = \delta_{\rm tot}^{\rm ch} = f_{\rm b}\delta_{\rm b}^{\rm ch} + f_{\rm dm}\delta_{\rm dm}^{\rm ch}, \tag{6}$$

where $\delta_{\rm b,dm}^{\rm ch}$ ($\delta_{\rm b,dm}^{E\delta}$) is the resulting linear overdensity from the fiducial calculation (E-δ model) for the baryons and DM, respectively. We then compare the equal δ model to our fiducial calculation. Fig. 1 shows the ratio between the fid ICs and the E-δ model for both the baryons and DM. We find that the E-δ model overestimates the baryon fluctuations by $\gtrsim$30 per cent on large scales ($k^{-1} \gtrsim$ 10 kpc) while the overestimate grows to a much larger factor on small scales.

Before recombination the baryons were tightly coupled to the radiation, resulting in suppression of the growth of their overdensity. However, the DM component, which is not affected by the photons, could basically grow once the fluctuation wavelength entered the Hubble horizon (in the linear regime, before equality, the dark mater fluctuations grew logarithmically with the scale factor, where after equality they grew linearly with the scale factor). Therefore, this resulted in a suppression of the baryonic overdensity by about three orders of magnitude compared to the DM at recombination (e.g. fig. 1 in Naoz & Barkana 2005). While the baryons subsequently fall into the potential wells of the DM, it takes them some time to catch up, and the baryon fluctuations are still suppressed even at lower redshifts. This point is often overlooked in simulations and analytical calculations.

2.3 Alternative model II – the mean sound speed approximation – 'mean cs'

Naoz & Barkana (2005) showed that the presence of spatial fluctuations in the sound speed modifies the calculation of perturbation growth significantly. Nevertheless, for completeness and as a case of comparison with previous results, we compare the simulation results with the results obtained using this approximation. Thus, we proceed by presenting the basic equations of the growth of density fluctuations, in this approximation of a uniform sound speed (hereafter 'mean $c_{\rm s}$'). The evolution of the density fluctuations is described by a different set of coupled second order differential equations:

$$\ddot{\delta}_{\rm dm} + 2H\dot{\delta}_{\rm dm} = \frac{3}{2}H_0^2\frac{\Omega_{\rm m}}{a^3}\left(f_{\rm b}\delta_{\rm b} + f_{\rm dm}\delta_{\rm dm}\right),$$
$$\ddot{\delta}_{\rm b} + 2H\dot{\delta}_{\rm b} = \frac{3}{2}H_0^2\frac{\Omega_{\rm m}}{a^3}\left(f_{\rm b}\delta_{\rm b} + f_{\rm dm}\delta_{\rm dm}\right) - \frac{k^2}{a^2}c_{\rm s}^2\delta_{\rm b}, \tag{7}$$

where $c_{\rm s}^2 = {\rm d}p/{\rm d}\rho$ is assumed to be spatially uniform (i.e. independent of k) and is thus calculated from the thermal evolution of a uniform gas undergoing Hubble expansion. With this assumption, the temperature fluctuations (as a function of k) are simply proportional at any given time to the gas density fluctuations:

$$\frac{\delta_{\rm T}}{\delta_{\rm b}} = \frac{c_{\rm s}^2}{k_{\rm B}\bar{T}/\mu} - 1. \tag{8}$$

Naoz & Barkana (2005) showed that this approximation leads to an underestimation of the baryon density fluctuations by up to 30 per cent at $z = 100$ and 10 per cent at $z = 20$ for large wavenumbers. Fig. 1 shows the ratio between the mean $c_{\rm s}$ initial conditions and the fiducial ones for both the baryons and DM. It agrees with our previous results, showing that the underestimate by the mean $c_{\rm s}$ model is greatest at $k^{-1} \sim 1$ kpc. The non-linear evolution resulting from these initial conditions will result in shallower potential wells compared to the fiducial calculation.

Even though it is clear that the precise baryon temperature fluctuations at high redshift are very significant, still many simulations use initial conditions that assume a uniform speed of sound in the Universe. As shown below this leads to significantly different estimates for the gas content of the early haloes.

3 THE SIMULATION

3.1 Basic parameters

We run a GADGET-2 simulation (Springel et al. 2001; Springel 2005) starting from redshift 99, for a total of 2×768^3 particles (768^3 particles each for the DM and baryon components) and our box size is 2 Mpc. We choose this box size so that a halo mass of 10^5 M$_\odot$ would have ~500 particles. This way according to Naoz, Barkana & Mesinger (2009) we are able to estimate the gas fraction in ~10^5 M$_\odot$ haloes correctly (see below for the halo definition). Our softening length is 0.2 comoving kpc.

For all runs, glass-like cosmological ICs were generated using the Zel'dovich approximation. The transfer functions were generated using the various models described above. We have used a glass file which was randomly displaced thus removing the coupling between nearby DM and gas particles. Using this randomization procedure we achieve essentially the same effect to that shown in Yoshida, Sugiyama & Hernquist (2003c). In generating the ICs, a convolution between the glass file and the transfer function from the different models was done, thus taking into account the different velocities of the DM and baryons (for the fiducial and mean cs models). We note that we have used the same phases for the DM and baryons, in all of the simulations.

We set the initial temperature to be 164.11 K (as derived from linear theory), and thus Gadget assumes neutral and monoatomic gas, and converts to thermal energy (i.e. adiabatic initial conditions). Although this work emphasizes the need for a precise calculation of the baryon overdensities resulting from temperature fluctuations, we actually neglect the temperature fluctuations in the initial conditions. This may not be a bad approximation since the haloes we

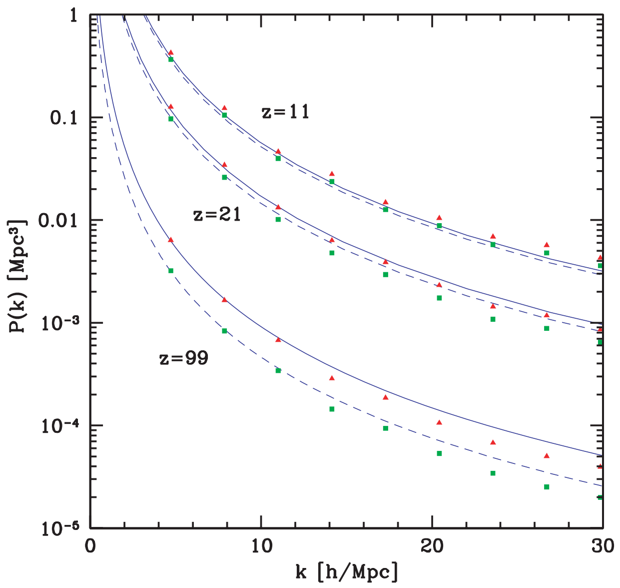

Figure 2. Comparison of the linear and non-linear power spectra. The linear power spectra (generated according to Naoz & Barkana 2005) are shown for the dark matter and baryon components (solid and dashed curves, respectively), while the corresponding non-linear spectra (as measured in the simulations) are shown as triangles and squares, respectively. We show results for the fid model at redshifts $z = 99$, 21 and 11.

