Semi-numerical Simulations of HII Bubble Growth During Reionization

Shuenn Patrick Ho

Office of Science, Science Undergraduate Laboratory Internship (SULI)

Princeton University

Stanford Linear Accelerator Center

Stanford, CA

August 15, 2008

Prepared in partial fulfillment of the requirements of the Office of Science, Department of Energy's Science Undergraduate Laboratory Internship under the direction of Marcelo Alvarez at the Kavli Institute for Particle Astrophysics and Cosmology, Stanford Linear Accelerator Center.

Participant: _____

Signature

Research Advisor: _____

Signature

TABLE OF CONTENTS

1 ABSTRACT

Semi-numerical Simulations of HII Bubble Growth During Reionization. SHUENN PATRICK HO (Princeton University, Princeton, NJ 08544) MARCELO ALVAREZ (Kavli Institute for Particle Astrophysics and Cosmology, Stanford Linear Accelerator Center, Stanford, CA 94025)

Related to such phenomena as cosmic structure formation and the birth of the first stars, the epoch of reionization provides a segue from the largely homogeneous universe described by the cosmic microwave background spectrum to the complexity we see today. While there are significant ongoing efforts to observe reionization, simulating reionization is an important and active area of research because it allows the testing and refining of analytical models. However, the radiative transfer simulations widely used in research, while providing a good qualitative picture of reionization, are too computationally taxing for such applications as probing a significant amount of the parameter space of reionization models. In the present study, we present an original code for simulating reionization much more efficiently based on a semi-numerical method involving a new algorithm for calculating smoothed overdensities in real space, as opposed to the fast Fourier transform (FFT) procedure used in previous semi-numerical studies. Using a cross-correlation coefficient analysis, we find that our simulations are successful at reproducing the results of previous simulations to a high degree of accuracy. Furthermore, we find our simulation code to be highly efficient, with a run time of $\sim$ 10-15 minutes for the 128^3-sized domains we simulate on in this study, and our design using a real-space smoothing algorithm promises a significant speed-up upon transfer to a graphics processing unit (GPU).

2 INTRODUCTION

The process of reionization describes a critical phase transition in the evolution of baryonic matter in the universe. In the early, cooling universe, baryons and electrons recombined to form neutral atomic hydrogen (HI), marking the beginning of the cosmological Dark Ages. However, structure formation continued through the amplification of density anisotropies (recorded at the moment of matter-radiation decoupling in the cosmic microwave background), creating regions massive enough to collapse through gravitational instability. This led to the appearance of the first luminous objects energetic enough to ionize hydrogen atoms, and the radiation created therewith caused the atomic hydrogen that pervaded the early universe to become largely composed of ionized hydrogen (HII). We refer to this period as the epoch of reionization (EOR), an important milestone in the evolution of the universe from the relative homogeneity following recombination to the complexity we see today.

Observational data provide some useful constraints on the EOR. One result is derived from the probing of the Gunn-Peterson absorption trough in the spectra of high-redshift quasars [1]. This technique provides a very sensitive measurement of the redshift at which the neutral HI fraction exceeds a threshold for near-complete absorption of Ly-α emission, suggesting that the end of reionization occurs at a redshift $z \sim 6$, roughly one billion years after the big bang. However, this technique is too sensitive — near-complete absorption begins at a neutral fraction $x_{HI} \sim 10^{-4}$ — for probing the EOR before its final moments. Further observational constraint comes from measurement of the cosmic microwave background (CMB) polarization. Thomson scattering of CMB photons off of free electrons generated at the onset of reionization can act as a source of secondary anistropy in the polarization of CMB light [2]. Measurements of the large-scale polarization of the CMB and comparison with the temperature anisotropy spectrum (used as an indication of primary polarization anisotropy) imply that reionization began as early as $z \sim 11 \pm 3$ [1]. Future studies will in-

vestigate the details of the EOR between these two constraints. One promising observational method is the measurement of the HI 21-cm line, an emission line resulting from forbidden transitions between the triplet and singlet states in atomic hydrogen. Measurements of this signal may lead to the detection and direct imaging of the neutral intergalactic medium and its evolution during the EOR. The data resulting from efforts such as 21-cm observation — currently underway at such facilities as the Mileura Widefield Array, the Primeval Structure Telescope, and the Low Frequency Array [1] — as well as measurements of the CMB secondary temperature anisotropy will lead to a clearer picture of the EOR.

