DRAFT VERSION JULY 14, 2015
Preprint typeset using LaTeX style emulateapj v. 5/2/11

FERMILAB-PUB-15-015-A
(accepted)

arXiv:1412.5607v2 [astro-ph.CO] 13 Jul 2015

COSMIC REIONIZATION ON COMPUTERS III. THE CLUMPING FACTOR

ALEXANDER A. KAUROV[1] AND NICKOLAY Y. GNEDIN[2,1,3]

Draft version July 14, 2015

ABSTRACT

We use fully self-consistent numerical simulations of cosmic reionization, completed under the Cosmic Reionization On Computers (CROC) project, to explore how well the recombinations in the ionized IGM can be quantified by the effective "clumping factor". The density distribution in the simulations (and, presumably, in a real universe) is highly inhomogeneous and more-or-less smoothly varying in space. However, even in highly complex and dynamic environments the concept of the IGM remains reasonably well-defined; the largest ambiguity comes from the unvirialized regions around galaxies that are over-ionized by the local enhancement in the radiation field ("proximity zones"). That ambiguity precludes computing the IGM clumping factor to better than about 20%. We also discuss a "local clumping factor", defined over a particular spatial scale, and quantify its scatter on a given scale and its variation as a function of scale.

Subject headings: cosmology: theory – methods: numerical – intergalactic medium

1. INTRODUCTION

Theoretical study of the epoch of reionization is becoming an increasingly important area of research, as the computational capabilities advance to the point when it becomes possible to run numerical simulations with self-consistent radiation transfer and star formation in cosmological size boxes (So et al. 2014; Norman et al. 2013; Hutter et al. 2014; Gnedin 2014; Gnedin & Kaurov 2014).

Even though it is common to discuss the sources of ionization radiation, the competing process, recombination, is not less important in the context of reionization. There are two distinct ways of thinking about recombinations during the epoch of reionization. The first one treats the gas density, temperature, and ionized fraction distributions in the universe as continuous fields, with the recombination rate having appropriate values everywhere in space; this is the approach used in advanced simulations.

The second approach is usually adopted in analytical models and is sometimes used in interpreting observations. It focuses on ionizing photons which, first, escape a host galaxy with some probability, then freely travel through the ionized Intergalactic Medium (IGM), and finally are absorbed either at ionization fronts between the ionized and neutral regions in the IGM or at Lyman Limit systems (LLS). Recombination inside the host galaxy are treated as a reduction in the source emission and are parametrized with the escape fraction (Gnedin et al. 2008; Kuhlen & Faucher-Giguère 2012); recombination in the ionized IGM are quantified by the so-called "clumping factor", and absorption by LLS are accounted for by an upper limit on the photon mean free path (Songaila & Cowie 2010; Kaurov & Gnedin 2013). Even though spontaneous recombination is the same physical process no matter where it occurs, three unrelated quantities – the escape fraction, the clumping factor, and the maximum mean free path – are used to describe it.

Our goal in this paper is to examine whether this "three-regime" approach is well defined in the fully self-consistent numerical simulations (which we consider as a plausible model of the real universe). Since the recombination rate varies smoothly in space, such a discrete separation of a continuous function into three distinct regimes would only make sense if the actual distribution of recombination rates is, in some particular way, tri-modal and a clear separation can be made between recombinations in the ISM (the escape fraction), IGM (the clumping factor), and in the LLS (the maximum mean free path).

In §3 we describe a useful phase diagram, which helps to define these three regimes in a well motivated (rather than based on some arbitrary density or ionized fraction thresholds) way. Hence, if carefully made, such a separation of cosmic recombinations into distinct regimes can indeed be reasonable, and the quantities such as the IGM clumping factor can be meaningfully defined.

Then, in §4, we focus on the ionized IGM and its clumpiness, since it occupies the majority of volume and defines the morphology of reionization. Due to its quadratic dependence on the density, the recombination rate inside a given volume V explicitly depends on the actual density distribution inside the volume. It is convenient to use the clumping factor

$$C = \langle n_i^2 \rangle / \langle n_i \rangle^2, \tag{1}$$

to factorize out that dependence and to express the recombination rate through the mean density inside the volume V,

$$R = \int_V n_i^2 \alpha dV \approx \alpha C \langle n_i \rangle_V^2 V, \tag{2}$$

where n_i is the number density of ionized hydrogen (we assume that the number of electrons is proportional to the number of ionized hydrogen, which is the case before He II reionization), R is the recombination rate inside the volume V, and $\alpha(T)$ is the recombination coefficient.

However, the lack of a common definition of the ionized IGM leads to problems with comparing the clumping factor between different numerical studies (Gnedin & Ostriker 1997; Iliev et al. 2005; Kohler et al. 2007; McQuinn et al. 2007; Trac & Cen 2007; Pawlik et al. 2008; Raičević & Theuns 2011; Finlator et al. 2012; Shull et al. 2012; Emberson et al. 2013; So et al. 2013;

[1] Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637 USA; kaurov@uchicago.edu
[2] Particle Astrophysics Center, Fermi National Accelerator Laboratory, Batavia, IL 60510, USA; gnedin@fnal.gov
[3] Kavli Institute for Cosmological Physics and Enrico Fermi Institute, The University of Chicago, Chicago, IL 60637 USA

Jeeson-Daniel et al. 2014) and analytical models (Sobacchi & Mesinger 2014; Kaurov & Gnedin 2014). We reexamine the value of the clumping factor with Gnedin (2014) simulation described in §2 and discuss its definition. In addition, we study its spatial inhomogeneity in §4.1 and its correlation with density in §4.2.

2. SIMULATION

As a physically plausible model of reionization we use numerical simulations from the Cosmic Reionization On Computers (CROC) project (Gnedin 2014; Gnedin & Kaurov 2014).

These simulations are suitable for our purposes for several reasons. CROC simulations include the whole range of physical processes required in order to model reionization fully self-consistently[4], from gas dynamics to fully coupled 3D radiative transfer, star formation, and stellar feedback. They match the existing observational constraints on the evolution of galaxy luminosity functions and on the full distribution function of Gunn-Peterson absorption in the spectra of high redshift quasars. Using the DC model formalism of Gnedin et al. (2011), the simulations account for the cosmic variance between several independent realizations. Finally, using Adaptive Mesh Refinement, CROC simulations achieve spatial resolution of 125 pc in simulation volumes of up to $40h^{-1}$ comoving Mpc.

Such high spatial resolution allows us to consider everything outside galaxies well resolved, and therefore, such quantities as clumping factor can be computed directly with no prior assumptions. The internal structure of galaxies is not well resolved with the spatial resolution of ~ 100 pc; therefore, we do not consider clumping of the ISM in this paper.

