The effect of the intergalactic environment on the observability of Lyα emitters during reionization

Ilian T. Iliev,[1,2]⋆ Paul R. Shapiro,[3] Patrick McDonald,[2] Garrelt Mellema[4]
and Ue-Li Pen[2]

[1]*Universität Zürich, Institut für Theoretische Physik, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[2]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St George Street, Toronto, ON M5S 3H8, Canada*
[3]*Department of Astronomy, University of Texas, Austin, TX 78712-1083, USA*
[4]*Stockholm Observatory, AlbaNova University Center, Stockholm University, SE-106 91 Stockholm, Sweden*

ABSTRACT

Observations of high-redshift Lyα sources are a major tool for studying the high-redshift universe and are one of the most promising ways to constrain the later stages of reionization. The understanding and interpretation of the data is far from straightforward, however. We discuss the effect of the reionizing intergalactic medium (IGM) on the observability of Lyα sources based on large simulations of early structure formation with radiative transfer. This takes into account self-consistently the reionization history, density, velocity and ionization structures and non-linear source clustering. We find that all fields are highly anisotropic and as a consequence there are very large variations in opacity among the different lines of sight. The velocity effects, from both infall and source peculiar velocity are most important for the luminous sources, affecting the line profile and depressing the bright end of the luminosity function. The line profiles are generally asymmetric and the line centres of the luminous sources are always absorbed due to the high density of the local IGM. For both luminous and average sources the damping wing effects are of similar magnitude and remain significant until fairly late, when the IGM is ionized between 30 and 70 per cent by mass.

The ionizing flux in the ionized patch surrounding a high-density peak is generally strongly dominated, particularly at late times, by the cluster of faint sources, rather than the central massive galaxy. Our results reproduce well the observed mean opacity of the IGM at $z \sim 6$. The IGM absorption does not change appreciably the correlation function of sources at high redshift. Our derived luminosity function assuming constant mass-to-light ratio provides an excellent match to the shape of the observed luminosity function at $z = 6.6$ with faint-end slope of $\alpha = -1.5$. The resulting mass-to-light ratio implies that the majority of sources responsible for reionization are too faint to be observed by the current surveys.

Key words: galaxies: formation – intergalactic medium – cosmology: theory – diffuse radiation – large-scale structure of Universe – radio lines: galaxies.

1 INTRODUCTION

The reionization of the universe was the last global transition of the intergalactic medium (IGM), from fully neutral after cosmic recombination at $z \sim 1100$ to fully ionized as we see it today, caused by the radiation from the first stars. Currently there are still only very few direct observational constraints on this epoch. The lack of Gunn–Peterson trough in the spectra of high-redshift sources indicates a low mean neutral fraction $x_{\mathrm{HI}} \lesssim 10^{-4}$ out to redshift $z \sim 6$,

which implies overlap was achieved sometime before that, at $z_{\mathrm{ov}} > 6$.

On the other hand, the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) 3-yr data (Spergel et al. 2007) yielded a fairly high value for the integrated Thomson electron scattering optical depth to the surface of last scattering, at $\tau_{\mathrm{es}} = 0.09 \pm 0.03$. This requires a significant ionized fraction out to high redshifts, $z > 12$, and thus implies an extended reionization. The optical depth by itself does not put very stringent constraints on the possible reionization histories, however. The reason for this is the self-regulated nature of the reionization process (Iliev et al. 2007a), whereby the Jeans-mass filtering of low-mass sources in the ionized regions results in the

⋆E-mail: iliev@physik.uzh.ch

τ_{es} and z_{ov}, the overlap redshift, being only loosely related. The overlap redshift z_{ov} is determined by the abundances and efficiencies of the high-mass sources, whose formation is not suppressed by reionization, while τ_{es} depends on both high- and low-mass sources. Thus, varying the ionizing efficiencies of the small sources yields a wide range of τ_{es} values for the same value of z_{ov}.

This relative lack of observational data is set to change dramatically in the near future due to a number of large observational projects which are currently under way. The 21-cm data from high redshifts contains potentially the richest set of information since the signal is inherently three dimensional, on the sky and in redshift/frequency (e.g. Madau, Meiksin & Rees 1997; Tozzi et al. 2000; Iliev et al. 2002, 2003; Zaldarriaga, Furlanetto & Hernquist 2004; Mellema et al. 2006b; Shapiro et al. 2006; Iliev et al. 2008a, see Furlanetto, Oh & Briggs 2006a for a detailed recent review). The features that could be derived include the full reionization history, geometry, statistics and individual bright features (Furlanetto et al. 2006a; Mellema et al. 2006b; Iliev et al. 2008a). There are significant challenges to be overcome; however, particularly related to precise subtraction of the very strong foregrounds present at low frequencies (e.g. Furlanetto, Oh & Briggs 2006a; Morales, Bowman & Hewitt 2006; Iliev et al. 2008a).

The patchiness of reionization creates secondary temperature anisotropies in the cosmic microwave background (CMB) through the kinetic Sunyaev–Zel'dovich effect (Gruzinov & Hu 1998; Hu 2000; Gnedin & Jaffe 2001; Santos et al. 2003; McQuinn et al. 2005; Iliev et al. 2007b), as well as polarization anisotropies (Hu 2000; Haiman & Holder 2003; Santos et al. 2003; Mortonson & Hu 2007; Doré et al. 2007). Unlike the 21-cm signal, the reionization signatures in the CMB are integrated over the reionization history and contain no frequency information. However, the typical scales of reionization are reflected in a characteristic peak of the kSZ anisotropy signal (Iliev et al. 2007b), and the shape of the power spectrum is dependent on the reionization parameters (source efficiencies and small-scale gas clumping). CMB anisotropy observations can therefore provide us with key information about the process of reionization and since its systematics are different it would be an important complement to the 21-cm studies. One could also combine these observations more directly, by using 21-cm observations to derive the Thomson optical depth fluctuations (Holder, Iliev & Mellema 2007). Small-scale CMB anisotropy and polarization measurements would be quite difficult due to the weakness of these signals, but are within the abilities of modern detectors (Doré et al. 2007; Iliev et al. 2007b).

Narrow-band searches for high-redshift Lyα emitters have been very successful at finding sources at ever higher redshifts, currently up to $z \sim 7$ (e.g. Hu et al. 2002; Kodaira et al. 2003; Rhoads et al. 2003; Malhotra & Rhoads 2004; Stanway et al. 2004; Hu et al. 2005; Taniguchi et al. 2005; Bunker et al. 2006; Kashikawa et al. 2006b; Shimasaku et al. 2006; Ota et al. 2008). Together with studies of the Lyα resonant absorption in the IGM (e.g. Fan et al. 2002; White et al. 2003; Fan 2004; Fan et al. 2006b) they provide important independent approaches for studying reionization (see Fan, Carilli & Keating 2006a, for a recent review). The optical depth for Lyα resonant absorption in neutral IGM at high redshifts is quite large (Shklovskii 1964; Gunn & Peterson 1965), thus absorption studies are mostly sensitive to very low hydrogen neutral fractions, typically below $x_{HI} \sim 10^{-4}$, otherwise the absorption saturates. This technique is thus best suited for studying highly ionized regions and the very end of reionization and has been very successful for setting low limits for the redshift at which reionization was completed. On the other hand, Lyα emitter surveys do not require a very low average

hydrogen neutral fraction and thus in principle can probe further into the reionization epoch (Malhotra & Rhoads 2004). Other related probes include damping wing measurements (e.g. Miralda-Escude 1998; Mesinger & Haiman 2004; Totani et al. 2006), quasar H II region sizes and their evolution (e.g. Wyithe & Loeb 2004; Fan et al. 2006b; Bolton & Haehnelt 2007a; Maselli et al. 2007; Mesinger & Haiman 2007), and sizes of dark gaps in high-z spectra (e.g. Fan et al. 2006b; Gallerani et al. 2008).

The correct interpretation of these data for reionization is far from straightforward, however. The strong dependence of Lyα absorption on the neutral fraction and gas density means that both should be modelled with certain precision. High-redshift sources are rare and strongly clustered (see e.g. Barkana & Loeb 2004b; Furlanetto, Zaldarriaga & Hernquist 2004b; Iliev et al. 2007a). As a result, H II regions generally contain multiple ionizing sources and grow much larger than the ones created by individual sources. This minimizes the effect of the damping wings and increases the transmission, allowing the detection of more and fainter sources than would be naively expected (Wyithe & Loeb 2005). However, while this is the generic expectation, the actual effect would be dependent on the exact geometry of reionization, e.g. sources close behind a neutral region will be damped even if they are inside a very large H II region. Simplified models typically assume spherical ionized regions and either ignore source clustering or assume linear bias (e.g. Cen & Haiman 2000; Furlanetto, Hernquist & Zaldarriaga 2004a; Santos 2004; Haiman & Cen 2005; Wyithe, Loeb & Carilli 2005; Kramer, Haiman & Oh 2006; Wyithe & Loeb 2007). In practice the large ionized regions form by local percolation of multiple smaller ones and as a consequence are highly non-spherical (Iliev et al. 2006a; Mellema et al. 2006c; Iliev et al. 2007a). The inhomogeneous cosmological density fields and non-equilibrium chemistry effects (particularly in recently ionized gas) further complicate the picture and point to the need of following the cosmological structure formation and the reionization history of a given region. A proper account of all these effects can only be done through detailed cosmological radiative transfer simulations.

A number of radiative transfer methods have been developed in recent years and now they are reaching a certain level of maturity and are producing fairly reliable results (Iliev et al. 2006b). However, performing large-scale reionization simulations, as required for Lyα studies is still technically very challenging. Recently McQuinn et al. (2007) used large-scale structure formation numerical simulations post-processed with radiation transfer to study the observability of Lyα emitters at high redshifts and what these can tell us about reionization. In order to achieve high dynamic range, these authors employed a subgrid model for the collapse of the smallest haloes. Another, seminumerical approach was used by Mesinger & Furlanetto (2008a), who took the linear density and velocity fields at early time (essentially the initial conditions for an N-body simulation) and used an excursion-set approach combined with a first-order Lagrangian theory to 'paint' the H II regions on the density field. This procedure provides large dynamic range at a low cost, but at the expense of making significant approximations and thus cannot fully replace full numerical simulations.

In this paper we use the results of large-scale numerical simulations to study the transfer of Lyα through the IGM. In Section 2 we briefly describe our simulation method. In Section 3 we describe the evolution and environment of a rare, massive source, similar to the ones that are currently observed. In Section 4 we address the observability of Lyα emitters, considering the reduction of the transferred line flux due to the absorption in the IGM and luminosity functions. In Section 5 we describe the effects of the patchiness

of reionization on the angular correlation function of Lyα emitters, and, finally, in Section 6 we sum up our conclusions.

2 SIMULATIONS

Our simulation results follow the full, self-consistent reionization history in a large volume of $(100\,h^{-1}\,\mathrm{Mpc})^3$ and were described in detail in Iliev et al. (2006a); Mellema et al. (2006b) and Iliev et al. (2007a). While this volume is too small to allow us to consider the rarest, most luminous sources like the SDSS QSOs (Fan et al. 2002, 2006b), we have sufficient resolution to locate the majority of sources responsible for reionization and take explicit account of their radiation, and to derive good-quality absorption spectra.

Of the range of simulations presented in Iliev et al. (2007a) we here consider one specific run, labelled f250C. Our simulations were performed using a combination of two very efficient computational tools, a cosmological particle-mesh code called PMFAST (Merz, Pen & Trac 2005) for following the structure formation, whose outputs are then post-processed using our radiative transfer and non-equilibrium chemistry code called C^2-RAY (Mellema et al. 2006a). The parameter $f_\gamma = 250$ characterizes the emissivity of the ionizing sources – how many ionizing photons per gas atom in the (resolved) haloes are produced and manage to escape from the host halo within ~ 20 Myr, which is the time between two consecutive density slices, equal to two radiative transfer time-steps, while 'C' indicates that this run models the gas clumping at small (subradiative transfer grid) scales based on a fit given by

$$C_{\mathrm{subgrid}}(z) = 26.2917\mathrm{e}^{-0.1822z+0.003\,505\,z^2}, \tag{1}$$

based on the used *WMAP*3 cosmology (a good fit for $6 < z < 30$). This fit to the small-scale clumping factor is a more precise version of the one we presented in Iliev, Scannapieco & Shapiro (2005). We derived it based on a PMFAST simulation with a computational mesh of 3248^3 and with particle number of 1624^3, with computational volume of $(3.5\,h^{-1}\,\mathrm{Mpc})^3$. These parameters correspond to particle mass of $10^3\,\mathrm{M}_\odot$, minimum resolved halo mass of $10^5\,\mathrm{M}_\odot$, and a spatial resolution of ~ 1 kpc comoving.

Our $(100\,h^{-1}\,\mathrm{Mpc})^3$ volume simulation resolves all haloes with mass above $2.2 \times 10^9\,\mathrm{M}_\odot$, higher than the mass above which atomic line cooling of hydrogen becomes effective, which is $\sim 10^8\,\mathrm{M}_\odot$. As a consequence, our treatment does not include the contribution of low-mass sources to reionization. Higher resolution, smaller box simulations which do include all ionizing sources above the atomic line cooling limit (Iliev et al. 2007a) indicate that the effects of low-mass sources are primarily confined to the earliest stages of reionization, when such sources are dominant. Throughout most of the reionization process, and especially during its late stages, the low-mass sources, which are strongly clustered around the high-density peaks, are heavily suppressed due to Jeans-mass filtering in the ionized regions and thus have limited effect on the reionization progress and large-scale geometry. Since the Lyα observations largely probe the later stages of reionization, where the neutral gas fraction is ~ 30 per cent or less (Malhotra & Rhoads 2004), we do not expect that our conclusions will be strongly affected by the absence of low-mass sources. Larger simulations, currently in progress (Iliev et al. 2008b; Shapiro et al. 2008), which resolve all atomically cooling haloes in ~ 100 Mpc boxes will settle these uncertainties.

Throughout this paper we assume a flat ($\Omega_{\mathrm{k}} = 0$) Lambda cold dark matter (ΛCDM) cosmology (Ω_{m}, Ω_Λ, Ω_{b}, h, σ_8, n) = $(0.24, 0.76, 0.042, 0.73, 0.74, 0.95)$ based on *WMAP* 3-yr results (Spergel et al. 2007), hereafter *WMAP*3. Here $\Omega_{\mathrm{m}}, \Omega_\Lambda$ and Ω_{b} are

the total matter, vacuum and baryonic densities in units of the critical density, σ_8 is the rms density fluctuations extrapolated to the present on the scale of $8\,h^{-1}$ Mpc according to the linear perturbation theory, and n is the index of the primordial power spectrum of density fluctuations.

