COSMOLOGICAL H II BUBBLE GROWTH DURING REIONIZATION

MIN-SU SHIN, HY TRAC, AND RENYUE CEN
Princeton University Observatory, Peyton Hall, Princeton, NJ 08544-1001

Received 2007 August 31; accepted 2008 March 13

ABSTRACT

We present general properties of ionized hydrogen (H II) bubbles and their growth based on a state-of-the-art, large-scale (100 Mpc h^{-1}) cosmological radiative transfer simulation. The simulation resolves all halos with atomic cooling at the relevant redshifts and simultaneously performs radiative transfer and dynamical evolution of structure formation. Our major conclusions include the following: (1) For significant H II bubbles, the number distribution is peaked at a volume of ~0.6 Mpc^3 h^{-3} at all redshifts. But at z \leq 10, one large, connected network of bubbles dominates the entire H II volume. (2) H II bubbles are highly nonspherical. (3) The H II regions are highly biased with respect to the underlying matter distribution, with the bias decreasing with time. (4) The non-Gaussianity of the H II region is small when the universe becomes 50% ionized. The non-Gaussianity reaches its maximum near the end of the reionization epoch z \sim 6. But at all redshifts of interest there is a significant non-Gaussianity in the H II field. (5) Population III galaxies may play a significant role in the reionization process. Small bubbles are initially largely produced by Population III stars. At z \geq 10 even the largest H II bubbles have a balanced ionizing photon contribution from Population II and Population III stars, while at z \leq 8 Population II stars start to dominate the overall ionizing photon production for large bubbles, although Population III stars continue to make a nonnegligible contribution. (6) The relationship between halo number density and bubble size is complicated, but a strong correlation is found between halo number density and bubble size for large bubbles.

Subject headings: cosmology: theory — early universe — intergalactic medium —
large-scale structure of universe — methods: numerical — radiative transfer

Online material: color figures

1. INTRODUCTION

Understanding the reionization process is a key challenge in cosmology. Current observations provide extremely useful but still limited information with respect to the epoch of reionization. On one hand, the absorption spectrum observations of high-redshift quasars from the Sloan Digital Sky Survey (SDSS) strongly suggest that the final reionization episode comes to completion at z \sim 6 (e.g., Becker et al. 2001; Fan et al. 2002; Barkana 2002; Cen & McDonald 2002). On the other hand, the Wilkinson Microwave Anisotropy Probe (WMAP) observations (Page et al. 2007; Spergel et al. 2007) infer that the intergalactic medium (IGM) is already significantly ionized by some very early time, possibly z \sim 15. In combination, it suggests that the reionization process may be quite complex and perhaps nonmonotonic (e.g., Cen 2003; Haiman & Holder 2003; Wyithe & Cen 2007). In addition, some other observations and analyses have also provided useful constraints on the reionization process. Theuns et al. (2002) and Hui & Haiman (2003) have shown that the observed temperature of the Lyα forest at redshift z = 2–4 requires the cosmological reionization to have occurred no earlier than z = 9–10, although He II reionization somewhat complicates this constraint. Analyses based on the SDSS quasar Strömgren sphere size suggest that the neutral hydrogen fraction at z = 6.3 is a few tens of percent (Wyithe & Loeb 2004; Mesinger & Haiman 2004), whereas analyses based on the evolution of Lyα emitters from z = 6.5 to 5.7 imply that a partially neutral IGM fraction of ~0.25 is consistent with observations (Malhotra & Rhoads 2004, 2006; Stern et al. 2005; Haiman & Cen 2005). But the recent observations of Lyα emitters begin to support complete reionization at z \sim 6, as the SDSS quasar observations imply (e.g., Kashikawa et al. 2006; Lidz et al. 2007a).

While a fully self-consistent reionization picture is far from being painted (see Barkana & Loeb 2001; Loeb et al. 2008; Fan et al. 2006 for a review), some key breakthroughs may rest in the redshift range z = 7–15, where upcoming observations, including 21 cm radio and cosmic microwave background observations, may be able to provide useful constraints. To understand the reionization process, several elegant analytical and semi-analytical models have been developed and predictions made (Miralda-Escudé et al. 2000; Barkana & Loeb 2002; Furlanetto et al. 2004, 2006; Furlanetto & Oh 2005; Zahn et al. 2007; Cohn & Chang 2007). These methods provide ways to economically explore the large parameter space and have provided very important insights with respect to H II bubble evolution, as well as the evolution of global quantities such as the ionization fraction. These methods, however, make necessary simplifying assumptions, which limit the scope of their applicability or the accuracy of the predictions.

We take a complementary approach by making detailed numerical simulations with fewer assumptions. Earlier simulations forced a choice to be made between simulating a large volume with very limited resolution or too small a volume with sufficient resolution (e.g., Gnedin 2000; Ciardi et al. 2000; Razoumov et al. 2002; Sokasian et al. 2003, 2004; Gnedin 2004). However, both a large volume (~100 Mpc h^{-1}) and an adequate resolution are necessary in order to follow the reionization sources and sinks properly. A large simulation box is required because the sources in question are extremely highly biased (Barkana & Loeb 2004; Trac & Cen 2007) and the bubbles have sizes of tens of Mpc prior to complete percolation (e.g., Furlanetto et al. 2004), whereas resolving the bulk of sources of halo mass ~10^8 M⊙ dictates, at least, the mass resolution of ~10^9 M⊙ and spatial resolution of a few kpc. With a unique hybrid (dark matter + baryons + stars + radiation) computational code (Trac & Cen 2007), we have crossed...
the threshold enabling us to simultaneously simulate a large volume and have an adequate resolution to identify the bulk of the ionizing sources at high redshift, as well as to have an approximate yet adequate treatment of radiation transfer. Recent direct simulations performed by other groups (Iliev et al. 2006; Zahn et al. 2007) have yielded very important results, but the inability to adequately resolve dark matter halos (DM halos) of $\sim 10^8 M_\odot$ renders the results uncertain. Additional features of our method include (1) simultaneously following the evolution of structure formation and radiative transfer (RT), instead of performing RT as a postprocessing step (Iliev et al. 2007) or adding unresolved halos analytically (McQuinn et al. 2007); and (2) a self-consistent, as a postprocessing step (Iliev et al. 2007) or adding unresolved formation and radiative transfer (RT), instead of performing RT to adequately resolve dark matter halos (DM halos) of $\sim 10^8 M_\odot$. Therefore, our simulations have more direct numerical treatments of radiation sources, small-scale clumping, and self-shielding than both postprocessing and semianalytical models.