In this paper we use three $40h^{-1}$ Mpc simulations with different DC modes (runs B40.sf1.uv2.bw10.A-C from Gnedin (2014)) as our fiducial set, with all presented quantities averaged over these three runs. In addition, we use two $20h^{-1}$ Mpc simulations: the "medium resolution" one (run B20.sf1.uv2.bw10.B from Gnedin (2014)) that matches our fiducial set in spatial and mass resolution and a higher resolution simulation (run B20HR.sf1.uv2.bw10.B from Gnedin (2014)) that we use to test numerical convergence.

3. PHASE DIAGRAM

In this section we describe our main tool (a kind of a "phase diagram"), which we use in the subsequent sections to classify the cosmic gas distribution into ISM, IGM, and LLS. Specifically, we plot the mass weighted distribution of all cells from a simulation in two dimensions: density, $(1+\delta)$, and *ionization state indicator*, $\varkappa$, which is defined as:

$$\varkappa \equiv (1+\delta)\,\frac{x_{\rm H\,II}^2}{x_{\rm H\,I}}, \tag{3}$$

[4] We purposedly distinguish terms "self-consistent" and "from the first principles". CROC simulations do include free parameters in the underlying physical model (gas depletion time due to star formation, delayed cooling time scale, effective emissivity of ionizing photons at the simulation resolution limit, etc), hence they are not "from the first principles" simulations. We call them "self-consistent", though, in a sense that all modeled physical processes are followed with the same spatial and temporal resolution, co-evolving and affecting each other as the simulation proceeds. In that sense the term "self-consistent" distinguishes CROC simulations from numerical models where radiative transfer is done in post-processing, or which track hydrodynamics and radiative transfer with two separate numerical schemes with widely divergent resolutions.

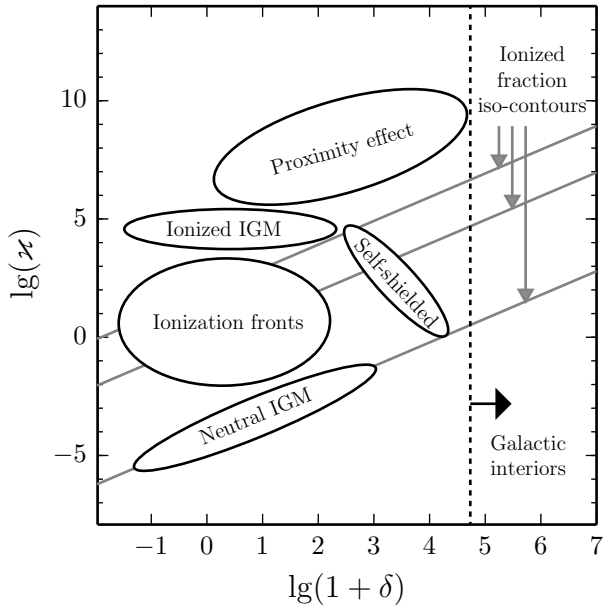

FIG. 1.— Schematic representation of the phase diagram introduced in §3.

where $x_{\rm H\,II}$ and $x_{\rm H\,I}$ are the fractions of ionized and neutral hydrogen relative to the total abundance of hydrogen.

This definition is motivated by a consideration that in the photoionization equilibrium

$$\Gamma\, n_{\rm H\,I} = n_{\rm H\,II}^2\, \alpha(T), \tag{4}$$

where Γ is ionizing background, and hence

$$\varkappa \propto \frac{\Gamma}{\alpha(T)}. \tag{5}$$

In particular, in the IGM with modest temperature fluctuations and approximately homogeneous ionizing background, $\varkappa \approx$ const.

Thus, the ionization state indicator is a good tracer of general IGM under the assumption of: (a) uniform ionizing background, (b) no collisional ionization [5], and (c) uniform temperature. Therefore, if the ionizing background is indeed uniform in ionized regions, then we should see a narrow distribution of $\varkappa$.

A few distinct features, which correspond to different states of gas, emerge from this phase diagram. We label these features in Figure 1 and discuss them individually, as well as the general properties of the phase diagram, in the following subsections. While the overdensity δ traces the large scale structure and gives an idea about where spatially these regions are located, the variations in the ionization state indicator show where the assumptions (a)-(c) break down.

3.1. *General properties of the $\delta-\varkappa$ phase diagram*

To give a better understanding of this type of a phase diagram, we outline its major properties using an optically thin (ionized) Lagrangian volume collapsing into a dense region as an example. Initially it is located in the region labeled as "Ionized IGM" in Figure 1. As the Lagrangian volume gets

[5] In case of presence of collisional ionization, the photoionization equilibrium does not hold. Thus, Equation 4 should have another term on the left hand side.