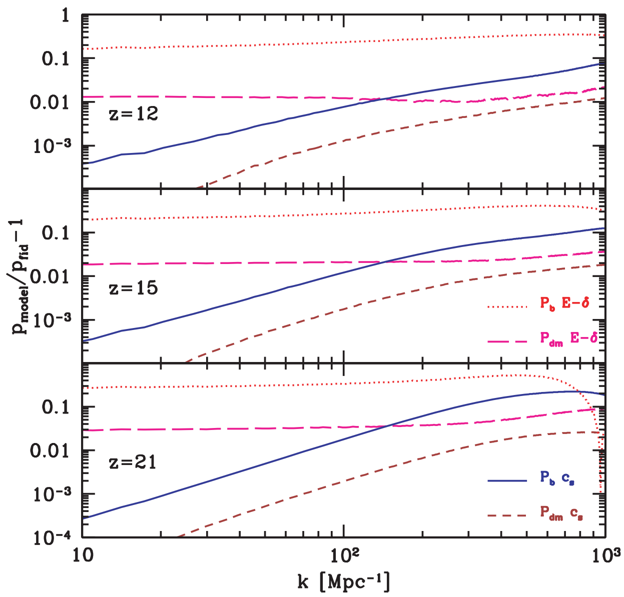

Figure 3. The ratio of the non-linear power spectra (specifically, $P_{\rm model}/P_{\rm fid} - 1$) at $z = 21$, 15 and 12 (from bottom to top); curves are denoted as in Fig. 1. Note that we have plotted here the absolute value; the mean c_s model underestimates and the E-δ model overestimates the power spectrum compared to the fid model.

study are already somewhat non-linear at our initial redshift, and the Compton heating is quite small compared to the adiabatic heating during non-linear gravitational collapse (see Appendix C). A more complete treatment would be to include in the simulation the precise temperature fluctuations, which we leave for future work. Nevertheless, even with the current treatment our results show consistency with linear theory.

3.2 Halo definition

We locate DM haloes by running a FOF group-finder algorithm with a linking parameter of 0.2. We then find the centre of mass of each halo and calculate the density profile of the DM and baryons, separately. In order to derive the density profile we assume a spherical halo, and divide it to 2000 shells. Combining these density profiles, we find the virial radius $r_{\rm vir}$ at which the overdensity is 200 times the background density, and the gas fraction of each halo.

Recently, Trenti et al. (2010) performed a resolution analysis in order to study the mass definition of haloes in simulations. Their conclusion (their fig. 2) is that using the FOF algorithm and assuming about 500 particles per spherical halo introduces an error of ∼15 per cent in the mass definition. In our gas fraction analysis we have chosen only haloes with a number of particles larger or equal to 500, i.e. we limit our errors in halo mass definition to below ∼15 per cent. Also, according to Naoz et al. (2009), this way we can estimate the gas fraction inside a halo accurately.

4 RESULTS AND COMPARISON AMONG THE MODELS

4.1 Non-linear power spectrum evolution

One way to probe cosmic structure particularly on small scales is through the non-linear power spectrum. We begin our simulation at $z = 99$ with linear initial conditions.[1] The main disagreement between the three models lies in the baryonic component (although the E-δ calculation also underestimates the DM overdensities by ∼10 per cent). This input difference is then modified by the non-linear evolution.

Following Yoshida et al. (2003c) we compared the linear power spectrum for the fid model, as computed from Naoz & Barkana (2005), for the DM and baryon components, with the non-linear power spectra from the simulation (see Fig. 2). The two power spectra agree well as expected in the linear regime. We note that the other two models approach the fid model at low redshifts (see Appendix A Fig. A1).

Fig 3 shows the differences among the fid, E-δ and mean c_s ICs, in terms of the non-linear power spectra at the later redshifts at which haloes were formed in our simulation. The mean c_s model maintains over time roughly the same level of discrepancy with the fid model, while in the E-δ model both the baryonic and DM differences decline slightly slower than with the inverse scale factor. As clearly can be seen from Fig. 3, the non-linear evolution of haloes is still strongly affected by the choice of initial conditions even at redshift 12. The fid ICs (Naoz & Barkana 2005) describe the linear evolution consistently and thus represent the best available prescription for the initial conditions.

4.2 The minimum gas rich mass

Studying the galaxy evolution and reionization either by using simulations [both adaptive mesh refinement (AMR) and smoothed particle hydrodynamics (SPH)] or by using analytical calculations relies on knowing the amount of gas within the DM haloes. The simplest assumption, often used in the literature, is that a DM halo has the mean cosmic fraction. This can lead to incorrect results, especially

[1] This is, of course, an approximation, since as shown in Naoz et al. (2006) at $z = 99$ overdensities are already slightly non-linear. The effect of starting the simulation at high redshifts is studied elsewhere (Naoz et. al, in preparation).

when one tries to study star formation, galaxy mergers, and related phenomena.

Consider the various scales involved in the formation of non-linear objects containing DM and gas. On large scales (small wavenumbers) gravity dominates halo formation and gas pressure can be neglected. On small scales, on the other hand, the pressure dominates gravity and prevents baryon density fluctuations from growing together with the DM fluctuations. The relative force balance at a given time can be characterized by the Jeans (1928) scale, which is the minimum scale on which a small gas perturbation will grow due to gravity overcoming the pressure gradient. As long as the Compton scattering of the CMB on the residual free electrons after cosmic recombination kept the gas temperature coupled to that of the CMB, the Jeans mass was constant in time. However, at $z \sim 200$ the gas temperature decoupled from the CMB temperature and the Jeans mass began to decrease with time as the gas cooled adiabatically. Any overdensity on a scale more massive than the Jeans mass at a given time can begin to collapse, due to a lack of sufficient pressure. However, the Jeans mass is related only to the evolution of perturbations at a given time. When the Jeans mass itself varies with time, the overall suppression of the growth of perturbations depends on a time-averaged Jeans mass.