3 LUMINOUS HIGH-REDSHIFT SOURCES AND THEIR ENVIRONMENT: PROPERTIES, EVOLUTION AND REIONIZATION HISTORY

The luminous sources at high redshift, the ones typically seen in current surveys, are hosted by rare, massive haloes which form at the location of the highest peaks of the density field. The statistics of Gaussian fields predicts that such high-density peaks are rare and highly clustered in space, more strongly so at high redshifts. As a consequence, each high-redshift, massive galaxy should be surrounded by numerous smaller ionizing sources. The self-consistent reionization history simulations of such regions require following a sufficiently large volume, in order to obtain the correct statistics and biasing of the rare peaks, while at the same time resolving all the low-mass haloes which are the main drivers of the reionization process. Our current radiative transfer simulations are able to achieve this. We also note that correctly modelling the non-linear bias of the rare peaks in semi-analytical modes is a difficult and still unsolved problem. As a consequence, the halo clustering in semi-analytical models is typically underestimated, and in some cases even ignored. This often yields incorrect results, e.g. in estimates of the suppression of low-mass sources by Jeans-mass filtering (e.g. Iliev et al. 2007a).

3.1 Mass accretion history of massive haloes at high redshift

In Fig. 1 we show the mass-growth history of the three most massive (at redshift $z = 6$) haloes found in our computational volume, with
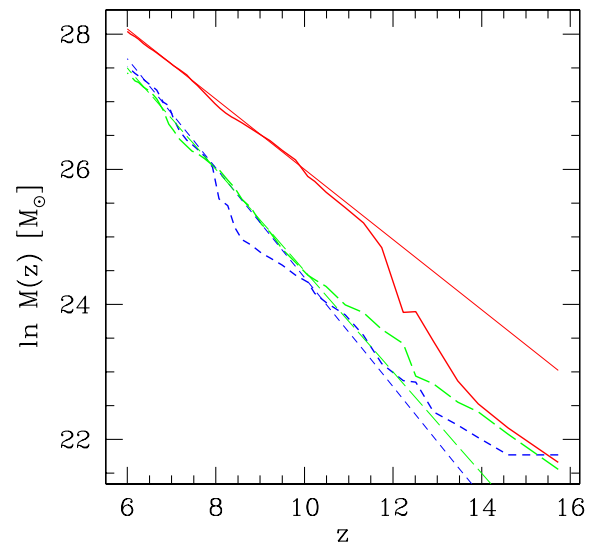

Figure 1. Mass accretion history of the three most massive haloes found in our computational volume at redshift $z = 6$. The mass growth is roughly exponential in redshift, and is well approximated by $M(z)[\mathrm{M}_\odot] = \exp(A - \alpha z)$, where $A = 31.2, \alpha = 0.52$ for most massive halo (solid, red), $A = 32.5, \alpha = 0.81$ for the next most massive halo (short-dashed, blue) and $A = 32.0, \alpha = 0.75$ for the third most massive halo (long-dashed, green). Fits are shown by the thin straight lines (with corresponding line types and colours).

final masses of $1.5 \times 10^{12}, 9 \times 10^{11}$ and $8.2 \times 10^{11}\,M_\odot$, respectively. All correspond to very rare, ≈ 4.5–5σ peaks of the density field. The first progenitors of these haloes resolved in our simulation ($M_{halo} > 2 \times 10^9\,M_\odot$) form very early, at $z \sim 16$. Thereafter, the halo masses grow roughly exponentially with redshift, $M \propto \exp(-\alpha\,z)$. This behaviour is in good agreement with previous results on the growth of dark matter haloes in hierarchical ΛCDM models (Wechsler et al. 2002), as well as with similar results obtained in more generic, idealized situations (Shapiro et al. 2004). The slopes α for our haloes exhibit much smaller scatter ($0.52 < \alpha < 0.81$; see Fig. 1 caption for the complete fitting formulae) than the ones found by Wechsler et al. (2002) ($0.4 < \alpha < 1.6$). The exponential fits are excellent during periods when no major mergers (i.e. mergers with haloes of similar mass) occur. This is the case e.g. at late times ($z < 11$) by which time these haloes have grown enough to dominate their surroundings and no similar-mass haloes remain in their vicinity. The exception here is the second most massive halo (short-dashed line). Its last major merger occurs at $z \sim 8$. At early times ($z > 12$) the mass evolution of all three haloes does not follow the overall exponential trend. Instead, the most massive halo is growing faster than its long-term trend, through series of rapid major mergers, while the other two haloes initially grow slower than their long-term trend.

3.2 Reionization history

In Fig. 2 (top and third row) we illustrate the main stages of the reionization history of a high-density peak and its intergalactic environment. The particular peak we show here is the one surrounding the largest mass halo found in our computational volume at redshift $z = 6$. For clarity and convenience for the reader we have shifted the peak to the box centre using the periodicity of our computational volume. The first resolved source in this region forms at $z \sim 16$. By redshift $z = 12.9$ (top left-hand panel; mass-weighted ionized fraction is $x_m = 0.001$ at this time) the central halo has already undergone several major mergers with nearby haloes and its total mass has grown to $1.5 \times 10^{10}\,M_\odot$. The combined effect of the central halo and the nearby members of the same source cluster is to create a substantial ionized region (of size $R \sim 1\,h^{-1}$ Mpc, defined as the radius at which the integrated continuum optical depth from the source at the Lyman limit reaches unity). At this time there are only a few ionized regions in our computational volume (and just two intersect the plane shown). By $z = 10.1$ (top middle; $x_m = 0.10$) many more H II regions appear and both the sources and the ionizing regions are strongly clustered in space. The central region still remains the largest one. By $z = 9.0$ (top right-hand panel; $x_m = 0.28$) many more haloes have formed, most of them in large clustered groups. The H II region surrounding the central peak remains among the largest ($R \sim 6\,h^{-1}$ Mpc), but several other ionized bubbles reach comparable sizes.

At redshift $z = 7.9$ (bottom left-hand panel; $x_m = 0.66$) some quite sizable regions, more than 10 Mpc across, have percolated, reaching local overlap. The central bubble has grown to a size of $R \sim 8\,h^{-1}$ Mpc. The reionization geometry becomes quite complex, with most ionized bubbles becoming interconnected, while leaving large neutral patches in between. By $z = 7.0$ (bottom middle; $x_m = 0.94$) the notion of isolated, quasi-spherical H II regions becomes rather meaningless, since all ionized regions have already merged into one topologically connected region, although substantial neutral patches still remain interspersed throughout our volume. The volume remains on average quite optically thick to ionizing continuum radiation. Finally, at $z = 6.0$ (bottom right-hand panel;

$x_m = 0.9999$; $R = 18.7\,h^{-1}$ Mpc) our volume is well beyond overlap (which we define by $x_m > 99$ per cent). Only by that time the volume becomes on average optically thin to ionizing radiation.

The corresponding panels in the second and fourth row of Fig. 2 show images of the (non-equilibrium) ionization rate distribution at the same redshifts. The distribution is highly inhomogeneous, following the patchiness and peaking strongly in the vicinity of large source clusters. The volume-averaged photoionization rates at redshifts $z = (12.9; 10.1; 9.0; 7.9; 7.0; 6.0)$ in units of $10^{-12}\,s^{-1}$ are $\Gamma_{-12} = (7.0 \times 10^{-2}; 1.3 \times 10^{-1}; 2.3 \times 10^{-1}; 7.6 \times 10^{-1}; 1.3; 4.9)$, growing strongly with time as larger fraction of the volume becomes ionized and ever more sources form. Except for the earliest times, the peak photoionization rate values on the grid (corresponding to the brightest points of the images) remain fairly constant with time, at the same redshifts they are $\Gamma_{-12} = (1.0 \times 10^2; 2.0 \times 10^3; 2.0 \times 10^3; 2.5 \times 10^3; 2.8 \times 10^3; 5.1 \times 10^3)$. The photoionization rate distributions and evolution are discussed further in Section 4.4.

To further characterize the shape and boundary sharpness of the central H II region in Fig. 3 we show a histogram of the radial distance R from the central source at which the cumulative continuum optical depth at the ionizing threshold of hydrogen along each line of sight (LOS) reaches unity (black) and 4.6 (red; this optical depth value corresponds to 1 per cent transmission). A spherically symmetric H II region would correspond to a single value for the radius for a given optical depth, regardless of the directionality, while any scatter around the peak would be a measure of the non-sphericity of the ionized region. Furthermore, a comparison between the distributions for the two optical depths measures the sharpness of the H II region boundary, as follows. A sharp transition from an ionized to a neutral gas along a given LOS results in both optical depth values being surpassed simultaneously (since the neutral gas is extremely optically thick) and in such situation the histograms would coincide. On the other hand, if the boundary of the ionized region is not clearly defined due to local percolation, gas is highly ionized and the optical depth along the LOS increases only slowly, thus the values of 1 and 4.6 are reached at very different distances from the source.

At redshift $z = 9$ ($x_m = 0.28$) the H II region surrounding the central source is largely spherical and its boundary is well defined, albeit with some modest spread around the peak, indicating some departures from sphericity, in agreement with what was seen in Fig. 2. At redshift $z = 8$ ($x_m = 0.62$) the H II region remains fairly spherical, slightly more so for $\tau = 1$. The $\tau = 4.6$ histogram has a long tail at large values of R, up to $\sim 25\,h^{-1}$ Mpc, meaning that a small percentage of the LOS reach 99 per cent opacity only at such fairly large distances. Both distributions are somewhat wider than before and start to depart from each other, reflecting the increasing 'fuzziness' of the bubble boundary as it merges with other nearby bubbles due to local percolation.

At $z = 7$ ($x_m = 0.94$) the $\tau = 1$ distribution still retains a well-defined peak, albeit one accompanied by a long high-R tail reaching out to tens of Mpc. However, the $\tau = 4.6$ histogram changes its character completely, becoming very broad, with only a low peak close to the $\tau = 1$ peak and several secondary peaks. A significant number (~ 14 per cent) of the LOS do not reach $\tau = 4.6$ within $50\,h^{-1}$ Mpc of the central source (these are collected in the last, $R = 50\,h^{-1}$ Mpc bin of the histogram). Finally, at $z = 6$ ($x_m = 0.9999$), well beyond the overlap epoch ($z \sim 6.6, x_m = 0.99$) there is no indication of a defined ionized bubble. Both distributions are very broad, with no clear peak. At this time the optical depth is determined by the tiny remaining neutral fraction, which in its turn is dictated by the density variations of the cosmic web. Either value
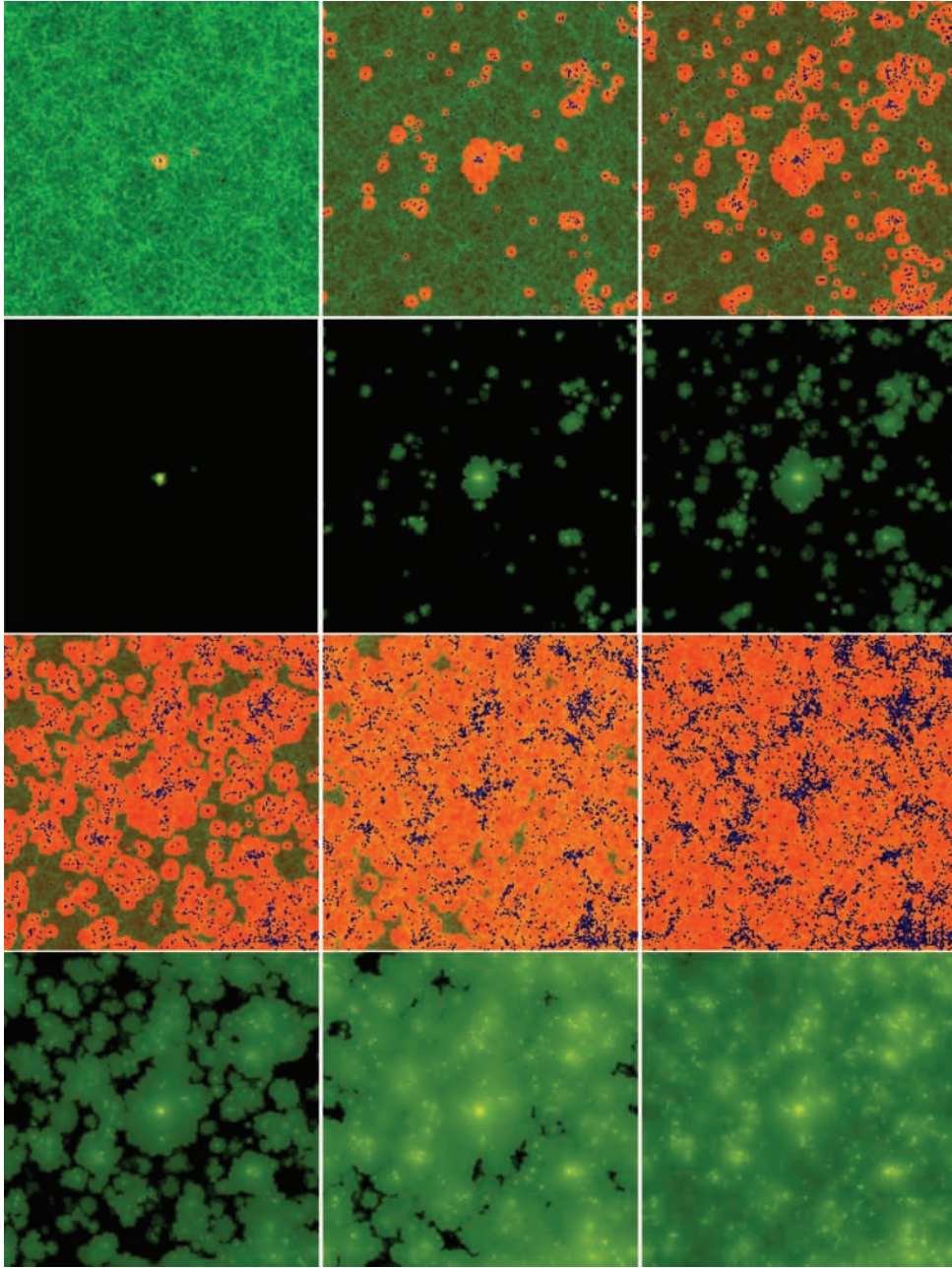
Figure 2. The reionization history of a high-density peak. The images are centred on the most massive (at $z = 6$) halo in our computational volume and are of size $100\,h^{-1}$ Mpc to the side. The snapshots are at: (top row) $z = 12.9$ (left-hand panels; global mass-weighted ionized fraction $x_{\mathrm{m}} = 0.001$), $z = 10.1$ (middle panels; $x_{\mathrm{m}} = 0.10$), $z = 9.0$ (right-hand panels; $x_{\mathrm{m}} = 0.28$) and (third row) $z = 7.9$ (left-hand panels; $x_{\mathrm{m}} = 0.66$), $z = 7.0$ (middle; $x_{\mathrm{m}} = 0.94$) and $z = 6.0$ (right-hand panels; $x_{\mathrm{m}} = 0.9999$). The underlying cosmological density field (dark, green) is superimposed with the ionized fraction (light, orange) and the cells containing ionizing sources (dark, blue dots), the slices thickness of $0.5\,h^{-1}$ Mpc (1 cell) in the density and ionized fraction fields and $10\,h^{-1}$ Mpc in terms of sources. The corresponding images of the (non-equilibrium) photoionization rates ($0.5\,h^{-1}$ Mpc thickness) are shown in the second and bottom rows.

of τ is reached at a wide range of distances in different directions, from $\sim 10\,h^{-1}$ Mpc to over $50\,h^{-1}$ Mpc. The IGM becomes largely optically thin to ionizing radiation and thus for most LOS neither $\tau = 1$ nor $\tau = 4.6$ is reached within $50\,h^{-1}$ Mpc from the source.