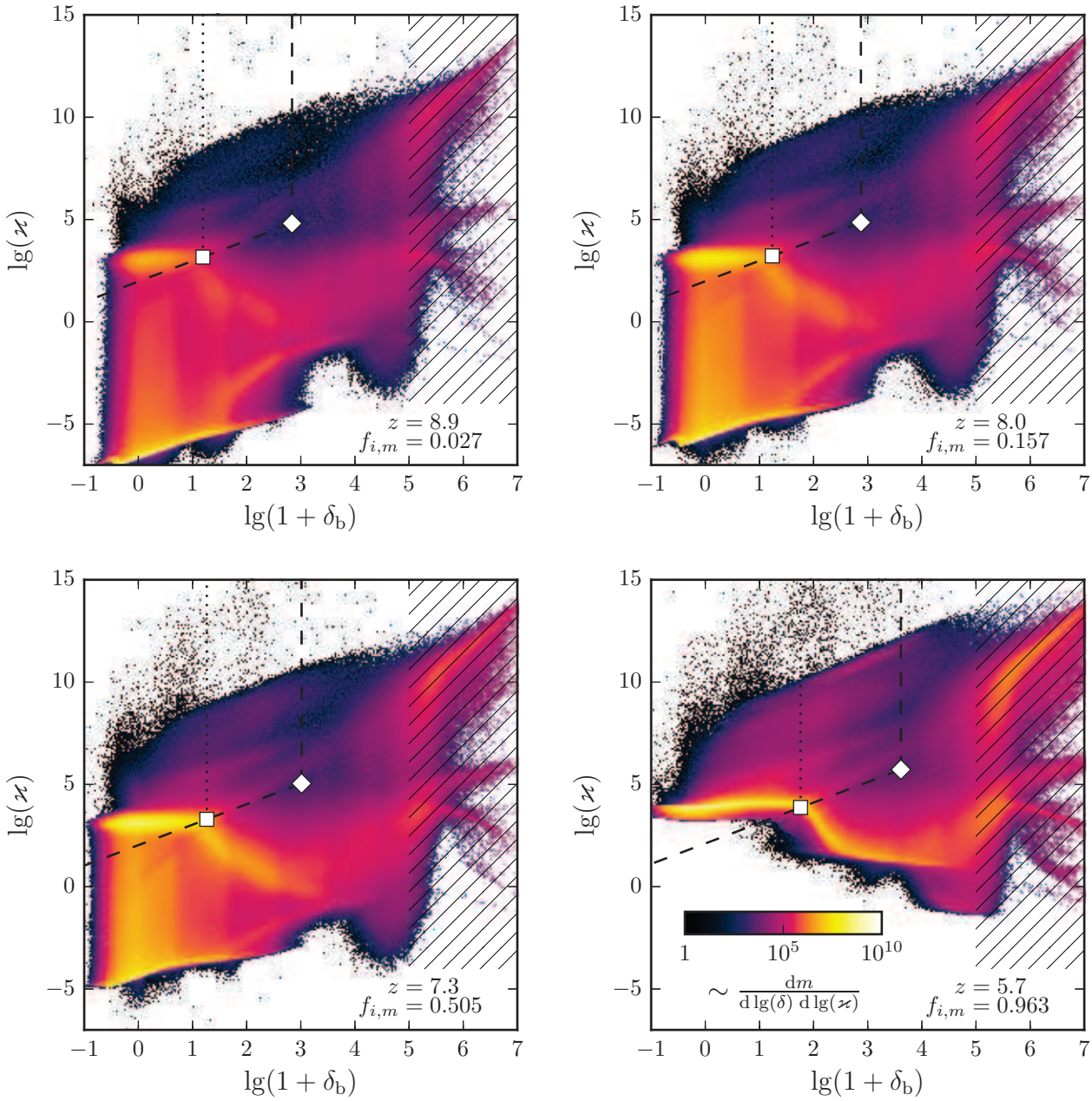

FIG. 2.— The $\delta-\varkappa$ phase diagram captured at four redshifts. The mass weighted ionized fraction. $f_{i,m}$, is shown in each panel. Dashed lines show density and ionized fraction thresholds presented in Figure 5. White diamond and square represent the first and the second pivot points we adopt for the definition of the ionized IGM. The hatched area corresponds to the high density regions inside galaxies (IGM), which is not well resolved in the simulation.

denser, it moves to the right in the phase diagram. However, the behavior of the $\varkappa$ value depends on the environment.

If the volume contracts slowly enough, it has enough time to recombine and to reach the equilibrium with the ionizing background. In this case the ionization state indicator does not change, and we see a horizontal movement in the phase diagram.

If the same volume collapses rapidly and contracts faster than it recombines, i.e. the ionized fraction does not change, then it moves along the line with the unity slope in $\delta-\varkappa$ space. Therefore, in the phase diagram that region moves up and may escape the "Ionized IGM" region into what we label as the "Proximity effect" part of the phase diagram.

Another mechanism for a volume to enter "Proximity" zone is to approach a source of ionizing radiation. It will locally boost radiation background and, if the medium remains optically thin, $\varkappa$ will also increase.

As the Lagrangian volume continues to collapse, at some point it enters the optically thick regime. The ionization rate inside it drops dramatically; consequently, the ionized fraction decreases, and the ionization state indicator $\varkappa$ decreases too. In the phase diagram it corresponds to the transition to the "Self-shielded" region in the phase diagram.

3.2. *Discussion of individual regions*

The actual phase diagrams from the CROC simulations at 4 different redshifts are shown in Figure 2. As reionization proceeds, the distribution of gas between various regions in the phase diagram changes.

Ionized and Neutral IGM. Before reionization the IGM has a small uniform ionized fraction left after the recombination epoch. Therefore, at redshifts 8.9, 8.0 and 7.3, when a significant fraction of low density gas is still neutral, we see it distributed along a line with the unity slope. The IGM progressively becomes more and more ionized as redshift approaches 5.7, and occupies different parts of the phase diagram. The value of $\varkappa$ increases, which represents the growth of the ionized fraction.

Ionization fronts are the intermediate stage between the ionized and neutral IGM. Since a particle spends relatively short time inside an ionization front, only a small fraction of mass is located in that phase at any given time. The fraction of total mass in ionization fronts is order of 5% at redshifts with 50% total ionization fraction and lower during other periods (and that is only an upper limit, as ionization fronts are not resolved in CROC simulations).

In individual ionized bubbles the ionization rate may vary significantly from bubble to bubble, due to random variations in the number of sources inside them. As ionized bubbles start to overlap, the ionization rate becomes more and more uniform. This is reflected in the phase diagram in the width of "Ionized IGM" region, which becomes much narrower at $z = 5.7$ compared to $z = 7.3$. The nuances of our specific IGM definition are discussed in §4.

Self-shielded regime. At higher densities (marked with a square in Figure 2) self-shielding becomes important. Spatial locations with no or very low star formation correspond to the sharp decline in $\varkappa$ from the typical values in the ionized IGM. As reionization proceeds, the ionizing background builds up, and burns deeper into these regions; therefore, the self-shielded region in the diagram moves to the right with time. In Figure 3 self-shielded regions manifest themselves as tiny dark spots in the ionization state indicator or temperature maps or bright spots in the neutral fraction map.

Proximity effect. After reionization, the scatter in $\varkappa$ in the optically thin regions can occur due to the proximity to ionizing sources, large temperature variations, and non-equilibrium effects.

By looking on spatial distribution of $\varkappa$ we can observe a few other effects. In Figure 3 we look at filament structure at redshift 5.7, however the same behavior is typical for any redshift inside ionized regions. Modest increase in $\varkappa$ is observed around filaments due to the rapid contraction of these regions. It leads to the increase of temperature, decrease of recombination coefficient, and, consequently, growth of ionization state indicator, which is inversely proportional to the recombination coefficient. Additionally, in these relatively sparse regions the characteristic contraction timescale may approach recombination time and, consequently, matter will be more ionized than it would be in the ionization equilibrium. So, the shell around filaments shows slight increase of $\varkappa$, but overall that effect is small.

In Figure 4 we zoom into a merging halo also at reshift 5.7. Regions with active merging stand out because of extreme temperatures. There, beside change in the recombination coefficient, collisional ionization of hydrogen also takes place, driving the neutral fraction away from a pure photoionization equilibrium. These regions are located in top left side of phase diagram – roughly at $10 < \delta < 10^3$ and $10^5 < \varkappa < 10^{12}$, and contain less than 1% of the total mass.