Numerical simulations may also play an important role in providing a convenient way to examine the structure evolution during the period that has heretofore been difficult to probe observationally (beyond $z \sim 6$). Recent efforts to examine the growth of HII regions numerically from first principles using radiative transfer simulations have shed much light on the behavior of HII regions during reionization and motivated analytical models for HII region growth [3]. The present study also seeks to produce large-scale simulations of reionization. However, instead of a costly radiative transfer method, we implement the reionization physics using a semi-numerical method inspired by Zahn, et al. (2007) [5]. This allows us to perform analyses that are prohibitively time-consuming and inefficient with costly radiative transfer simulations. For example, we are able to more efficiently survey a large parameter space in the properties of the ionizing sources. In particular, in this study we present a semi-numerical simulation code implementing a real-space smoothing algorithm as opposed to the fast Fourier transform (FFT) algorithms used by prior studies. This is because we hope to adapt our code to run on a graphics processing unit (GPU), and FFT algorithms are very inefficient on such processors. However, with a real-space algorithm, running on a GPU promises a significant speed-up in run time (as much as 40 to 100 times CPU speed)[6].

3 METHODS

A. The Semi-numerical Method

The semi-numerical method begins with the simple assumption that each collapsed halo of mass m_{halo} (which is assumed to contain a galaxy) produces a spherical ionized region around it, the size of which is directly proportional to the galaxy mass. This assumption simplifies all of the complex recombination and radiative transfer physics into a proportionality constant, which we call the efficiency factor ζ [3]. Thus,

$$m_{ion} = \zeta m_{halo} \ , \tag{1}$$

where m_{ion} and m_{halo} are the masses of the ionized region and the halo, respectively. For example, the efficiency factor can encode the escape efficiency of photons from the galaxy f_{esc}, the star formation efficiency within the galaxy f_*, the number of ionizing photons produced in stars per baryon $N_{\gamma/b}$, and the average number of recombinations for a hydrogen atom n_{rec} by writing $\zeta = f_{esc} f_* N_{\gamma/b}/(1 + n_{rec})$ [3].

It is straightforward to extend this assumption to a condition stipulating whether a region of mass m is fully ionized. First, we can consider Eq. 1 equivalent to a relation between the ionized fraction f_{ion} and the collapse fraction f_{coll}: $f_{ion} = \zeta f_{coll}$. This then translates to a minimum condition on the collapse fraction for full ionization ($f_{ion} = 1$) to occur [3]:

$$f_{coll} \geq f_x \equiv \zeta^{-1} \tag{2}$$

We can then use the extended Press-Schechter model to relate this condition on the collapse fraction to a condition on the mean mass overdensity smoothed over a region of scale m,

commonly denoted δ_m:

$$f_{coll} = \text{erfc}\left[\frac{\delta_c(z) - \delta_M}{\sqrt{2[\sigma^2(M_{min}) - \sigma^2(M)]}}\right] , \tag{3}$$

where $\sigma^2(M)$ is the mass variance if the density i smoothed on the scale M, $\delta_c(z)$ is the critical overdensity for collapse, and M_{min} is the minimum mass for which a halo is considered to be a source of ionizing radiation [4]. With these relations, we can convert Eq. 2 into a condition on the mass overdensity itself,

$$\delta_M \geq \delta_x(M, z) \equiv \delta_c(z) - \sqrt{2}\left[\sigma^2(M_{min}) - \sigma^2(M)\right]^{1/2} \text{erf}^{-1}(1 - \zeta^{-1}) , \tag{4}$$

where we have clearly defined the minimum overdensity $\delta_x(M, z)$ above which a region will self-ionize.

