Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **424,** 1335–1345 (2012) doi:10.1111/j.1365-2966.2012.21318.x

Impact of the relative motion between the dark matter and baryons on the first stars: semi-analytical modelling

Anastasia Fialkov,[1⋆] Rennan Barkana,[1] Dmitriy Tseliakhovich[2] and Christopher M. Hirata[3]

[1] *Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel*
[2] *California Institute of Technology, M/C 249-17, Pasadena, CA 91125, USA*
[3] *California Institute of Technology, M/C 350-17, Pasadena, CA 91125, USA*

Accepted 2012 May 15. Received 2012 May 15; in original form 2011 October 11

ABSTRACT

Recently the initial supersonic relative velocity between the dark matter and baryons was shown to have an important effect on galaxy formation at high redshift. We study the impact of this relative motion on the distribution of the star-forming haloes and on the formation redshift of the very first star. We include a new aspect of the relative velocity effect found in recent simulations by fitting their results to obtain the spatially varying minimum halo mass needed for molecular cooling. Thus, the relative velocities have three separate effects: suppression of the halo abundance, suppression of the gas content within each halo and boosting of the minimum cooling mass. We show that the two suppressions (of gas content and of halo abundance) are the primary effects on the small minihaloes that cannot form stars, while the cooling mass boost combines with the abundance suppression to produce order unity fluctuations in stellar density. We quantify the large-scale inhomogeneity of galaxies, finding that 68 per cent of the star formation (averaged on a 3 Mpc scale) is confined to 35 per cent of the volume at $z = 20$ (and just 18 per cent at $z = 40$). In addition, we estimate the first observable star to be formed at redshift $z = 65$ ($t \sim 33$ Myr) which includes a delay of $\Delta z \sim 5$ ($\Delta t \sim 3.6$ Myr) due to the relative velocity.

Key words: galaxies: formation – galaxies: high-redshift – cosmology: theory.

1 INTRODUCTION

In the present era of 'precision cosmology' and rapidly advancing observational capabilities it is important to make precise theoretical predictions for future observations. Among the major goals of observational cosmology in the near future are to collect data on structure at high redshifts (including the first galaxies), detect the 21-cm line of intergalactic hydrogen and study the cosmic reionization history. A deep understanding of structure formation on small scales and at high redshifts is crucial for making reliable predictions that will help us explore this observational frontier.

The linear perturbation theory of structure formation in the framework of the flat Λ cold dark matter (ΛCDM) model is well understood. It allows us to follow the evolution of structure starting from tiny perturbations. The large-scale perturbations are $\mathcal{O}(10^{-5})$ of the background quantities at cosmic recombination, $z \sim 1100$ (Komatsu et al. 2011), and may have been produced during an early period of inflation. Structure on the smaller scales on which haloes form evolves non-linearly. In order to make reliable predictions, it is important to verify when we can trust the results of the linear perturbation theory and on which scales the non-linear effects must be accounted for.

Linear theory separates different scales, so that each density perturbation mode at a given wavenumber k evolves independently. Thus, non-linear terms that couple the large-scale velocity to the small-scale density perturbations are neglected in linear perturbation theory. However, recently it was shown (Tseliakhovich & Hirata 2010) that such terms might be of the same order of magnitude as the linear terms exactly at the time and on the scales on which the first baryonic objects formed. Specifically, the photon–baryon coupling before recombination left the dark matter and baryonic fluids with large relative velocities. These velocities impede the gravitational perturbation growth on small scales, leading to a spatially variable suppression in the abundance of haloes (Tseliakhovich & Hirata 2010). Moreover, haloes that later form cannot accrete the gas as it shoots past the collapsing dark matter (Dalal, Pen & Seljak 2010; Tseliakhovich, Barkana & Hirata 2011). Thus, the relative velocity effect reopens basic questions regarding the formation and properties of the first stars.

⋆E-mail: anastasia.fialkov@gmail.com

The formation of the first baryonic objects (in particular the first stars) was an important milestone in the history of the Universe. It marked the transition between the cold, neutral, metal-free Universe (the epoch called the 'dark' cosmological ages that started right after recombination) and the modern ionized, hot and metal-rich Universe. The formation of the very first stars is expected to be relatively simple; this is due to the primordial chemistry before stars produced heavy elements, and the simplified gas dynamics in the absence of dynamically relevant magnetic fields and feedback from luminous objects (Tegmark et al. 1997; Barkana & Loeb 2001).

Since molecular hydrogen line emission is the lowest temperature coolant in metal-free gas, the first stars are expected to have formed in haloes with total mass above $\sim 10^5\,\mathrm{M}_\odot$ (Tegmark et al. 1997). More generally, if the mass of a dark matter halo is higher than a threshold referred to as the minimum cooling mass ($M_{\rm cool}$), the collapsing gas is heated to a high enough temperature that it emits radiation. It then cools and condenses, allowing a star to form. The threshold can also be described as a minimum circular velocity ($V_{\rm cool}$) via the standard relation $V_{\rm c} = \sqrt{GM/R}$ for a halo of mass M and virial radius R.

This scenario of the earliest star formation has been confirmed by numerical simulations using both adaptive mesh refinement (AMR) and smooth particle hydrodynamics (SPH) codes (e.g. Fuller & Couchman 2000; Abel, Bryan & Norman 2002; Bromm, Coppie & Larson 2002; Yoshida et al. 2003a,b, 2006; Bromm & Larson 2004; Reed et al. 2005; Gao et al. 2006; Wise & Abel 2007; Maio et al. 2010, 2011a; Petkova & Maio 2011; Turk et al. 2011; Stacy, Greif & Bromm 2012a). All these simulations, though, did not account for the initial relative velocities between the baryons and the dark matter.

In this paper we study the impact of the relative velocities on the distribution of the star-forming haloes at high redshift and on the redshift of formation of the very first star. In particular, we include an aspect of the relative velocity effect that has not been previously accounted for, and which is critical for understanding the overall impact of the velocities on the distribution of star formation. Namely, recent small-scale numerical simulations (Greif et al. 2011; Stacy, Bromm & Loeb 2012b) found that the relative velocity substantially increases the minimum halo mass in which stars can form from gas that cools via molecular hydrogen cooling.

The effect of the velocities was first simulated by Maio, Koopmans & Ciardi (2011b), using an SPH code to follow 320^3 particles each in gas and dark matter within a 1-Mpc box. They found a reduction in the star formation rates, abundance and gas fractions of haloes, but did not consider the minimum cooling mass. Stacy et al. (2012b) used an SPH code to follow 128^3 particles of each type within a $0.1\,h^{-1}$ Mpc box, and Greif et al. (2011) used a moving mesh (hereafter MMH) code to follow 256^3 particles in a 0.5-Mpc box; however, once they identify a halo they run a zoomed-in simulation which achieves a much higher resolution than the other simulation papers. To model star formation, the three simulation papers mentioned here tracked the abundance and the cooling of the chemical components that filled the early Universe, along with the dark matter and gravity. The relevant chemical network includes the evolution of H, H^+, H^-, H_2^+, H_2, He, He^+, He^{++}, e^-, D, D^+, D^-, HD and HD^+, which is determined by processes such as H and He collisional ionization, excitation and recombination cooling, bremsstrahlung, inverse Compton cooling, collisional excitation cooling via H_2 and HD and H_2 cooling via collisions with protons and electrons. More recently, Naoz, Yoshida & Gnedin (2012) simulated the effect of the velocities, carefully controlling numerical resolution and statistical uncertainties, but focusing on the abundance of haloes (i.e. not including gas cooling).

While numerical simulations can successfully form early stars, they face a great difficulty at high redshift, since they must resolve the then-typical tiny haloes while at the same time capture the global galaxy distribution which is characterized by strong fluctuations on surprisingly large scales (Barkana & Loeb 2004). The relative velocities are correlated up to scales above 100 Mpc, and they are important at high redshifts where star formation is dominated by very small haloes. Cosmological simulations that cover this range of scales are not currently feasible.

In this paper we use the simulation results to find the minimal cooling mass for star formation versus formation redshift and the relative velocity; we then use semi-analytical methods to determine the large-scale distribution of haloes and average over cosmological volumes. This allows us to statistically account for all the rare fluctuations in the overdensity and estimate the formation redshift of the first star. For the input from simulations we focus on the simulation results (Greif et al. 2011; Stacy et al. 2012b) that indicate the boost in the minimum cooling mass of haloes, since this is a new effect that has not been included in previous analytical studies. We fit the simulation results and apply this fit to estimate the formation time of the first star.