We also note that these results on the H II region boundaries are valid for soft, stellar (either Population II or Population III) spectra. Should hard sources, e.g. QSOs, contribute significantly to reionization the ionized region boundaries will inevitably be thicker and the transition from ionized to neutral smoother, due to the longer mean free paths of the hard photons. However, most current observational indications are that stellar sources dominate reionization, in which case the main effect of the relatively few hard photons present will be to heat the neutral IGM, but not to ionize it significantly.

3.3 Halo clustering in the vicinity of a luminous source

Taking a closer look at the halo clustering in the immediate vicinity of a high-density peak in Fig. 4 we show the projected spatial distribution of haloes around the most massive halo at $z = 7$. There
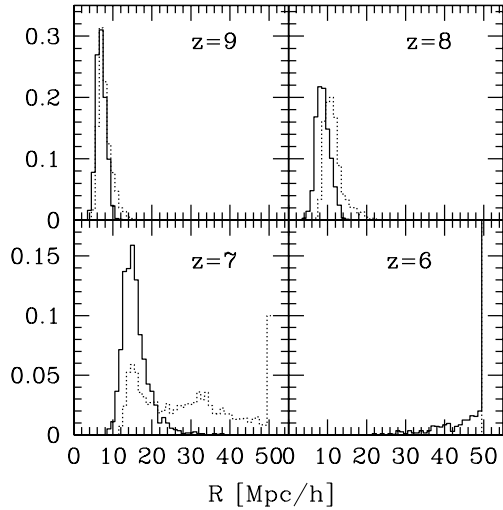
Figure 3. Histogram of the bubble size distributions along LOS. Shown are the distances from the source (the most massive halo in our volume at $z = 6$) at which continuum optical depth (at the hydrogen ionization threshold) surpasses $\tau = 1$ (solid lines) or $\tau = 4.6$ (dotted lines), at several illustrative redshifts, as labelled.

Figure 5. Histograms in radial bins of the total cumulative emissivity from all haloes within distance R, $N_{\rm ph, tot}(R)$, measured from the central object, here the most massive object in the computational volume, in units of the central object's emissivity, $N_{\rm ph, central}$. Shown are (bottom to top on the right-hand side) $z = 12.9, 9, 8, 7$ and 6.25 ($x_{\rm m} = 0.001, 0.28, 0.62, 0.94$ and 0.998). The last radial bin contains the total emissivity of the computational box at that time. The vertical lines on top (corresponding line types and colours) indicate the average size of the H II region at the corresponding redshift.

are 31 resolved haloes within $1\ h^{-1}$ Mpc from the most massive halo and 360 resolved haloes within $5\ h^{-1}$ Mpc (both including the largest halo itself). The area of each circle is proportional to the mass (and thus, in accordance to our source model, to the luminosity) of the corresponding halo. Haloes are distributed very anisotropically, concentrating preferentially along the density filaments and sheets of the local cosmic web.

The mass of the most massive halo is $9 \times 10^{11}\,{\rm M}_\odot$ at that time, well above any other halo in its vicinity. None the less, the low-mass, but numerous haloes surrounding the peak contribute significantly to the total ionizing emissivity. In order to quantify this point further, in Fig. 5 we show the cumulative emissivity versus radial distance from the central halo for several redshifts spanning the whole range of interest here. At all redshifts only within $\sim 2\,h^{-1}$ Mpc from itself does the most massive halo dominate the total photoionizing emission (i.e. it contributes more than 50 per cent of the cumulative total emission coming from that region). The exception is the highest redshift ($z = 12.9$), in which case the surrounding cluster of sources, rather than the central source, dominate the total flux even within $1\,h^{-1}$ Mpc of the peak. The rea-
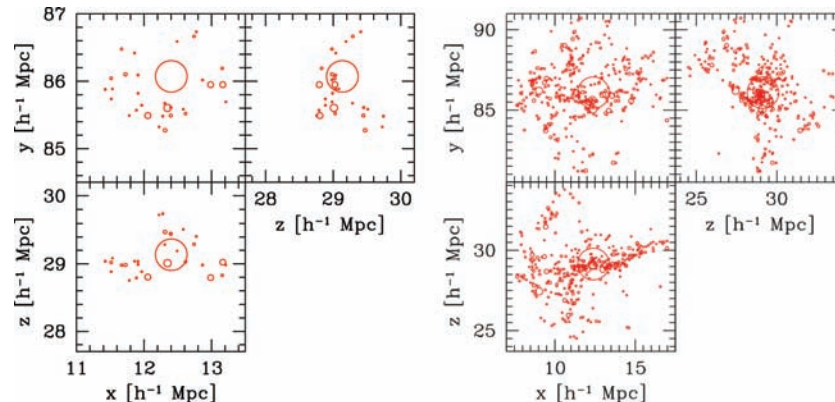
son for this is the presence nearby of another halo of almost the same mass as the central one, which shortly thereafter merges with it. The cumulative emission is dominated by the small halo contribution beyond that distance. Within $10\,h^{-1}$ Mpc the central source contributes only 10–30 per cent of the total emission. As more and more haloes form and the lowest mass ones become relatively common the fractional contribution of the most luminous source to the total emissivity gradually decreases in time for all radii larger than a few comoving Mpc. At $30\,h^{-1}$ Mpc the central source is dominated by the rest by 1.5–2 orders of magnitude. As fraction of the total emissivity of all sources in our computational volume (shown in the last bin of the histogram) the most luminous one contributes only ~ 2 per cent of the total at $z = 12.9$, decreasing to ~ 0.1 per cent at $z = 6.25$.

In Fig. 5 we also indicate the current H II region mean radius (vertical lines of corresponding line types and colours). Within its own bubble the central source contributes ~ 50 per cent of the

Figure 4. Projected distributions at $z = 7$ of haloes within $1\,h^{-1}$ Mpc comoving (left-hand part) and within $5\,h^{-1}$ Mpc comoving from the most massive object in the box. Circled areas are proportional to the mass of the corresponding halo.
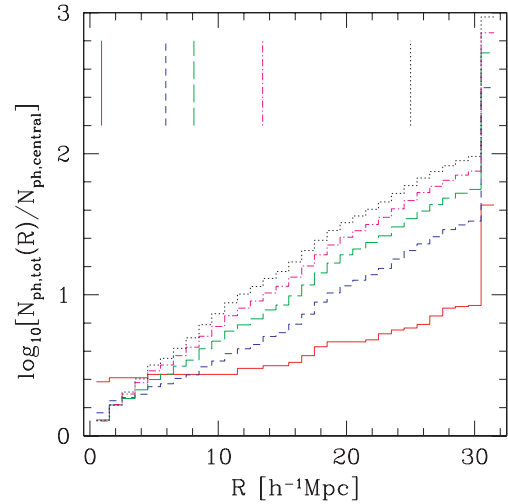
emission at $z = 12.9$, decreasing to $\sim$10 per cent at $z = 7$ and $\sim$1 per cent at $z = 6.25$.

Recently, Wyithe & Loeb (2005) presented a semi-analytical model of the source clustering at high redshift. They found that a central galaxy at $z = 7$ and with a velocity dispersion similar to our most massive source, $\sigma_V \approx 200\,\mathrm{km\,s^{-1}}$) contributes $\sim$40–70 per cent of the H II region radius, or $\sim$10–35 per cent of its volume (see their fig. 2, right-hand panel), a factor of a few larger compared to our results where we find that the central galaxy contribution to the total flux is $\sim$1–10 per cent. This discrepancy is most probably due to underestimate of source clustering in their bias model, and possibly also to the differences in the assumed source efficiencies. Furthermore, we have only considered a single massive source, which of course does not account for random statistical variations from source to source. Nevertheless, this result underlines the importance of considering the luminous sources individually (as opposed to considering an averaged 'mean' source) and using simulations in order to account for non-linear bias effects.

3.4 IGM environment of luminous sources at high redshift

The Lyα absorption is strongly influenced by the local IGM environment – density, velocity and ionization structure – around the source. As an illustrative example, let us consider a particular redshift, $z = 6.6$, similar to the highest redshifts at which currently there are Lyα source detections. At this time the universe in our simulation is already highly ionized, with a mean mass-weighted ionized fraction of $x_m = 0.939$, and a volume ionized fraction of $x_v = 0.925$. In Fig. 6 we show the spherically averaged profiles around the central source of the gas number density, n, neutral fraction, x_{HI}, integrated continuum optical depth (τ, at $h\nu = 13.6$ eV) and radial velocity, v_r. The radiative transfer cell which contains the central source is highly overdense, with $\delta = n/\bar{n} - 1 = 256$, which indicates that this cell roughly coincides with the source halo itself. The radial density profile declines steeply away from the source. The overdensity is $\delta = 3.3$ at $R \sim 1\,h^{-1}$ Mpc, decreasing to $\delta = 1$

at $R \sim 2.5\,h^{-1}$ Mpc, and approaching the mean density at distances beyond $R \sim 10\,h^{-1}$ Mpc.

The radial velocity profile shows an extended ($R \sim 20\,h^{-1}$ Mpc) infall region, with the mean radial velocity peaking at $\sim$150 km s^{-1} before dropping to zero inside the source halo itself. The proximity region of the central source ($R < 15\,h^{-1}$ Mpc) is highly ionized, with neutral fraction $x_{HI} < 10^{-4}$. The rest of the volume still has an appreciable neutral fraction ($\sim$0.1–1 per cent), however, and is thus on average still optically thick, with the mean optical depth reaching $\tau = 63$ at $R = 50\,h^{-1}$ Mpc.

However, the spherically averaged quantities provide only a limited information about the state of the IGM surrounding each source. All quantities are distributed highly anisotropically, and thus affect the Lyα emission differently along each LOS. In particular, the effect of the relative velocities of the IGM and source is relatively poorly studied at present. An optically thick medium at rest with respect to a Lyα source would absorb the blue wing of the line and transmit its red wing, at longer wavelengths than the line centre at $\lambda_0 = 1215$ Å (e.g. Miralda-Escude 1998). A relative motion of the IGM gas and the source along the LOS would result in either more or less transmission, depending on the motion direction. For example, gas infall towards the source along the LOS from the observer would redshift it, resulting in some absorption of the red wing of the line. We note that in order to evaluate these velocity effects with any precision much higher resolution simulations (and ones including outflows) will be required. Our radiative transfer grid has cell size of $\sim$0.5 h^{-1} Mpc comoving, or $\sim$70 h^{-1} kpc physical at $z = 6$, which roughly corresponds to the size of the largest haloes found in our box. The velocity and density fields used are at twice higher resolution ($\sim$0.25 h^{-1} Mpc comoving). Therefore, our results here should be considered as a guidance, illustrating that the velocity effects are quite important and should not be ignored.

In Fig. 7 we show the average IGM gas velocity relative to the source (line) versus distance from it and the variance of that average velocity (error bars). As noted above, in the vicinity of the source the IGM on average infalls towards the halo. However, there are large variations, of the order of hundreds of km s^{-1} around this mean. For example, a velocity offset of 200 km s^{-1} at $z = 6$ corresponds to $\Delta\lambda \sim 6$ Å, of the same order as the typical observed linewidths, and thus a relative motion of the IGM and source of this order could have
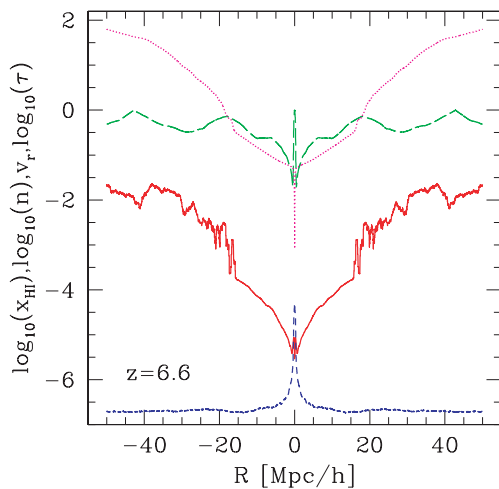

Figure 6. Spherically averaged profiles of (bottom to top on the left-hand/right-hand side): the comoving number density of hydrogen (in cm^{-3}; short-dashed, blue), the neutral hydrogen fraction (solid, red), radial velocity, v_r (in 100 km s^{-1}; long-dashed, green) and continuum optical depth at $h\nu = 13.6$ eV integrated from the source ($R = 0$) outward, all at redshift $z = 6.6$ ($x_m = 0.99$) versus R, the comoving radial distance from the most massive galaxy.
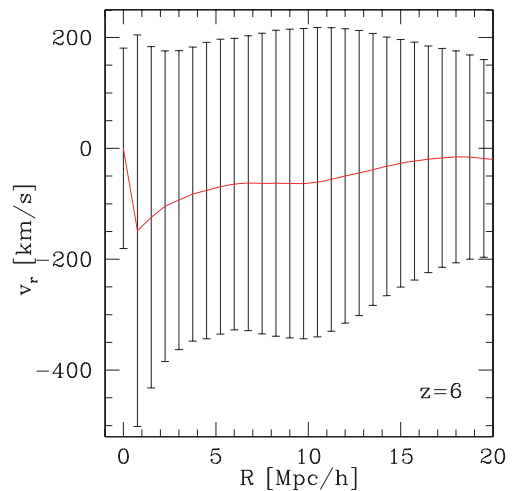

Figure 7. Mean radial velocity at $z = 6$ of the IGM with respect to the source (red solid line, negative means towards the halo, positive – away) and its rms variation (error bars), both plotted versus (comoving) radius from the source.

a very significant effect on the observed line. In the next section we would quantify the effect of peculiar velocities on the Lyα line.