This formulation provides a simple method to answer the question of whether a region about a point is self-ionizing at some redshift z. A region is self-ionizing if the condition in Eq. 4 is satisfied over any scale M. Similarly, we can also calculate when a region first becomes self-reionizing. We define the reionization redshift at a position z_{reion} as the maximum redshift (the earliest time) at which the region about that position is self-ionizing at any scale M. Inverting Eq. 4 to solve for z, we find an expression for the earliest redshift a region of scale M becomes self-ionizing:

$$z_{ion}(M) = \frac{7D_6}{\delta_{c,0}}\left[\delta_M + Z(M, \zeta)\right] - 1 , \tag{5}$$

where have explicitly written $\delta_c(z) = \frac{\delta_{c,0}(1+z)}{7D_6}$ by assuming that the critical overdensity evolves as the linear growth factor $1/D(z)$ where $D(z) \equiv D_6\left(\frac{7}{1+z}\right)$, and where we have defined the quantity $Z(M, \zeta) \equiv \sqrt{2}\left[\sigma^2(M_{min}) - \sigma^2(M)\right]^{1/2} \text{erf}^{-1}(1 - \zeta^{-1})$. Calculating the reionization redshift for any given position thus amounts to looping through scales M about

that position to find $z_{reion} = \max[z_{ion}(M)]$.

B. The Simulation

Our semi-numerical simulation produces a field of z_{reion} throughout a large-scale ($\sim$ 200 Mpc/h) domain, calculating for each point its corresponding reionization redshift. Such a field provides a very compact representation of the time evolution of HII bubbles. In producing such a field, our method computes z_{ion} using Eq. 5 for a range of scales M at each point, as similar studies have done in the past. However, our method differs in that, where past studies have employed k-space FFT procedures for computing the overdensity δ_M, we have devised an efficient, hierarchical algorithm to compute the overdensity δ_M with a smoothing procedure in real space [5]. Inspired by the *marching cubes* algorithm for generating three-dimensional computer graphics, our algorithm is much faster than a brute-force averaging.

Given an input density field $\rho(\vec{r})$, we first translate this to the density contrast,

$$\delta(\vec{r}) = \frac{\rho(\vec{r}) - \bar{\rho}}{\bar{\rho}} \ , \tag{6}$$

In preparation for our optimized algorithm for calculating δ_M, we then derive from this initial field a series of coarser fields, each 2^{-3} the size of the last. These coarser fields are populated with increasingly lower-resolution density values drawn recursively from the field before it, with each cell in a coarse field representing the mean value of 2^3 cells in the finer field from which it was drawn. For example, if the input field is of size 128^3, the field one coarser is of size 64^3, and the cell (i, j, k) in the coarser field corresponds to the average of the points $(2i, 2j, 2k)$, $(2i+1, 2j, 2k)$, $(2i, 2j+1, 2k)$, etc. in the finer field. Given an input field of size $2^{(3N_i)}$ and a coarsest field of size $2^{(3N_{min})}$, a field of coarseness order n is then defined to be of size $2^{3(N_{min}+n)}$. Performing this "downgridding" process sequentially on the input field and

storing the coarser fields preliminarily allows for an expedited smoothing algorithm. The extent of this downgridding is at the discretion of the user (more downgridding provides more efficiency); in the simulations presented in this study, we downgrid from an initial field of 128^3 down through a coarsest field of size 8^3.

In calculating mean overdensity, we smooth over δ about a point using a spherical top hat in real-space,

$$\delta_M = \frac{1}{V_M} \int_{V_M} \delta(\mathbf{x}) dV \ , \tag{7}$$

where $R = [3M/4\pi\bar{\rho}]^{1/3}$ is the smoothing radius for a scale M and $V_M = (4\pi/3)R^3$ is the corresponding smoothing sphere volume. Fig. 1 diagrams our hierarchical "cascading" algorithm for calculating δ_M in contrast to the inefficient "brute force" method. The brute force method involves simply searching through all cells in the domain and averaging only those within the radius R of the center of the smoothing sphere to calculate the value δ_M. Our algorithm, while more complex, is significantly more efficient. Beginning on the coarsest field, the hierarchical algorithm includes cells within the smoothing sphere of radius R, cascading down to finer fields where higher resolution is needed near the boundary. If a cell is completely outside R, it is discarded; if it is completely inside, it is volume-averaged into the total smoothing value; and if it is on the sphere, we go to one finer scale and repeat the process on the eight finer cubes within that cell. This process is done recursively until we reach the finest scale, that of the input field, at which point the brute force method is performed on any cells remaining to be examined; that is, they are either included if their midpoints are within R or excluded otherwise. As a comparison between Fig. 1a and 1b shows, this results in identical volumes being covered. In all tests, relative error between the cascade method and the brute force method was exactly zero, while efficiency (measured by the number of cells calculated on) was improved by roughly an order of magnitude.