This paper is organized as follows. In Section 2 we briefly review the results of Tseliakhovich & Hirata (2010) and Tseliakhovich et al. (2011). In Section 3 we summarize the results of the recent simulations that include the effect of the relative velocity on the formation of the first stars via molecular cooling. We use the simulation results to find the behaviour of the minimal cooling mass versus redshift and magnitude of the relative velocity. In Section 4 we study in detail the probability distribution of the gas fraction in haloes at high redshift, separating out and comparing the importance of the various effects of the bulk velocity. In Section 5 we then estimate the redshift of the very first star accounting for the relative velocity effect. Finally, in Section 6 we summarize our results and also give a complete list of differences compared to three previous papers: Tseliakhovich & Hirata (2010), Dalal et al. (2010) and Tseliakhovich et al. (2011).

Our calculations are carried out in a flat ΛCDM universe with cosmological parameters taken from the 7-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*7) results [*WMAP*7+baryon acoustic oscillations (BAO)+H_0 maximum likelihood fit from Komatsu et al. 2011]: the dark matter density today $\Omega_{\rm c,0} = 0.2265$, the baryon density $\Omega_{\rm b,0} = 0.0455$, the vacuum energy density $\Omega_\Lambda = 0.728$, the Hubble constant $H_0 = 70.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and the spectral index $n_{\rm s} = 0.967$. We normalize the power spectrum to give a present value of $\sigma_8 = 0.81$ (Komatsu et al. 2011).

2 REVIEW OF THE RELATIVE VELOCITY EFFECT

In this section we briefly review the non-linear effect of the relative velocities between the baryons and dark matter, as discussed in Tseliakhovich & Hirata (2010) and Tseliakhovich et al. (2011), the latter of which we closely follow in our subsequent calculations.

The initial conditions at recombination include significant relative velocities between the baryons and the CDM (which we denote $v_{\rm bc}$). Before the baryons kinematically decouple from the radiation (around $z = 1100$), they are carried along with the photons, while the dark matter moves according to the gravitational growth of fluctuations which has been advancing since matter-radiation equality ($z \sim 3200$). At decoupling, the baryonic speed of sound drops

precipitously, and the relative velocity then becomes a substantial effect.

In the standard picture of Gaussian initial conditions (e.g. from a period of inflation), the density and the components of relative velocity are Gaussian random variables. The velocity and density are spatially correlated (at different points) since the continuity equation relates the velocity divergence to the density. Indeed, this equation gives an extra factor of $1/k$ in the velocity (where k is the wavenumber), making the velocity field coherent on larger scales than the density. Specifically, velocity fluctuations have significant power over the range $k \sim 0.01$–$0.5\,\mathrm{Mpc}^{-1}$.

The relative velocity is thus coherent on scales smaller than $\sim$3 comoving Mpc. We therefore analyse probability distributions in such coherent patches, and refer to the uniform relative velocity within each patch as the 'bulk' or 'streaming' velocity. The magnitude of the bulk velocity in each coherence patch at recombination is distributed according to a Maxwell–Boltzmann distribution function:

$$p_{v_{\rm bc}}(v_{\rm bc}) = \left(\frac{3}{2\pi\sigma^2_{v_{\rm bc}}}\right)^{3/2} 4\pi v_{\rm bc}^2 \exp\left(-\frac{3v_{\rm bc}^2}{2\sigma^2_{v_{\rm bc}}}\right), \quad (1)$$

where $\sigma_{v_{\rm bc}} \sim 30\,\mathrm{km\,s}^{-1}$ is the root-mean-square velocity at recombination. Just like any peculiar velocity, the bulk velocity $v_{\rm bc}$ decays as $(1 + z)$ with the expansion of the Universe. In addition to the bulk velocity, within each patch there are small-scale peculiar velocities of the baryons and dark matter related to the evolution of perturbations (and formation of haloes) within the patch.

As was shown in the above references, inside each coherent region the linear evolution equations for density and velocity perturbations are modified. For example, on small scales the non-linear term in the continuity equation that couples the local density to the velocity field, $a^{-1}\boldsymbol{v} \cdot \nabla\delta$, is comparable to linear terms such as the velocity term $a^{-1}\nabla \cdot \boldsymbol{v}$. The leading contribution of the non-linear term comes from the bulk motion ($a^{-1}\boldsymbol{v}_{\rm bc} \cdot \nabla\delta$) and this contribution is then linear in terms of the perturbations within the patch. As a result, the evolution equations for the perturbations inside a coherent patch are still linear but dependent on the bulk $v_{\rm bc}$. The resulting velocity-dependent terms were previously neglected but must be included when structure on small scales and at high redshifts is considered.

The relative velocity effect is particularly important for the formation of the first stars and galaxies. As the first baryonic objects try to form, they must do so in a moving background of the dark matter potential wells. This relative motion means that the dark matter's gravity must work harder in order to trap the baryons. As a result, the formation of the first bounded baryonic objects is delayed. The effect, though, is less relevant for structure formation today, since the relative velocity decays with time while the typical mass of galactic host haloes increases. However, the relative motion may shift slightly the positions of the BAO peaks and produce a unique signature in the bispectrum of galaxies (Yoo, Dalal & Seljak 2011).

3 CALIBRATION OF THE MINIMUM HALO COOLING MASS WITH SIMULATIONS

While it is difficult for numerical simulations to capture the full range of scales involved in galaxy formation at high redshift (see Section 1), they remain the best tool for studying the complex, non-linear formation of haloes on small scales. The scales relevant to the formation of the small haloes that host the first stars are well below the coherence scale of the relative velocity field. Therefore it is possible to simulate halo formation in small patches of uniform $v_{\rm bc}$. We focus on recent SPH (Stacy et al. 2012b) and MMH (Greif et al. 2011) simulations that studied the impact of the relative streaming velocity $v_{\rm bc}$ on the mass reached by a halo when it first allows a star to form, i.e. when it first contains a high-density gas core formed out of gas that cooled and collapsed. The results show a substantially increased halo mass in regions with a significant relative velocity.

This is a different effect from the suppression of the amount of gas, which implies a smaller number of stars in the halo at a given time; instead in this case there is a substantial delay in the formation of the first star within the halo. Moreover, this effect is not simply related to the total amount of accreted gas, since in the cases with a bulk velocity, even if we wait for the halo to accrete the same total gas mass as its no-velocity counterpart, it still does not form a star (even within the now deeper potential of a more massive host halo); the delay is substantially longer than would be expected based on a fixed total mass of accreted gas. Instead, it appears that the explanation lies with the internal density and temperature profiles of the gas, which are strongly affected by the presence of the streaming motion. A plausible explanation for the resulting delay in star formation is that the first star forms from the gas that would have accreted early and formed the dense central cores in which stars form; this gas tries to accrete early (when $v_{\rm bc}$ is still very large) into a still-small halo progenitor, so it is affected most strongly by the suppression of gas accretion due to the bulk velocity.

The simulations yield a minimum halo cooling mass at various redshifts, so we fit the results to find the dependence of the minimum halo mass on the redshift of collapse and on the bulk velocity, $v_{\rm bc}$, in the patch. This will then allow us to study the effect of the relative velocity on the formation of the first stars using statistical methods that average over large cosmological regions that cannot be directly simulated.

Stacy et al. (2012b) and Greif et al. (2011) state apparently contradictory conclusions, one claiming a negligible effect on star-forming haloes and the other a large effect. In order to meaningfully compare their results, it is important to put them both on the same scale. We express the cooling threshold as a halo circular velocity, since simulations (cited above) without the bulk velocity find an approximately redshift-independent threshold $V_{\rm cool,0}$; this is naturally expected since molecular cooling turns on essentially at a fixed gas temperature, and the halo circular velocity determines the virial temperature to which the gas is heated. Thus, the limit of zero bulk velocity simply gives a fixed threshold $V_{\rm cool,0}$. When we add the relative velocities, in principle the minimum circular velocity in a patch may be a separate function of two parameters: the redshift z and the bulk velocity at halo formation $v_{\rm bc}(z)$. The history of $v_{\rm bc}$ at earlier redshifts cannot introduce additional parameters, since given both z and $v_{\rm bc}(z)$, the full history of $v_{\rm bc}$ is determined, i.e. at any other redshift z', $v_{\rm bc}(z') = v_{\rm bc}(z)\,(1 + z')/(1 + z)$. In particular, we frequently use the value of $v_{\rm bc}$ at recombination to parametrize the evolution history. We note that the value of $v_{\rm bc}(z)$ in a particular place results from a combination of two separate inputs, i.e. the initial value of the streaming velocity at recombination (determined by the random initial conditions) and the redshift. Thus a given value of the parameter $v_{\rm bc}(z)$ may result from different combinations of initial velocity and redshift.