In this paper we analyze the evolution of H II bubbles in a 100 Mpc $h^{-1}$ simulation box. In particular, we study how H II bubbles grow and what physics determines the growth. Understanding the evolution of H II bubbles is closely related to predicting future observation results of the redshifted 21 cm line and cosmic microwave background (see Fan et al. 2006 for a review). Among large-volume reionization simulations, the morphology of H II regions was studied in Iliev et al. (2006), Zahn et al. (2007), and McQuinn et al. (2007). In the details of our simulation are described. Results of simulations are given in § 3, followed by discussion and conclusions in § 4. We use the following cosmological parameters based on the WMAP3 results (see Spergel et al. 2007 and references therein): $\Omega_m = 0.26, \Omega_b = 0.044, h = 0.72, \sigma_8 = 0.77$, and $n_s = 0.95$. Throughout the paper, both length and volume are given in comoving units.

2. SIMULATIONS

Our simulation is generally based on the numerical methodology described in Trac & Cen (2007) and similarly utilizes an N-body algorithm for dark matter, a star formation prescription, and a RT algorithm for ionizing radiation. However, we have taken some simpler approaches in this initial step to satisfy the computational challenge of large-volume, high-resolution simulations of cosmic reionization. In particular, we use an alternative RT algorithm and do not use a halo model for prescribing baryons and star formation. Here we summarize the main components and describe the modifications.

In an $L = 100$ Mpc $h^{-1}$ simulation box, a high-resolution N-body calculation for 2880$^3$ dark matter particles on an effective grid with 11,520$^3$ cells is performed using a particle multimesh code (Trac & Pen 2006). The particle mass resolution is $3.02 \times 10^6 M_\odot$, and approximately 33 particles make up a $10^6 M_\odot$ $h^{-1}$ halo. The comoving grid spacing is 8.68 kpc $h^{-1}$, and approximately 12 cells make up the virial volume of a $10^7 M_\odot$ $h^{-1}$ halo. We identify dark matter halos in postprocessing rather than during the course of the simulation.

We do not use the halo model of Trac & Cen (2007) for prescribing baryons and star formation. Instead, an alternative approach is taken, where we calculate the local matter density $\rho_m$ and velocity dispersion $\sigma_v$ for each particle. The baryons are assumed to trace the dark matter distribution on all scales, and we obtain the local baryon density $\rho_b = \rho_m(\Omega_b/\Omega_m)$ and gas temperature $T = \mu m_p k_B/3k$ (3K). Star formation is only allowed to occur in particles with densities $\rho_m > 100 \rho_{crit}(z)$ and temperatures $T > 10^4$ K. This cut in the temperature-density phase space effectively restricts star formation to regions within the virial radius of halos where efficient atomic line cooling allows the gas to dissipate energy and further collapse to very high densities. We also differentiate between the first-generation Pop III stars and the second-generation Pop II stars by following the chemical enrichment of the interstellar medium and IGM as described in Trac & Cen (2007).

Our RT algorithm for ionizing radiation is based on a photon-advection scheme, which is much less computationally expensive than ray tracing. Particles, sources, and photons are collected on a RT grid with 360$^3$ cells. Each cell spans 278 comoving kpc $h^{-1}$, and the RT time step is set by the light-crossing time. For a source cell, the excess source photons are propagated to the 26 neighboring cells. The advection is photon-conserving, and the isotropic redistribution uses a weighting function that is proportional to $1/r^2$. However, for a nonsource cell with excess radiation, the advection is generally anisotropic. For a H II region, photons originating from a central source propagate in the direction coinciding with decreasing radiation flux. Therefore, we propagate photons by looking for gradients in the radiation field. Consider a nonsource cell with cell indices $(i,j,k)$ and radiation density $n_s(i,j,k)$. Radiation can propagate to an adjacent cell with indices $(i + di, j + dj, k + dk)$ if either of the gradient terms,

$$\Delta_i = n_s(i + di, j + dj, k + dk) - n_s(i, j, k),$$

$$\Delta_j = n_s(i, j, k) - n_s(i - di, j - dj, k - dk),$$

is negative. The first gradient term indicates the downstream expansion direction, while the second term indicates the upstream direction. For the 26 possible neighbor cells, we count how many cells satisfy the above criteria and redistribute the excess photons equally among them. If none of the 26 neighbor cells satisfy the above criteria, then the central cell is a convergent point, and we redistribute the photons equally among the neighbors.

For an isolated H II region with one central source, if one gradient term is false, the other is generally false too. However, this may not be the case near the interface of merging or overlapping H II regions. If a weaker radiation field is trying to expand into the vicinity of a stronger radiation field, then the first gradient term is positive even though the second gradient term is negative. Therefore, it is necessary to use both of the gradient terms to determine the direction of radiation propagation. In the Appendix we compare the photon-advection scheme with the ray-tracing scheme of Trac & Cen (2007) and show that this simpler approach correctly captures the RT for a significant majority of the reionization epoch.

For each particle, we store 12 floating-point variables: three-dimensional position, three-dimensional velocity, matter density, baryon density, temperature, stellar fraction, H i fraction, He i fraction, and He ii fraction. We calculate the ionization and recombination for each particle individually rather than using the lower resolution density field defined on the RT grid. This allows us to correctly account for the clumping factor and self-shielding of dense gas down to small scales of several comoving kpc $h^{-1}$.

The simulation was run at the National Center for Supercomputing Applications on a shared-memory SGI Altix with Itanium 2 processors. We used 512 processors, 2 Tbyte of memory, and approximately 80,000 CPU hours. With nearly 24 billion particles, this is the largest cosmological N-body simulation run to date. This is also the first reionization simulation with an $L = 100$ Mpc $h^{-1}$ simulation box that can directly resolve dark matter halos down to virial temperatures of $10^4$ K.