4 OBSERVABILITY OF HIGH-z LYα SOURCES

4.1 Absorption spectra of luminous sources

Much of the information about high-redshift Lyα sources and IGM is based on absorption spectra. We thus start our discussion of the observable signatures by presenting and discussing some sample spectra intersecting the position of the most luminous source in our computational volume. Figs 8–11 we show sample spectra along three random LOSs at a few selected redshifts spanning the complete range of interest here. On the left-hand panels we show the distributions of Lyα Gunn–Peterson optical depth, τ_{GP}, neutral fraction, $x_{HI} = 1 - x$ (multiplied by 10^5 for clarity), and gas density in units of the mean, $\Delta = n/\bar{n}$. On the right-hand panels we show the corresponding Gunn–Peterson transmission spectra for flux level of unity, $\exp(-\tau_{GP})$. The horizontal lines on the left-hand panels

indicate the optical depth of $\tau = 4.6$, which corresponds to 1 per cent transmission. For reference, this value is roughly equal to the optical depth of a hydrogen gas with neutral fraction of $x_{HI} = 10^{-5}$ at the mean density at redshift $z = 6.6$. All quantities shown are in redshift/wavelength space and in the observer ($z = 0$) frame. For reference, on the top axis of each figure we show the approximate corresponding distances in real space. Finally, in Fig. 12 we show the mean, averaged over all random LOS, transmission spectra at the same redshifts, for the most massive source (left-hand panel) and mean over all sources (right-hand panel).

The Lyα absorption as a function of wavelength is computed using the standard procedure (e.g. Theuns et al. 1998). The optical depth and transmission results include the redshift-space distortions due to the local peculiar velocities, relative to the peculiar velocity of the source (i.e. after applying peculiar velocity distortions, the whole spectrum has been shifted slightly so the source is returned to its real-space position). The temperature of the gas is assumed to be 10^4 K when computing thermal broadening, consistent with the assumption adopted for the simulations.
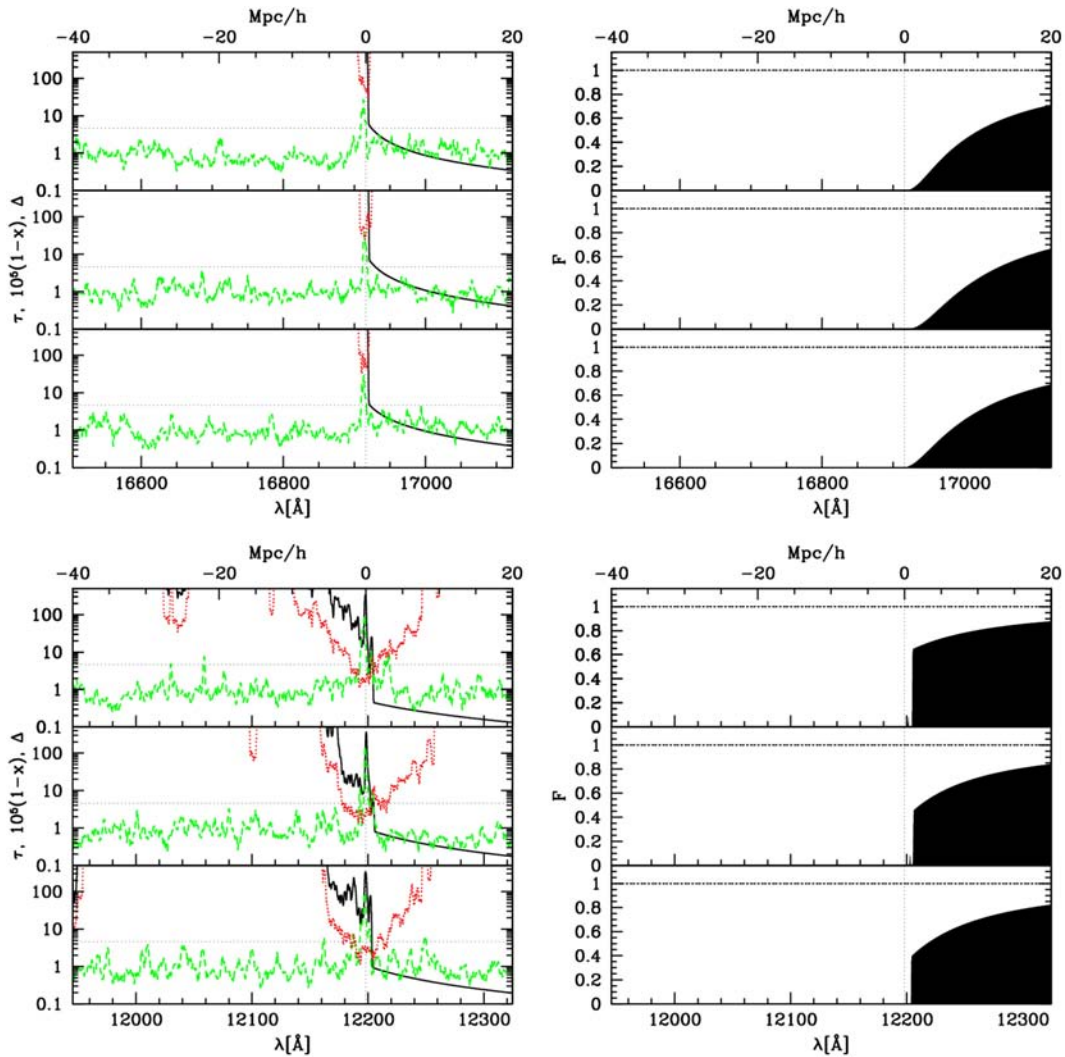

Figure 8. Sample LOS at redshifts $z = 12.9$ (top; $x_m = 0.001$) and 9.0 (bottom; $x_m = 0.28$) versus λ/comoving distance from the most massive galaxy. Shown are (left-hand panels) the optical depth (solid), neutral fraction $x_{HI} = 1 - x$ ($\times 10^5$; dotted) and density in units of the mean (dashed), and (right-hand panels) the corresponding transmission. The vertical lines show the position of the central source (in redshift space, i.e. accounting for its peculiar velocity along the LOS). The horizontal lines on the left-hand side indicate the optical depth equivalent to 1 per cent transmission. On the right-hand side, the shaded region is the transmission in the case where the unabsorbed spectrum is flat (the horizontal dotted line).
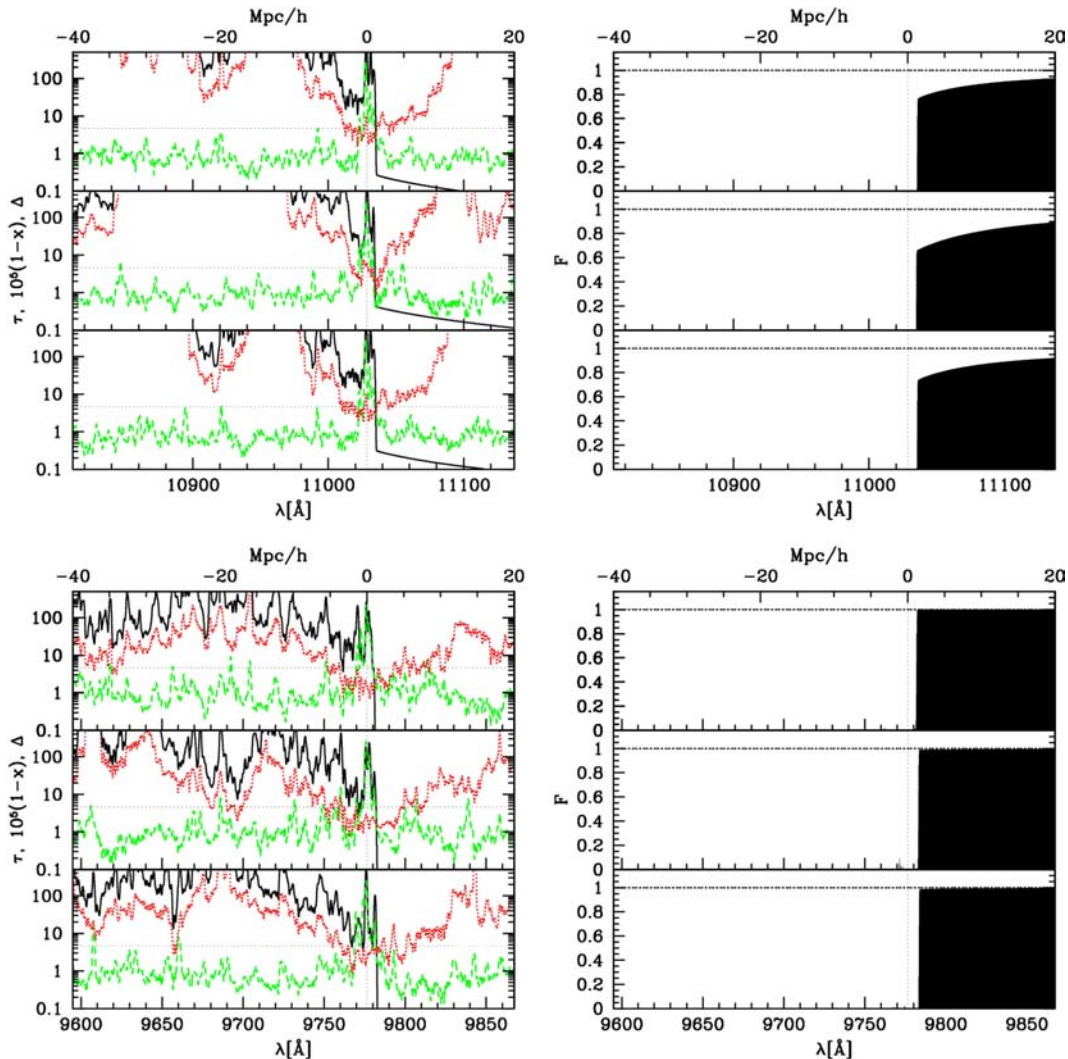
Figure 9. Same as Fig. 8, but at redshifts $z = 8.1$ (top; $x_m = 0.62$) and 7.0 (bottom; $x_m = 0.94$).

The nominal resolution of our spectra is $R \sim 6000$–$12\,000$ at $z = 6$ (higher at higher redshifts), based on our grid resolution of 203^3 (radiative transfer) and (406^3 density and velocity fields). This resolution roughly corresponds to the one for medium-resolution observed spectra. In reality the situation is more complicated. Due to the non-linear transformations between our raw simulation data and the final spectra the limited simulation resolution can affect the results even if it were better than the observational resolution. A separate issue pointing in the same direction is the fact that the real data have effectively infinite resolution in the transverse direction (i.e. the width of the light beam), and thus in order for us to be accurate we have to ensure we are resolving essentially all the transverse structure, which is not the case for the current simulations. As a result, our spectra should not be considered completely realistic predictions, but rather as a guidance showing some important features to be expected from real spectra, as discussed below. Future higher resolution, more detailed simulations will be better suited to make realistic predictions of the actual detailed spectral properties.

At early times ($z = 12.9$; Fig. 8, top) the H II region surrounding the most massive source is still quite small (see also Figs 2, 3 and 5) and the Lyα emission of the source is completely suppressed by the

damping wing of the Lyα line profile, rendering it unobservable. The ionized region grows quickly after that and by $z = 9$ reaches size of $\sim 6\,h^{-1}$ Mpc (Fig. 8, bottom). Regardless, the Lyα optical depth even within the source proximity region remains quite high, at few up to ~ 10, which allows through only a very weak transmission. The damping wing slightly weakens compared to the higher redshifts, but is still quite substantial, and still depresses most of the red wing of the emission line.

Some of the continuum immediately behind the source in redshift space (within a few Å) is absorbed due to gas infall towards the density peak. Because of that additional velocity towards the source the gas in front of it is slightly redshifted and thus absorbs at wavelengths redward of the source position. The line centre at the luminous source's position is completely dark due to the high density in the middle of the density peak, regardless of the very low neutral fraction there. These features persist throughout the evolution and are very characteristic for all luminous sources, since these always associated with high-density peaks and surrounded by infall.

At redshift $z = 8.1$ (Fig. 9, top) a weak damping wing is still present and essentially no transmission occurs on the blue side of the line. The sources at this time may be potentially visible with very deep observations. Only by redshift $z = 7$ (Fig. 12) the ionized
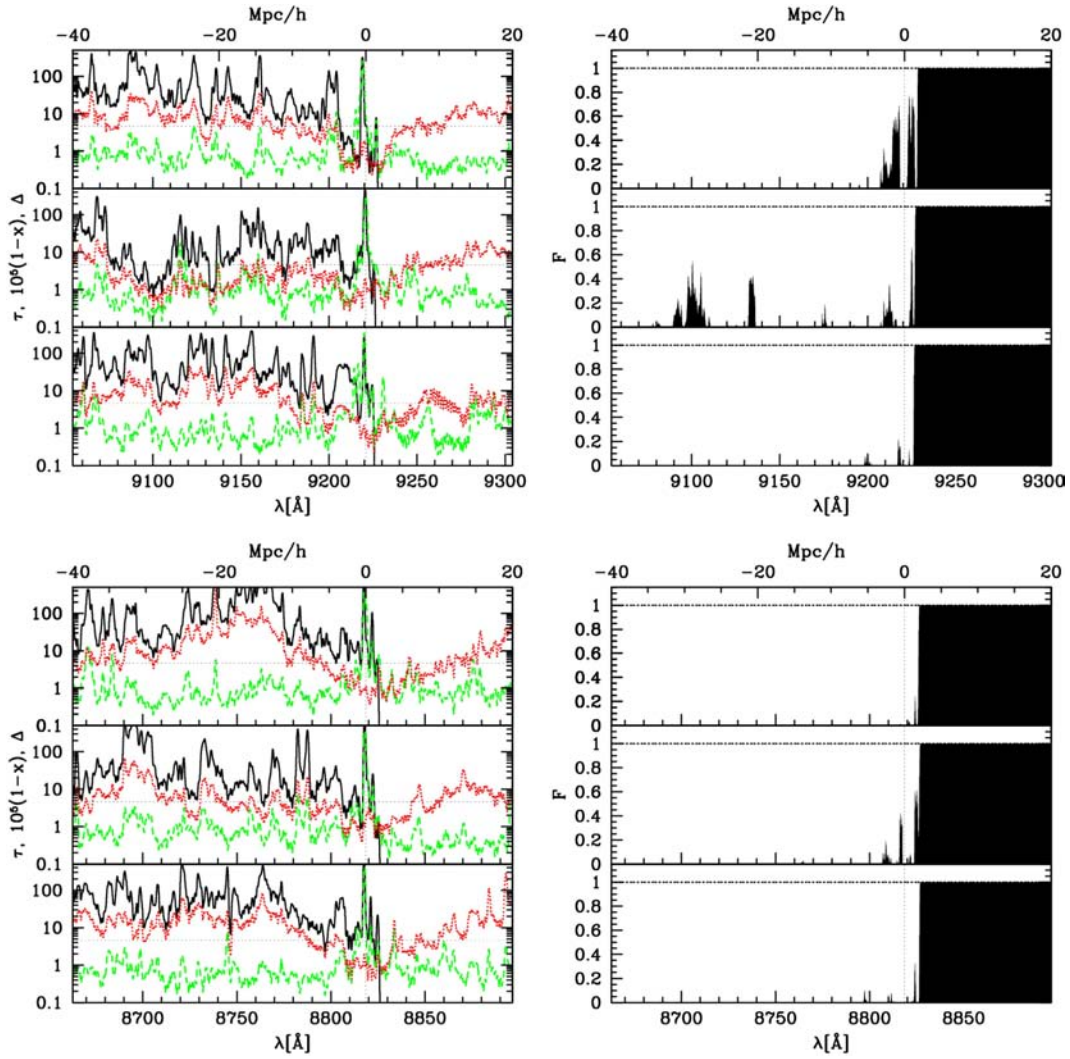
Figure 10. Same as Fig. 8, but at redshifts $z = 6.6$ (top; $x_m = 0.99$) and 6.25 (bottom; $x_m = 0.998$).
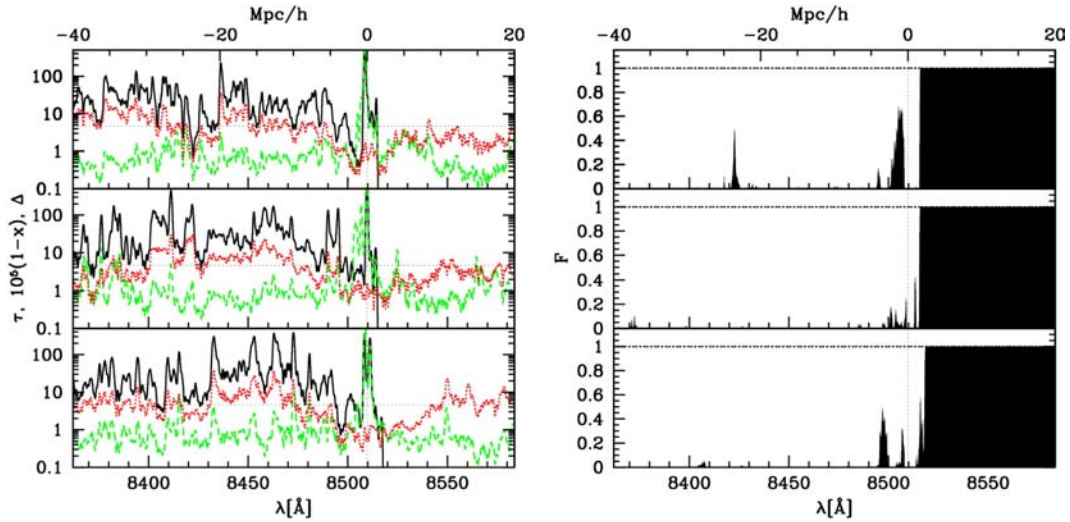