Consider now the limit of a very high bulk velocity, $v_{\rm bc}(z) \gg V_{\rm cool,0}$, so that the effect of $V_{\rm cool,0}$ is negligible. For simplicity, consider for a moment a constant $v_{\rm bc}$ versus redshift, fixed at its final value $v_{\rm bc}(z)$ at the halo formation redshift z. In this case there is only one velocity scale in the problem. As in a Jeans mass analysis, in the reference frame of a collapsing dark matter halo with a circular

velocity V_c, clearly gravity will be able to pull in the gas (which streams by at the velocity $v_{bc}(z)$) if $V_c \gtrsim v_{bc}(z)$. Now, in the real case where $v_{bc}(z')$ is higher during the formation of the halo, we would expect to get a threshold that is higher than $v_{bc}(z)$, but by a fixed factor, because the physics is scale-free: on one side, v_{bc} scales in a simple way with redshift, and on the other side, halo formation (in the high-redshift, Einstein–de Sitter universe) also scales in a simple way, as we know from spherical collapse; e.g. turnaround for a halo that forms at redshift z always occurs at z' where $1 + z' = 1.59(1 + z)$ so that $v_{bc}(z') = 1.59 v_{bc}(z)$. The only new scale that enters is from v_{bc} at recombination, but as long as we consider haloes that form long after recombination, this should be insignificant.

Thus, the threshold circular velocity V_{cool} should change continuously between two limits, $V_{cool} = V_{cool,0}$ when $v_{bc}(z) \ll V_{cool,0}$, and $V_{cool} = \alpha v_{bc}(z)$ when $v_{bc}(z) \gg V_{cool,0}$ (in terms of a fixed, dimensionless parameter α). When V_{cool} is expressed as a function of $v_{bc}(z)$, there is no additional dependence on z in these two limits, so we might naturally expect this to be true in the intermediate region as well. Indeed, the above argument suggests more generally that halo formation and $v_{bc}(z)$ scale together so that the effect of the bulk velocity should not depend separately on redshift; also the effect of molecular cooling is a redshift-independent threshold. Thus, when both effects act together, the result should still depend on just one parameter.

We expect the dependence on velocity to be smooth and well behaved for vector $\boldsymbol{v}_{bc}(z)$ near zero, i.e. as a function of the velocity components. This suggests a quadratic dependence on $[v_{bc}(z)]^2 = [\boldsymbol{v}_{bc}(z)]^2$ rather than e.g. a linear dependence on $v_{bc}(z)$. We thus propose a simple ansatz for the minimum cooling threshold of haloes that form at redshift z:

$$V_{cool}(z) = \left\{ V_{cool,0}^2 + [\alpha v_{bc}(z)]^2 \right\}^{1/2}. \quad (2)$$

The dependence of the circular velocity V_{cool} on redshift only through the final value $v_{bc}(z)$ implies that the star formation threshold in a patch with a statistically rare, high value of v_{bc} at low redshift is the same as the threshold in a patch with the same (but now statistically more typical) value of v_{bc} at high redshift. This should be the case during the era of primordial star formation, before metal enrichment and other feedbacks complicate matters.

We summarize the results of the two simulations together with the best fits to each of them (with $V_{cool,0}$ and α as free parameters) in Fig. 1 (top panel). We obtain four data points from Stacy et al. (2012b) with non-zero velocities (and two more at $v_{bc}(z) = 0$), and three points from Greif et al. (2011) (plus three more at $v_{bc}(z) = 0$). The best-fitting parameters are (1) $V_{cool,0} = 3.640\,\mathrm{km\,s^{-1}}$ and $\alpha = 3.176$ for the results of Stacy et al. (2012b); (2) $V_{cool,0} = 3.786\,\mathrm{km\,s^{-1}}$ and $\alpha = 4.707$ for Greif et al. (2011).

We note that despite the small numbers of haloes, we would not necessarily expect as large a scatter in the measured $V_{cool}(z)$ as in other measurements of halo properties; for example, in a sample with a large number of haloes of various masses at each redshift, we would expect a large range of redshifts for the first star formation within a halo, but if we only take haloes that first formed a star at a given redshift z, their masses at z might span a narrow range, all near the minimum cooling mass for that redshift (since any halo well above the cooling mass at z would already have formed a star earlier). In any case, our ansatz fits each set of simulation results reasonably well, but there is some scatter and also a systematic difference between the two sets (with Greif et al. 2011 indicating a stronger effect of the bulk velocity). Because of the small number of simulated haloes, it is difficult to separate the possible effects of different numerical resolutions, other differences in the gravitational or hydrodynamical solvers and real cosmic scatter among haloes. However, the difference between the two simulation sets is at least a rough indication of the possible systematic uncertainty due to these various sources.

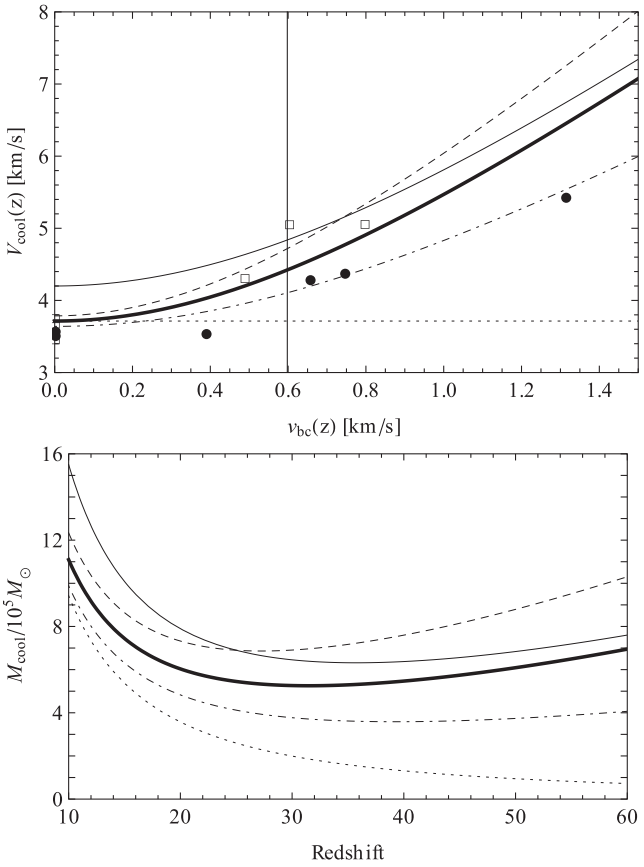

Figure 1. Top panel: the minimum halo circular velocity for gas cooling via molecular hydrogen versus the bulk velocity $v_{bc}(z)$ when the halo virializes. We use the data from fig. 2 of Stacy et al. (2012b) (•) and from fig. 3 of Greif et al. (2011) (□), where we calculate $V_{cool}(z)$ and $v_{bc}(z)$ from the minimum cooling mass, the initial streaming velocity and the redshift of star formation. We show our fits to each set of simulation results (dot–dashed and dashed, respectively). We also show our 'optimal' fit to the SPH and MMH simulations (thick solid line), the 'fit' to AMR simulations (regular solid line) and the case of no streaming velocity (dotted line, based on our optimal fit). The vertical solid line marks the root-mean-square value of $v_{bc}(z)$ at $z = 20$. Bottom panel: we show the minimum halo mass for molecular cooling versus redshift, in a patch with the root-mean-square value of $v_{bc}(z)$ at each redshift z, for each of the fits from the top panel; in particular, we show (dotted line) the case of no relative motion based on our optimal fit (i.e. $V_{cool} = V_{cool,0} = 3.714\,\mathrm{km\,s^{-1}}$).

Given the systematic offset, we do not simultaneously fit both sets of points, but instead average the best-fitting parameters of the SPH and MMH simulation sets. We mostly use this fit, which we refer to as our *optimal fit*, in the following sections:

$$V_{cool}(z) = \left\{ \left(3.714\,\mathrm{km\,s^{-1}}\right)^2 + [4.015\, v_{bc}(z)]^2 \right\}^{1/2}. \quad (3)$$

There is some discrepancy in the value of $V_{cool,0}$ found in AMR and SPH simulations. In order to test the full current uncertainty range including different types of simulations, we also consider the average value $V_{cool,0} \sim 4.2\,\mathrm{km\,s^{-1}}$ found in AMR simulations (Yoshida et al. 2006; Turk et al. 2011). Thus, we combine this value of $V_{cool,0}$ with α from our optimal fit to obtain what we refer

to as a 'fit' to AMR simulations. In other words, we assume that the discrepancy between the two simulation methods is only in the cooling process (due to systematic entropy differences in dense cores), but that they would agree on the effect of the bulk motion. Regardless of which fit we use, Fig. 1 shows that the relative motion has a large effect on the minimum circular velocity.