Gnedin & Hui (1998) defined a ‘filtering mass’ that describes the highest mass scale on which the baryonic pressure still manages to suppress the linear baryonic fluctuations significantly. Gnedin (2000) suggested, based on a simulation, that the filtering mass also describes the largest halo mass whose gas content is significantly suppressed compared to the cosmic baryon fraction. The latter mass scale, in general termed the ‘characteristic mass’, is defined as the halo mass for which the enclosed baryon fraction equals half the mean cosmic fraction. Thus, the characteristic mass distinguishes between gas-rich and gas-poor haloes. Many semi-analytical models of dwarfs galaxies often use the characteristic mass scale in order to estimate the gas fraction in haloes (e.g. Bullock, Kravtsov & Weinberg 2000; Benson et al. 2002b,a; Somerville 2002). Theoretically this sets an approximate minimum value on the mass that can still form stars.

4.2.1 *Prediction from linear theory*

In linear theory the filtering mass, first defined by Gnedin & Hui (1998), describes the highest mass scale on which the baryon density fluctuations are suppressed significantly compared to the DM fluctuations. Naoz & Barkana (2007) included the fact that the baryons have smoother ICs than the DM (see Naoz & Barkana 2005) and found a lower value of the filtering mass (by a factor of 3–10, depending on the redshift). Following Naoz & Barkana (2007), the filtering scale (specifically, the filtering wavenumber k_F) is defined by expanding the ratio of baryonic to total density fluctuations to first order in k^2:

$$\frac{\delta_b}{\delta_{tot}} = 1 - \frac{k^2}{k_F^2} + r_{LSS}, \qquad (9)$$

where k is the wavenumber, and δ_b and δ_{tot} are the baryonic and total (i.e. including both baryons and DM) density fluctuations, respectively. The parameter r_{LSS} (a negative quantity) describes the relative difference between δ_b and δ_{tot} on *large scales* (Naoz & Barkana 2007), i.e.

$$r_{LSS} \equiv \frac{\Delta}{\delta_{tot}}, \qquad (10)$$

where $\Delta = \delta_b - \delta_{tot}$, (see also Barkana & Loeb 2005). The ratio r_{LSS} is independent of k, and its magnitude decreases with time approximately $\propto 1/a$, since Δ is roughly constant and δ_{tot} is dominated by the growing mode $\propto a$ (see fig. 1 top panel in Naoz & Barkana 2007).

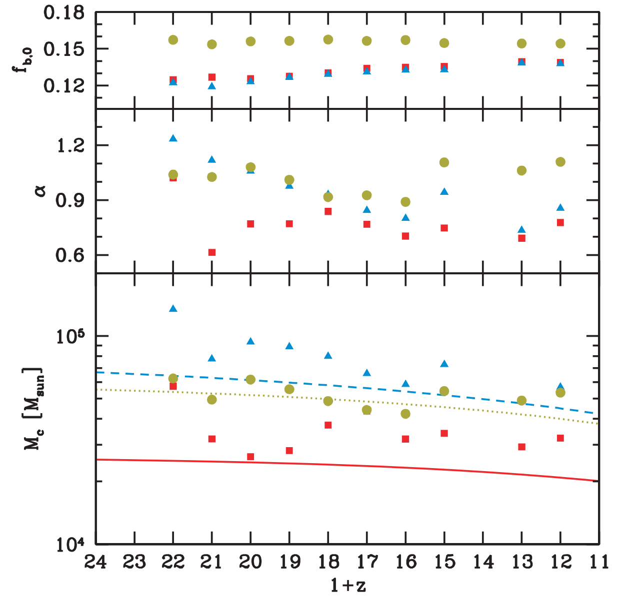

Figure 4. The parameters of the best fits in the form of equation 12; different panels show M_c, α and $f_{b,0}$. We consider the fiducial calculation, mean c_s approximation and the E-δ model (boxes, triangles and circles, respectively), where we fit equation (12) to all data points from haloes with at least 500 particles. In the bottom panel we also show the analytical calculation following Naoz & Barkana (2007), for all the models, assuming the same ICs as in the simulations (solid, dashed and dotted curves for fid, mean c_s and E-δ, respectively). We note that at $1 + z = 13$ the mean c_s and the E-δ models have the same value of M_c, and that the fid model and the E-δ overlap at $1 + z = 17$. We also note that the data for $1 + z = 14$ were unavailable due to a computer failure.

The filtering mass is defined from k_F simply as

$$M_F = \frac{4\pi}{3}\bar{\rho}_0 \left(\frac{1}{2}\frac{2\pi}{k_F}\right)^3, \qquad (11)$$

where $\bar{\rho}_0$ is the mean matter density today. This relation is one eighth of the definition in Gnedin (2000; who also used a non-standard definition of the Jeans mass). In Fig. 4 (bottom panel) we show the filtering mass (solid curve) resulting from equation (11), as calculated in Naoz & Barkana (2007) Naoz & Barkana (2007, see also their fig. 3).

For each of the models we calculate the filtering mass as described here, assuming the model’s initial conditions. Since the simulation is limited in box size, all of the perturbations on large scales are effectively frozen in the simulation. Therefore, we do not extract r_{LSS} directly from the simulations, but instead calculate it based on the initial conditions as $r_{LSS} = \Delta_{in}/(\delta_{tot,in} a)$, where the subscript ‘in’ refers to initial. Thus, for example, for the E-δ case, $r_{LSS} = 0$. Fig. 4 (bottom panel) shows the analytical results of the filtering mass for the fid calculation, the mean c_s approximation and E-δ (solid, dashed and dotted curves, respectively). Since the fid calculation is the most consistent calculation, we compare the two other models to it.

The filtering mass represents the competition between gravity and pressure, as it measures the largest scale at which pressure has had a significant overall effect on halo formation. Since it measures an integrated effect over the formation, this mass scale is also very sensitive to the evolution history and the initial conditions (as shown in Naoz & Barkana 2007). In the mean c_s model, the temperature fluctuations are greatly overestimated on all relevant scales (see Naoz & Barkana 2005), while in reality the coupling to the CMB (in the fid model) keeps the temperature fluctuations highly suppressed for some time after recombination. Moreover, as mentioned in Section 3.1 (and see also Appendix C), we do not include explicitly the effect of initial temperature fluctuations in the simulations. However, the temperature fluctuations from higher redshifts influence the baryon density at the initial redshift (see Fig. 1) and suppress the baryon density on small scales. As demonstrated in Naoz & Barkana (2007) the system remembers the initial conditions. In other words, the initially enhanced filtering mass (compared to the fid model) helps maintain a higher filtering mass even at moderately low redshift.