Galactic interiors. In the simulation that we used, the galactic interiors are not well resolved. Therefore, we do not discuss the features at $\delta \gtrsim 10^5$ in the phase diagram - that would require higher resolution simulations of individual galaxies.

4. IGM CLUMPING FACTOR

Before calculating the clumping factor of the ionized IGM, one first needs to define what the ionized IGM is. The most common approach of defining the ionized IGM is based on two thresholds: an upper limit on the gas density and a lower limit on the ionization fraction. These cuts exclude neutral dense matter and the ISM from contributing to gas clumping. We adopt the notation $C_{\delta, x_{\rm H\,II}}$ for the clumping factor calculated over the volume with overdensity below δ and the ionization fraction above $x_{\rm H\,II}$. The choice of these thresholds is somewhat arbitrary, which makes such a definition not well motivated. Here we propose a physically-motivated fix, based on the phase diagram introduced in the previous section.

The $\varkappa - \delta$ phase diagram contains a few features, which can be used as pivot points. The most prominent one is the transition between "Ionized IGM" and "Self-shielded" regions (a square symbol in Figure 2). At $z = 5.7$ it is easily identifiable along the yellow ridge that contains most of mass in low density regions as a point when the ridge turns down (towards more neutral gas). However, it is less pronounced at high redshifts and also the corresponding overdensity threshold does not fully include "Proximity" zone. Instead, we can use another pivot point located at a local minimum between "Proximity effect", "Self-shielded" and "Galactic interiors" zones (a diamond symbol in Figure 2). It is prominent at all redshifts, and includes "Proximity" zone, and mathematically well defined, making it much easier to find with a simple algorithm. For each redshifts we determine this point and record corresponding overdensity and ionized fraction (the actual values are plotted in Figure 5).

At redshifts where both these points are well defined, they always lie at the same value of the ionized fraction. It is not clear whether this is a universal property of the $\varkappa - \delta$ phase diagram or a mere coincidence, but that fact does not appear to be important enough to warrant a targeted study. Hence, using that property, we can define the transition into the "Self-shielded" regime at all redshifts, as a local maximum in the phase diagram along the line of constant ionized fraction (dashed line) passing through the diamond pivot point.

The values of the pivot points can be used in defining physically-motivated thresholds in the definition of the ionized IGM, although one can, in principle, define the ionized IGM as an arbitrary region in the $\varkappa - \delta$ plane rather than a region bounded by constant density and ionized fraction constraints. The thresholds for the two definitions of the IGM (corresponding to two pivot points) are shown with dotted and dashed lines in Figure 2.

The proposed definition, in contrast to the choice of arbitrary fixed density and ionized fraction thresholds, uses the specific distribution of gas in the $\varkappa$-δ plane to define the ionized IGM - which makes the density and the ionization fraction cuts time dependent (and, potentially, different in different simulations). Therefore, the applied thresholds account for the evolving ionizing background and for structure formation.

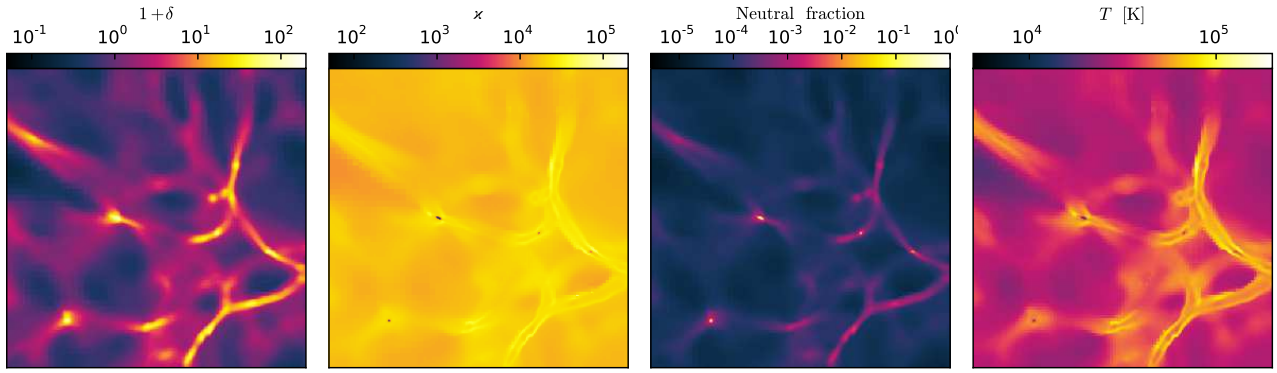

FIG. 3.— $2h^{-1}$ Mpc $\times$ $2h^{-1}$ Mpc slice at $z = 5.7$ showing filament structure. From left panel to right panel: gas density, ionization state indicator, neutral fraction of hydrogen, and temperature.

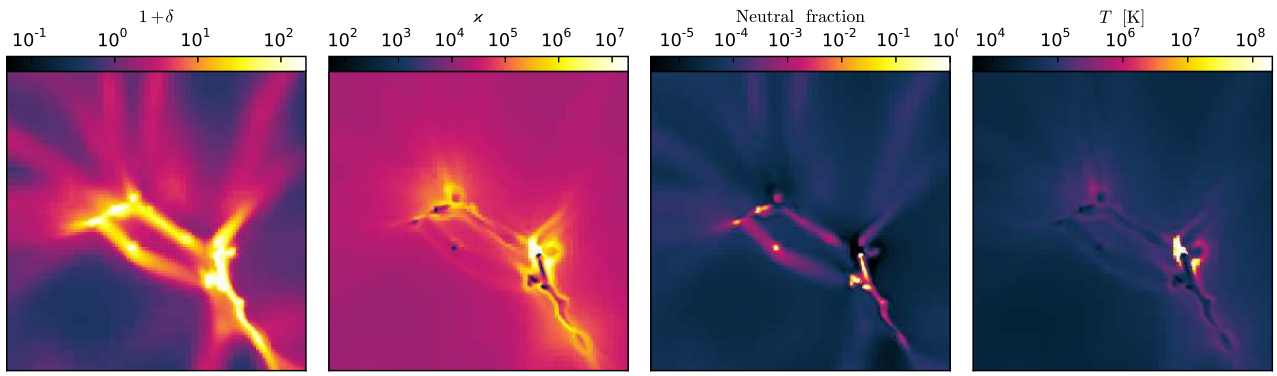

FIG. 4.— $1h^{-1}$ Mpc $\times$ $1h^{-1}$ Mpc slice at $z = 5.7$ showing a region with active merging. From left panel to right panel: gas density, ionization state indicator, neutral fraction of hydrogen, and temperature.