In postprocessing, we have identified dark matter halos using a friends-of-friends (FoF) algorithm (Davis et al. 1985) with a standard linking length of $b = 0.2n^{-1/3}$, where $n$ is the mean
particle number density. Figure 1 compares our mass functions with the analytical prediction of Press & Schechter (1974) and the empirical prediction of Warren et al. (2006). For redshifts $z < 10$, our results are in very good agreement with recent works on the mass function of high-redshift dark matter halos (Reed et al. 2007; Lukic et al. 2007; Cohn & White 2008). For higher redshifts $z \gtrsim 15$, we systematically underresolve halos because our starting redshift of $z = 60$ is too late to capture the nonlinear gravitational collapse. Reed et al. (2007) have suggested that simulations must start $\approx 10-20$ expansion factors before the redshift at which results converge. We have found that we can accurately capture the formation of halos at $z \approx 15-20$ with a smaller simulation starting at $z = 300$ (Trac & Cen 2007).

Our star formation prescription resembles those used in hydrodynamic simulations (e.g., Springel & Hernquist 2003), and we obtain very similar results. In Figure 2 we compare the star formation rate $\dot{\rho}(z)$ with an analytical model calibrated using hydrodynamic simulations (Hernquist & Springel 2003), corrected for our cosmology. The overall shape is in very good agreement at all relevant redshifts. At $z = 6$, our amplitude is 1.25 times larger, and this difference is simply due to the fact that we have chosen a coarse value of $c_s = 0.06$ for the star formation efficiency. For our purposes, the overall amplitude of the star formation rate is unimportant, since it is degenerate with the radiation escape fraction. We have correspondingly chosen a radiation escape fraction of $f_{esc} = 0.15$ in order to have the redshift of complete overlap occur at $z \approx 6$.

3. RESULTS

The first reionization region appears when areas of dense baryons turn on star formation. Once the radiation sources begin ionization around them, the global ionization regions appear as shown in Figure 3. At the early stage of reionization, isolated ionization regions are easily found (see $z = 11.228$ in Fig. 3). As more radiation sources develop, the isolated bubbles get connected along filaments. The process of H ii bubble merging is very complex and difficult to treat without detailed simulation. Visually, in Figure 3 the computed H ii bubbles do not appear to be close to spheres, as is confirmed quantitatively below. After $z \approx 10$ one large connected network of H ii bubbles begins to dominate the simulation volume, and the H ii percolation process enters the “cannibalistic” phase, where the dominant H ii bubble

---

Figure 1.—Dark matter halo mass functions. Dark matter halos are identified using a FoF algorithm with a standard linking length of 0.2 times the mean interparticle spacing. Our mass functions are in better agreement with Warren et al. (2006; solid line) than with Press & Schechter (1974; dashed line) for $z \lesssim 10$. At $z \gtrsim 15$, we underresolve the halos because of the late starting redshift of $z = 60$. A smaller simulation starting at $z = 300$ correctly captures the abundance of high-redshift halos (open circles).
In the following subsections, we present quantitative analysis of the found $H_\alpha$ bubbles. We present the size distribution of $H_\alpha$ bubbles that are not limited by our radiation transfer resolution and should be real. The peak at $\sim 0.6$ Mpc$^3$ $h^{-3}$ probably represents typical mature $H_\alpha$ bubbles produced by individual galaxies (probably with their satellite galaxies), which may subsequently merge with other bubbles. This peak is maintained at all times until the percolation permeates the entire volume to complete the reionization, as seen from the difference between $z = 6.3$ and 5.7 in Figure 4. The properties of these characteristic sizes are further analyzed and understood in §§ 3.1.3 and 3.1.4.

The appearance of few dominant large bubbles is easily found by the change of volume fraction for different sizes of bubbles. As shown in Figure 4, the volume fraction is dominated by a few large bubbles after $z = 13.5$. This result is also visually recognized in Figure 3. In the plot of $z \sim 11.2$, large bubbles begin to appear while many small bubbles around newly formed sources dominate the number fraction in Figure 4. Therefore, the characteristic sizes of 0.6, 0.03, and 0.006 Mpc$^3$ $h^{-3}$ can be maintained until the scale of large bubbles is big enough to percolate small bubbles that are distributed between large bubbles.

The volume distribution of $H_\alpha$ bubbles in Iliev et al. (2006) can be compared to our results. Both simulations use the same definition of $H_\alpha$ bubbles, and the bubbles are found by the FoF method. The first difference is that in our simulation, large bubbles of $\geq 10^4$ Mpc$^3$ $h^{-3}$ appear later than $z \sim 15$ when the global volume-weighted ionization fraction is about 0.04. But Iliev et al. (2006) found the existence of those large bubbles around $z \sim 15$ when in their simulation the global volume-weighted ionization fraction is 0.05. This difference is also related to the different characteristic sizes found in our simulation. In our simulation, most $H_\alpha$ bubbles are much smaller than $\geq 10^4$ Mpc$^3$ $h^{-3}$. Even the found characteristic size of intermediate-size bubbles is $\sim 0.6$ Mpc$^3$ $h^{-3}$. Iliev et al. (2006) generally show larger bubbles for all ranges of the global ionization fraction and formation of large bubbles at earlier times than ours. Moreover, the largest fraction of bubbles is explained by intermediate-size bubbles in Iliev et al. (2006), while our results show the dominance of small bubbles for almost the entire simulation time.

We think this is probably due to the difference in the adopted cosmological models in which the model used by Iliev et al. (2006) has a higher $\sigma_8$, and the fact that they require the universe to be ionized much earlier, such that their computed $t_e$ matches the results from the first-year WMAP results. The low mass resolution of Iliev et al. (2006) also does not permit the formation of small bubbles, because it cannot capture the formation of low-mass DM halos (Zahn et al. 2007). Finally, the implementation of baryon physics like ionizing sources in their simulation is quite different from our prescription given in § 2. The importance of this difference is explained in §§ 3.1.3 and 3.1.4.

In Figure 4 we also present the bubble size distribution based on the definition of Zahn et al. (2007) for comparison. The same bubble boundary threshold of 90% is adopted in this measurement. For global volume-weighted ionization fractions similar to those of Zahn et al. (2007), the size distribution from our simulation has its peak at a smaller radius than their size distribution. The existence of many resolved small bubbles results in the difference in the size distribution, even though we adopt the same definition as Zahn et al. (2007). For example, we find two peaks of size distribution at $\sim 0.1$ and 0.5 Mpc $h^{-1}$ for the volume-weighted.