In the E-δ model, the baryon perturbations start out much higher than in the other models, so one might expect that the final baryon fraction in haloes would tend to be higher as well; here, however, it is important to separate two issues. The high initial baryon perturbations in the E-δ model are present at all scales, so they affect even high-mass haloes that are unaffected by pressure. This can explain why the simulation with the E-δ ICs produced the highest baryon fraction in high-mass haloes (see the top panel of Fig. 4). However, when we consider the differences between large and small scales, the high baryon perturbations produce a large pressure term, increasing the effect of pressure relative to gravity and producing a higher filtering mass in the E-δ model than in the fid model. Note that the filtering mass is particularly sensitive to the importance of pressure at the very highest redshifts (above 100), since at lower redshifts the gas cools and the Jeans mass decreases, reducing the contribution of these redshifts to the final filtering mass.

We note that in Naoz & Barkana (2007) the calculation of the filtering mass in the fiducial model was compared to the time integrated filtering mass in a model that assumes the mean speed of sound model, neglects the $r_{\rm LSS}$ factor, and starts out with initial conditions as in the E-δ model. Here, we have separated our discussion into several different cases.

4.2.2 The non-linear characteristic mass

There is no a priori reason to think that the filtering mass can also accurately describe properties of highly non-linear, virialized objects. For haloes, Gnedin (2000) defined a characteristic mass $M_{\rm c}$ for which a halo contains half the mean cosmic baryon fraction $f_{\rm b}$. In his simulation he found the mean gas fraction in haloes of a given total mass M, and fitted the simulation results to the following formula:

$$f_{\rm g,calc} = f_{\rm b,0}\left[1 + \left(2^{\alpha/3} - 1\right)\left(\frac{M_{\rm c}}{M}\right)^{\alpha}\right]^{-3/\alpha}, \qquad (12)$$

where $f_{\rm b,0}$ is the gas fraction in the high-mass limit. In this function, a higher α causes a sharper transition between the high-mass (constant $f_{\rm g}$) limit and the low-mass limit (assumed to be $f_{\rm g} \propto M^3$). Gnedin (2000) found a good fit for $\alpha = 1$, with a characteristic mass that in fact equaled the filtering mass by his definition. By our definition, the claim from Gnedin (2000) is that $M_{\rm c} = 8 \times M_{\rm F}$.

Naoz et al. (2009) found that, given their errors, the filtering mass from linear theory is consistent with the characteristic mass fitted from the simulations, for two (pre-reionization) scenarios that they tested: the NoUV case (i.e. no stellar heating) and the Flash case (i.e. after a sudden flash of stellar heating). For clarity, we emphasize that this statement ($M_{\rm c} = M_{\rm F}$) refers to our definition of $M_{\rm F}$ in equation (11).

The characteristic mass is essentially a non-linear version of the filtering mass, and so it also measures the competition between gravity and pressure. At high masses, where pressure is unimportant, $f_{\rm g} \to f_{b,0}$, while the low mass tail is determined by the suppression of gas accretion caused by high baryonic pressure.

4.2.3 Comparison between the simulation and the theoretical predictions

A major conclusion of the simulation results is that different ICs result in different gas fractions in the final haloes. Specifically, we measure these differences through the characteristic mass at various redshifts. varies for different ICs. We determine for each simulation output the characteristic mass and the parameter α using a two-dimensional fit to equation (12), with $f_{\rm b,0}$ separately fixed to equal the average of the highest few mass bins (see Appendix B for a complete description of the fitting process, together with the 1σ errors). In Fig. 4 we show $f_{\rm b,0}$, α and $M_{\rm c}$, for all the simulated cases. The characteristic mass clearly depends on the initial conditions, with the mean c_s model and E-δ model both yielding gas suppression at systematically higher halo masses then for the fid model. The parameter α shows a less clear pattern with redshift, but it is generally lowest for the fid model. Overall, the most important implication is that the gas fraction in haloes is highly sensitive to the assumed initial conditions.

Comparing to linear theory allows us to understand some of these results. As noted in section 4.2.1, we calculated the filtering mass from linear theory for each of the ICs, and the linear calculation allows us to understand the relative importance of pressure in the various IC models, at least during the linear evolution. Although the simulation results come from non-linear, viralized haloes, we find an approximate agreement (typically to within ∼20 per cent) between the filtering mass, as defined here and in our previous work (Naoz & Barkana 2007; Naoz et al. 2009), and the characteristic mass as measured in the simulation, for all the models. In particular, the relative sizes of $M_{\rm c}$ among the various models, and the slow decline of all the characteristic masses with time, are well matched by the corresponding $M_{\rm F}$ values predicted from linear theory. This close match can be understood from the fact that while both gravity and pressure increase during the non-linear evolution, their relative strength only changes by a relatively small factor as a halo undergoes non-linear collapse and virialization. Haloes in which pressure had a large effect during the early, linear evolution stage, keep sufficient pressure to maintain the suppressed baryon content all through the final collapse. On the other hand, in more massive haloes in which gravity overcame pressure early on, the baryons keep up with the collapse of the DM and the pressure never has a major role.

For the E-δ alternative model, we find that the resulting characteristic mass is higher than the result in the fid model. Specifically, at $z = 20$ we find $M_{\rm c} \sim 5 \times 10^4$ $M_{\odot}$ and $\alpha \sim 1$). This can be understood since setting the gas fluctuations to be equal to the DM's means that the pressure of the gas is higher compared to the fid model. As can also be seen from comparison to linear theory, the system retains the memory of the pressure, due to the time integrated nature of the filtering mass. Therefore, the higher pressure translates to a higher filtering/characteristic mass.

The mean c_s approximation starts with effectively smoother ICs than in the fid model (∼20 per cent underestimate of the small-scale baryon overdensity). Thus, the baryonic components lag behind the DM collapse, and the pressure is always overestimated for a given baryon overdensity (due to the overestimated temperature fluctuations), resulting in a lower gas fraction for any given halo mass, i.e. the characteristic mass is higher than in the fid model. Specifically, at $z = 20$ we find $M_c \sim 7 \times 10^4$ M$_\odot$ and $\alpha \sim 1$. This can be compared with $M_c \sim 3 \times 10^4$ M$_\odot$ and $\alpha \sim 0.6$ for the fid ICs.

Recently, Hoeft et al. (2006) and Okamoto, Gao & Theuns (2008) showed that the characteristic mass scale does not agree with the Gnedin & Hui (1998) filtering mass in the low-redshift, post-reionization regime. However, it is important to note that at these low redshifts, the heating/cooling and other feedback mechanisms are complex and highly inhomogeneous, so that the 'filtering mass' calculated from linear theory is not really precisely defined, and the comparison of the linear and non-linear results cannot really be considered a direct and precise test. In contrast, Naoz et al. (2009) found that the filtering mass gives a good approximation to the characteristic mass, even in the presence of a 'flash' heating event (see also Mesinger, Bryan & Haiman 2006) that is physically somewhat contrived but allows for a clear comparison of the linear and non-linear results.