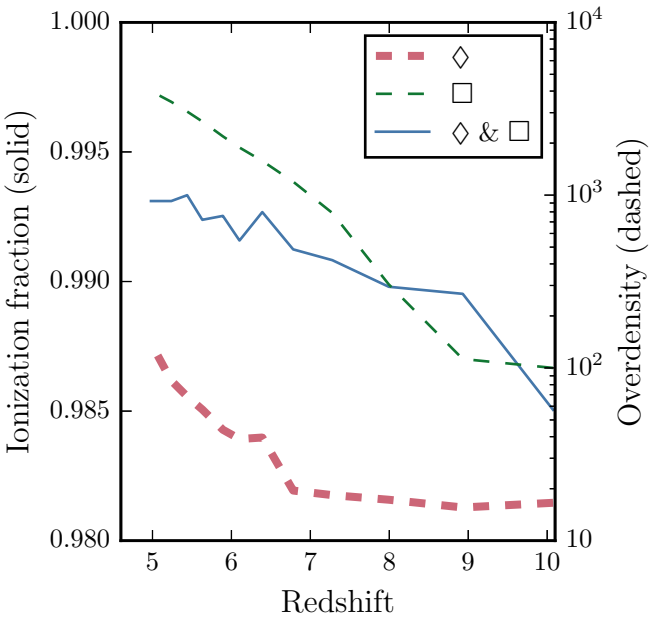

FIG. 5.— Density thresholds used for definition of $C_\diamond$ (thin green dashed line) and $C_\square$ (thick red dashed line), along with the ionization threshold used in both definitions (blue solid line).

Once we have defined the region in the $\delta - \varkappa$ phase space that we identify with the ionized IGM, we can calculate its clumping factor. Given a definition of the ionized IGM, we calculate the average squared ionized hydrogen density and divide it by the squared average ionized hydrogen density over all simulation cells that fall within the IGM definition,

$$C_{\rm IGM} = \frac{\sum_{i\in{\rm IGM}} V_i \sum_{i\in{\rm IGM}} V_i (1+\delta_i)^2 x_{{\rm H\,II},i}^2}{\left(\sum_{i\in{\rm IGM}} V_i (1+\delta_i) x_{{\rm H\,II},i}\right)^2}, \qquad (6)$$

where V_i, δ_i and $x_{{\rm H\,II},i}$ are the volume, the overdensity and the ionized fraction of the i-th cell correspondingly.

In order to distinguish various definitions of the ionized IGM that we discuss above, we will use symbols $C_\diamond$ and $C_\square$ to label the definitions of the IGM based on the thresholds from Figure 5 and a symbol $C_{\delta,x_{\rm H\,II}}$ for the definition of the IGM from fixed thresholds in density and ionized fraction.

It is worth mentioning that all the information necessary for calculating the clumping factor is contained in the phase diagram. Hence, instead of iterating over cells in the simulation box, one can integrate over the area S of the phase diagram, which is identified with the ionized IGM,

$$C_{\rm IGM} = \frac{\int_S N_M(1+\delta,\varkappa)\times(1+\delta)\,{\rm d}(1+\delta){\rm d}\varkappa}{\int_S N_M(1+\delta,\varkappa)/(1+\delta)\,{\rm d}(1+\delta){\rm d}\varkappa}, \qquad (7)$$

where N_M is mass weighted 2D histogram shown in Figure 2.

The result of such calculations is presented in Figure 6. Notice, that the adaptive thresholds may change for a different reionization model, to reflect the actual onset of self-shielding in the IGM. By accident, for our particular model, $C_\square$ and $C_{\delta<100,\,x_{\rm H\,II}>0.99}$ end up very similar.

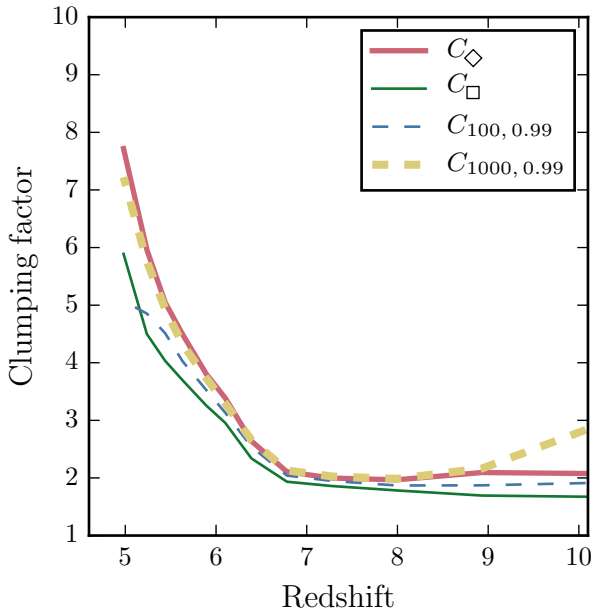

FIG. 6.— Clumping factor of the IGM as a function of redshift. Solid lines represents our two adaptive definitions of the ionized IGM, while dashed lines show models with fixed thresholds.

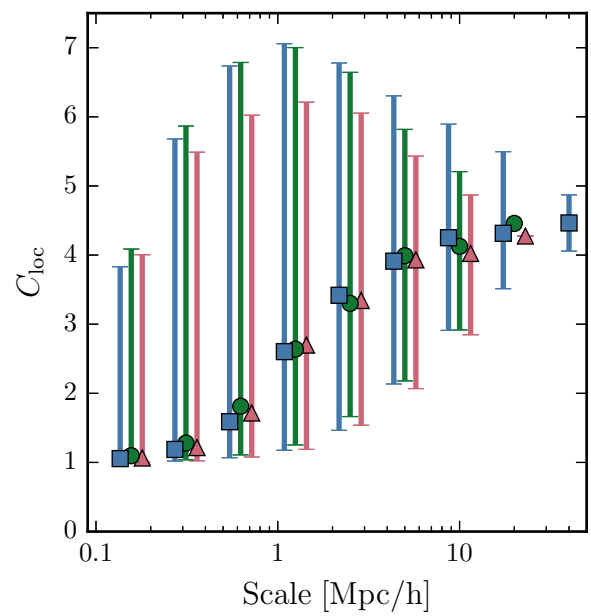

FIG. 7.— The distribution of local clumping factors at $z = 5.7$ versus scale it is defined over. The marker shows the median value and errorbars trace its 1%-99% scatter range. Blue squares correspond to all three $40h^{-1}$ Mpc simulations; the rightmost point and its errorbar correspond to the mean and the standard deviation of the clumping factor between the three independent realizations. Other two sets of points correspond to two $20h^{-1}$ Mpc simulations with identical initial conditions but different resolutions. Simulation which correspond to the red triangle has the same resolution as our $40h^{-1}$ Mpc runs, and the simulation shown with green circles has 8 times higher mass resolution and 2 times higher spatial resolution in the IGM.