3.1. Size of $H_\alpha$ Bubbles
3.1.1. Characteristic Size

The size of $H_\alpha$ bubbles is a basic quantity that has received significant attention (Furlanetto & Oh 2005; Furlanetto et al. 2006; Iliev et al. 2006; McQuinn et al. 2007; Zahn et al. 2007). We define $H_\alpha$ bubbles as follows. First, we mark all the cells that are ionized at a $>50\%$ level. Then, a bubble is defined by all such cells that are connected by at least one face; this is practically done by grouping cells using a FoF algorithm with a linking length equal to the simulation data output cell size, i.e., $\sim 0.14$ Mpc $h^{-1}$.

In addition, at least two cells are needed to make up a single bubble.

Since the shape of $H_\alpha$ bubbles is not close to a sphere, as shown in § 3.2, we present in Figure 4 the volume distribution of the found $H_\alpha$ bubbles instead of the radius of a sphere. Note that in previous analytical studies $H_\alpha$ bubbles are often assumed to be spherical (e.g., Pritchard et al. 2007) or have a characteristic size (e.g., McQuinn et al. 2007). We present the size distribution of $z = 5.7, 6.3, 8.9, 13.5, 17.2$, and 20.6.

In this paper we define the characteristic size of bubbles based on their number fraction, not their volume fraction as used in Iliev et al. (2006). Therefore, our characteristic sizes well describe the existence of many small nonconnected bubbles. Meanwhile, the definition of Zahn et al. (2007) is affected by the volume occupied by $H_\alpha$ bubbles because it measures an ionization fraction within a certain smoothing radius.

The volume distribution by the number fraction shows that there are three characteristic sizes of $H_\alpha$ bubbles before the simulation box is dominated by a single bubble. The characteristic volumes are $0.6, 0.03$, and $0.006$ Mpc$^3$ $h^{-3}$. These characteristic volumes do not change even though more regions become ionized. But the least volume of the possible bubbles is $\sim 0.005$ Mpc$^3$ $h^{-3}$, because our bubble-finding method defines the smallest bubble as at least two connected simulation cells. Therefore, the smallest volume peak in the $H_\alpha$ bubble distribution may be limited by our radiation cell size and should be considered as an upper limit for the peak at the smallest scales.

The other two characteristic sizes of $\sim 0.03$ and $\sim 0.6$ Mpc$^3$ $h^{-3}$ are real features of the size distribution. The peak at $\sim 0.03$ Mpc$^3$ $h^{-3}$ represents $H_\alpha$ bubbles that are not limited by our radiation transfer resolution and should be real. The peak at $\sim 0.6$ Mpc$^3$ $h^{-3}$ probably represents typical mature $H_\alpha$ bubbles produced by individual galaxies (probably with their satellite galaxies), which may subsequently merge with other bubbles. This peak is maintained at all times until the percolation permeates the entire volume to complete the reionization, as seen from the difference between $z = 6.3$ and 5.7 in Figure 4. The properties of these characteristic sizes are further analyzed and understood in §§ 3.1.3 and 3.1.4.

Fig. 2.—Comoving star formation rate. The total star formation rate (solid line) from Pop III (long-dashed line) and Pop II (short-dashed line) is consistent with that of Hernquist & Springel (2003; dotted line).
Fig. 3.—Distributions of ionization fraction at redshifts $z \approx 13.5, 11.2, 9.3, \text{and } 7.2$. Highly ionized regions are shown in black, while the ionization fraction below 50% is shown in white. This is a plot of one slice of the simulation box. Each side of the plot is 100 Mpc $h^{-1}$. The global ionization fractions are $\sim 10\%, 30\%, 50\%, \text{and } 90\%$ for $z \approx 13.5, 11.2, 9.3, \text{and } 7.2$, respectively.
ionization fraction of 0.12, while Zahn et al. (2007) found a single peak at \(C_{24}0.7 \, \text{Mpc} \, h^{-1}\) for the volume-weighted ionization fraction of 0.11. Comparing the results for the volume-weighted ionization fraction of 0.52, our distribution shows a single peak size, but the size is \(C_{24}2.5 \, \text{Mpc} \, h^{-1}\), which is smaller than the \(C_{24}4.5 \, \text{Mpc} \, h^{-1}\) of Zahn et al. (2007). In particular, the resolved small bubbles make the peak probability of the size distribution, i.e., the height of the peak, insensitive to the change of a global ionization fraction. Therefore, even with the same definition of bubble size, our result shows a slightly different result, although the general result is consistent.

3.1.2. Bubble Merger History

As it turns out, the percolation of bubbles is also related to the transition of major ionizing photon production from Pop III to Pop II stars. To shed light on the overall reionization process, we follow the change of ionizing environment at four different locations in the simulation box that are occupied by four individual bubbles that all started at \(z = 21\), shown in Figure 5. In the plot, we track the size change of the bubbles that occupy the four locations. We see that these bubbles initially formed at about the same time but at four different locations and display a variety of histories, clearly due to the rich, disparate structure formation histories and the evolution of the ionizing environment. One striking and very clear signature is some intermittent, rapid downturns in the \(H_{\text{II}}\) volume during an individual location’s evolution history. This is due to a rapid transition of the IMF from Pop III to Pop II in the bubble, which in general occurs in the redshift range \(z = 10–15\). This is consistent with the overall contributions of Pop II and Pop III stars as a function of redshift shown in Figure 6, where we see that ionizing photons from Pop II stars begin to become more dominant between \(z = 10\) and 15. But this transition occurs spatially at different times, such that each bubble position has a

---

**Fig. 4.**—Number and accumulated volume fraction distributions of \(H_{\text{II}}\) bubble volumes (left) and bubble size distribution (right). At redshifts \(z = 20.6, 17.2, 13.5, 8.9, 6.3, \) and 5.7, the number (solid line) and accumulated volume distribution (dashed line) are given with the total number of bubbles, \(N_{\text{tot}}\), and the global volume-weighted mean ionization fraction, \(<x>\). Before large bubbles that are comparable to the simulation box size appear around a redshift of 6, bubble size distributions show three peaks that have volumes of about 0.6, 0.03, and 0.006 \(\text{Mpc}^3 \, h^{-3}\). The domination of one large bubble appears at \(z < 10\), as shown in the accumulated volume fractions of bubbles. In the right panel, we present the bubble size distributions that are measured as in Zahn et al. (2007) at \(z = 13.2, 12.1, 10.5, 9.0, \) and 8.0, corresponding to the global volume-weighted mean ionization fractions of 0.12, 0.18, 0.32, 0.52, and 0.75 from the small to large peak size.