Summarizing our results, we find a good agreement between the characteristic mass and the filtering mass in all the models. Fig. 4 shows the best fitted parameters at various redshifts for M_c and α, and our value for $f_{b,0}$, for all models [the 1σ (68 per cent) confidence regions are listed in table B1]. It is important to emphasize that in this statement we are referring to our definition in equation (11), which is one eighth of the original definition which Gnedin (2000) claimed was a good fit to the characteristic mass. While we have been careful to select haloes with at least 500 particles, based on the results of Naoz et al. (2009), we do not have the even higher mass resolution needed to perform a resolution convergence test as they did. Our main conclusion is that at least in the redshift range $z = 11$–21 the filtering mass provides a fairly good estimate for the characteristic mass. This extends the redshift range of the agreement between the filtering mass and the characteristic mass found in Naoz et al. (2009)Naoz et al. (2009, $z = 20$–25). Another significant result from this agreement is that previous work (either analytical, semi-analytical or using simulations) that used the filtering or characteristic mass without accounting for the correct initial conditions resulted in inaccurate results. This is due to the significant (factor of 2–3) variation among the predictions of the filtering/characteristic mass in the various models. Since this mass scale is of prime importance in early structure formation it is imperative to calculate it accurately.

5 DISCUSSION

We have used three-dimensional hydrodynamical simulations to investigate the effect of different initial conditions on the gas fraction in haloes in the early Universe. Specifically, we studied the minimum 'gas-rich' mass defined to have half of the mean cosmic baryon fraction. We tested three different models for the initial conditions (see text for more details).

(i) 'fid' ICs: this model is based on the linear evolution from Naoz & Barkana (2005), which allows the baryonic speed of sound to spatially vary as a result of the Compton scattering with the CMB.

(ii) 'E-δ' ICs: in this model, the linear evolution from Naoz & Barkana (2005) is modified to match a common assumption in the literature, where the linear initial overdensity of the baryons is taken to be equal to that of the DM, i.e. $\delta_b = \delta_{dm} = \delta_{tot}$, while conserving σ_8 from the fid model.

(iii) 'mean c_s ICs': this model assumes that the baryonic speed of sound is spatially uniform. Although Naoz & Barkana (2005) showed that this assumption yields an inaccurate evolution of the baryon density and temperature perturbations, it is still often used in codes that generate initial conditions for simulations.

For all of the tests we used a total of 2×768^3 particles of DM and baryons with a box size of 2 Mpc, starting at $z = 99$.

There are two major findings from the analysis we present here. The first, shown throughout the paper, is the importance of assuming the correct initial conditions, both for analytical calculations and numerical simulations. Structure formation (both in the linear and non-linear regime) and halo gas fractions are very sensitive to the initial conditions even at relatively low redshifts ($\sim$10). The second major finding is the apparent agreement between the filtering mass and the characteristic mass (to within $\sim$20 per cent). This suggests, as a broader implication, that one can use linear theory in order to predict the overall trend of highly non-linear behaviour (at least in the case of determining the gas fraction of haloes).

The the fiducial calculation, which was presented in Naoz & Barkana (2005), follows the time evolution of the linear overdensities correctly. However, the other ICs produce different results for the baryonic structure formation. For instance, the non-linear power spectrum (Fig. 3) shows that the system still remembers its initial condition differences even at redshift 15. In particular, the c_s model underestimates the non-linear baryonic fluctuations by about 10 per cent while the E-δ model overestimates them by 40 per cent on small scales.

The mean c_s approximation and the E-δ model are often used to set the initial conditions in simulations, e.g. the CMBFAST code (Seljak & Zaldarriaga 1996) assumes the mean c_s approximation while Eisenstein & Hu (1999) is used with the E-δ assumption. We have shown that the non-linear evolution is very sensitive to the initial conditions (Fig. 3) and they affect the gas fraction in small haloes down to redshift $\sim$10 (Fig. 4). Our results emphasize the importance of the differences between the DM and baryons and of the spatial sound speed fluctuations, in both the linear calculation and the initial conditions of the simulations.

It is important to emphasize that although Compton heating is not included in the GADGET code that we used in this analysis (GADGET-2), the fiducial calculation still describes fairly well the non-linear behaviour. Actually, the Compton heating contribution to the heating of the gas in non-linear objects is negligible compared to the adiabatic heating due to the gravitational collapse (see Appendix C). Also, as noted above, much of the contribution to the filtering mass comes from the highest redshifts, above our simulation starting redshift of 99, since the Jeans mass is highest then and so the pressure has the greatest impact at that early time.

In each simulation, we calculated the characteristic mass for which a halo keeps most of its baryons (equation 12). We found that the fid calculation gives the lowest value, which suggests that with these correct ICs, pressure plays only a moderate role in galaxy formation. In particular, the characteristic mass of $\sim 3 \times 10^4$ M$_\odot$ is significantly below the minimum mass for molecular hydrogen cooling, so the gas content is not strongly suppressed even in the smallest star-forming haloes. In other words we find that before significant heating took place the baryon fraction in haloes is (equation 12 with $M_c \sim 3 \times 10^4$ M$_\odot$ and $\alpha \sim 0.64$)

$$M_{\rm b} \sim M_{\rm tot} f_{\rm b,0} \left[1 + 0.16 \left(\frac{3 \times 10^4 {\rm M}_{\odot}}{M} \right)^{16/25} \right]^{-75/16}. \quad (13)$$

The other alternative models give incorrect higher value for the characteristic mass, closer to the minimum mass for forming stars. Even with the fid ICs, pressure does strongly limit the amount of gas in minihaloes below the molecular hydrogen cooling mass. We note that this value of 3×10^4 $M_{\odot}$ assumes adiabatic evolution, in particular with no stellar heating. This value is consistent with the results of Naoz et al. (2009) for a somewhat higher redshift range.

We find that the theoretical linear filtering mass (as defined in Section 4.2.1) is in fairly good agreement with the characteristic mass. This finding is true for all the models tested here, throughout a significant redshift range, so this may imply more generally a close relation between linear theory and non-linear halo formation. In addition, this is consistent with the findings by Naoz et al. (2009) from AMR simulations, where the filtering mass and the characteristic mass agreed in the 'E-δ' model, even when a sudden heating was introduced.