The implications for the minimum cooling mass as a function of redshift are also shown in Fig. 1 (bottom panel). In a patch with no relative motion, the mass drops rapidly with redshift, since at higher redshift the gas density is higher and a given halo mass heats the infalling gas to a higher virial temperature. However, in a region at the root-mean-square value of $v_{\rm bc}$[1] the higher bulk velocity at high redshift implies that a higher halo mass is needed for efficient molecular cooling. In particular, at redshift 20 a patch with $v_{\rm bc} = 0$ will form stars in $3.6 \times 10^5\,{\rm M}_\odot$ haloes, while a patch with the root-mean-square value of $v_{\rm bc}$ has a minimum cooling mass of $6.0 \times 10^5\,{\rm M}_\odot$ according to the optimal fit, or a range of $(4.8–7.3) \times 10^5\,{\rm M}_\odot$ from the other fits. At $z = 60$ these numbers become 7.2×10^4, 7.0×10^5 and $(4.1–10.3) \times 10^5\,{\rm M}_\odot$, respectively. In patches with low bulk velocity we expect stars to form earlier, since the haloes with lower masses are more abundant and form earlier in the hierarchical picture of structure formation. This is the basis of the discussion that follows.

4 GAS FRACTION IN THE FIRST BOUND BARYONIC OBJECTS

In this and the following sections we use the local small-scale results we have discussed (e.g. equation 3) to study the impact of the streaming motion on the overall star formation in the universe. We can do this as the coherence length of the streaming velocity is much larger than the scale of an individual star-forming halo. Thus, we divide the universe into patches (of order the coherence scale of the streaming velocity), and apply the small-scale results to each patch. This allows us to compute the global one-point distributions of various quantities of interest.

In this section we study in detail the probability distribution of the gas fraction in haloes at high redshift accounting for the bulk velocity. Following Tseliakhovich et al. (2011) we solve the linear evolution equations for density and velocity perturbations, which also include the non-linear term (discussed in Section 2) and include the effect of Compton heating from the cosmic microwave background (CMB) on the sound speed and fluctuations in the temperature distribution (Naoz & Barkana 2005). Since we are interested in star formation, we solve the system of equations on small scales assuming a constant local value for the background relative velocity $v_{\rm bc}$. We use the CAMB sources linear perturbation code (Lewis & Challinor 2007) to generate initial conditions at recombination (specifically, at $z = 1020$ and 970 in order to obtain the needed derivatives). We solve the system of equations to find the power spectra of CDM and baryons for various values of the relative velocity and use them to calculate relevant quantities, e.g. the gas fraction and mass fraction of various haloes, as a function of $v_{\rm bc}$.

The population of gas-filled haloes at high redshift divides naturally into two major categories. The first category consists of large haloes in which the gas can cool (via molecular hydrogen cooling); these are presumed to be the sites of formation of the first stars, and are obviously most important since the stellar radiation is in principle observable, and it also produces feedback on the intergalactic medium and on other nearby sites of star formation. Also interesting, though, is the second category, namely the smaller haloes ('minihaloes') in which the gas accumulates to roughly virial density and yet cannot cool. The latter may affect the epoch of reionization by acting as a sink for ionizing photons (e.g. Haiman, Abel & Madau 2001; Barkana & Loeb 2002; Ciardi et al. 2005; Iliev, Scannapieco & Shapiro 2005) and may generate a 21-cm signal from collisional excitation of H I (e.g. Iliev et al. 2003; Furlanetto & Oh 2006).

We find the fraction of the baryon density contained in haloes with mass larger than the minimum cooling mass $M_{\rm cool}$:

$$f_{\rm gas}(>M_{\rm cool}) = \int_{M_{\rm cool}}^{\infty} \frac{M}{\bar{\rho}_0} \frac{{\rm d}n}{{\rm d}M} \frac{f_{\rm g}(M)}{f_{\rm b}} {\rm d}M, \quad (4)$$

where $\bar{\rho}_0$ is the mean matter density today, ${\rm d}n/{\rm d}M$ is the comoving abundance of haloes of mass M, $f_{\rm b} \equiv \Omega_{\rm b}/\Omega_{\rm m}$ is the mean cosmic baryon fraction and $f_{\rm g}(M)$ is the fraction of the total halo mass which is in the form of gas. The gas fractions $f_{\rm g}(M)$ depend on the filtering mass, which measures the scale at which the baryon fluctuations differ substantially from those in the dark matter. In each patch, the filtering mass depends on the bulk velocity, and thus so do the gas fractions. Since the baryons contribute to the total power spectrum, the halo abundance ${\rm d}n/{\rm d}M$ (which depends on fluctuations in the total matter density) varies as well with $v_{\rm bc}$. We use the halo mass function of Sheth & Tormen (1999) with the proper critical density of collapse $\delta_{\rm c}(z)$ of Naoz & Barkana (2007) for our flat ΛCDM cosmology.

Our numerical routine is similar to the one used in Tseliakhovich et al. (2011). Therefore we refer the interested reader to Sections 2 and 3 of that paper for the full details.

4.1 Global average

In this subsection we apply the result we found for the minimum cooling mass to find the redshift evolution of the gas fraction in these two categories. In the following subsections we explore the probability distribution function (PDF) of the gas fraction, beginning with its dependence on the bulk velocity. In addition, though, in each patch of coherent velocity the mean density is slightly different, varying as a result of random density fluctuations on scales larger than the patch size. We thus also study the full PDF as determined by the joint dependence of the gas fraction in haloes on the bulk velocity and the local overdensity in each patch.

We begin by recalculating some of the results of Tseliakhovich et al. (2011). We show in Fig. 2 the redshift evolution of the globally averaged gas fraction in star-forming haloes or in gas minihaloes. Compared with fig. 8 of Tseliakhovich et al. (2011), our gas fractions are substantially lower, e.g. the gas fraction in haloes above the minimum cooling mass is lower by a factor of ∼3 at redshift $z = 20$, with a spread of ±7 per cent for the different fits. The lower gas fraction is due to our higher $M_{\rm cool}$ and lower power spectrum normalization (see Section 6 for a full discussion of our differences with previous papers). Note that the gas fraction in haloes above the minimum cooling mass is proportional to the stellar mass density, assuming a fixed star formation efficiency (averaged over each 3-Mpc patch).

In general, the importance of the relative velocities increases with redshift. Comparing the two categories of haloes, we find that the relative suppression of the minihaloes is larger than that of the

[1] Since $v_{\rm bc}$ decays as $1 + z$ throughout the universe, a patch that has the root-mean-square value of $v_{\rm bc}$ at one redshift will have the root-mean-square value of the relative velocity at every redshift, and in particular $v_{\rm bc} = 30\,{\rm km\,s^{-1}}$ at recombination.

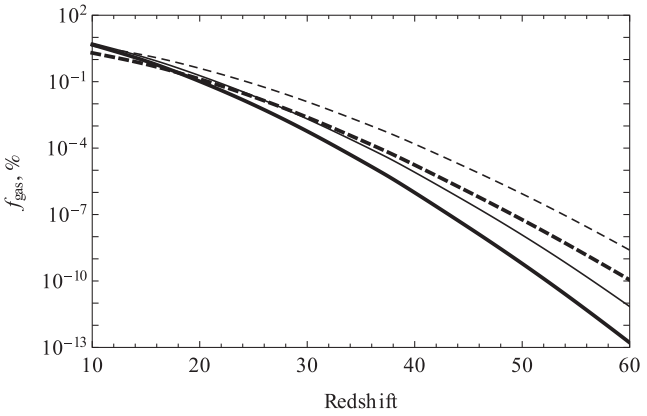

Figure 2. The global mean gas fraction in star-forming haloes (solid curves) and in minihaloes, i.e. haloes below the cooling threshold (dashed curves). The results, based on our optimal fit (equation 3), are shown after averaging over the distribution of relative velocity (thick curves), or in the case of no relative motion, i.e. for $v_{\rm bc}(z) = 0$ (thin curves).

star-forming haloes at low redshift; however, the relative suppression of the star-forming haloes increases faster with redshift, and eventually it becomes larger than that of the minihaloes (beyond $z \sim 50$). At $z = 20$, the bulk velocities reduce the mean gas fraction in star-forming haloes by a factor of 1.8 and that in minihaloes by 3.1.

Unlike previous analytical studies, in our calculations the relative velocities produces three distinct effects (equation 4): suppression of the halo abundance (dn/dM), suppression of the gas content within each halo ($f_{\rm g}(M)$), and boosting of the minimum cooling mass ($M_{\rm cool}$, determined by $V_{\rm cool}(z)$). Note that this separation into three distinct effects is natural within our model, but this does not preclude the possibility that they are physically correlated or mutually dependent. In order to gain a better physical understanding, and for easier comparison with previous papers, we investigate the relative importance of each effect in Fig. 3. For the star-forming haloes, the suppression of gas content is always the least significant effect (e.g. suppression by a factor of 1.13 on its own at $z = 20$), while the cooling mass boost is most important above $z = 28.5$ (factor of 1.26 on its own at $z = 20$), and the halo abundance cut is most important at lower redshifts (factor of 1.43 on its own at $z = 20$). For the minihaloes, the boosting of the minimum cooling mass acts as a (small) positive effect, since it moves gas from the star forming to the minihalo category (e.g. boost by a factor of 1.10 on its own at $z = 20$), while the other two effects are larger and comparable (e.g. at $z = 20$ the suppression of gas content would give a reduction by a factor of 2.17 on its own, and the halo abundance cut would give a suppression factor of 1.74).