---

**Fig. 5.**—Bubble merger history. The evolution of the bubble size is given for four positions that are occupied by the four randomly selected smallest bubbles at \(z = 21.5\). The bubbles that occupy the four positions are tracked, and their volumes are measured. This bubble merger history is well explained by the change of a dominant stellar population and its ionizing photon production rate between \(z \sim 10\) and \(\sim 15\), as shown in Fig. 7. During that transition, bubbles experience reconnection to other bubbles.
slight different history. We note that this transition time is not the same as the epoch when the star formation rate is dominated by Pop II stars, since Pop III stars are more efficient producers of ionizing photons than Pop II stars for a given star formation rate. McQuinn et al. (2007) pointed out that the nature of the ionizing sources plays an important role in shaping the bubbles, and our results quantitatively confirm their conclusion by being able to perform a detailed calculation that allows for a spatially resolved transition from Pop III to Pop II stars.

The merger history of the bubbles is also consistent with the changing total number of bubbles and the dominance of the largest bubble over small bubbles after $z < 10$. As percolation of the bubbles proceeds, the total volume of H ii bubbles increases while the number of individual bubbles decreases, as shown in Figure 7. That is, percolation of the bubbles becomes so important that the largest network of connected bubbles dominates the simulation box. Although the number of small bubbles also begins to decrease after $z < 10$, the fraction of small bubbles increases because the size of the bubbles is not large enough to connect to other bubbles (see Fig. 4).

3.1.3. Dark Matter Halo and Bubble Size Distribution

Basically, the clustering of radiation sources affects the percolation domination and the spatially varying transition epoch from Pop III star formation to Pop II star formation. Because the two small characteristic size peaks found in Figure 4 are maintained by the formation of new radiation sources, the correlation between the bubble sizes and the DM halos can be expected to explain the physical reason behind the bubble size distribution.

In Figure 8 we show the number density of halos as a function of H ii bubble volume. All DM halos whose centers are inside the bubble are said to belong to the bubble. We separate the member DM halos into three separate mass ranges: $M_{\text{halo}} < 10^8 M_{\odot} h^{-1}$, $10^8 M_{\odot} h^{-1} \leq M_{\text{halo}} < 10^9 M_{\odot} h^{-1}$, and $10^9 M_{\odot} h^{-1} \leq M_{\text{halo}}$. We see that there is an upturn of halo number density for all mass ranges for the smallest bubbles. This indicates that these bubbles are rapidly expanding due to newly formed sources while the number of halos stays approximately constant. These smallest bubbles do not contain DM halos of $M_{\text{halo}} > 10^9 M_{\odot}$, corresponding to the peaks of $\sim 0.11$ and $0.19$ Mpc $h^{-1}$ as found in $\S$ 3.1.1. The halo number density decreases as one moves to the right and then levels out at approximately $-1$ to 0 of the displayed x-axis. The most interesting feature is that there is a sudden rise in halo number density for all masses toward the largest bubbles, indicating that percolation is expected to continue with time, consistent with our analysis of actual bubble size evolution. The boundary between the percolation dominance and the source formation dominance is the characteristic volume of $\sim 0.6$ Mpc$^3 h^{-3}$. This feature is described as the beginning of the flat DM halo density distribution in Figure 8. A large drop in the minimum number density is caused by differences of bubble age among DM halos of the same mass. For example, the same mass DM halos of $10^{8.5} M_{\odot} h^{-1}$ collapse at different epochs, which results in the differentially matured H ii bubbles around them. Thus, the number density of DM halos can have a large dispersion, in particular, for the intermediate-mass DM halos.

As characteristic sizes of H ii bubbles do not change much until $z \sim 6$, the DM halo distribution does not change, as shown for $z \sim 8$ and 10. Therefore, we find that the size distribution of H ii bubbles is mainly regulated by an overall mass density that determines the evolution of baryons. The latter property indicates highly biased galaxy formation and suggests that stellar mass density is expected to rapidly rise in high-density regions. Whether this property may be preserved requires more detailed analysis.

3.1.4. Dependence of Bubble Size on Ionizing Photon Production

The rate of ionizing photon production per unit volume is likely the main physical factor in determining the evolution of
H II bubbles, while the broad correlation between star formation rate and DM halo density has likely produced some of the interesting features presented in the previous subsection, because the latter is essentially proportional to the total, integrated ionizing photon number density. An ionizing photon production rate inside H II bubbles is estimated to be proportional to the star formation rate times the ionizing photon production efficiency that is a function of the IMF. Figure 9 shows the ionizing photon production rate, separately from Pop III and Pop II stars, as a function of the bubble size at z = 8 and 10, respectively. The flat part of the average ionizing photon production rate represents the value, which is approximately set by the average baryon number density and recombination rate inside the bubbles. It is clear that for the smallest bubbles with size \( \lesssim 0.3 \) Mpc \( h^{-1} \), Pop III stars dominate the ionizing photon production rate at both \( z = 10 \) and 8. This indicates that these bubbles are produced by the first generation of stars formed in those locations which are likely relatively removed from large halos where star formation has gone through more than one generation and the gas has been significantly enriched with metals, consistent with Figure 8. We expect that Pop III galaxies may be present in small bubbles until \( z \sim 6 \). For the large bubbles (size \( \gtrsim 10 \) Mpc \( h^{-1} \)) the contributions from Pop III and Pop II stars seem balanced at \( z \sim 10 \). But, by \( z \sim 8 \), these large bubbles begin to be dominated by Pop II stars, although Pop III stars continue to make significant contributions of ionizing photons.

We show that careful consideration of this transition is needed to understand the reason we find the bubble size distribution given in Figure 4. We believe that the overall evolution of H II bubbles depends on detailed modeling of Pop II and Pop III stars. Neglect or nondetailed treatment (e.g., Zahn et al. 2007; Furlanetto & Oh 2005; Iliev et al. 2007) of the Pop III stars continue to make significant change in the characteristics of bubble evolution.