Figure 11. Same as Fig. 8, but at redshift $z = 6.0$ ($x_m = 0.9999$).
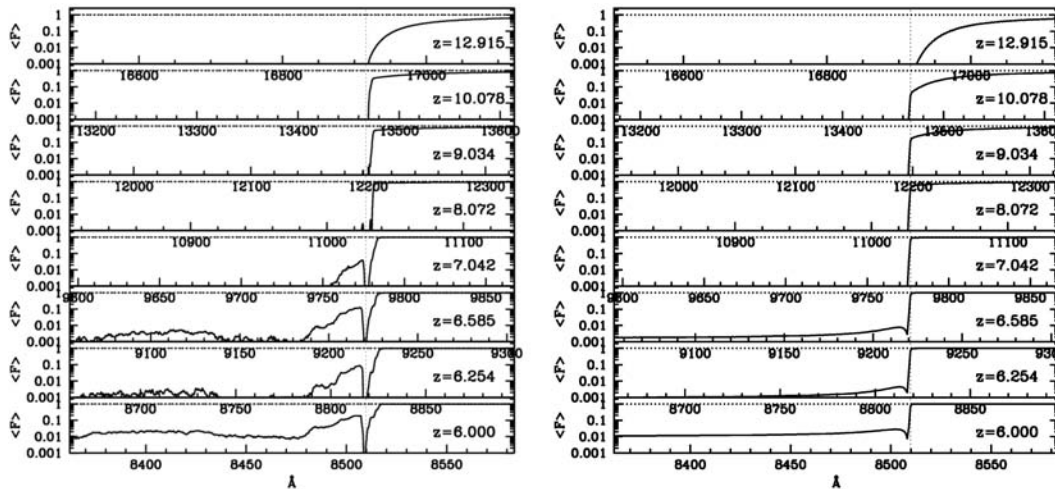
Figure 12. Mean radially averaged transmission around the most massive source (left-hand panel) and average over all sources (right-hand panel) at several representative redshifts, as labelled. The corresponding ionized fractions by mass are $x_m = 0.001, 0.105, 0.279, 0.618, 0.939, 0.992, 0.9986$ and 0.9999.

region is sufficiently large for the damping wing to effectively disappear. However, the gas in the ionized region still has a significant neutral fraction, and is sufficiently dense to absorb all photons on the blue side of the Lyα line along most LOS. On average a weak transmission at a few per cent level starts coming through in the proximity region of the source (Fig. 12). As more and more sources form and the ionizing flux rises the first transmission gaps start to appear at $z < 7$, both in the mean IGM away from the peak and in the source proximity (Fig. 10). At redshift $z = 6.6$, our nominal overlap time (Figs 10 and 12), the proximity region extends for ~30 Å and has become fairly optically thin, allowing up to 30–40 per cent transmission. Most of the volume still remains optically thick, but some substantial transmission regions appear in the IGM away from the peak. We also note that there are significant variations between the different LOS, with some allowing for much more transmission than others. The size and properties of the proximity region also vary due to its asymmetry and the anisotropy of nearby structures. Finally, during the post-overlap epoch [Figs 10 (bottom), 11 and 12] the IGM slowly becomes more optically thin to Lyα and gradually approaches the state of the Lyα forest. There are no more clearly defined H II bubbles. Only a few isolated low-density regions remain neutral. This is due to the inside-out character of the reionization process, whereby the high-density regions are preferentially ionized first, while the voids, where structure formation is delayed are ionized last.

Comparing the radially averaged mean transmission around a luminous source (high-density peak) and the average for all sources (i.e. the mean behaviour around a typical source; Fig. 12) we see both some similarities and several notable differences. The mean damping wing has similar evolution with redshift in the two cases, strong at high redshift ($z > 10$), gradually becoming weaker ($z \sim$ 8–9) and finally disappearing at later times ($z < 7$). Naively, one might expect that compared to an average source the luminous ones would suffer from weaker damping since such sources typically reside in the middle of large H II regions, far from their boundaries. In fact, this expectation proves only partially correct. The damping is somewhat more pronounced around an average source than around a luminous one, but the differences we observe are rather modest, reflecting the fact that weaker sources are strongly clustered around the same density peaks that host the luminous ones, thus largely share the damping (or its lack) with the central source. For the same

reason the damping becomes irrelevant at about the same time in both cases, which would not have been the case if the weak sources were residing in smaller, isolated bubbles.

The infall around the high-density peak leads to some redshifted absorption that appears behind the redshift-space position of the source. The same behaviour is not seen for a typical source. Such smaller haloes tend to move more in tandem with their surrounding IGM, often towards the nearest high-density peak. Some local infall should exist also for these haloes, but this is at very small scales, unresolved here. However, these scales are small compared to the typical emission linewidth (see next section) and thus we do not expect that such local infall has significant effect on the emission line.

One final important difference between a luminous and an average source is that the latter does not typically have a proximity transmission region on the blue side of the line. The spectra of the luminous sources, on the other hand exhibit extended high-transmission (10–60 per cent transmission) regions within 5 Mpc h^{-1} (~30 Å). This behaviour is again due to the high source clustering around the density peak. The combined effect of all sources is to achieve much lower neutral gas fraction regardless of the higher gas density there. The line centre coinciding with a high-density peak remains optically thick; however, unlike the line centre of a typical source. Away from the proximity region the absorption is largely saturated, but there are a number of transmission gaps with up to a few per cent transmission. Future work would quantify the statistics of these features and its evolution.

4.2 Emission line shape and its evolution

In order to study the effect of IGM absorption on the line profile shape we model the intrinsic Lyα line as a Gaussian with an rms width of $160 \, \text{km s}^{-1}$ and peak amplitude normalized to unity. In Figs 13–15 we show sample results for several LOS through the most luminous source at redshifts $z = 8, 6.6$ and 6, respectively. These examples are picked to illustrate the typical cases for the observed line shape. On the left-hand panels we show the assumed intrinsic (black) transmitted (red) emission line, while on the right-hand panels we show the corresponding distributions of Lyα Gunn–Peterson optical depth, τ_{GP}, neutral fraction, $x_{HI} = 1 - x$ (multiplied
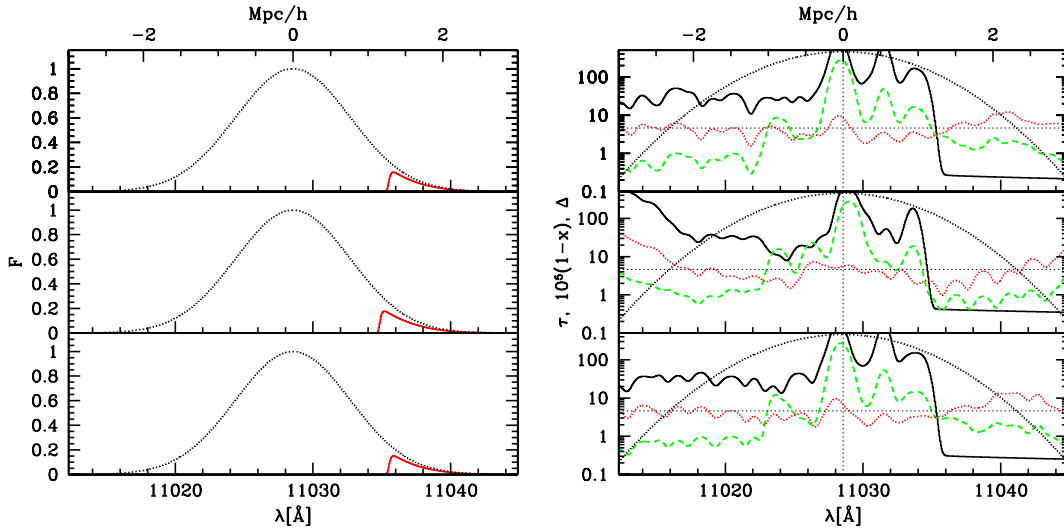
Figure 13. Emission lines: (left-hand panels) intrinsic line, assumed a Gaussian with rms width of $160\,km\,s^{-1}$ (black, top), and transmitted one (red, bottom) for three sample LOS at redshift $z = 8.1$ ($x_m = 0.62$), and (right-hand panels) the corresponding optical depth (solid, black), neutral fraction $x_{HI} = 1 - x$ ($\times 10^5$; dotted, red) and density in units of the mean (dashed, green), all in redshift space (i.e. accounting for the relative velocities). The intrinsic emission line is also shown for reference.
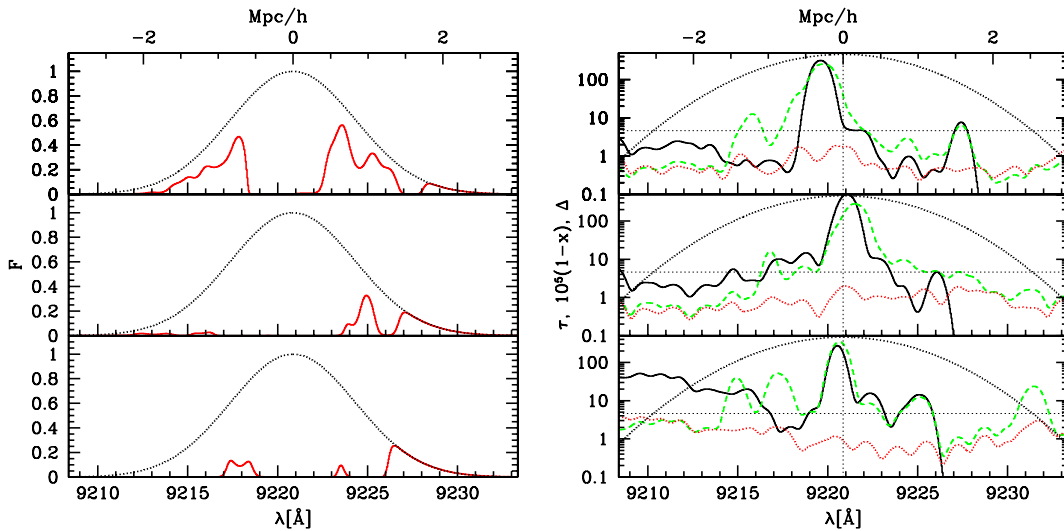
Figure 14. Same as Fig. 13, but at redshift $z = 6.6$ ($x_m = 0.99$).

by 10^5 for clarity), and gas density in units of the mean, $\Delta = n/\bar{n}$ and again (for reference) the intrinsic emission line.

At redshift $z = 8.1$ (Fig. 13) the emission line shape is fairly regular and does not vary strongly for different LOS. The blue wing is generally highly absorbed and no appreciable flux comes through and much of the red wing is absorbed as well. The reasons for this behaviour become clear from the right-hand panels. The neutral fraction is fairly uniform throughout this region, at $x_{HI} \sim 4 \times 10^{-5}$, thus the GP optical depth largely follows the local density field. The high density of the peak and its immediate vicinity results in optical depths of $\tau_{GP} > 10$ everywhere and $\tau > 100$ at the peak itself. The gas infall towards the peak leads to a significant absorption of the red wing. The only transmitted flux comes from the far red side of the line (slightly depressed by the weak remaining damping wing).

At overlap (1 per cent global neutral fraction by mass; $z = 6.6$) the line shape becomes much more irregular and varies significantly

between the different LOS (Fig. 14). Significantly larger fraction of the flux is transmitted, both on the red and on the blue side of the line. The neutral fraction in the vicinity of the source is still fairly uniform, but much lower than at the earlier time, at $x_{HI} \lesssim \times 10^{-5}$. Thus the GP optical depth again largely follows the density distribution. The effect of the gas infall towards the peak is still present and some of the red wing is absorbed, but much less so than at higher redshifts since the infalling gas is more highly ionized. The line centre remains absorbed and all damping wing effects have disappeared.