4.2 Inhomogeneous gas fraction due to the dependence on the relative velocity

The gas fractions shown in Figs 2 and 3 are globally averaged. However, in reality the Universe is highly inhomogeneous on small cosmological scales. We can divide it into patches that have various bulk velocities and densities. In this section we consider just the variation with velocity, i.e. averaged over all density fluctuations. In other words, we look at the contribution of velocity fluctuations to fluctuations in the gas fraction in haloes. If we consider patches that are still small enough to have a coherent $v_{\rm bc}$ (e.g. cubes of 3 comoving Mpc on a side), then the absolute value of the bulk

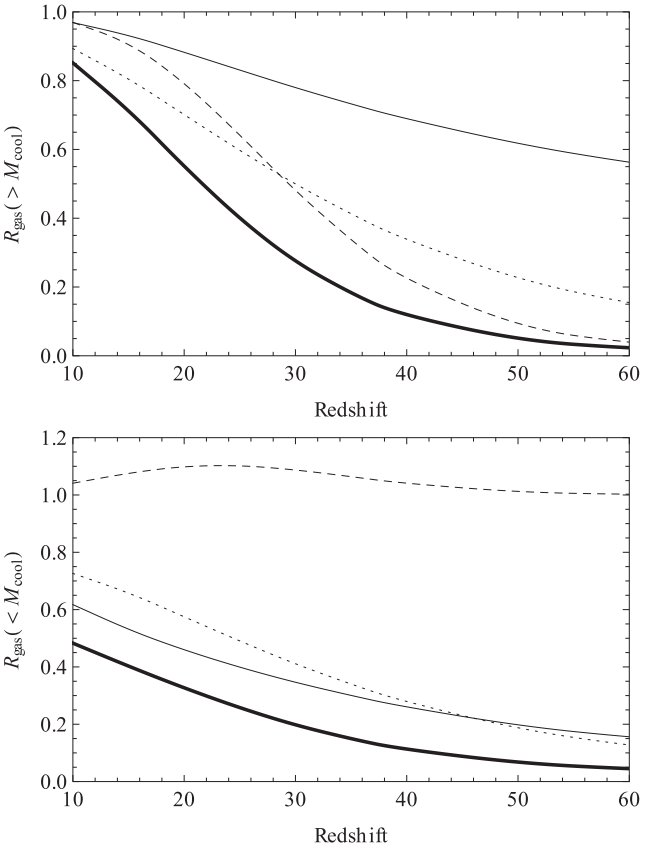

Figure 3. The ratio (compared to the $v_{\rm bc} = 0$ case) by which the bulk velocities change the global mean gas fraction in haloes above the cooling mass (top panel) and in starless minihaloes (bottom panel). We consider four different cases: the full effect of the velocities (thick solid curves); the effect of $v_{\rm bc}$ in boosting the cooling mass only (dashed curves); the effect of $v_{\rm bc}$ in suppressing the halo abundance only (dotted curves) and the effect of $v_{\rm bc}$ in suppressing the gas fraction only (thin solid curves).

velocity in each one follows a Maxwell–Boltzmann distribution (equation 1).

Consider the contributions of patches of various velocities to the total amount of star formation. At a given redshift, the gas fraction in star-forming haloes is lower in the patches with a high value of the relative velocity, because all three velocity effects (see the previous subsection) tend to reduce this gas fraction. On the other hand, patches with zero bulk velocity do not contribute much, simply because they are rare. As shown in the top panel of Fig. 4, the most common bulk velocity is $v_{\rm bc} \sim 0.82\sigma_{v_{\rm bc}}$, where $v_{\rm bc}$ and $\sigma_{v_{\rm bc}}$ are both measured at the same redshift (recombination or any other z). If the stellar density were independent of the bulk velocity, then the contribution of regions of various velocities would be proportional to the velocity PDF. Instead, the velocity suppression effect shifts the contribution to stellar density (assumed proportional to the gas fraction in star-forming haloes) towards lower $v_{\rm bc}$, with the relative change (compared to the Maxwell–Boltzmann distribution) increasing strongly with redshift. Thus, the biggest contribution to stellar density comes from $v_{\rm bc} = 0.67\sigma_{v_{\rm bc}}$ patches at $z = 20$, and from $v_{\rm bc} = 0.23\sigma_{v_{\rm bc}}$ patches at $z = 60$. We compare the contributions of the three distinct effects of the velocity to the shift in the distribution of star formation (Fig. 4, bottom panel). As in the top panel of Fig. 3, we find that the suppression of halo gas content has the least significant effect on star-forming haloes at $z = 20$ (typically, a ∼10 per cent effect on the distribution), while the other two

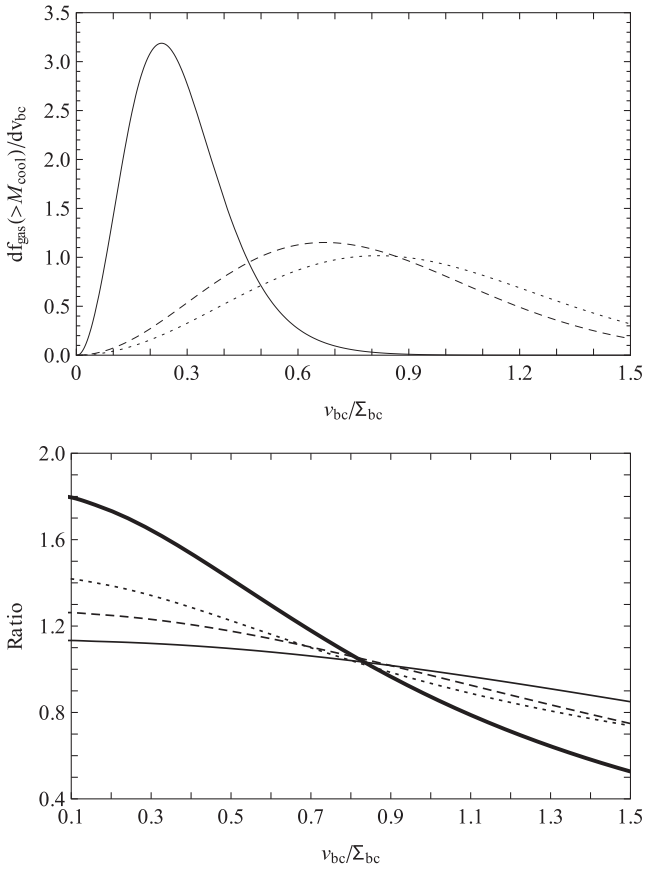

Figure 4. Top panel: the relative contribution of regions with a given streaming velocity to the global gas fraction in haloes above the cooling mass, i.e. $df_{gas}(>M_{cool})/dv_{bc}$ normalized to an area of unity. The dependence is shown for $z = 60$ (solid curve) and 20 (dashed curve). We also show the Maxwell–Boltzmann distribution of the bulk velocity (dotted curve). The velocity is expressed in units of its root-mean-square value $\sigma_{v_{bc}}$. Bottom panel: the ratio at $z = 20$ between the quantity shown in the top panel (the relative contribution of regions with a given streaming velocity to the gas fraction in star-forming haloes) and the Maxwell–Boltzmann distribution. If star formation were independent of bulk velocity, this ratio would equal unity. We consider this ratio for the same four cases as in Fig. 3: the full velocity effect (thick solid curve); the boost in the cooling mass only (dashed curve); the suppression of halo abundance only (dotted curve) and the suppression of the gas fraction only (thin solid curve).

effects (halo abundance suppression and cooling mass boost) have a ∼20–30 per cent effect each.