Finally, we emphasize that our results are valid only in the pre-reionization era. At the end of the reionization, Mesinger & Dijkstra (2008) concluded that the characteristic mass is likely to be close to the atomic-cooling threshold of $\sim 10^8$ $M_{\odot}$, which is also close to the values found by Hoeft et al. (2006) and Okamoto et al. (2008).

Recently Tseliakhovich & Hirata (2010) argued that the initial velocity difference between the baryons and DM after recombination has not been fully accounted for, because of a higher order contribution that is not included in the linear theory approach. They estimated this higher order effect within the mean $c_{\rm s}$ approximation and found that it causes an additional suppression of the small-scale power spectrum, in turn affecting the formation of the first structures. This effect should be further investigated as in our detailed approach here, although this would be more difficult (analytically, it is a higher order and anisotropic term, and to simulate it directly would require starting at quite high redshifts).

ACKNOWLEDGMENTS

We thank the anonymous referee for useful and helpful comments. We thank Ikkoh Shimizo for supplying the code for calculating the non-linear power spectrum. We also thank Nick Gnedin, Andrey Kravtsov, Matt McQuinn and Michele Trenti for useful discussions. SN also expresses special thanks to Yoram Lithwick for interesting discussions. The authors acknowledge financial support by the Grants-in-Aid for Young Scientists (S) 20674003 by the Japan Society for the Promotion of Science. SN acknowledges NASA ATP grant NNX07AH22G and in part the German-Israeli Project Cooperation (DIP) grant STE1869/1-1.GE625/15-1 and the generous support of the National Post Doctoral Award Programme for Advancing Women in Science (Weizmann Institute of Science). RB is grateful for the support of Israel Science Foundation grant 823/09.

REFERENCES

Abel T., Bryan G. L., Norman M. L., 2002, Sci, 295, 93
Barkana R., Loeb A., 2002, ApJ, 578, 1
Barkana R., Loeb A., 2004, ApJ, 609, 474
Barkana R., Loeb A., 2005, MNRAS, 363, L36
Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002a, MNRAS, 333, 156
Benson A. J., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2002b, MNRAS, 333, 177
Bouwens R. J. et al., 2010, ApJ, 709, L133
Bromm V., Loeb A., 2006, ApJ, 642, 382
Bromm V., Coppi P. S., Larson R. B., 1999, ApJ, 527, L5
Bromm V., Coppi P. S., Larson R. B., 2002, ApJ, 564, 23
Bromm V., Yoshida N., Hernquist L., McKee C. F., 2009, Nat, 459, 49
Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517
Ciardi B., Scannapieco E., Stoehr F., Ferrara A., Iliev I. T., Shapiro P. R., 2006, MNRAS, 366, 689
Clark P. C., Glover S. C. O., Klessen R. S., Bromm V., 2011, ApJ, 727, 110
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, MNRAS, 405, 2161
Eisenstein D. J., Hu W., 1999, ApJ, 511, 5
Furlanetto S. R., Oh S. P., 2006, ApJ, 652, 849
Gao L., Yoshida N., Abel T., Frenk C. S., Jenkins A., Springel V., 2007, MNRAS, 378, 449
Gnedin N. Y., 2000, ApJ, 542, 535
Gnedin N. Y., Hui L., 1998, MNRAS, 296, 44
Gnedin N. Y., Kravtsov A. V., Chen H.-W., 2008, ApJ, 672, 765
Greif T. H., Glover S. C. O., Bromm V., Klessen R. S., 2010, ApJ, 716, 510
Hoeft M., Yepes G., Gottlöber S., Springel V., 2006, MNRAS, 371, 401
Iliev I. T., Scannapieco E., Martel H., Shapiro P. R., 2003, MNRAS, 341, 81
Iliev I. T., Scannapieco E., Shapiro P. R., 2005, ApJ, 624, 491
Jeans J. H., 1928, Cambridge Univ. Press, Cambridge
Komatsu E. et al., 2009, ApJS, 180, 330
Komatsu E. et al., 2011, ApJS, 192, 18
Kuhlen M., Madau P., Montgomery R., 2006, ApJ, 637, L1
Lin W. P., Jing Y. P., Mao S., Gao L., McCarthy I. G., 2006, ApJ, 651, 636
Lin L., Liang E., Zhang S., 2010, ScChG, 53, 64
Ma C., Bertschinger E., 1995, ApJ, 455, 7
McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D. P., Targett T. A., Ellis R. S., 2010, MNRAS, 126
McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, MNRAS, 377, 1043
Mesinger A., Dijkstra M., 2008, MNRAS, 390, 1071
Mesinger A., Bryan G., Haiman Z., 2006, ApJ, 648, 835
Naoz S., Barkana R., 2005, MNRAS, 362, 1047
Naoz S., Barkana R., 2007, MNRAS, 377, 667
Naoz S., Barkana R., 2008, MNRAS, 385, L63
Naoz S., Bromberg O., 2007, MNRAS, 380, 757
Naoz S., Noter S., Barkana R., 2006, MNRAS, 373, L98
Naoz S., Barkana R., Mesinger A., 2009, MNRAS, 377, 667
Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920
Ricotti M., Gnedin N. Y., Shull J. M., 2002, ApJ, 575, 49
Salvaterra R. et al., 2009, Nat, 461, 1258
Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437
Shapiro P. R., Iliev I. T., Raga A. C., 2004, MNRAS, 348, 753
Shapiro P. R., Ahn K., Alvarez M. A., Iliev I. T., Martel H., Ryu D., 2006, ApJ, 646, 681
Somerville R. S., 2002, ApJ, 572, L23
Spergel D. N. et al., 2003, ApJS, 148, 175
Spergel D. N. et al., 2007, ApJS, 170, 377
Springel V., 2005, MNRAS, 364, 1105
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Tanvir N. R. et al., 2009, Nat, 461, 1254
Trenti M., Stiavelli M., 2009, ApJ, 694, 879
Trenti M., Stiavelli M., Michael Shull J., 2009, ApJ, 700, 1672
Trenti M., Smith B. D., Hallman E. J., Skillman S. W., Shull J. M., 2010, ApJ, 711, 1198
Tseliakhovich D., Hirata C., 2010, Phys. Rev. D, 82, 083520
Yamamoto K., Sugiyama N., Sato H., 1997, Phys. Rev. D, 56, 7566
Yamamoto K., Sugiyama N., Sato H., 1998, ApJ, 501, 442
Yoshida N., 2006, New Astron. Rev., 50, 19
Yoshida N., 2010, COSP, 38, 2366
Yoshida N., Abel T., Hernquist L., Sugiyama N., 2003a, ApJ, 592, 645
Yoshida N., Sokasian A., Hernquist L., Springel V., 2003b, ApJ, 598, 73
Yoshida N., Sugiyama N., Hernquist L., 2003c, MNRAS, 344, 481
Yoshida N., Omukai K., Hernquist L., Abel T., 2006, ApJ, 652, 6
Yoshida N., Omukai K., Hernquist L., 2008, Sci, 321, 669

APPENDIX A: σ_8 CONSERVATION

We have defined the two different models such that they conserve $\sigma_8(z = 0)$. From linear theory we do not expect the evolution of the mean c_s model to be significantly different from that of the fid model (in terms of halo abundance and total power spectrum). This is indeed the case for the evolution in time of the total fluctuations of the mean c_s model compared to the fid model on large scales (small k), as shown in Fig. A1 (lower set of thin curves).