3.2. Shape of H II Bubbles

The shape of H II bubbles is far from a sphere, unlike what is assumed in most analytical studies. Our result is the first effort to present a quantitative measurement of H II bubble shape in the

Fig. 8.—Number density distribution of dark matter halos inside H II bubbles at \( z \sim 8 \) and 10. Triangles, circles, and squares represent maximum, average, and minimum number densities of halos for a given bubble size, respectively. The ranges of DM halo masses are \( M_{\text{halo}} < 10^8 M_\odot h^{-1} \), \( 10^8 M_\odot h^{-1} \leq M_{\text{halo}} < 10^9 M_\odot h^{-1} \), and \( 10^9 M_\odot h^{-1} \leq M_{\text{halo}} \). [See the electronic edition of the Journal for a color version of this figure.]
research of cosmic reionization. In order to quantify the shape of the bubbles as ellipsoid, we first calculate the inertia tensor,

\[
I_{ij} = \sum (x_i - x_c) (x_j - x_c)^{1/2},
\]

where \( x \) is the coordinate of every cell that is included in each H II bubble and \( x_c \) is the geometrical center of the bubble. From the above-defined \( I_{ij} \), we obtain the square roots of the three eigenvalues, \( a > b > c \). We measure the shape of the bubbles that have at least six member cells. Therefore, the smallest volume of the bubble is \( \sim 0.016 \text{ Mpc}^3 \text{ h}^{-3} \). The shape is expressed as two parameters, \( e_1 \) and \( e_2 \),

\[
e_1 = \sqrt{1 - \frac{b^2}{a^2}}, \quad e_2 = \sqrt{1 - \frac{c^2}{b^2}}.
\]

After percolation becomes a main process in the growth of H II bubbles, the above measurement of shape is useless. At late stages of reionization, the complicatedly connected bubbles are not similar to ellipsoids. Moreover, neutral hydrogen regions can be surrounded by H II bubbles in a grapelike structure. Therefore, we derive \( e_1 \) and \( e_2 \) for the bubbles before the largest bubble is comparable to the simulation volume.

The bubble volume-weighted shape distribution is estimated in the following way:

\[
P(e_1, e_2) = \int P(e_1, e_2 | V) P(V) dV.
\]

Figure 10 shows the concentration of volume-weighted H II bubbles in the \((e_1, e_2)\) plane. Note that a spherical bubble would have \( e_1 = e_2 = 0 \). In Figure 11 we show the bubble shape as a function of bubble size, with the largest bubble separately shown as a star, for \( z = 13.8 \) and 10. The shape of the bubbles follows a complicated dependence on their sizes. The shapes of small bubbles are, however, less accurately calculated because of limited resolution. Clearly, most of the volume in H II bubbles is far from
Fig. 10.—Distribution probability of H II bubble shapes expressed as $e_1$ and $e_2$ at $z = 13.8$ and 10.0. Most bubbles are far from a sphere, i.e., $e_1$ and $e_2 \gg 0$. This trend does not change even though reionized volume increases. The distribution is derived from eq. (5). [See the electronic edition of the Journal for a color version of this figure.]

Fig. 11.—Distribution probability of bubble shape parameters and volume at $z = 13.8$ and 10. The bubble shape has a distribution that does not change even though the fraction of reionized regions increases between $z = 13.8$ and 10. While small bubbles have the narrow range of shapes around $e_1$ or $e_2 \sim 0$ or 1, the bubbles of intermediate size show the broad range of shapes. The stars represent the shape of the largest bubble. [See the electronic edition of the Journal for a color version of this figure.]
spherical. The fact that $e_1$ is quite close to 1 and $e_2$ has a wide range indicates that bubbles tend to have shapes ranging from a cigar to a ruler. Note that at $z < 12$ the shape distribution is basically dominated by the largest bubble. For example, at $z = 10.0$, the most probable shape parameter is close to $e_1 \sim 0.9$ and $e_2 \sim 0.6$ of the largest bubble, corresponding to axial ratios $b/a = 0.4$ and $c/b = 0.8$. This highly complex nonsphericality would introduce substantial inaccuracies in calculations based on analytical modeling of observables of the high-redshift universe (see Fan et al. 2006 for a review). The complicated shape of H II bubbles is largely determined by the clustering of ionizing sources, as bubbles grow around large-scale structures, primarily filamentary networks. Figure 12 gives three orthogonal projections of a single randomly chosen bubble which shows characteristics of bubbles resulting from mergers of smaller bubbles.

3.3. Non-Gaussianity of Reionization

We sample 5000 spheres randomly placed in the simulation box and measure the mass-weighted ionization fraction of the spheres. That is, we measure $x_{\text{bubble}} = (\text{mass of H II})/(\text{mass of H})$ for each sampled sphere. It is a slightly different way to find a probability distribution function compared to that of Iliev et al. (2006), in which they sample independent subcubes in their simulation box. From these sampled spheres, we estimate mean ($\mu$), variance ($\sigma^2$), skewness ($\gamma_1$), and kurtosis ($\gamma_2$). We use three different sphere sizes of comoving radius 5, 10, and 20 Mpc $h^{-1}$. Figure 13 shows measurements of the underlying H II probability distribution functions. First of all, we see that the mean ionization fraction reaches 50% at $z \sim 9.5$ (hereafter $z_{50\%}$). Because the size of the sampled spheres is much larger than the characteristic bubble sizes, the...
measured $\mu$ reflects the global mean ionization fraction. The largest dispersion of the ionization fraction appears at $z \sim 8.9$, close to $z_{50\%}$. It is easily seen that the degree of non-Gaussianity decreases with increasing size of the sphere (bottom panels), as expected. While the variance starts to fall at $z < 8.9$, the non-Gaussianity as measured by $\gamma_1$ and $\gamma_2$ continues to rise, peaking at $z \sim 5.7$ when the largest bubble starts to fill the entire simulation box. Interestingly, it might be of some comfort for analytical modelers that at $z_{50\%}$, the $H II$ field is weakly non-Gaussian with $\gamma_1 = (0.22, 0.17, 0.01)$ and $\gamma_2 = (-0.66, -0.26, -0.46)$ for $r = (5, 10, 20)$ Mpc $h^{-1}$, respectively. But $\gamma_1$ and $\gamma_2$ are closest to zero at $z \sim 9$ and $\sim 10$, respectively. Because the boundary between fully neutral and fully ionized regions is much smaller than the sizes of the sampling spheres, the measured non-Gaussianity is quite small around $z_{50\%}$ (Wyithe & Morales 2007; Lidz et al. 2007b).