Thus, at the highest redshifts, the star formation is concentrated in low-velocity regions which are rare, i.e. at the low-probability v_{bc}^2 end of the Maxwell–Boltzmann distribution function. The universe at these epochs is very inhomogeneous, with a few bright regions filled with stars, while in all other regions the relative velocity is too high to allow significant star formation. As the universe expands, the relative velocity decays, and in more and more patches across the universe the relative velocity drops enough to allow for star formation. As a result, the stellar distribution becomes increasingly homogeneous. To quantify the degree of inhomogeneity caused by the dependence of stellar density on the bulk velocity, we plot the fraction of the volume of the universe (at lowest velocity, i.e. at highest stellar density) that contains 68 or 95 per cent of the star-forming haloes (Fig. 5). The effect of volume concentration is mild at $z = 20$ (68 per cent of the stars are in 54 per cent of the volume, and 95 per cent in 89 per cent of the volume), while it becomes very

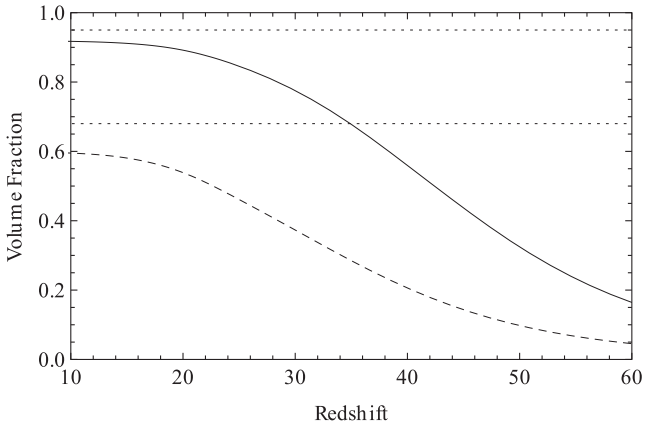

Figure 5. The fractional volume of the universe that contains 68 (dashed curve) or 95 per cent (solid curve) of the star-forming haloes as a function of redshift, where we consider just the contribution of velocity fluctuations to the inhomogeneity of star formation on 3 Mpc scales.

strong at $z = 60$ (68 per cent of stars in 4.6 per cent of the volume, and 95 per cent in 16 per cent of the volume).

4.3 Inhomogeneous gas fraction due to velocity and density fluctuations

In order to quantify the full degree of inhomogeneity and concentration of star formation, we must include the effect of density fluctuations as well. In this section we thus consider the full PDF of the halo gas fraction within 3 Mpc patches, where the fluctuations result from a combination of the relative velocity distribution considered in the previous section and density fluctuations. Specifically, the average density in a patch varies due to fluctuations on scales larger than its size. This average density follows a Gaussian distribution and is independent of the relative velocity within the same patch.

To find the modified halo mass function within a patch of a given overdensity δ_R and bulk velocity v_{bc}, we use the hybrid prescription (which combines the Sheth & Tormen 1999 mass function with the extended Press–Schechter model) introduced by Barkana & Loeb (2004) and generalized by Tseliakhovich et al. (2011) to include v_{bc}. The dependence of the gas fraction in haloes above the cooling mass on the two independent variables is illustrated in Fig. 6. The dependence on both δ_R and v_{bc} (each measured in terms of its root-mean-square value) is stronger at higher redshifts. At a given redshift, the dependence on δ_R is stronger (i.e. the slope is higher) when v_{bc} is higher, since in this case the large haloes (above the high cooling mass) are rarer and their abundance is more sensitive to the overdensity of the patch. If we consider the total range between 0 and 2σ, we find that density and velocity fluctuations make comparable contributions to the star formation fluctuations on the 3 Mpc scale. The relative importance of velocity increases with redshift and it will also increase if we consider larger scales. Even at $z = 20$ the velocity causes order unity fluctuations in the stellar density, and these fluctuations should be present at the large (100 Mpc) scales spanned by the velocity correlations.

The resulting full PDF of the halo gas fraction is shown in Fig. 7 (top panel), both for the star-forming haloes, and the starless gas minihaloes. The main effect of the bulk velocities is to shift the distributions towards lower gas fractions. At redshift 20, the effect is larger on the minihaloes. In Fig. 7 (bottom panel) we show the fraction of the volume of the universe (at the high gas fraction end

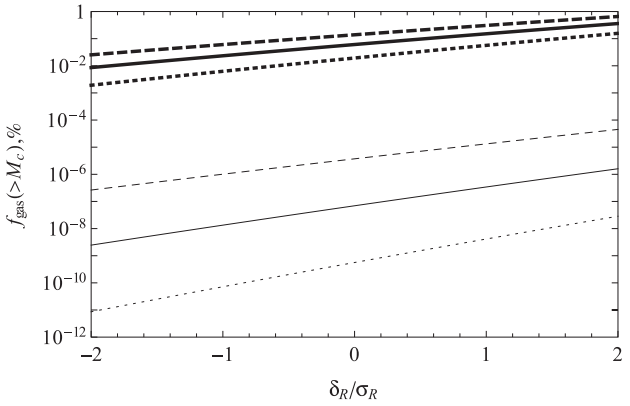

Figure 6. The percentage of the gas fraction in star-forming haloes at redshifts $z = 20$ (thick curves) and 40 (thin curves) as a function of the average overdensity δ_R in the 3-Mpc patch (normalized by its root-mean-square value σ_R) for various values of the relative velocity: no relative motion (dashed); $v_{\rm bc} = \sigma_{v_{\rm bc}}$ (solid) and $v_{\rm bc} = 2\sigma_{v_{\rm bc}}$ (dotted).

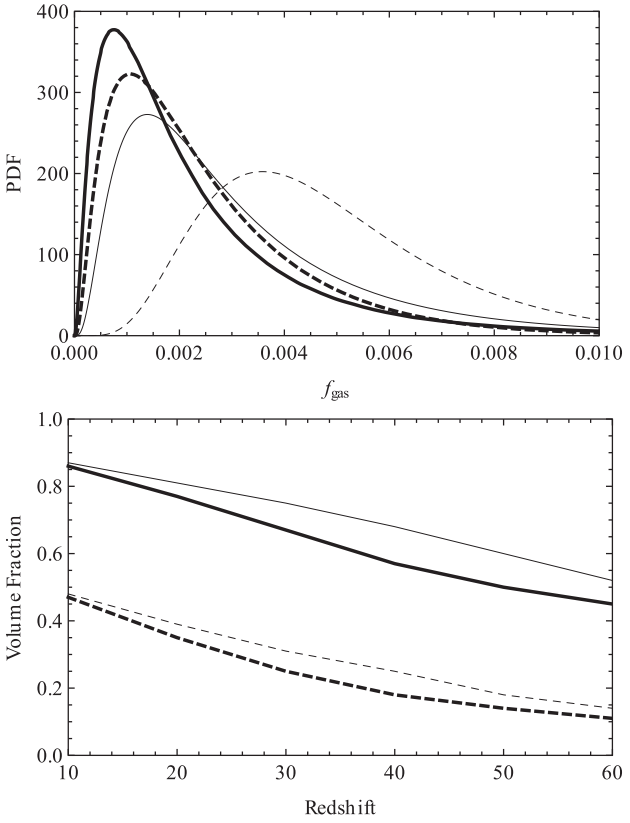

Figure 7. Top panel: the full PDF of the gas fraction at redshift $z = 20$. We show the PDF of the gas fraction in haloes above the cooling mass (solid curves) and the PDF of the gas fraction in starless minihaloes (dashed curves). We consider two cases: randomly distributed $v_{\rm bc}$ and δ_R (thick curves) and $v_{\rm bc} = 0$ but random δ_R (thin curves). Bottom panel: the fractional volume of the universe that contains 68 (dashed curves) and 95 per cent (solid curves) of the star-forming haloes, where we consider the full PDF in 3 Mpc patches. In each case we consider including the relative motion (thick curves) or not ($v_{\rm bc} = 0$, thin curves).

of the full PDF) that contains 68 or 95 per cent of the stars, with and without the velocity effect.

The volume concentration of star formation is a result of a complex interplay of the two sources of fluctuations. The global star formation is highest in the rare regions with both low bulk velocity and high overdensity, but more generally, one of these can compensate for the other. The effect of $v_{\rm bc}$ on star-forming haloes vanishes by $z \sim 10$, in agreement with our previous results, leaving just the effect of the local density. Even at somewhat higher redshifts (up to $z \sim 35$), the concentrating effect of the velocities on their own (Fig. 5) remains weaker than that of the densities alone (no-velocity case in Fig. 7), so at these redshifts the full case is dominated by the densities, and the concentrating effect of density is enhanced by including the velocities (which steepen the dependence on density; Fig. 6). At redshifts above ∼35, velocities dominate, and then including the density fluctuations (compared to averaging over them at each velocity) actually reduces the concentration since it allows low-velocity regions to contribute relatively more volume with high gas fractions (due to the steeper density dependence at high bulk velocity).

Specifically, at $z = 20$, density fluctuations alone (i.e. setting $v_{\rm bc} = 0$) would concentrate 68 per cent of the stars into 39 per cent of the volume and 95 per cent into 81 per cent of the volume. The addition of the bulk velocity provides a mildly increased concentration into 35 and 77 per cent of the volume, respectively. At redshift 60 the results are that 68 per cent of the stars are in 11 per cent of the volume and 95 per cent in 45 per cent (which is higher than in Fig. 5), compared to 14 and 52 per cent of the volume, respectively, at zero bulk velocity. The effect of the velocities should be more clearly apparent on scales larger than our 3 Mpc pixels, i.e. in addition to the small additional concentration that they cause (as seen in Fig. 7), their effect is to redistribute the star-forming regions to produce larger coherent regions of either high star formation or low star formation (voids).