A more delicate treatment is needed for the E-δ model (see Section 2.2). In this case, at high redshift (such as the initial $z = 99$), the baryons are in the process of falling into the DM potential. This results in a faster growth of the total fluctuations compared to the case in which there is a relative velocity between the DM and the baryons (such as in the case of the mean cs and fid models, where the relative velocity for the E-δ model are negligible); see Fig. A1 dotted thick curve. At later times, the baryon fluctuations approach the DM fluctuations, and the large scale behaviour (i.e. on linear scales) deviates from the fid model by less then 0.7 per cent (see the solid curve in Fig. A1).

We also note that we have checked the overall effect of $\sigma_8(z = 0)$ on the main results. We have performed two additional simulations for the E-δ model, where we increased or decreased σ_8 by 5 per cent. We found that the calculated M_c is within the fit errors (see Appendix B and Table B1) at $z > 12$. At $z \leq 12$, the difference in the best fitted value is below 0.5 per cent.

APPENDIX B: FIT PROPERTIES

For each redshift snapshot for each run we find the characteristic mass and α using a two dimensional fit. In Fig. B1 we consider two example redshifts (high, $z = 19$ and low, $z = 12$) for which we show the binned data points and the resulting fit. In table B1 we show our best fit parameters. We note that we have checked that the fits give consistent results if we lower the condition on the minimum number of particles per halo to 300 (instead of 500). We also note that our determination of M_c relies on an extrapolation (via the fit) below our simulations' resolution limit

The parameter $f_{b,0}$ in equation (12) is an average of the gas fraction values in the few highest mass bins. In our simulation the high-end tail of the masses has large scatter in the estimated gas fraction because of the low number of haloes (each bin among the last 3 or 4 in Fig. B1 represents just 1 or 2 haloes), thus we have to average over this scatter to get a reasonable result. This scatter is in part a result of assuming that the haloes are spherical, and thus haloes that are undergoing a major merger deviate greatly from a spherical shape and are treated inaccurately in our analysis. We have tested the resulting $f_{b,0}$ when taking a linking parameter of 0.1, which indeed resulted in more high-mass haloes, but in any case was consistent with the value of $f_{b,0}$ we found with the 0.2 linking parameter. Thus, in this paper, we use the standard value of 0.2.

As expected at high redshift, where we have fewer haloes, the errors become quite large. We also tried, following Naoz et al.

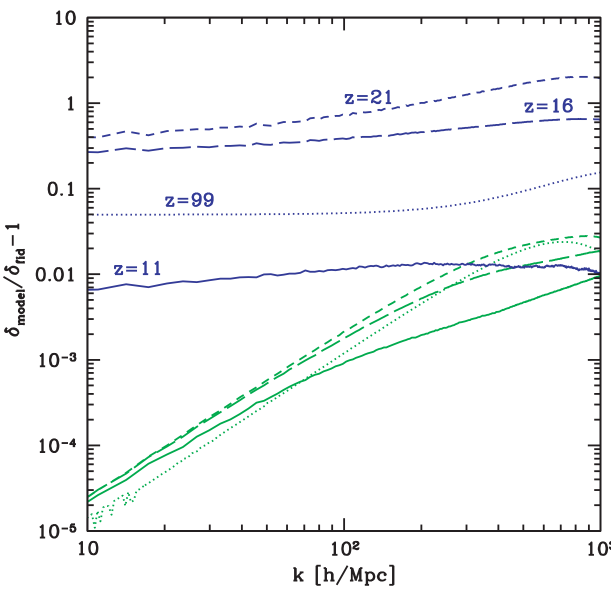

Figure A1. Demonstration of σ_8 conservation. The ratio of the *total* non-linear fluctuation for different redshifts (specifically $\delta_{tot,model}/\delta_{tot,fid} - 1$). We show the E-$\delta$ (upper set of four curves) and the mean c_s model (lower set of four thin curves). We consider redshift $z = 99$, 21, 16 and 11; dotted, short-dashed, long-dashed and solid curves, respectively (labelled for the E-δ model).

Table B1. The best-fitting parameters from equation (12).

Redshift	M_c (10^4 M$_\odot$)	α
	Fiducial calculation	
21	$5.7^{+9.9}_{-5.3}$	$0.7^{+0.45}_{-0.45}$
20	$3.2^{+5.6}_{-1.9}$	$0.61^{+0.39}_{-0.39}$
19	$2.6^{+5.4}_{-2.1}$	$0.77^{+0.39}_{-0.5}$
18	$2.8^{+2.3}_{-2.4}$	$0.77^{+0.03}_{-0.23}$
17	$3.7^{+1.6}_{-1.6}$	$0.84^{+0.02}_{-0.24}$
16	$4.4^{+2.4}_{-2.2}$	$0.77^{+0.21}_{-0.14}$
15	$3.2^{+1.3}_{-1.3}$	$0.7^{+0.1}_{-0.24}$
14	$3.4^{+0.4}_{-1.2}$	$0.75^{+0.12}_{-0.09}$
12	$2.9^{+0.2}_{-0.2}$	$0.69^{+0.1}_{-0.15}$
11	$3.2^{+0.1}_{-0.1}$	$0.78^{+0.1}_{-0.14}$
	Mean c_s	
21	$13.4^{+10.7}_{-7.5}$	$1.23^{+0.52}_{-0.72}$
20	$7.2^{+5}_{-5.5}$	$1.18^{+0.06}_{-0.92}$
19	$9.4^{+5.5}_{-4.9}$	$1.07^{+0.32}_{-0.9}$
18	$8.9^{+5.5}_{-6}$	$0.98^{+0.62}_{-0.31}$
17	$8^{+3.4}_{-3.3}$	$0.92^{+0.3}_{-0.2}$
16	$6.6^{+3}_{-2.6}$	$0.69^{+0.26}_{-0.53}$
15	$5.9^{+2.1}_{-2}$	$0.69^{+0.1}_{-0.26}$
14	$7.3^{+1.6}_{-1.6}$	$0.94^{+0.12}_{-0.09}$
12	$4.9^{+0.8}_{-0.1}$	$0.74^{+0.06}_{-0.04}$
11	$5.7^{+0.7}_{-0.7}$	$0.86^{+0.06}_{-0.03}$
	E-δ	
21	$6.3^{+12.2}_{-4.8}$	$1.04^{+1.5}_{-0.56}$
20	$5^{+9.6}_{-4.8}$	$1.03^{+1.4}_{-0.51}$
19	$6.2^{+5}_{-4.8}$	$1.08^{+1.2}_{-0.45}$
18	$5.5^{+4.5}_{-4.8}$	$1.01^{+1}_{-0.3}$
17	$4.9^{+3.2}_{-3.1}$	$0.92^{+0.42}_{-0.21}$
16	$4.4^{+2.5}_{-2.4}$	$0.93^{+0.31}_{-0.18}$
15	$4.2^{+1.7}_{-1.7}$	$0.89^{+0.19}_{-0.12}$
14	$5.4^{+1.5}_{-1.7}$	$1.11^{+0.13}_{-0.19}$
12	$4.9^{+0.8}_{-0.7}$	$1.06^{+0.06}_{-0.06}$
11	$5.3^{+2.2}_{-0.8}$	$1.11^{+0.05}_{-0.05}$