3.4. Power Spectrum of Ionized Hydrogen Density Fluctuation

We present the three-dimensional power spectrum of the ionized hydrogen density fluctuation field ($\delta_{HI}$), neutral hydrogen density fluctuation field ($\delta_{HII}$), and matter density fluctuation field ($\delta_{mat}$) in Figure 14. We note that the power spectrum is not measured for ionization fraction fluctuation but for ionized density fluctuation. The dimensionless power spectrum is defined here as

$$\Delta_k^2 = \frac{V}{(2\pi^2)^3} 4\pi k^3 P(k).$$

We see that the ionized hydrogen density field shows a strong bias with respect to the matter power spectrum, where the bias is strongest on small scales and decreases toward large scales. In general, the bias of ionized regions always decreases with time, whereas the bias of neutral regions does the opposite. We find that $\Delta_{HII}^2$ matches $\delta_{mat}$ at the early time when most hydrogen is still neutral, as expected, whereas $\Delta_{HII}^2$ becomes equal to $\delta_{mat}$ when the universe gets completely ionized. This is in agreement with the underlying physics that the earlier bubbles are produced by highly biased galaxies formed inside DM halos of high-$\sigma$ peaks, whereas only those highly biased, and hence large, galaxies host most of the neutral hydrogen when the universe becomes highly ionized. It is interesting to note that at around $z_{50\%}$, $\Delta_{HII}^2$ and $\Delta_{HII}^2$ appear to have roughly the same shape, although the latter has somewhat more power than the former, while both have more power than the total matter, suggesting a significant non-Gaussian distribution of both. We also note that there are significant differences between the nonlinear (actual) power spectrum of total matter and the linear power spectrum at $k \geq 0.3$ at $z \leq 10$, indicating the necessity of taking nonlinear effects into account.

4. CONCLUSIONS AND DISCUSSION

Using the state-of-the-art, largest cosmological simulation of box size 100 Mpc $h^{-1}$ with detailed radiative transfer of ionizing photons, we compute the evolution of $H II$ bubbles during cosmological reionization from $z \sim 25$ to 6. Our simulation resolves galaxy sources that produce most of the ionizing photons in the universe in the concerned redshift range. Here are a few major findings from our analysis.

1. We find that, for significant $H II$ bubbles, their number distribution is peaked at a volume of $\sim 0.6$ Mpc$^3$ $h^{-3}$ at all redshifts. But, at $z \leq 10$, one large, connected network of bubbles dominates the entire $H II$ volume.

2. $H II$ bubbles are highly nonspherical. This result is not totally unexpected, since we are quite familiar with the generic filamentary nature of cosmic structure formation. The mergers of $H II$ bubbles blown by galaxies formed along filaments consequently produce filamentary (fatter) $H II$ bubbles.

3. The $H II$ regions are highly biased with respect to the underlying matter distribution, with the bias decreasing with time as more and more less-biased structures begin to dominate the ionizing photon production rate. The opposite is true for the neutral hydrogen region in the sense that the bias of the neutral regions increases with time.

4. The universe becomes 50% ionized at redshift $z \sim 9.5$, when the $H II$ region is actually the least non-Gaussian in the redshift range concerned. The non-Gaussianity of the $H II$ region reaches its maximum near the end of the reionization epoch $z \sim 6$. But at all redshifts of interest there is a significant non-Gaussianity in the $H II$ field.

5. Pop III galaxies play a very important role in the reionization process. With our spatially resolved treatment of metal enrichment, we show that small bubbles are initially largely produced by Pop III stars. At $z \geq 10$ even the largest $H II$ bubbles have a balanced ionizing photon contribution from Pop II and Pop III stars. At $z \leq 8$, however, Pop II starts to dominate the overall ionizing photon budget for large bubbles, although Pop III stars continue to make a nonnegligible contribution.

6. The relationship between halo number density and bubble size is complicated, although there is a strong correlation for large bubbles in which larger bubbles tend to have higher halo number density. In addition, we find that only the large bubbles (size $\geq 1$ Mpc $h^{-1}$) contain halos of mass in excess of $10^9 M_\odot h^{-1}$.

Some of the results depend on modeling the formation of Pop III stars and the transition from Pop III to Pop II star formation that needs more careful study (e.g., Schneider et al. 2006; Smith & Sigurdsson 2007). For one thing, the correct ionizing photon generation rate from the unit mass of Pop III or Pop II stars is as important as a careful treatment of the recombination process in reionization simulation. As shown in § 3.1.4, the formation of small bubbles is closely connected to the ionizing...
efficiency of the Pop III stars, which is subject to their uncertain IMF. Our reionization simulation is the first attempt to imitate the transition of Pop III star formation to Pop II star formation in a large simulation box of $100 \, \text{Mpc} \, h^{-1}$. By improving our understanding of chemical enrichment during reionization (e.g., Greif et al. 2007), the transition of a dominant radiation source will be more certainly simulated in future.

In addition to the uncertain physics of Pop III stars, the numerical treatment of self-shielding and recombination has to be improved in future simulations. We use information from simulation particles for calculating the recombination rate that has better resolution than the grid-based approach. But our approach also has a limited ability to resolve small-scale physics like self-shielding. That must be improved in a reionization simulation with a subgrid model.

We thank Paul Bode for allowing us to use his FoF code and Marcelo Alvarez for his bubble size measurement code. We would also like to thank the referee for many constructive comments that improved this paper. This research is supported in part by grants AST 04-07176 and NNG06GI09G. H. T. is additionally supported in part by NASA grant LTSA-03-000-0090.

APPENDIX

RADIATIVE TRANSFER: PHOTON ADVECTION VERSUS RAY TRACING

We compare the photon-advection scheme with the ray-tracing scheme of Trac & Cen (2007) by applying these RT algorithms to two reionization simulations with identical initial conditions and source prescriptions. The test is conducted in a $25 \, \text{Mpc} \, h^{-1}$ box with $720^3$ dark matter particles and $90^3$ RT cells and has the same spatial, mass, and temporal resolution as the large $100 \, \text{Mpc} \, h^{-1}$ simulation. We find that the photon-advection scheme produces very similar results in terms of the spatial and temporal evolution of \H into a substantial majority of the reionization epoch. Figure 15 shows very good agreement in the redshift evolution of the volume-averaged...
HI fraction $f_{\rm HI}$, and deviations are only found when the box is already significantly ionized. The timing starts to differ at late stages near complete overlap. There are small delays of $\Delta z \sim 0.1$ and 0.2 when the neutral fraction drops to 0.1 and 0.01, respectively.