We note that the assumption that the local overdensity on large scales δ_R and the streaming velocity $v_{\rm bc}$ are statistically independent is not perfectly accurate. A patch with a high local overdensity has expanded less than other patches, so that the peculiar velocity $v_{\rm bc}$ has not declined as much compared to the expansion. Indeed, we expect that $v_{\rm bc} \rightarrow v_{\rm bc}(1 + \delta_R/3)$. However, we have found that this correction makes only a small difference to the PDF (up to a 4 per cent relative error at $z = 60$, and less at lower redshifts).

5 THE FIRST STAR

In the previous sections we have discussed the conditions needed to initiate star formation. The main condition is that the halo mass must be large enough to allow molecular cooling. Given a large enough initial density fluctuation, a halo with a sufficiently large mass will form relatively early. The very first stars depend on extremely rare fluctuations, hence we need to average over the volume of the observable Universe $(14\,{\rm Gpc})^3$ in order to have the full statistical range needed to accurately estimate the formation time of the first star.

Because of computational limitations, numerical simulations can form stars only in a very limited cosmological context. For instance, Greif et al. (2011) studied star formation in a $(500\,{\rm kpc})^3$ volume and Stacy et al. (2012b) were limited to $(100\,h^{-1}\,{\rm kpc})^3$. In a small volume the chance of getting a rare high density fluctuation is quite small. Therefore the formation redshift of the first stars in simulations is greatly underestimated, with most simulations forming their first star below redshift 30 (i.e. when the Universe was >100 Myr old). The highest redshift where a star has formed in a simulation is $z = 47$ (∼53 Myr after the big bang; Reed et al. 2005).

Naoz, Noter & Barkana (2006) first applied these statistical considerations in order to predict the redshift of the first observable star (i.e. in our past light cone) analytically. They estimated the

redshift of the first star to be $z = 65$ (i.e. when the Universe was only 32.9 Myr old), using the 3-year *WMAP* set of cosmological parameters (Spergel et al. 2007) and assuming a minimum circular velocity for cooling of $V_{\rm cool} = 4.5\,{\rm km\,s^{-1}}$. In this section we generalize their method in order to account for the bulk velocities and estimate their impact on the epoch of the first star formation. This problem is particularly relevant since the effect of the relative velocity on star formation increases with redshift, and is thus at its maximum when we consider the very first star. We also study the sensitivity of the first-star redshift to various uncertainties.

Following Naoz et al. (2006) we calculate the mean expected number $\langle N(>z)\rangle$ of star-forming haloes that formed at redshift z or higher, but where the halo abundance is now averaged over the bulk velocity distribution at each redshift. This number is the ensemble-averaged number of stars, but we have only one Universe to observe. Hence, we expect Poisson fluctuations in the actual observed numbers. The probability of finding at least one star is then $1 - \exp[-\langle N(>z)\rangle]$, and (minus) the redshift derivative of this gives the probability distribution $p_*(z)$, where the probability of finding the first star between z and $z + {\rm d}z$ is $p_*(z)\,{\rm d}z$.

As shown in Fig. 8 (top panel), we find that in the absence of the bulk velocities, the first star would be most likely to form at $z = 69.9$, with a median $z = 70.3$ (corresponding to $t = 29.3$ Myr after the big bang). The difference with Naoz et al. (2006) is due to the changes in the cosmological parameters between *WMAP*3 and *WMAP*7, specifically the increased power on the relevant scales (since the increased spectral index has a larger effect than the reduced σ_8), and the decreased cooling mass in the $v_{\rm bc} = 0$ case compared to the value assumed by Naoz et al. (2006).

The relative velocity effect delays star formation, where for the very first star we find a delay of $\Delta z = 5.3$ (i.e. by $\Delta t = 3.6$ Myr). The first star is now most likely to form at $z = 64.6$, with a median $z = 65.0$ (corresponding to $t = 32.9$ Myr) that has a 1σ (68 per cent) confidence range $z = 63.9$–66.5 due to the Poisson fluctuations. In addition, the redshift of the first star is uncertain due to the current errors in the cosmological parameters and the uncertainty in the cooling mass. Regarding the cosmological parameters, the redshift of the first star is sensitive to the amount of power on the scale of the first haloes. The uncertainly of *WMAP*7 (Komatsu et al. 2011) in the amplitude of the primordial fluctuations (parametrized by σ_8) is $\Delta\sigma_8 = \pm 0.024$, which implies (for our optimal fit) an uncertainty of $\Delta z = \pm 2.2$ in the median redshift of the first star. The larger is σ_8, the earlier will the first star form. More generally, we include the current correlated errors in the full suite of standard cosmological parameters, and find a resulting $\Delta z = \pm 5.1$.

In order to estimate the impact of the current uncertainty in the effect of the bulk velocity on the minimum cooling mass, we estimate the redshift of the first star for each of the fits discussed in Section 3. We find in Fig. 8 (bottom panel) that the range of the SPH and MMH simulations is a $\Delta z = 1.5$, and the discrepancy between them and the AMR simulations is comparable. Thus, we conclude that the delay due to the bulk motion is substantial, but there are still significant uncertainties in it. In summary, we find the median redshift of the first star in our observable Universe to be

$$z = 65.0^{+1.5}_{-1.1}({\rm Poisson})^{+0.8}_{-1.5}({\rm simulations}) \pm 5.1({\rm cosmology}), \qquad (5)$$

or equivalently

$$t = 32.9^{+0.8}_{-1.1}({\rm Poisson})^{+1.1}_{-0.6}({\rm simulations})^{+4.2}_{-3.5}({\rm cosmology})\,{\rm Myr}.$$

Thus, current uncertainties in the values of the cosmological parameters dominate over the differences in the simulations and the irreducible Poisson fluctuations.

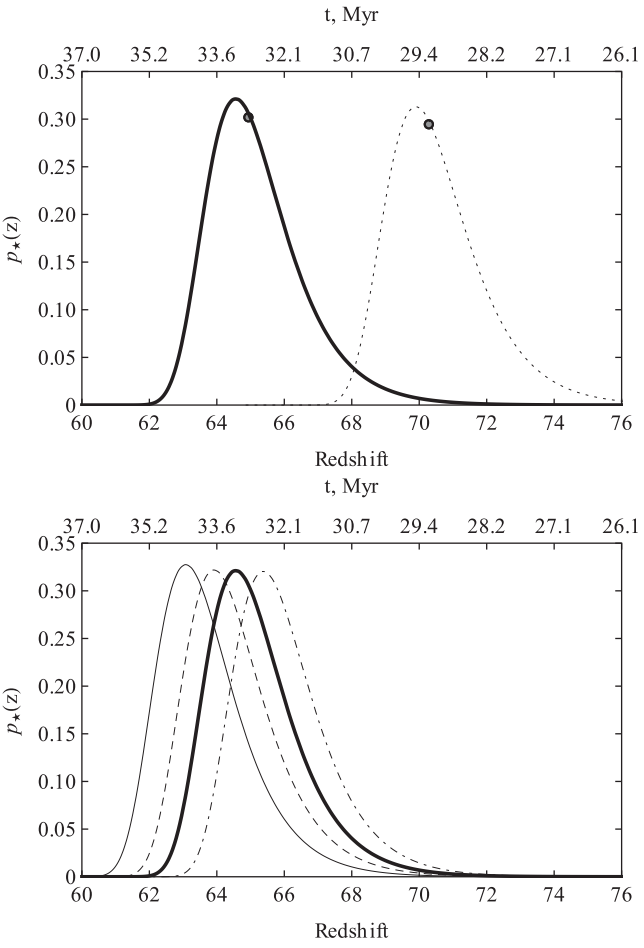

Figure 8. Top panel: the impact of the relative velocity on the redshift of the very first observable star. We plot the probability density of seeing the first star at a given redshift, including the effect of relative velocity for our optimal fit (solid curve), or without the effect of the velocity (i.e. for the same fit but with $v_{\rm bc} = 0$, dotted curve). The formation of the first star is delayed by $\Delta z = 5.3$ ($\Delta t = 3.6$ Myr) due to the relative velocity effect. We mark the median redshift of the first star for each distribution (•), which is $z = 65.0$ (corresponding to $t = 32.9$ Myr) in the case of the optimal fit to the SPH and MMH simulations and $z = 70.3$ ($t = 29.3$ Myr) in the no-velocity case. Bottom panel: the probability density of the redshift of the first star calculated for each of the fits of Fig. 1. The median redshifts of the first star (from left to right) are $z = 63.5$ ('fit' to the AMR simulations), $z = 64.3$ (fit to Greif et al. 2011), $z = 65.0$ (the optimal fit to the SPH and MMH simulations) and $z = 65.8$ (fit to Stacy et al. 2012b).