Utilizing this real-space smoothing method for calculating δ_M, our simulation can be

summarized as consisting of five major steps. Diagrammed in Fig. 2, these steps are: (1) read in the initial density field; (2) downgrid over a range of coarser fields; (3) read in a pre-calculated array of $\sigma(M)$ over a range of scales M (or, equivalently, R); (4) calculate δ_M and then $z_{reion}(M)$ over a range of scales M; and (5) look up values of $z_{reion}(M)$ and return the highest z_{reion} value at each point.

The loops described in steps (4) and (5) for calculating and then looking up z_{reion} values are described in more detail in Fig. 3. For optimization reasons, our algorithm for the calculation of z_{reion} calculates and stores z_{reion} values at lower resolutions when smoothing over larger scales. As summarized in Fig. 3a, we produce another array of grids of sequential coarseness orders similar to those generated by the downgridding of the density fields. Each grid of coarseness order n corresponds to a range of smoothing scales $R : \tau 2^{N_i - n} \leq R < \tau 2^{(N_i - n + 1)}$, where we set $\tau = \sqrt{3}$, such that the threshold corresponds to the diagonal of a cube in grid n. Then, when looping over smoothing scales, the highest z_{reion} values within each range of R are then calculated only over the grid-points of the corresponding grid. Our algorithm for the look-up of z_{reion} uses these values to calculate the highest z_{reion} at each point on the finest level. As summarized in Fig. 3b, the look-up algorithm accomplishes this at each position by looping through grids of all coarseness levels and, where necessary, interpolating between points stored at lower resolution. The highest of these values is taken to be the z_{reion} of that particular position in our domain. We believe that we can employ this method with little error because δ_M values calculated on larger smoothing scales are not sensitive to small shifts in the position of the center of the smoothing sphere, and so we do not need to calculate δ_M at high resolution on higher smoothing scales.

For the simulations presented in this draft, we calculated z_{reion} on a grid of size 128^3 corresponding to a cube with side length 200 Mpc/h given a pre-generated field of random Gaussian density perturbations at $z = 19$. Run on a single-processor desktop computer, the algorithm takes approximately 10-15 minutes to finish. Other parameters used in these

simulations include: the critical overdensity at $z = 0$: $\delta_{c,0} = 1.686$, the linear growth constant: $D_6 = 0.1917$, the Hubble constant: $h = 0.7$, the minimum collapse mass: $M_{min} = 10^8 M_\odot$, and the ionization efficiency: $\zeta = 200$. All FFT simulations presented for comparison were generated with identical parameters.

4 RESULTS

We evaluate the accuracy of our newly developed simulation code by testing its results against those of a previously developed code that employs the aforementioned FFT smoothing procedure. We compare the different data using the cross correlation coefficient between the FFT and real-space data, defined as:

$$\xi_{FR}(k) = \frac{P_{FR}(k)}{[P_{FF}(k)P_{RR}(k)]^{1/2}} \tag{8}$$

where P is the power spectrum:

$$P_{\alpha\beta} = \langle \delta_{k,\alpha} \delta^*_{k,\beta} \rangle \tag{9}$$

and $\delta_{k,\alpha}$ are the Fourier coefficients of $\delta_{\vec{r},\alpha} = \frac{\alpha - \bar{\alpha}}{\bar{\alpha}}$ for the quantity α. In our case, the two quantities being cross-correlated are the z_{reion} values generated by the two respective simulations. The cross-correlation coefficient is exactly 1 for perfectly correlated data, -1 for perfectly anti-correlated data, and 0 for no correlation.