Our two adopted definitions share the same ionization threshold, which does not change much with redshift in this particular model of reionization (see Figure 5). On the other hand, the density threshold evolves significantly, and the one associated with $C_\diamond$ is about two order of magnitude higher than the one used in $C_\square$. Both definitions account for ionized IGM; however, $C_\square$, in contrast to $C_\diamond$, includes less volume associated with the proximity effect (see Figure 2). This difference has a moderate influence on the global clumping factor at redshifts $z \gtrsim 6$ (see Figure 6), but reaches about 25% at $z \sim 5$. Therefore, if $C_\diamond$ definition is used, one needs to account for recombinations in the proximity zones by increasing proportionally the effective escape fractions from ionizing sources.

In contrast to previous works (Shull et al. 2012; So et al. 2013; Jeeson-Daniel et al. 2014) where fixed density threshold is used for defining clumping factor, we propose to derive it directly from a simulation. In McQuinn et al. (2011) the connection between ionizing background and such a density threshold has been studied and a power law connection has been found. This result has been used to determine the density threshold in Finlator et al. (2012) study. Similar approach has been adopted in Emberson et al. (2013), where the authors derive the critical density cut-off from their simulation and also find power law dependence. Our findings can be interpreted in a similar manner. The fact that the ionization fraction threshold (Figure 5) does not change significantly means that both our markers lie on the same line in $\varkappa - \delta$ phase diagram at all redshifts. Therefore, taking into account Equation 5, the density threshold δ has a power law dependence on average ionizing background Γ.

4.1. *Local variations of the clumping factor*

In addition to the temporal evolution of the global, averaged over the whole universe clumping factor, one may also be interested in its (in)homogeneity. We now explore spatial variations of the clumping factor at fixed redshift $z = 5.7$, when the IGM is already highly ionized. In this section we use our definition of the clumping factor ($C_\diamond$), however all general trends are expected to be similar for other definitions of the clumping factor as well.

In order to measure the spatial variations in the clumping factor in the simulations, we split the simulation box into cubic sub-boxes with sizes from $0.15h^{-1}$ Mpc up to $20h^{-1}$ Mpc (half the box size), and calculate the local clumping factor in each sub-box. We define the local clumping factor as:

$$C_{\rm loc} = \frac{\langle n_i^2 \rangle_{\rm loc}}{\langle n_i \rangle_{\rm loc}^2} = \frac{\langle n_i^2 \rangle_{\rm loc}}{\bar{n}_{i_{\rm Universe}}^2 (1+\bar{\delta}_{\rm loc})^2}, \qquad (8)$$

where n_i is number density in ionized regions, $\langle\rangle_{\rm loc}$ is the average inside each cubic sub-box, and $\bar{n}_{i_{\rm Universe}}$ and $\bar{\delta}_{\rm loc}$ are the cosmic mean density of ionized regions and the average overdensity of the sub-box. Notice, that there is the $(1+\bar{\delta}_{\rm loc})^{-2}$ dependence in our definition that accounts for the given sub-box being over- or under-dense.

The local clumping factor versus the spatial averaging scale is shown in Figure 7. A couple of trends can be observed. First, the median value of the clumping factor approaches unity at small scales, and the scatter is lower in smaller sub-boxes. Reduced clumping factor implies that smaller sub-boxes contain less substructure; when the clumping factor approaches unity, the averaging scale becomes comparable to a smoothing scale, below which the density field is uniform. There are two candidates for this smoothing scale: physical – the pressure smoothing scale, and numerical – the finite resolution of the simulation. The first one is the scale over which the pressure of the photo-ionized gas erases baryonic fluctuations (Gnedin & Hui 1998; Gnedin et al. 2003; Kulkarni et al. 2015). Therefore, the gas distribution in sub-boxes with sizes comparable to the pressure smoothing scale scale is more-or-

less uniform, and, consequently, the clumping factor is close to unity.

The exact value of the pressure smoothing scale depends on the variety of factors, such as thermal history and the degree of nonlinearity in the distribution of the particular subset of gas under consideration (namely, the gas that falls inside the cuts in Fig. 2). In the post-reionization epoch ($z \sim 2-4$) the linear pressure smoothing scale (often called "filtering scale") is about $50-70$ kpc (Gnedin et al. 2003); the nonlinear pressure smoothing scale, however, is somewhat larger, closer to 100 kpc at $z \sim 3$ (Kulkarni et al. 2015), and getting even larger (up to 200 kpc) as the IGM temperature increases to close to 10^5 K (i.e. closer to reionization). The latter value is similar to (albeith still smaller than) the scale in Fig. 7 at which the local clumping factor approaches unity. The exact comparison between these two quantities, however, would require a numerically expensive focused study, whose value would be largely academic.

However, simulations do not have infinitely fine resolution, so it is also possible that the simulation does not resolve this scale, and observed smoothing is just a resolution effect. In order to exclude the latter possibility, we perform the same analysis on two $20h^{-1}$ Mpc runs with identical initial conditions, but with different mass and spatial resolutions. The result is presented in Figure 7. The clumping factor and its scatter are only insignificantly larger in a higher resolution simulation; hence the behavior of the clumping factor shown in Figure 7 is real and not a numerical artifact.

The second obvious trend in Figure 7 is that, as the averaging scale increases, the median gets closer to the global clumping factor, and the scatter also decreases. Even if the actual values of the clumping factor on large scales are affected by the finite size of the simulation box, the qualitative behavior is as expected, since at the largest scales the universe is approaching homogeneity.

The scatter of the clumping factor peaks at the intermediate scale of a few h^{-1} Mpc. This scale is in the same order-of-magnitude range as several physical scales in the problem (galaxy clustering scale, typical size of ionized bubbles, the photon mean free path due to LLS, etc), so the reason behind the increase of scatter would be virtually impossible to isolate.

4.2. *Local clumping and density correlation*

The scatter in the local clumping factor from Figure 7 is not necessarily random. It can correlate with several properties of sub-boxes, of which the mean density is the primary candidate. In Figure 8 slices of density and the clumping factor in $1.25h^{-1}$ Mpc sub-boxes are presented. There is an obvious correlation between them. It motivates us to study this dependence in more detail.