After overlap (Fig. 15) the neutral fraction gradually declines, decreasing the optical depth and allowing ever more flux to be transmitted. The neutral fraction is still fairly uniform and thus the optical depth mostly follows the density fluctuations. The line shapes remain rather irregular, with significant variations between the different LOS. There is transmission in both red and blue wing of the line.
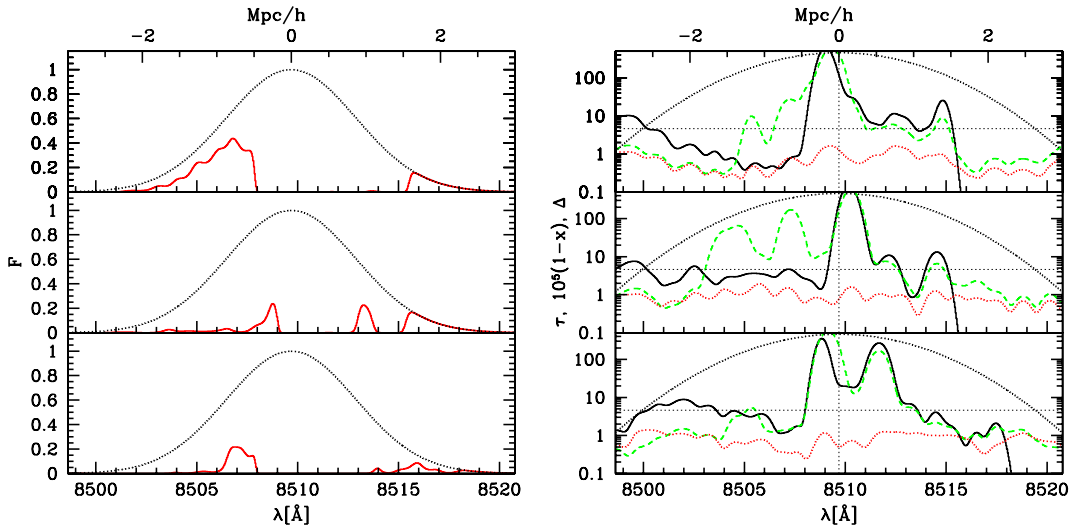
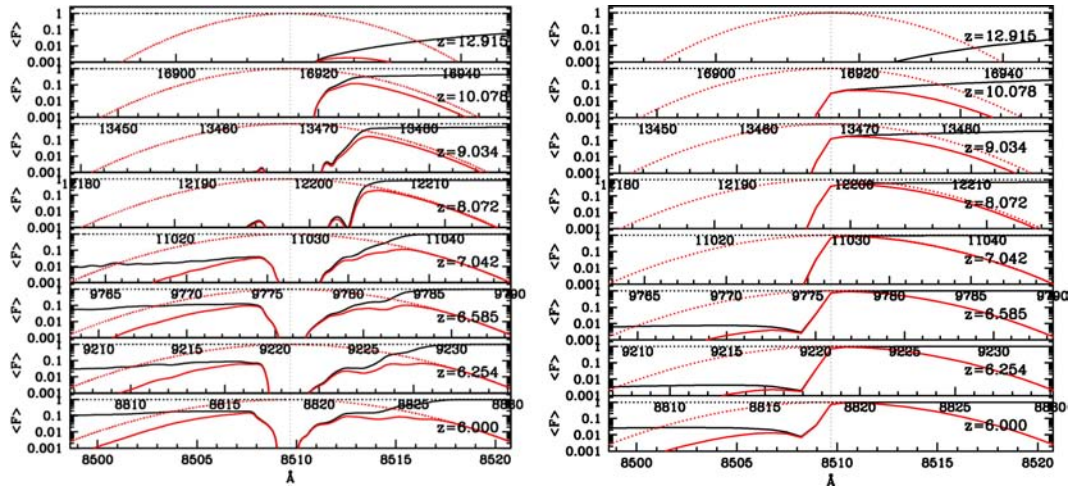**Figure 15.** Same as Fig. 13, but at redshift $z = 6$ ($x_m = 0.9999$).

Figure 16. Evolution of the mean emission lines for most massive source (left-hand panels) and average over all sources (right-hand panels). Shown are the intrinsic emission line (dotted, red; assumed a Gaussian with rms width of $160\,\mathrm{km\,s^{-1}}$, normalized to one at the peak), the transmitted line (solid, red), and mean absorption (solid, black) for three sample LOS several redshifts, as labelled.

In Fig. 16 we show the evolution of the mean (i.e. averaged over all LOS) observed emission line shape for the most luminous source in our volume (left-hand panel) and average over all sources (right-hand panel). In both cases the line starts completely damped ($z = 12.9$). The later evolution of the mean line shape differs significantly, however. By redshifts $z = 9$–10 a significant fraction of the red wing of the line is transmitted for the luminous source, except for the line centre, while much less of the red wing is transmitted for an average source. At later times ($z \sim 7$–8) this situation is reversed – practically all of the red wing of the line is transmitted for an average source, but much of the flux is still absorbed for the luminous source due to the high-density peak in the middle and its surrounding infall. As a word of caution we should note that some of this effect is in fact not physical but numerical, since our simulations do not resolve well the detailed structure around the smaller haloes. However, as we also mentioned above, these resolution effects should be modest considering that the emission line is fairly wide and thus reasonably well resolved and any corrections due to smaller scale structures will not affect much of the line.

On the blue side of the line there are further important differences between the luminous and average sources. The strong clustering of sources around the density peaks result in very high fluxes and thus a more pronounced highly ionized proximity region blueward of the line centre. Thus, significantly more of the blue wing of the luminous source line is transmitted, up to 10 per cent on average at $z = 6$, versus only ∼1–2 per cent for an average source.

An interesting consequence of the very high absorption observed at the line centre for massive sources and the redshift-space distortions due to the gas infall (the latter similar to the one studied theoretically in a more idealized setup by Barkana & Loeb 2004a) is that the Lyα line naturally takes a double-peaked (or even multiple-peaked in some cases) observed profile. This suggests that in principle it might be possible to use the line profiles of bright Lyα sources to study the infall surrounding their host haloes. In practice this might be difficult due to a number of complications. The line structure is quite different along different LOS, partly (as we pointed above in Section 2.4) due to the very anisotropic velocity structure surrounding the source, as well as its own peculiar motion.
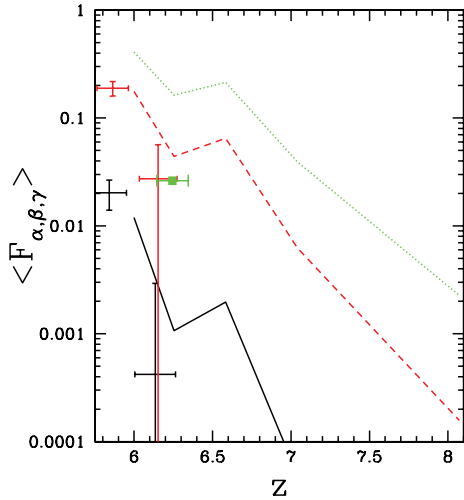
Figure 17. Overall mean transmission fraction (not necessarily near a source) in Lyα (solid, black), Lyβ (dashed, red) and Lyγ (dotted, green). The points with horizontal error bars show the measurements of (Fan et al. 2006b). The (black, red), (lower, upper), points with vertical error bars show Ly(α, β), while the green dot shows Lyγ (see text).

Figure 18. PDF of the photoionization rate at $z = 10.1$ (blue; $x_m = 0.105$), $z = 7.0$ (red; $x_m = 0.94$) and $z = 6.0$ (green; $x_m = 0.9999$). We show the actual, non-equilibrium rates (solid) and the corresponding equilibrium rates (dotted, same colour at each redshift). All PDFs are normalized to have an area of unity below the curve.

Furthermore, our analysis only takes into account the effects of the IGM on the line shape, while realistic ones will be affected also by the host galaxy's internal structure, outflows, etc. Modelling all those effects correctly will require high-resolution radiative-hydrodynamic simulations, which is well beyond the scope of this paper.

4.3 Evolution of the mean transmissivity

Fig. 17 shows the mean transmission fraction as a function of redshift, for the Lyα, β and γ transitions. The latter two are, respectively, 6.2 and 17.9 times weaker than Lyα, so transmission can be seen in cases where Lyα would be opaque (Fan et al. 2006b). At the low-redshift end of our simulation data these quantities have been observed in the highest redshift quasar spectra (Fan et al. 2006b). Fig. 17 shows the (Fan et al. 2006b) (points with error bars). At $z > 6$, the Lyα and Lyβ transmission measurements are only upper limits. It is unclear exactly what error bar should be assigned to the Lyγ point. Fan et al. (2006a) presented two upper limits and a detection in Lyγ, which, taken together, support a detected mean transmission level (plotted in the figure) well below our prediction, with relatively small (∼40 per cent) errors; however, with only these three points we cannot be sure that the sample variance is not substantially larger. It appears that the reionization model in our simulations reproduces roughly the correct tail end of reionization, but the data favour somewhat less transmission than is present in the model, i.e. a weaker radiation background and possibly slightly later reionization.

4.4 Photoionization rates

In Fig. 18 we show the normalized probability density functions (PDFs) of the non-equilibrium photoionization rates for all cells in our computational volume at three representative redshifts: $z = 10.1$ ($x_m = 0.105$), 7.0 ($x_m = 0.94$) and 6.0 ($x_m = 0.9999$) (early times, late times and well after overlap) in linear (top) and log (bottom) scales. For comparison we also plot the photoionization rates if the corresponding cells were in ionization equilibrium.
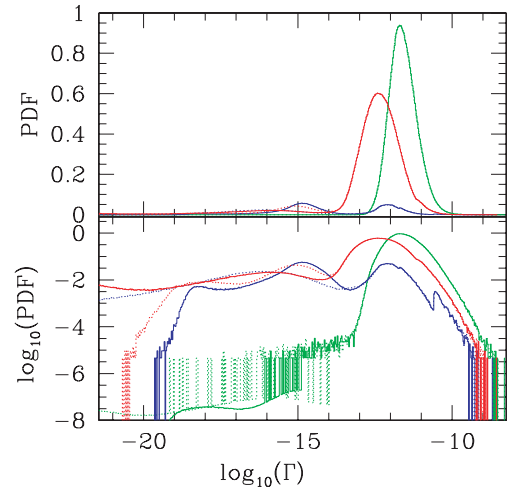
The right-hand peak of each PDF distribution reflects the most common photoionization rate values in the ionized regions. The peak position remains at $\Gamma_{-12} \sim 1$ throughout the evolution, with only slight shifts. The distributions are quasi-Gaussian, but with long non-Gaussian tails at both high and (especially) low values of Γ. As could have been expected, the fraction of high-Γ cells, all of which either contain sources or are in the immediate vicinity of a source, grows strongly with time, as ever. The highest photoionization rate values we find reach $\Gamma_{-12} \sim 10^3 - 10^4$. This peak value rises over time, as a consequence of the growth of galaxies and the large number of sources forming in and around the density peaks. In the ionized regions the equilibration time is short and photoionization equilibrium is generally a good approximation. The cells with lower values of Γ (below $\Gamma_{-12} \sim 0.01$–0.1 correspond to the ionization fronts or neutral regions. There are many more cells in I-fronts at $z = 10.1$ than at later times, when most of the IGM is already ionized. The equilibrium and actual rates differ widely in those regions, indicating that the assumption of ionization equilibrium would be a very poor approximation there.

The photoionization rate–density correlations at $z = 9$ and 6 are shown in Fig. 19 as contour plots. At high densities there is a clear, and fairly tight, correlation between the density and the photoionization rate. This reflects the fact that many more sources form in overdense regions, resulting in higher photoionization rates. At densities just above the mean the correlation becomes much less tight and around the mean densities it becomes nonexistent. At high-redshift regions with $1 + \delta \sim 1$ can have any value of Γ_{-12} from 10 (for cells close to sources) down to essentially 0 (in neutral and shielded cells). There are three broad peaks of the distribution, at $\Gamma_{-12} \sim 1$ (the H II regions), $\Gamma_{-12} \sim 10^{-3}$ (cells at and around I-fronts) and $\Gamma_{-12} \sim 0$ (neutral regions). By $z = 6.0$, which is well after overlap both the neutral and self-shielded regions and the I-fronts have mostly disappeared and $\Gamma_{-12} > 0.1$ almost everywhere, rising with time as more galaxies form. These values are fairly high compared to the photoionization rate values found from the Lyα forest at $z \sim 2$–4 (Cen & McDonald 2002; Tytler et al. 2004; Fan et al. 2006b; Bolton & Haehnelt 2007b), in agreement with the high mean transmitted flux we found in Section 4.3. Both point to
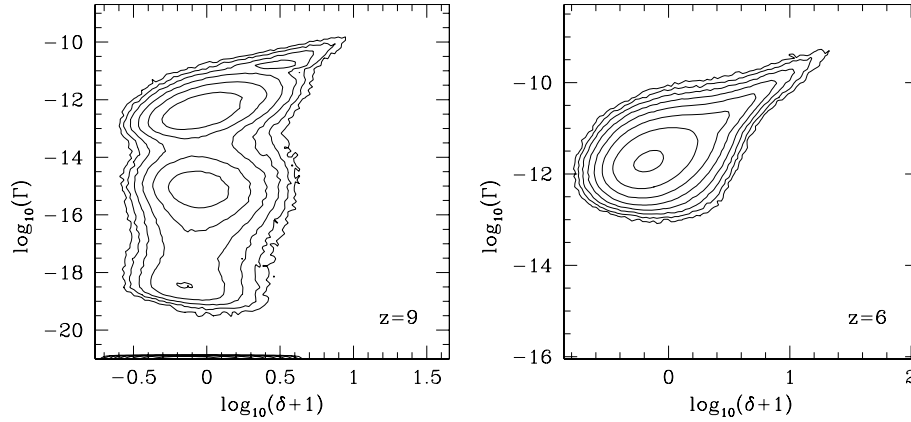
Figure 19. Photoionization rate–overdensity correlation at $z = 9$ (left-hand panel; $x_m = 0.28$) and $z = 6$ (right-hand panel; $x_m = 0.9999$). Contours are logarithmic, from 10 cells up every 0.5 dex.

somewhat lower ionizing source efficiencies than the ones assumed here and to a correspondingly later end of reionization.

4.5 Luminosity function of high-z Lyα sources

The luminosity function is an important statistical measure of the properties of high-redshift galaxies. It depends on both the intrinsic luminosity of the galaxies and the absorption in the surrounding IGM. In Fig. 20 we show our results of the size of the effect of absorption on a luminosity function of high z objects. For this fiducial case we assume that the Lyα luminosity is simply proportional to the mass of our haloes (similar to our model for the ionizing sources). We compute the reduction in luminosity of each halo due to absorption (Figs 8–11 show examples of this suppression). We assume that the intrinsic Lyα emission line is a Gaussian with an rms of $160 \, \mathrm{km \, s^{-1}}$. Luminosity function of high-redshift sources without (in this fiducial case this is just the halo mass function; solid line) and with absorption included (dashed) at redshifts $z = 9, 7$ and 6. For reference, the dotted line shows the result if each source is assumed 50 per cent absorbed, which would be the case if e.g. all of the blue wing of the emission line were absorbed, while all of the red wing were transmitted. Top panels show the bin-by-

bin ratios of the observed to the intrinsic luminosity function. Note that due to the binning at fixed luminosity (intrinsic or observed) this ratio is not the same as the average suppression per source of a given mass (i.e. the absorption shifts the curve both down and to the left-hand side).