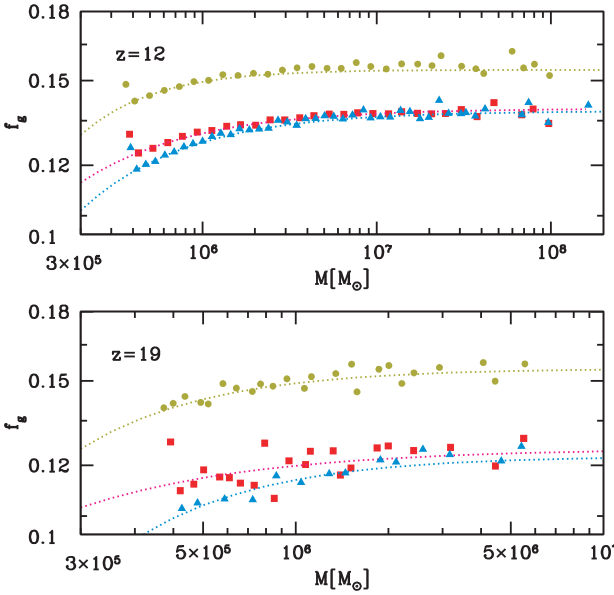

Figure B1. Two redshift examples of fitting the characteristic mass (z = 19 and 12). We consider the fiducial calculation, mean c_s approximation and the E-δ model (boxes, triangles and circles, respectively), where we fit equation 12 to all data points from haloes with at least 500 particles. We also show the fits from table B1 (dotted curves).

(2009), to bin the data and to perform the fit for the binned data with the 1σ weight for each bin. For the redshifts for which we had more than ∼1000 haloes we got that the binned analysis gave results within the non-binned fit errors,and with comparable errors.

We also tried the approach of taking $f_{b,0}$ to be a free parameter, but this produced very problematic fits.[2] This is mainly because of the large scatter at the high mass end, so that a three-parameter fit could not strongly constrain the parameter values. We also note the fact that $f_{b,0}$ is lower than the mean cosmic fraction $\bar{f}_b$, by about 20 per cent – 12 per cent for the fid and mean c_s models, and ∼5 per cent for the E-δ model (see Fig. 4 top panel). The result in the fid and mean c_s models may reflect the real suppression of the large-scale baryon fluctuations in these models; the difference in linear theory is ∼6 per cent at $z = 20$ (Naoz & Barkana 2007), but the non-linear evolution may increase this effect. The discrepancy in the E-δ model may reflect a limitation of the simulation; we note that in Naoz et al. (2009) $f_{b,0}$ was also lower than $\bar{f}_b$ and even lower by 10 per cent from our results at the overlapping redshifts (where we compare the E-δ model in both cases). This might be due to the fact that gas shocks in AMR are sharper than in Gadget simulations, and thus AMR may produce a more realistic gas profile, although the result is still below the universal cosmic baryon fraction (Lin et al. 2006). In our simulation, going to a larger radii can result in a more realistic value, but we used R_{200} for consistency with the common definition.

[2] Naoz et al. (2009) also found that treating $f_{b,0}$ as a free parameter was unproductive.

APPENDIX C: HEATING OF NON-LINEAR HALOES

The fiducial model follows correctly the baryon density and temperature perturbations due to Compton scattering on the residual free electrons after recombination. While this is fully incorporated in our fid ICs, our simulation does not take into account Compton heating. Below we show that for non-linear objects the heating is actually negligible compared to the adiabatic heating due to the gravitational collapse of baryons into the DM potential wells. Therefore, it is sufficient to include Compton heating in the linear stage only.

The heating of the gas $Q_{\rm comp}$ due to Compton heating from the CMB (Naoz & Barkana 2005) during the free-fall time $1/\sqrt{G\rho}$ of gravitational collapse is

$$Q_{\rm comp} \propto 4\frac{\sigma_T\, c}{m_{\rm e}}\, k_{\rm B}(T_\gamma - T)\rho_\gamma x_{\rm e}(t)\frac{1}{\sqrt{G\rho}}, \qquad \text{(C1)}$$

where $\sigma_{\rm T}$ is the Thomson scattering cross-section, ρ_γ is the photon energy density, T_γ and T are the CMB and gas temperature and $x_{\rm e}(t)$ is the electron fraction out of the total number density of gas particles at time t.

The virial theorem gives a relation in collapsed objects between the thermal energy $E_{\rm th}$ and the gravitational energy $E_{\rm gr}$, i.e. $E_{\rm th} = -E_{\rm gr}/2$. Thus, for a halo mass M with virial radius $r_{\rm vir}$ the thermal energy can be expressed as

$$E_{\rm th} \sim \frac{1}{2}\frac{GM^2}{r_{\rm vir}}. \qquad \text{(C2)}$$

For all relevant redshifts and mass scales we find that $Q_{\rm comp}/E_{\rm th} \ll 1$. Therefore, neglecting the contribution of the Compton heating during the non-linear evolution is justified. However, as we have shown, neglecting the Compton heating in the linear evolution and in the initial conditions leads to inaccurate values for the gas fraction in haloes.

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