We group sub-boxes by density and measure the distributions of clumping factors within sub-boxes of the same density. In Figure 9 we show the scatter of local clumping factor as a function of the mean density in $1.25h^{-1}$ Mpc sub-boxes. The correlation is apparent, but the scatter of the clumping factor at fixed density is still over a factor of 3-5. Nevertheless, this dependence can be used for sub-grid modeling of clumping factor in low resolution simulations.

By definition, the clumping factor is nothing else but the second moment (variance) of the probability distribution function (PDF) of density. Therefore we take a look at PDFs within sub-boxes with a given mean density, in order to explore what features in the PDF lead to the increased clumping factor in denser regions.

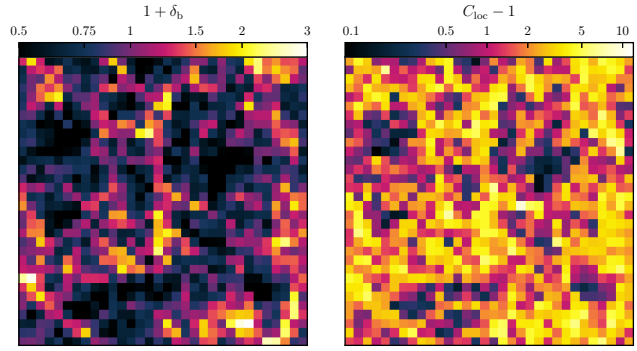

FIG. 8.— The slice of $40h^{-1}$ Mpc simulation at redshift 5.7. Colors represents the baryon density field (left panel), the local clumping factor (right panel). Both fields are defined on $1.25h^{-1}$ Mpc scale.

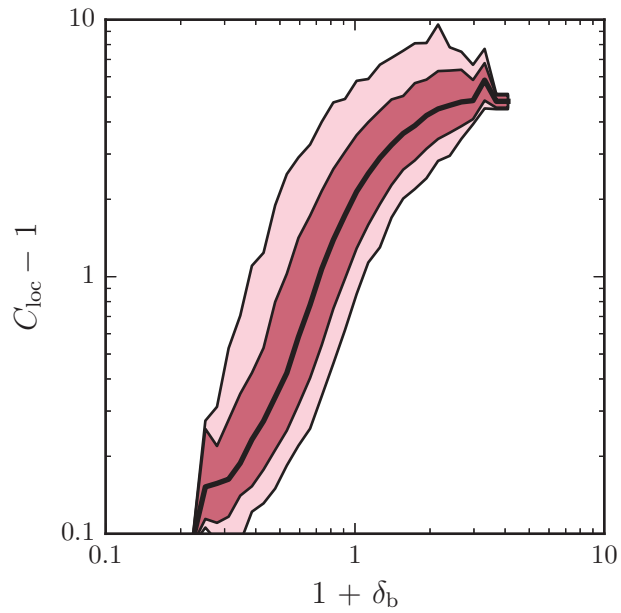

FIG. 9.— Local clumping factor versus average local density at $z = 5.7$ defined at scale of $1.25h^{-1}$ Mpc. Black solid line represents median and contours show 10%-90% and 1%-99% intervals.

These PDFs are presented in the left panel of Figure 10. The observed shapes correspond to neither normal nor log-normal distributions, and reveal a power law slope at high densities, followed by a partial break at densities where the self-shielding sets in (as a reminder, we only consider densities of the ionized gas, since only they contribute to the clumping factor, Equation 1). In sub-boxes of higher mean densities the self-shielding sets in at proportionally higher densities, so that when the densities in each sub-box are scaled by the mean sub-box density, the self-shielding threshold remains roughly constant, $(1+\delta)/(1+\bar{\delta}) \approx 100$ (at this redshift and for this reionization model).

From the shapes of PDFs in the left panel of Figure 10 it is not immediately clear which range of densities contributes most to the clumping factor. Therefore, we show in the right panel of Figure 10 the cumulative clumping factor as a function of the maximum density for the PDFs from the left panel. As one can see, almost all of the contribution to the clumping factor comes from modest densities, $(1+\delta)/(1+\bar{\delta}) \lesssim 10-100$,

well below the self-shielding threshold. Hence, the dominant contribution to the clumping factor comes form the densities around the peak of the PDF, and not from the high density tail. Hence, the increase of the clumping factor with the density is not unexpected: denser regions, being analogous to denser universes, are more evolved and, hence, have a wider density distribution.

5. CONCLUSIONS

Spontaneous recombination of ionized hydrogen takes away ionizing photons, and, hence, is an important physical process during and after cosmic reionization. Even though recombination is the same physical process no matter where it occurs, it is customarily quantified in analytical studies and in not-fully-self-consistent simulations with three separate quantities: the escape fraction, the IGM clumping factor, and the maximum mean free path.

Using fully self-consistent numerical simulations of cosmic reionization, we explore whether such a separation is physically motivated and robust. To that end, we use a convenient physical quantity, the "ionization state indicator" $\varkappa$ and a $\varkappa-\delta$ "phase diagram", to introduce a well-motivated and approximately robust definition of the IGM, and to compute its clumping factor. In comparing to previous work, we find that our physical definition of the IGM is reasonably well approximated by simple fixed thresholds in ionized fraction ($x_{\rm H\,II} > 0.99$) and density ($\delta \gtrsim 10^2 - 10^3$).

The largest ambiguity in the definition of the ionized IGM comes from the unvirialized regions around galaxies that are over-ionized by the local enhancement in the radiation field ("proximity zones"). That inherent ambiguity imposes a "systematic error" on the value of the clumping factor of about 20% (somewhat smaller during reionization but increasing to $\approx$25% at lower redshifts).

The fact that self-shielded neutral regions separate cleanly from the ionized IGM allows one to account for them separately. In ionized IGM recombinations are proportional to density squared, and therefore the clumping factor is a convenient descriptor. Recombinations on the surface of self-shielded regions only compensates for the ionizations from external radiation, and in that case the maximum mean free path of ionizing photons is an appropriate quantity.

The last regime of "three-regime" approach is galactic interiors. Even though our simulations do not resolve internal structure of galaxies, the compactness of galaxies and their mutual separation allow to consider them as isolated systems. Therefore, characterizing the escape of photons with a single number (the escape fraction) and neglecting angular inhomogeneity may be sufficient for many studies.

We also explore the scale-dependence of the clumping factor over the range of scales, faithfully represented in our simulations. We find that the clumping factor computed in sub-boxes of a given size correlates strongly, but not perfectly, with the mean density in such sub-boxes. This correlation is driven by the increase in the PDF width in denser sub-boxes (and not by their high density tails), which, being analogous to denser universes, are more evolved and, hence, have a wider density distribution.