The resulting data indicate that our simulations are robustly accurate in comparison with the FFT simulations. Plots of ξ_{FR} for data generated at a number of different values for the ionization efficiency ζ and for the resolution in the smoothing scales N_{sm} show this accuracy to be insensitive to changes in these parameters (Fig. 4). The cross-correlation coefficient is well above 0.9 for $k < 1$ and remains above 0.85 even to high k values. This suggests that earlier and later reionization occurs largely in the same places in both simulations, and thus

they generate very similar pictures of the time-evolution of reionization. As suggested by the higher cross-correlation coefficient at $k < 1$ than at higher k, this is especially true for large-scale structure and deteriorates a bit for small-scale structure. For visual comparison of the strikingly similar data, Fig. 5 shows one example of slices through the $x - y$ planes of data generated by both the FFT and real-space simulations, with parameters $\zeta = 157$ and $N_{sm} = 60$. Note that the real-space simulation is biased slightly towards higher values of z_{reion} and thus earlier reionization in some regions — the FFT simulation has reionization starting at $z = 25.2$ and ending at $z = 11.2$, while the real-space simulation runs from $z = 28.9$ to $z = 11.3$. However, qualitatively the pictures of reionization generated by the two are the same, as long as the same regions reionize before other regions.

Using our semi-numerical method, we can also derive the time-evolution of reionization using just the one grid of z_{reion} data generated by our simulation. By plotting surfaces of constant z_{reion}, we can retrieve the positions of the ionization fronts at a given redshift z and thus easily image the extent of the HII bubbles. Fig. 6 shows an example of a sequence of such images. Our simulation produces a similar picture of the process of reionization to that produced by the FFT simulation, and both reflect the accepted model for reionization: the evolution of individual HII bubbles, their coalescence, and the rapid overlap phase leading to complete reionization.

Significantly, the optimizations we have coded into our simulation have allowed it to run very efficiently. Each run described in this study finished on the order of $\sim 10 - 15$ minutes. And since we have successfully produced this code using a real-space-based smoothing algorithm, there is the promise of a significant speed-up when we compile this code for the graphics processing unit.

5 CONCLUSIONS

We have presented results from an original simulation code, based on a semi-numerical method for calculating z_{reion} throughout a domain and employing a new algorithm for calculating δ_M in real space. In comparison to previously developed FFT simulations using identical input parameters, we find that our simulations produce very similar results. As diagrammed in Fig. 4, we find high cross-correlation coefficients between our data and the FFT simulation-generated data for both large and small scales, and this indicates to us that our simulations are very accurate in relation to simulations that have been used in previous studies. Significantly, our simulation code is very efficient and inexpensive in terms of runtime ($\sim 10-15$ minutes for the simulations in this study), and is also different from previous simulations in that, being based on a real-space smoothing algorithm, it is optimally prepared for a significant speed-up when compiled for a GPU.

Some further work can be conducted on exploring the accuracy and efficiency of this simulation code. We have established in this study that the code is relative robust to changes in parameters such as ionization efficiency (ζ) and resolution in smoothing scales (N_{sm}). Using the same cross-correlation coefficient analysis as described above, we can evaluate the effect of computational parameters such as the threshold for transferring between grids of different coarseness (τ) in order to investigate the accuracy-efficiency trade-off and operate optimally in the future. In the very near future, we expect that this simulation code can be easily adapted to run on a GPU, from which we expect a significant decrease in computation time. We hope to then apply these simulations to probe the parameter space of our reionization model and refine the picture of reionization physics.

6 ACKNOWLEDGMENTS

Thanks are due to my mentor, Marcelo Alvarez, for guiding me through this project, to the Kavli Institute and SLAC for hosting my internship, and to the DOE Office of Science for funding the SULI program and, by extension, this study.

REFERENCES

[1] Fan, X., Carilli, C.L., and Keating, B., Observational Constraints on Cosmic Reionization, Annual Review of Astronomy and Astrophysics, **44**, Page 415–462, 2006.

[2] Loeb, A. and Barkana, R., The Reionization of the Universe by the First Stars and Quasars, Annual Review of Astronomy and Astrophysics, **39**, Page 19–66, 2001.

[3] Furlanetto, S., Zaldarriaga, M., and Hernquist, L., The Growth of HII Regions During Reionization, The Astrophysical Journal, **613**, Page 1–15, 2004.

[4] Coles, P. and Lucchin, F., *Cosmology: The Origin and Evolution of Cosmic Structure*, John Wiley & Sons, Chichester, 1995.

[5] Zahn, O, et. al., Simulations and Analytic Calculations of Bubble Growth During Hydrogen Reionization, The Astrophysical Journal, **654**, Page 12–26, 2007.