In order to quantify the agreement in the spatial evolution of $H\,^1$, we cross-correlated the $f_{\rm HI}(x)$ fields from the two RT simulations and plotted the results in Figure 16. From the two power spectra, $P_{\rm adv}(k)$ and $P_{\rm ray}(k)$, and the cross-power spectrum $P_{\rm adv-ray}(k)$, we can define the bias function,

$$b(k) \equiv \sqrt{\frac{P_{\rm adv}(k)}{P_{\rm ray}(k)}}$$

Fig. 15.—Comparison of the volume-weighted neutral hydrogen fraction $f_{\rm HI}$, from the photon-advection scheme (solid line) and the ray-tracing scheme (dashed line) for RT. The photon-advection scheme correctly captures the reionization process up until it is $\sim 75\%$ completed by volume. The magnitude of the difference in $f_{\rm HI}$, (dotted line) is generally very small and only reaches a maximum value of 0.05 near complete overlap. [See the electronic edition of the Journal for a color version of this figure.]

Fig. 16.—Comparison of the photon-advection scheme and ray-tracing scheme for RT using the bias $b(k)$ and cross-correlation $r(k)$ of the neutral hydrogen fraction $f_{\rm HI}(x)$ field. At $z = 8.2$ (solid line) and 7.5 (long-dashed line), both simulations have volume-averaged neutral fractions of 0.5 and 0.25, respectively, and the spatial evolution of HI is very similar. At $z = 7.1$ (short-dashed line), the correlation is poor because reionization has progressed slightly faster in the ray-tracing scheme. However, when the two simulations are compared at the same neutral fraction of 0.1 (dotted line), the correlation is much better, particularly on large scales. [See the electronic edition of the Journal for a color version of this figure.]
and cross-correlation function,

\[ r(k) = \frac{P_{\text{adv-ray}}(k)}{\sqrt{P_{\text{adv}}(k)P_{\text{ray}}(k)}} \quad (A2) \]

and use them to quantify the statistical correlation between the two fields. In both simulations, the neutral fraction drops to 0.5 and 0.25 at \( z = 8.2 \) and 7.5, respectively. At these two stages of reionization, the bias \( b(k) \) is very close to unity and only deviates by a maximum of \( \sim 0.1 \) near the RT grid Nyquist frequency \( k = 11.3 \text{ Mpc} h^{-1} \). Similarly, the cross-correlation \( r(k) \) shows very good agreement even down to the smallest scale. In the ray-tracing simulation, the neutral fraction drops to 0.1 at \( z = 7.1 \), but \( f_{\text{H}}^i \) is slightly higher by 0.04 in the photon-advection case. At this redshift, the bias and cross-correlation functions show that the two fields differ appreciably at all scales. However, when the two simulations are compared at the same neutral fraction of 0.1, the correlation is much better, particularly on large scales. The deviations from unity at the smallest scales are due to the appearance of new sources in the photon-advection simulation taken at a slightly later redshift.

In summary, we find that the photon-advection scheme correctly captures the reionization process up until it is \( \sim 75\% \) completed by volume. At earlier stages, the radiation field is highly nonuniform, even within the \( \text{H} \) regions, and the propagation of photons in the direction of decreasing radiation flux is a good description. However, at later stages of reionization near complete overlap, the radiation field is much more uniform, and the weak gradients in the radiation density do not provide accurate directions for photon propagation. We conclude that the photon-advection scheme provides a cost-effective approach to RT for a significant majority of the reionization epoch. However, the stages just before and after reionization should be simulated using more accurate approaches, such as ray tracing. Our results in this paper are valid at all stages when the ionization fraction, rather than the redshift, is used as an indicator of the progress of reionization.

**REFERENCES**

Barkana, R. 2002. NewA, 7, 85
Barkana, R., & Loeb, A. 2001. Phys. Rep., 349, 125
———. 2002, ApJ, 578, 1
———. 2004, ApJ, 609, 474
Becker, R. H., et al. 2001, AJ, 122, 2850
Cen, R. 2003, ApJ, 591, 12
Cen, R., & McDonald, P. 2002, ApJ, 570, 457
Ciardi, B., Ferrara, A., Governato, F., & Jenkins, A. 2000, MNRAS, 314, 611
Cohn, J. D., & Chang, T.-C. 2007, MNRAS, 374, 72
Cohn, J. D., & White, M. 2008, MNRAS, 385, 2025
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415
Fan, X., Narayanan, V. K., Strauss, M. A., White, R. L., Becker, R. H., Pentericci, L., & Rix, H.-W. 2002, AJ, 123, 1247
Furlanetto, S. R., McQuinn, M., & Hernquist, L. 2006, MNRAS, 365, 115
Furlanetto, S. R., & Oh, S. P. 2005, MNRAS, 363, 1031
Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004, ApJ, 613, 1
Gnedin, N. Y. 2000, ApJ, 535, 530
———. 2004, ApJ, 610, 9
Greif, T. H., Johnson, J. L., Bromm, V., & Klessen, R. S. 2007, ApJ, 670, 1
Haiman, Z., & Cen, R. 2005, ApJ, 623, 627
Haiman, Z., & Holder, G. P. 2003, ApJ, 595, 1
Hernquist, L., & Springel, V. 2003, MNRAS, 341, 1253
Hui, L., & Haiman, Z. 2003, ApJ, 596, 9
Illiev, I. T., Mellema, G., Pen, U.-L., Merz, H., Shapiro, P. R., & Alvarez, M. A. 2006, MNRAS, 369, 1625
Iliev, I. T., Mellema, G., Pen, U.-L., Merz, H., Shapiro, P. R., & Alvarez, M. A. 2006, MNRAS, 369, 1625
Iliev, I. T., Mellema, G., Shapiro, P. R., & Pen, U.-L. 2007, MNRAS, 376, 534
Kashikawa, N., et al. 2006, ApJ, 648, 7
Lidz, A., McQuinn, M., Zaldarriaga, M., Hernquist, L., & Dutta, S. 2007a, ApJ, 670, 39
Lidz, A., Zahn, O., McQuinn, M., Zaldarriaga, M., Dutta, S., & Hernquist, L. 2007b, ApJ, 659, 863
Loeb, A., Ferrara, A., & Ellis, R. S. 2008, First Light in the Universe (Berlin: Springer)