Nevertheless, the correlation between the local clumping factor and the mean density over the scale it is computed is not perfect, and other factors introduce significant (factor of 3-5) scatter in the relation. In principle, numerical simulations would allow us to further explore that additional dependence; however, we do not engage in such study in this work, as its practical need is not presently clear.

Fermilab is operated by Fermi Research Alliance, LLC, under Contract No. DE-AC02-07CH11359 with the United States Department of Energy. This work was also supported in part by the NSF grant AST-1211190 and by the NASA grant NNX-09AJ54G. This work made extensive use of the NASA Astrophysics Data System and `arXiv.org` preprint server. This work was done with significant usage of YT package (Turk et al. 2011).

REFERENCES

Emberson, J., Thomas, R. M., & Alvarez, M. A. 2013, Astrophys.J., 763, 146
Finlator, K., Oh, S. P., Özel, F., & Davé, R. 2012, MNRAS, 427, 2464
Gnedin, N. Y. 2014, Astrophys.J., 793, 29
Gnedin, N. Y., Baker, E. J., Bethell, T. J., et al. 2003, Astrophys.J., 583, 525
Gnedin, N. Y., & Hui, L. 1998, MNRAS, 296, 44
Gnedin, N. Y., & Kaurov, A. A. 2014, Astrophys.J., 793, 30
Gnedin, N. Y., Kravtsov, A. V., & Chen, H.-W. 2008, ApJ, 672, 765
Gnedin, N. Y., Kravtsov, A. V., & Rudd, D. H. 2011, ApJS, 194, 46
Gnedin, N. Y., & Ostriker, J. P. 1997, Astrophys.J., 486, 581
Hutter, A., Dayal, P., Partl, A. M., & Müller, V. 2014, MNRAS, 441, 2861
Iliev, I. T., Scannapieco, E., & Shapiro, P. R. 2005, Astrophys.J., 624, 491
Jeeson-Daniel, A., Ciardi, B., & Graziani, L. 2014, Mon.Not.Roy.Astron.Soc., 443, 2722
Kaurov, A. A., & Gnedin, N. Y. 2013, ApJ, 771, 35
Kaurov, A. A., & Gnedin, N. Y. 2014, Astrophys.J., 787, 146
Kohler, K., Gnedin, N. Y., & Hamilton, A. J. 2007, Astrophys.J., 657, 15
Kuhlen, M., & Faucher-Giguère, C.-A. 2012, MNRAS, 423, 862
Kulkarni, G., Hennawi, J. F., OÃ±orbe, J., Rorai, A., & Springel, V. 2015
McQuinn, M., Lidz, A., Zahn, O., et al. 2007, Mon.Not.Roy.Astron.Soc., 377, 1043
McQuinn, M., Oh, S. P., & Faucher-Giguère, C.-A. 2011, The Astrophysical Journal, 743, 82
Norman, M. L., Reynolds, D. R., So, G. C., & Harkness, R. P. 2013, ArXiv e-prints
Pawlik, A. H., Schaye, J., & van Scherpenzeel, E. 2008, MNRAS, 394, 1812
Raičević, M., & Theuns, T. 2011, MNRAS, 412, L16
Shull, J. M., Harness, A., Trenti, M., & Smith, B. D. 2012, ApJ, 747, 100
So, G. C., Norman, M. L., Reynolds, D. R., & Harkness, R. P. 2013, ApJ, 789, 149
So, G. C., Norman, M. L., Reynolds, D. R., & Wise, J. H. 2014, ApJ, 789, 149
Sobacchi, E., & Mesinger, A. 2014, MNRAS, 440, 1662
Songaila, A., & Cowie, L. L. 2010, ApJ, 721, 1448
Trac, H., & Cen, R. 2007, Astrophys.J., 671, 1
Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, ApJS, 192, 9

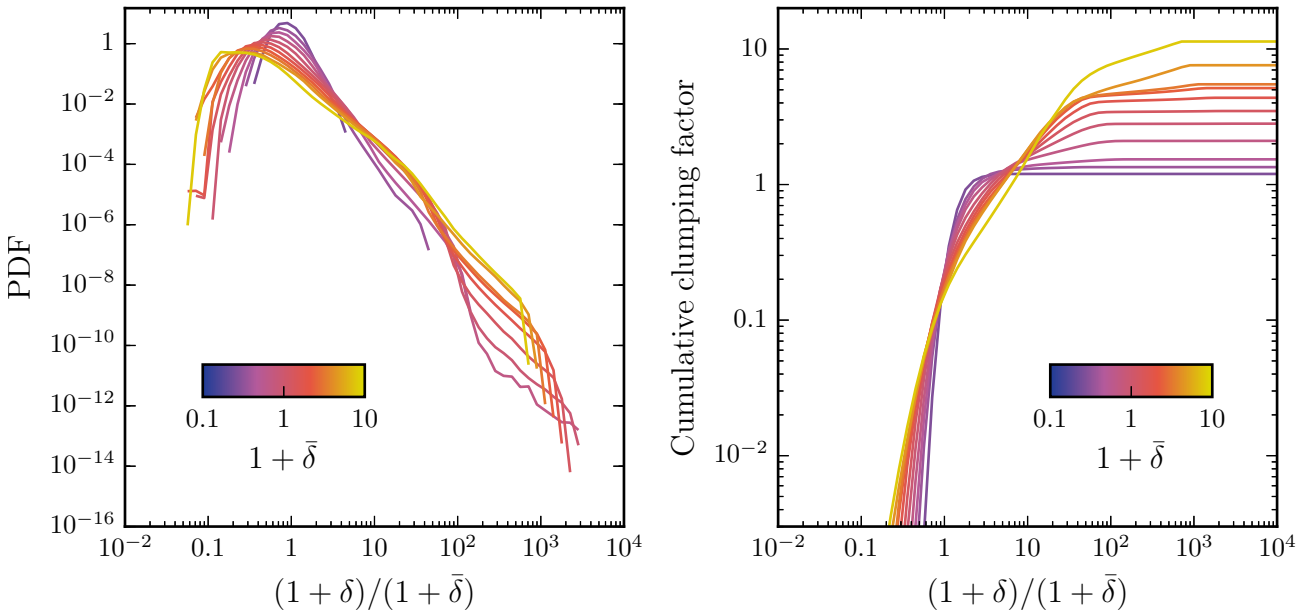

FIG. 10.— Left panel: probability distribution functions (PDF) of ionized baryon density in $1.25h^{-1}$ Mpc sub-boxes of given mean density (color coded) **at** $z = 5.7$. Right panel: the cumulative clumping factor as a function of density in the same sub-boxes (i.e., the variance (second moment) of the PDFs in the left panel as a function of the maximum density of integration).