The luminosity function exhibits clear evolution from high to low redshift. As we discussed in Section 4.1, at high redshift the damping wings are strong and thus not only the blue side, but also significant part of the red side of the line is absorbed, as evidenced by the significant difference between the dashed and dotted lines. As a result of this absorption, the number of sources per luminosity bin drops by one to two orders of magnitude. During the later stages of reionization the damping wing effectively disappears. As a consequence, the change in the faint end of the luminosity function due to IGM absorption is on average well represented by simply reducing each source luminosity by 50 per cent, which would be the case if the blue half of the line were absorbed and the red half were not. The situation is different at the bright end of the luminosity function; however, where on average significantly more than half of the intrinsic flux is absorbed at both $z = 7$ and 6. The shape of the luminosity function shows some evolution, as well, which is in part due to an evolution in the shape of the halo
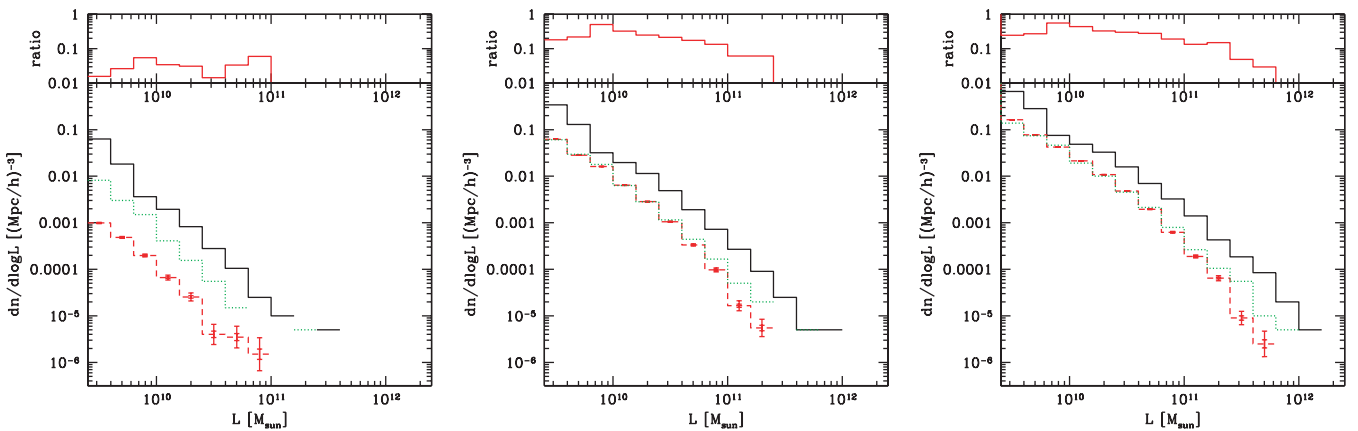

Figure 20. Bottom panels: Lyα luminosity function of high-redshift sources without (black) and with absorption included (red) at redshifts $z = 9$ (left-hand side; global mass-weighted ionized fraction $x_m = 0.28$), $z = 7$ (middle; $x_m = 0.94$) and $z = 6.0$ (right-hand side; $x_m = 0.9999$). For reference, the green, dotted line shows the result if each source is assumed 50 per cent absorbed, which would be the case if e.g. all of the blue wing of the emission line were absorbed, while all of the red wing were transmitted. The error bar in each bin reflects the number of sources in that bin found in our computational volume. Top panels: Bin-by-bin ratio of the observed to the intrinsic luminosity function.
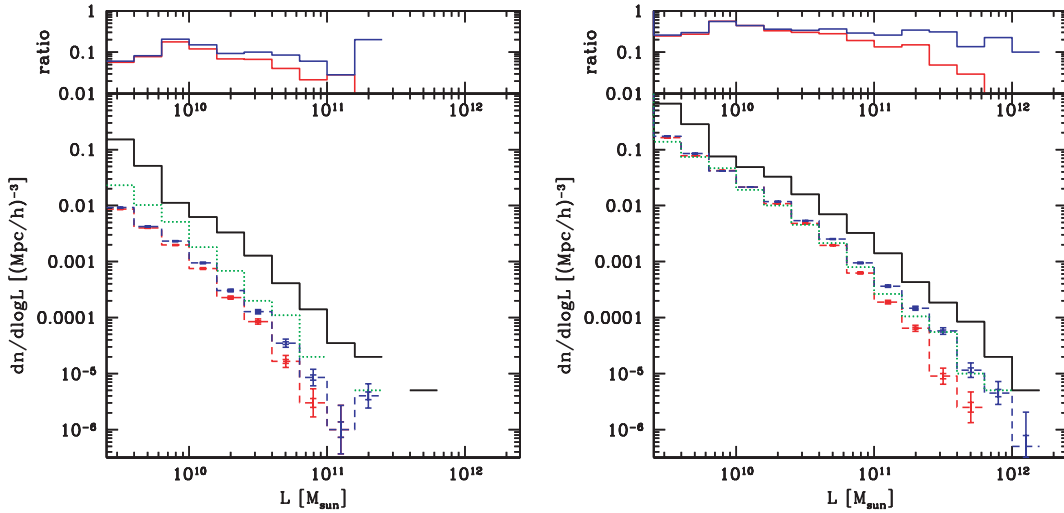
Figure 21. Luminosity function at redshifts $z = 8.1$ (left-hand panel; $x_{\rm m} = 0.62$) and $z = 6$ (right-hand panel; $x_{\rm m} = 0.9999$) if peculiar velocities are ignored (blue, long-dashed). For reference, we also show the data from Fig. 20, same notation.

mass function and in part to the higher mean absorption for the more massive sources. The higher average absorption levels for the luminous sources are consequence of the infall which surrounds the high-density peaks they are in. In order to demonstrate this, we recalculated the luminosity function at $z = 8.1$ and 6 using exactly the same data, but setting all peculiar velocities to zero. The results are shown in Fig. 21. With no peculiar velocities present the intrinsic emission of all sources is absorbed on average at roughly the same level, by a factor of $\sim$10 at $z = 8.1$ and by a factor of 2 at $z = 6$. The resulting luminosity function at late times agrees well with the one where we simply assumed 50 per cent absorption. Early on this is not the case as a consequence of the still-present damping wing.

How important is our assumption that all sources have same rms width of their intrinsic emission line profile? To check this, we replaced this assumption with one where the rms linewidth varies with the halo mass as $133\,{\rm km\,s^{-1}}\,(M/10^{11}\,{\rm M_\odot})^{1/3}$ (Dijkstra, Lidz & Wyithe 2007). Results, again at $z = 6$ and in head-to-head comparison with our fiducial case of fixed linewidth, are shown in

Fig. 22. The faint end of the luminosity function proves insensitive to the linewidth, which is easy to understand. As we have shown above, on average the IGM completely absorbs the blue half of the line for the weaker sources and completely transmits the red half. However, the variable linewidth has some effect on the absorption of bright sources. Their lines become wider under this assumption and thus are less affected by the absorption due to the infalling gas, resulting in higher transmission by a factor of $\sim$2.

The probability distribution of the transmission fraction per source of a given mass/luminosity is shown in Fig. 23. There are 10 random LOS per source and we required at least 50 LOS in each bin for sampling the distribution properly. We also plot the bin-by-bin average transmission. For the luminosity bins which do not contain our minimum number of LOS we plot only the mean. Several interesting trends emerge. The distributions are fairly wide at all times and for all sources, reflecting the large variations in opacity from source to source and from LOS to LOS. The former is due to the different environments sources are found in, while the later reflects the anisotropies around each source. At all redshifts the mean and
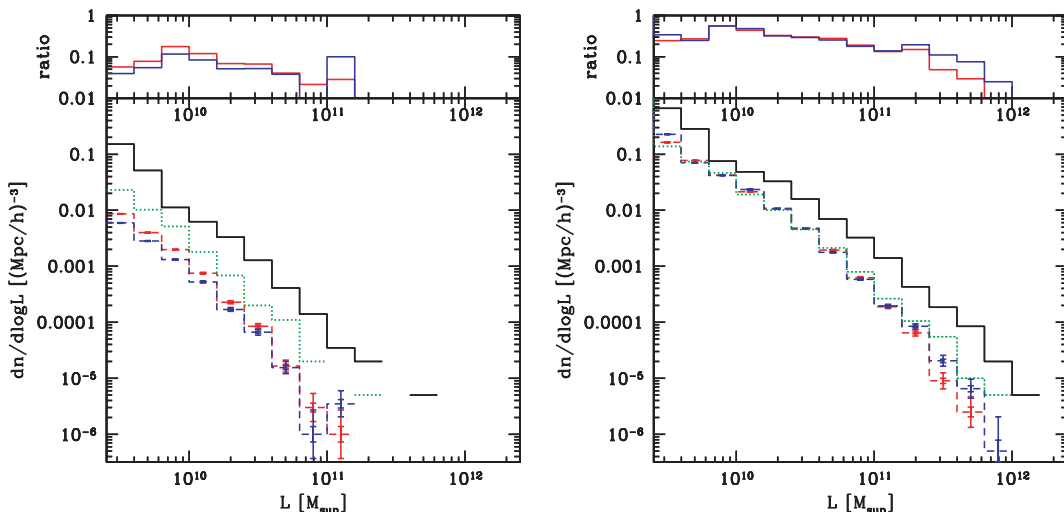

Figure 22. Luminosity function at redshifts $z = 8.1$ (left-hand panel; $x_{\rm m} = 0.62$) and $z = 6$ (right-hand panel; $x_{\rm m} = 0.9999$) for variable emission linewidth (blue, long-dashed; see text for details). For reference, we also show the data from Fig. 20, same notation.
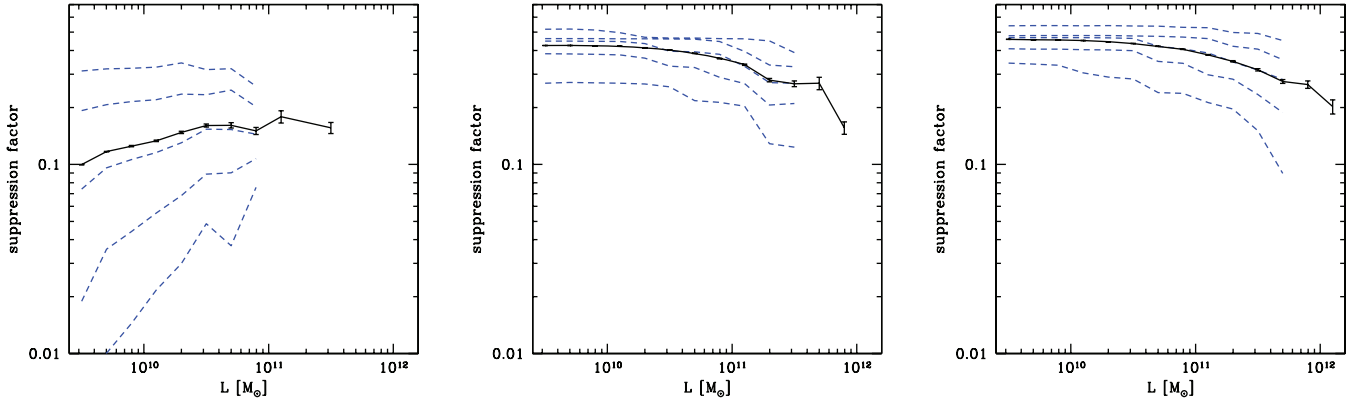
Figure 23. Transmission fraction as a function of luminosity at redshifts $z = 9.0$ ($x_{\mathrm{m}} = 0.28$; left-hand panel), $z = 7.0$ ($x_{\mathrm{m}} = 0.94$; centre) and $z = 6.0$ ($x_{\mathrm{m}} = 0.9999$; right-hand panel). Shown are (dashed lines, top to bottom) 0.023th, 0.16th, 0.5th, 0.84th and 0.977th percentiles (e.g. 2.3 per cent of points have suppression less than the uppermost line). The centre line is the median of all the LOS. We required at least 50 LOS in each bin for sampling the distribution properly. There are 10 random LOS per source. We also plot the mean (solid line) for all LOS in each bin.
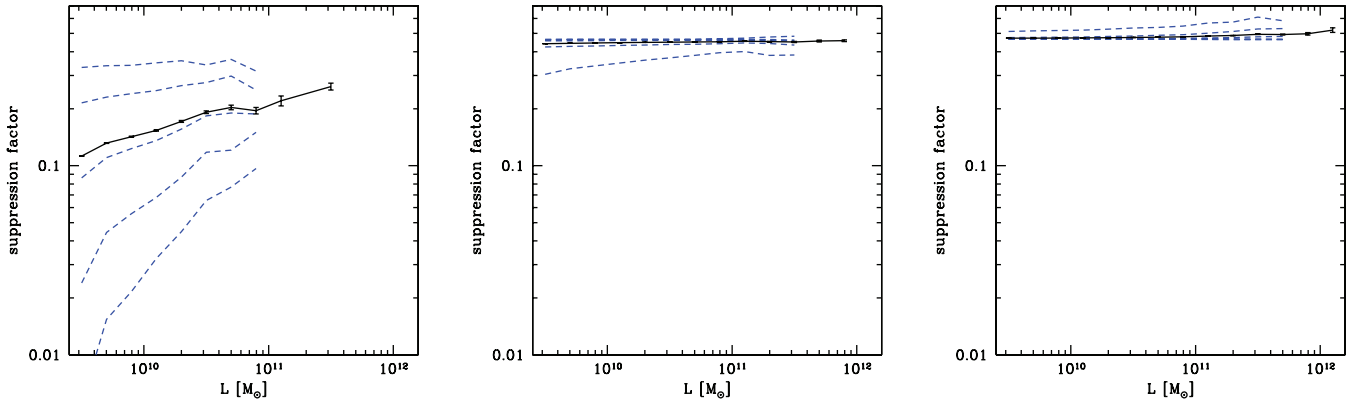

Figure 24. Same as Fig. 23, but ignoring any peculiar velocities of the haloes and the IGM.

the median curves are very similar for all bins, and are essentially identical at the bright end.

The distribution itself changes its character as the evolution progresses. At early times the distribution is much wider for the fainter sources and the mean and median are gently rising towards the bright end, reflecting the fact that bright sources are found in the middle of larger H II regions, while fainter sources are found in a variety of environments. Thus, during the early evolution the main factor shaping the distribution is the local variation of the neutral fraction around each source. At late times ($z < 7$), however, the situation changes to the opposite, with the distribution becoming wider at the bright end and the mean and median decreasing there as well. By that time the IGM is already largely ionized and the main environmental dependence is due to the anisotropies of the density field and, even more so, of the infall around the bright peaks, as discussed above. This is clearly demonstrated in Fig. 24, where we show the distributions as they would be if there were no peculiar velocities. At early times the results are largely unchanged, while at late times the variations between the different LOS essentially disappear and all the curves become flat, showing that the distribution is shaped mainly by the effects of the peculiar velocities.

Finally, in Fig. 25 we show the effect of varying intrinsic linewidth on the distributions. Compared to our fiducial case of constant linewidth, at high redshift the distributions are hardly affected, except for slightly higher absorption of the faintest sources. However,

at later times the varying linewidth has a more significant effect. At $z = 7$ the mean and the median values for the majority of sources decrease from ∼40–45 per cent down to ∼30 per cent, but the brightest sources are affected much less. As a result the curves for the mean and median become largely flat rather than decreasing towards the bright end. Furthermore, the distribution becomes wider at the faint end, with many more LOS being absorbed by a factor of 10 or more. A similar effect is seen at $z = 6$, except in this case the brightest sources are even less absorbed due to their wider emission lines, in agreement with what we observed in the luminosity functions.