[6] "CUDA Zone – The resource for CUDA developers." NVIDIA, (2008) http://www.nvidia.com/object/cuda_home.html

FIGURES

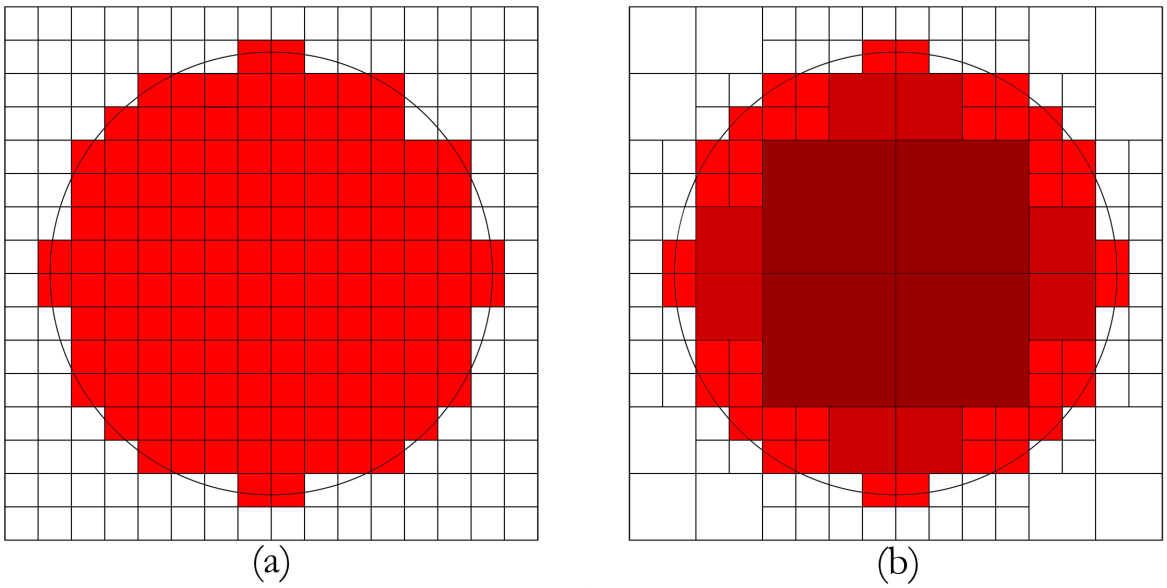

Figure 1: Diagram of smoothing algorithm by (a) brute force and (b) cascade methods.

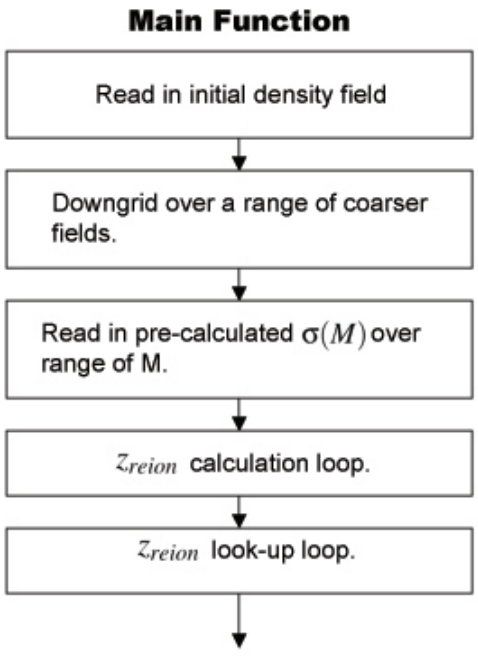

Figure 2: Flow chart of main function for simulation.