6 DISCUSSION

We have studied the impact of the relative motion between the gas and the dark matter on the formation of the first stars. We included a new effect found in recent small-scale hydrodynamic simulations. In particular, we fit their results to a physically motivated ansatz that expresses the minimum circular velocity of gas-cooling haloes as a simple function of the local bulk velocity when the halo forms. This result implies that in contrast to previous expectations, the minimum mass of star-forming haloes does not decrease with redshift, except in regions with very low values of the bulk velocity.

This result implies that the relative velocities produce three distinct effects: suppression of the halo abundance, suppression of the gas content within each halo and boosting of the minimum halo mass required for cooling. Quantitatively, we found that the halo abundance cut has a large effect on the two categories of haloes

(star-forming haloes and starless minihaloes), while the cooling mass boost primarily affects star-forming haloes and the suppression of gas content primarily affects the minihaloes. In total, at $z = 20$ the bulk velocities reduce the mean gas fraction in star-forming haloes by a factor of 1.8 and that in minihaloes by 3.1. Thus, even at $z = 20$ the velocity causes order unity fluctuations in the stellar density, and these fluctuations should be present at the large (100 Mpc) scales spanned by the velocity correlations.

The velocity dependence of the gas fraction tends to concentrate the global star formation into regions of low bulk velocity. In particular, at $z = 20$, 68 per cent of the stars are in the 54 per cent of the volume with the lowest velocity, and 95 per cent are in 89 per cent of the volume. Adding in the effect of density fluctuations tends to concentrate the global star formation into regions of both low bulk velocity and high overdensity. As a result, at $z = 20$, 68 per cent of the stars form within 35 per cent of the volume and 95 per cent in 77 per cent of the volume. This concentration effect becomes much stronger at higher redshifts.

The formation of the very first star is delayed by $\Delta t = 3.6$ Myr due to the bulk velocities. Given the updated cosmological and astrophysical parameters, the first star is now most likely to form at $t = 33.2$ Myr, with a median formation time $t = 32.9$ Myr after the big bang. In other words, the formation time is delayed by 11 per cent on average over the Universe with respect to the cosmic time at that redshift (which itself is 0.23 per cent of the present age of the Universe). Because of the combination of density and velocity fluctuations, the formation of stars begins at different times in different regions. This leads to a very inhomogeneous early universe. Although by redshift 20 most of the Universe is populated, the age of the oldest stars in each region is significantly different.

To make the novelty of our work clear, we now make a full comparison of the ingredients of our calculations with those in the previous literature. We start with Tseliakhovich & Hirata (2010), who discovered that the relative velocity effect is important. They only calculated the impact on the halo abundance, but this was sufficient for them to deduce the implication of large-scale fluctuations. However, their calculations had a number of simplifying assumptions: they calculated the baryon perturbations under the approximation of a uniform sound speed, and used the old Press–Schechter halo mass function.

Dalal et al. (2010) were the first to point out the effect of the relative velocity on suppressing the gas content of haloes. However, they made a number of simplifying approximations that we have relaxed here. These include the following.

(i) We have calculated the filtering mass ($M_{\rm F}$) from linear theory, while they took the effective value found in simulations in the standard (no relative velocity) case, and then multiplied it by a simple $v_{\rm bc}$-dependent ansatz.

(ii) We have allowed for a smooth transition between gas-rich haloes at $M \gg M_{\rm F}$ and gas-poor haloes at $M \ll M_{\rm F}$ as is suggested by simulations, rather than applying a step-function cut-off.

(iii) We have simultaneously included the dependence of the gas fraction in haloes on the large-scale matter overdensity δ_R and relative velocity $v_{\rm bc}$. This combines both the 'traditional' biasing model (which includes δ_R but not $v_{\rm bc}$) and the Dalal et al. (2010) treatment (which includes $v_{\rm bc}$ but not δ_R). We found that both effects are important (compare Sections 4.2 and 4.3).

(iv) We included the effect of $v_{\rm bc}$ on the halo mass function (Tseliakhovich & Hirata 2010), which Dalal et al. (2010) did not.

(v) Most importantly, we incorporated a cooling criterion for star formation, rather than scaling by the total gas content in haloes. The vast majority of the gas is in minihaloes that cannot cool, and because of their low circular velocities their ability to collect baryons is much more affected by $v_{\rm bc}$ than the star-forming haloes. This suggests that the effect of relative velocities on early star formation might be less than found by Dalal et al. (2010). However, we find that the inclusion of the other effects (mass function and cooling threshold, in addition to baryon fraction) does restore the expectation for order unity fluctuations, with exciting implications for observational 21-cm cosmology.

In part of this paper we closely followed Tseliakhovich et al. (2011). However, we fixed two inaccuracies in their power spectrum (in the normalization and the spectral slope) that gave substantially too much power on small scales. Then, our main goals were to include the new effect on the cooling mass based on simulations, to extend the calculations to the highest redshifts of star formation and to quantify the degree of concentration of star-forming haloes. With there now being three distinct effects of the bulk velocity, we also carefully studied the relative importance of these various effects.

ACKNOWLEDGMENTS

This work was supported in part by European Research Council grant 203247 (for AF) and by Israel Science Foundation grant 823/09 (for RB). DT and CMH were supported by the National Science Foundation (AST-0807337) and the US Department of Energy (DE-FG03-92-ER40701). CMH was also supported by the David and Lucile Packard Foundation.

REFERENCES

Abel T. L., Bryan G. L., Norman M. L., 2002, Sci, 295, 93
Barkana R., Loeb A., 2001, Phys. Rep., 349, 125
Barkana R., Loeb A., 2002, ApJ, 578, 1
Barkana R., Loeb A., 2004, ApJ, 609, 474
Bromm V., Larson R. B., 2004, ARA&A, 42, 79
Bromm V., Coppi P. S., Larson R. B., 2002, ApJ, 564, 23
Ciardi B., Scannapieco E., Stoehr F., Ferrara A., Iliev I., Shapiro P., 2005, MNRAS, 366, 689
Dalal N., Pen U.-L., Seljak U., 2010, J. Cosmol. Astropart. Phys., 11, 7
Fuller T. M., Couchman H. M. P., 2000, ApJ, 544, 6
Furlanetto S., Oh S. P., 2006, ApJ, 652, 849
Gao L., Abel T., Frenk C. S., Jenkins A., Springel V., Yoshida N., 2006, MNRAS, 378, 449
Greif T., White S., Klessen R., Springel V., 2011, ApJ, 736, 147
Haiman Z., Abel T., Madau P., 2001, ApJ, 551, 599
Hirata C. M., 2006, MNRAS, 367, 259
Iliev I., Scannapieco E., Martel H., Shapiro P., 2003, MNRAS, 341, 81
Iliev I., Scannapieco E., Shapiro P., 2005, ApJ, 624, 491
Komatsu E. et al., 2011, ApJS, 192, 18
Lewis A., Challinor A., 2007, Phys. Rev. D, 76, 083005
Maio U., Ciardi B., Dolag K., Tornatore L., Khochfar S., 2010, MNRAS, 407, 1003
Maio U., Khochfar S., Johnson J. L., Ciardi B., 2011a, MNRAS, 414, 1145
Maio U., Koopmans L. V. E., Ciardi B., 2011b, MNRAS, 412, L40
Naoz S., Barkana R., 2005, MNRAS, 362, 1047
Naoz S., Barkana R., 2007, MNRAS, 377, 667
Naoz S., Noter S., Barkana R., 2006, MNRAS, 373, L98
Naoz S., Yoshida N., Gnedin N. Y., 2012, ApJ, 747, 128
Petkova M., Maio U., 2011, MNRAS, 422, 3067
Reed D. S., Bower R., Frenk C. S., Gao L., Jenkins A., Theuns T., White S. D. M., 2005, MNRAS, 363, 393
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Spergel D. N. et al., 2007, ApJS, 170, 377
Stacy A., Greif T. H., Bromm V., 2012a, MNRAS, 422, 290

Stacy A., Bromm V., Loeb A., 2012b, MNRAS, 413, 543
Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, ApJ, 474, 1
Tseliakhovich D., Hirata C. M., 2010, Phys. Rev. D, 82, 083520
Tseliakhovich D., Barkana R., Hirata C. M., 2011, MNRAS, 418, 906
Turk M. J., Clark P., Glover S. C. O., Greif T. H., Abel T., Klessen R., Bromm V., 2011, ApJ, 726, 55
Wise J. H., Abel T., 2007, ApJ, 665, 899
Yoo J., Dalal N., Seljak U., 2011, JCAP, 07, 018
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., Omukai K., Hernquist L., Abel T., 2006, ApJ, 652, 6

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