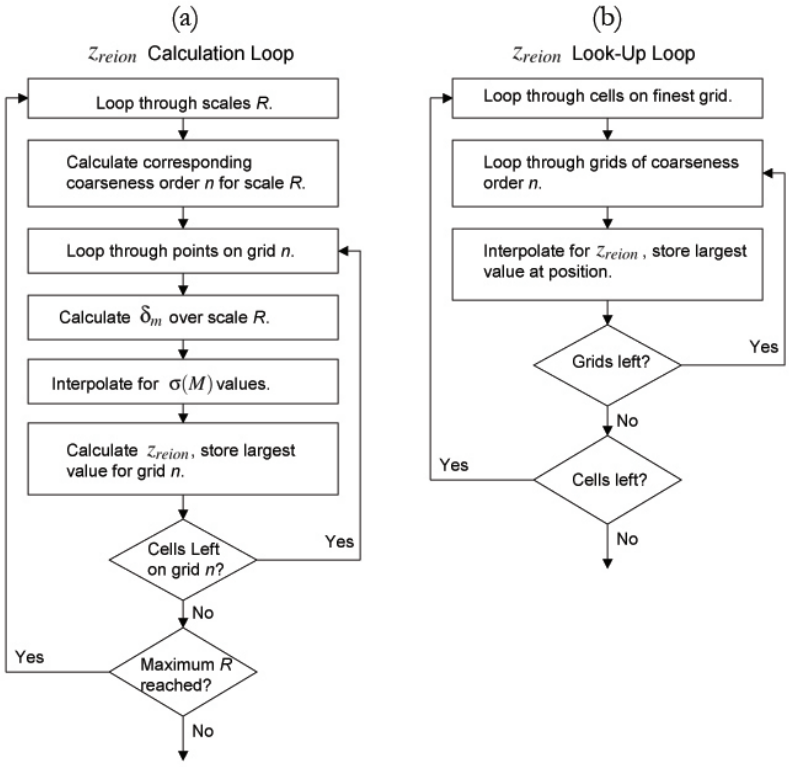

Figure 3: Flow charts for z_{reion} (a) calculation and (b) look-up loops.

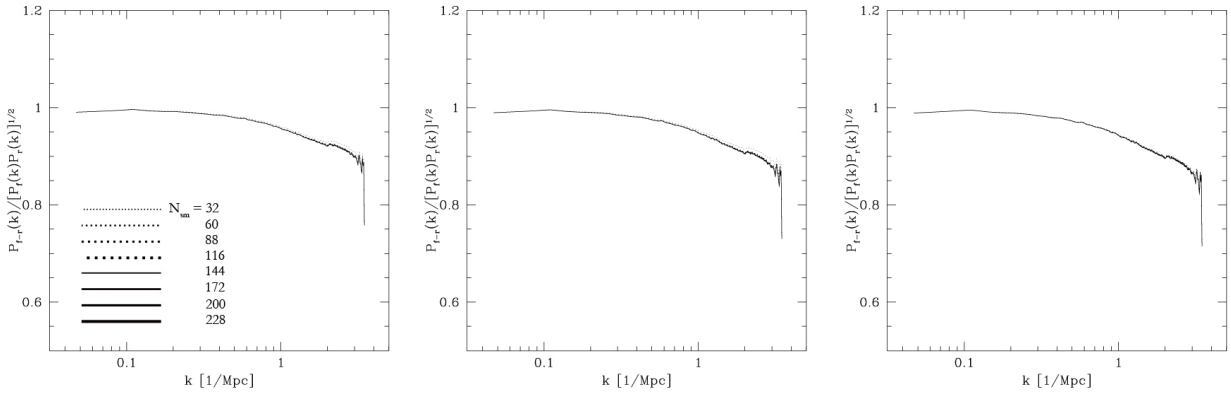

Figure 4: Plots of cross-correlation coefficient for simulations generated with ζ values of, from left to right, $\zeta = 50, \zeta = 157, \zeta = 300$, and with N_{sm} values of 32, 60, 88, 116, 144, 172, 200, 228, as shown in the legend.

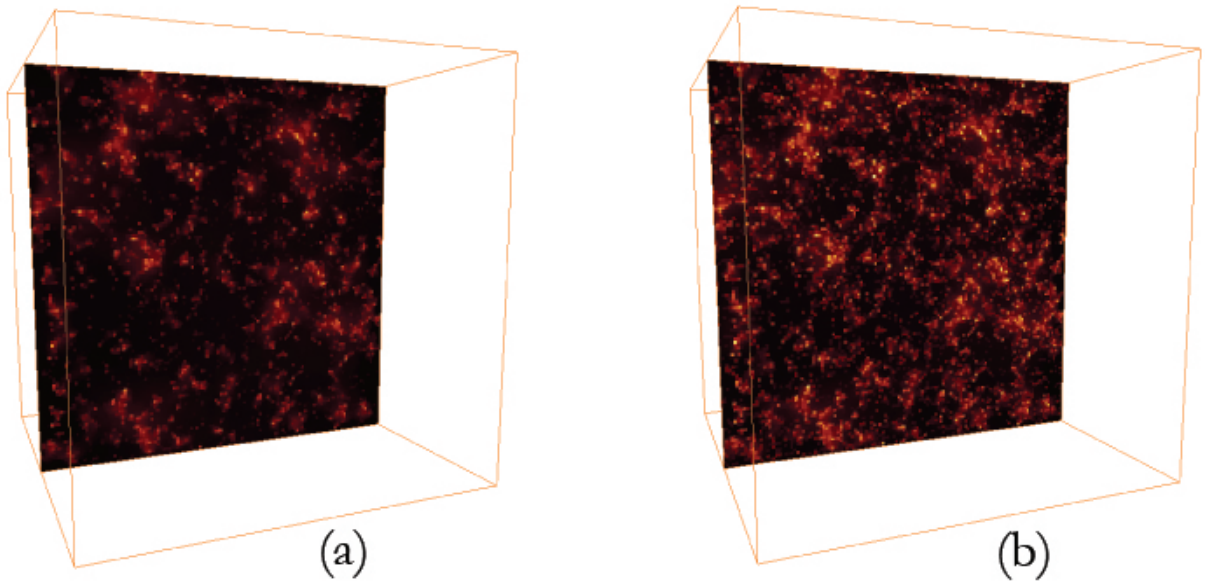

Figure 5: Slices through the $x-y$ planes of z_{reion} data generated by (a) the FFT simulation and (b) our real-space simulation.

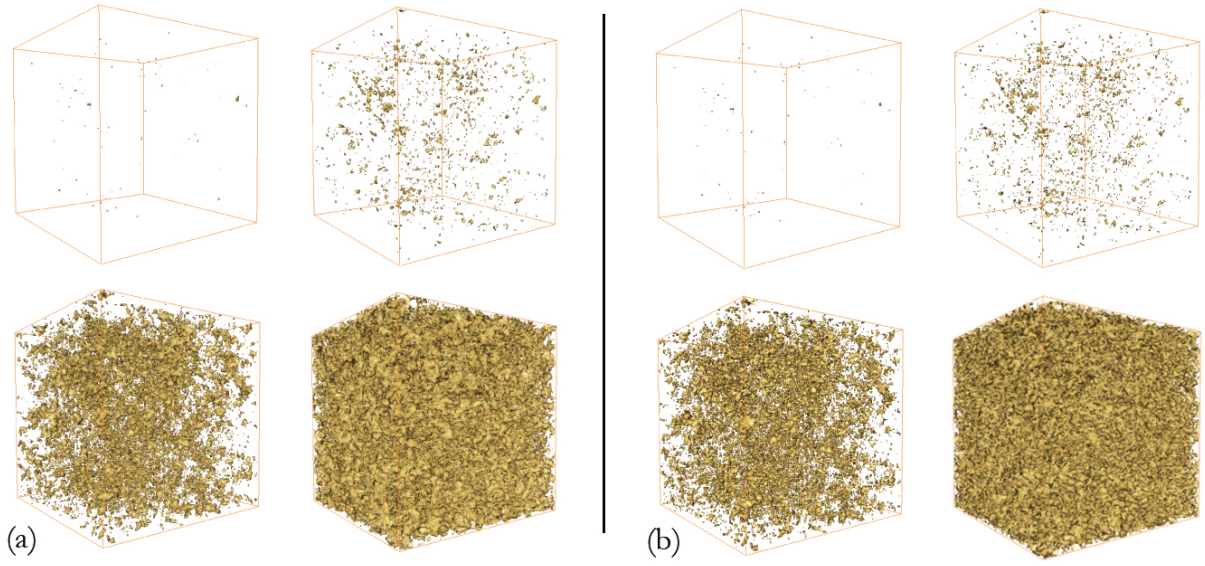

Figure 6: Time-evolution images of ionization fronts in (a) the FFT simulation at $z =$ 19.3, 16.7, 14.3, 12.0 and (b) our real-space simulation at $z = 21.4$, 18.4, 15.7, 12.4, with parameters $\zeta = 157$ and $N_{sm} = 60$.