Our derived luminosity functions assume a constant mass-to-light ratio and are in arbitrary units (halo mass × absorption by the IGM), proportional to a yet undetermined mass-to-observed light ratio. We can roughly determine the latter by comparing the number densities of observed and simulated objects. Kashikawa et al. (2006a) currently provide the best set of data at $z > 6$. They have provided fits to a Schechter function:

$$\phi(L)\mathrm{d}L = \phi_* \left(\frac{L}{L_*}\right)^\alpha \exp\left(-\frac{L}{L_*}\right) \frac{\mathrm{d}L}{L_*}. \tag{2}$$

Since the high-redshift data still have large uncertainties, particularly in terms of the faint-end slope, the data are fitted by assuming $\alpha = (-2, -1.5, -1)$, with best-fitting parameters at $z = 6.56$ given by $\log(L_*/\mathrm{h}_{70}^{-2}\,\mathrm{erg\,s^{-1}}) = (42.74, 42.60, 42.48)$ and $\log(\phi_*/\mathrm{Mpc}^{-3}\,h_{70}^2) = (-3.14, -2.88, -2.74)$, respectively. We plot these
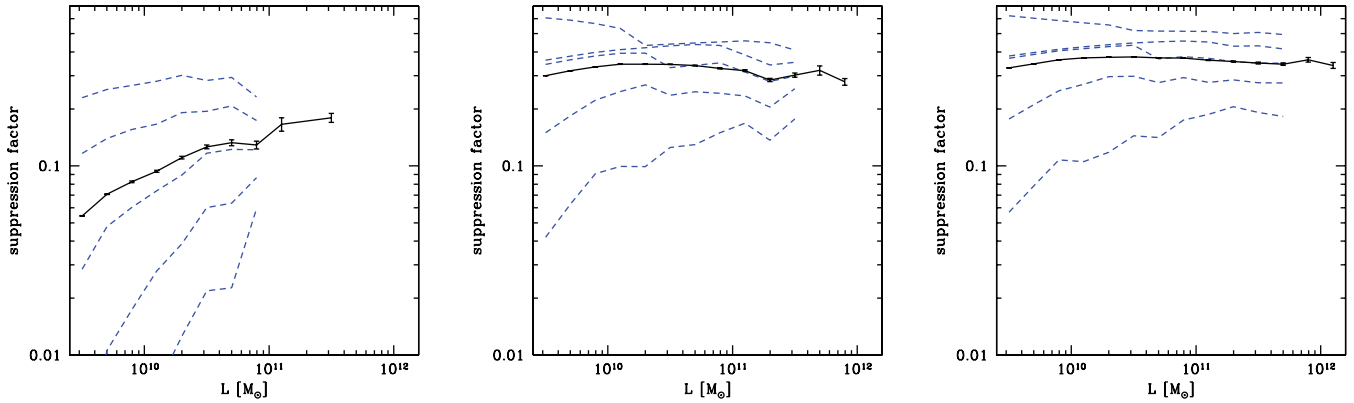
Figure 25. Same as Fig. 23, but for a variable Lyα emission linewidth, as discussed in the text.
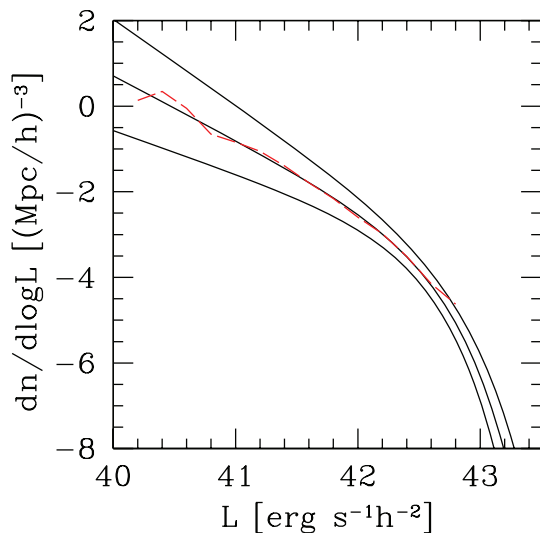

Figure 26. Simulated luminosity function at $z = 6.6$ (dashed; $x_m = 0.99$) versus best fits of Kashikawa et al. (2006a) (at $z = 6.56$) for (top to bottom) faint-end slopes of $\alpha = -2, -1.5, -1$ (solid).

fits in Fig. 26 against our derived luminosity function. The latter was obtained by rescaling our arbitrary luminosity units to physical ones using a constant ratio,

$$L = L(\mathrm{M}_\odot) \times 10^{30.9}, \tag{3}$$

so as to match it to the observed luminosity function for the same number densities of objects.

The fit assuming $\alpha = -1.5$ provides by far the best match to our luminosity function. The two agree in both amplitude and shape over the whole available range. The differences at both the luminous and the faint end should both be expected, the former due to cosmic variance and the latter due to both numerical resolution and lack of reliable observational data. Matching the other two faint end slopes would require us to relax our constant mass-to-light ratio assumption.

Taking equation (3) at face value, we can now make an approximate correspondence between observed luminosities and masses of the underlying haloes. The sources observed with Subaru at $z = 6.56$ Kashikawa et al. (2006a) have luminosities $\sim 10^{42}$–$10^{42.7}$ erg s^{-1} h^{-2}. This corresponds to the luminous end of our luminosity function, for effective masses (halo mass × absorption by the IGM) of the order of $10^{11}\,\mathrm{M}_\odot$ or larger, or relatively rare haloes. The ob-

servations are not yet sufficiently sensitive to detect the faint end, which contributes most of the ionizing emissivity during reionization. Hence, claims that observations show that there are not enough ionizing photons at $z \sim 6$ to reionize the universe appear premature.

5 CORRELATION FUNCTIONS

As we discussed above, the high-redshift haloes are highly clustered in space, and Lyα sources should be clustered as well. The latter clustering has been recently observed at $z \sim 5.7$ (Murayama et al. 2007). An interesting question to ask is if the absorption due to the surrounding IGM affects this clustering. If this were the case, then measuring the correlation function of high-redshift Lyα sources can give us information about the state of the IGM at that time.

We derive the correlation functions as follows. First, we calculate the total luminosity of each source with and without IGM absorption using the same method as above, but instead of random LOS directions we only consider parallel LOS, as would be seen by far away observer. For simplicity we consider LOS parallel to the axes of our computational box. We mock a flux-limited survey by imposing a cut-off on the observed luminosity. We compare the resulting correlation function to the one obtained for the same number, but now of the brightest sources based on their intrinsic luminosity (i.e. the ones hosted by the most massive haloes, thus ignoring IGM absorption in this latter case). We calculate the 3D correlation functions by direct summation over all pairs of haloes as described in Martel (1991). The results at redshift $z = 9$ [with cut-off $L_{\min}(M) = 10^{9.5}\,\mathrm{M}_\odot$] and $z = 7$ [with cut-off $L_{\min}(M) = 10^{10}\,\mathrm{M}_\odot$] are shown in Figs 27 and 28. In both cases the luminosity cut-offs were chosen so as to maximize the difference in the correlation functions with and without IGM absorption, while at the same time allowing for sufficient number of haloes above the cut-off to reduce the noise of the correlation.

In Fig. 27 we show the projection at $z = 9$ of the two source distributions on to the YZ plane (left-hand panel), and the corresponding 3D correlation functions and their ratio (right-hand panel). The projection shows that the two source populations differ, but cluster in the same spatial regions. This visual impression is confirmed by the correlation functions. We find that IGM absorption introduces only small variations in the correlation of sources. The difference is largest at small scales, for separations below 1 comoving Mpc. Even there they never exceed 10 per cent. At intermediate scales, $R \sim 2$–10 Mpc the departures are up to 5 per cent. There are no appreciable differences at large scales.
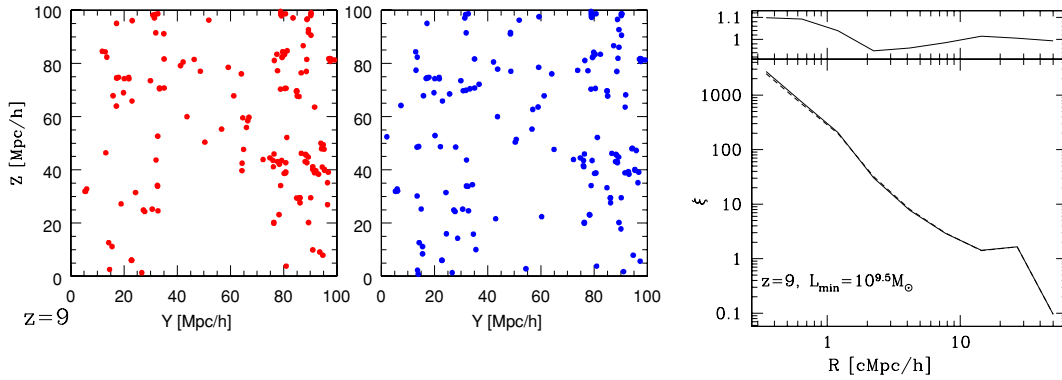
Figure 27. Projection of the sources as seen by a mock flux-limited survey with $L > 10^{9.5}\,M_\odot$ (198 sources in total) at $z = 9$ (left-hand panel; $x_m = 0.28$) and sources with the same number density if the IGM absorption were ignored (middle panel) and the two-point 3D correlation functions (right-hand panels) of the distribution with IGM absorption (solid) and without (dashed) and their ratio (top).

Figure 28. Projection of the sources as seen by a mock flux-limited survey with $L > 10^{10}\,M_\odot$ (1617 sources in total) at $z = 7$ (left-hand panel; $x_m = 0.94$) and sources with the same number density if the IGM absorption were ignored (middle panel) and the two-point 3D correlation functions (right-hand panels) of the distribution with IGM absorption (solid) and without (dashed) and their ratio (top).
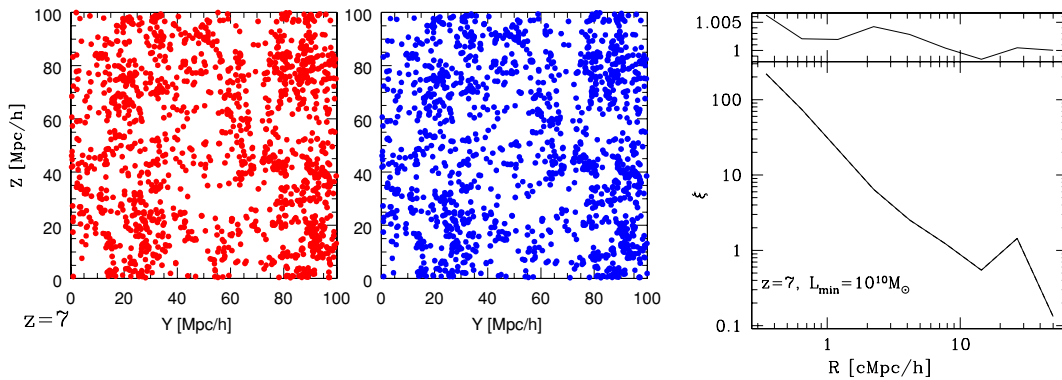
At redshift $z = 7$, close to overlap (Fig. 28) there are many more sources, even with the low-luminosity cut-off raised to $10^{10}\,M_\odot$. The two source distributions remain different, but are still clustered in a very similar way, resulting in largely identical correlation functions, which never differ by more than 0.7 per cent.

The small effect of IGM absorption on the correlation function appears counterintuitive. The reason for this is that the sources at high redshift, especially the most massive/luminous ones are strongly clustered and the ionization field is closely correlated with the galaxy field. The luminous sources are typically found in the inner parts of the largest ionized bubbles, where they are typically unaffected by damping from the remaining (few) neutral patches. This is also supported by the profiles in Fig. 12 which show the damping wing effect being important over the same time interval for both luminous and average sources. The damping affects average sources more strongly, since these are more often found closer to a neutral patch than the more massive sources. However, this does not change the correlation function significantly. Basically, the reionization patchiness has little effect on the source clustering properties. The reason is that while the IGM absorption diminishes the flux from all sources, the same source clusters (although not necessarily the same individual sources) are seen as would be without IGM absorption.

At late times any clearly defined ionized regions have already disappeared and the effect of IGM absorption is to replace some sources above the luminosity cut-off with other ones essentially at random, due to small local variations of the residual neutral fraction and the gas velocities. Therefore, the reionization patchiness ultimately has little effect on the correlation function. That of course does not mean that the source population is not modified by the IGM – a flux-limited survey will see many fewer sources than if IGM absorption were not present, but the clustering properties of those sources are almost the same as with no absorption for the same number density of sources.

Recently, McQuinn et al. (2007) claimed that reionization patchiness has a significant effect on observed source clustering, in apparent contradiction to our results (see also Furlanetto et al. 2006b; Mesinger & Furlanetto 2008b). However, they compared the clustering properties of Lyα sources with and without (i.e. intrinsic) IGM absorption for *a fixed* luminosity cut-off, rather than at fixed number density of sources, as we did. When the IGM absorption is accounted for, the sources remaining above the imposed flux limit are many fewer than in the case without absorption and they are also much brighter on average, hosted by more massive haloes. This naturally results in higher bias of the observed sources compared to all sources with intrinsic luminosity above that cut-off. However, it is not necessarily related to reionization patchiness. For example, source bias will increase also if the small residual neutral fraction in the ionized IGM increases, boosting the IGM opacity and causing dimmer, less clustered sources to fall below the luminosity cut-off,

thus no neutral patches are required for this to happen. In fact, our simulation results show almost no neutral patches below $z \sim 6.6$ (at which time the global mean ionized fraction by mass is below 1 per cent). Therefore, while our conclusions disagree with the ones of McQuinn et al. (2007), at least some of the differences could be attributed to the different comparisons we make, but some of the variations are possibly real, since McQuinn et al. (2007) claimed that the trend is present even for fixed number densities, albeit with no quantitative details. We would like to stress, however, that we observe the same qualitative trend of enhanced clustering of Lyα sources due to patchiness during the early stages of reionization, but the quantitative level of the effect is different, being much weaker in our case.

There are a number of possible explanations. In particular, in our simulations the pre-absorption clustering between ionized regions and sources appears to be more important than in the simulations of McQuinn et al. (2007). There are also notable differences between our and their modelling of the Lyα sources and the IGM absorption. McQuinn et al. (2007) (as many other current studies do) assume complete resonance absorption of the blue side of the Lyα line, and only study the suppression due to the IGM damping wing and do not include velocity effects in their analysis. They also assume a particular duty cycle for their Lyα emitters (that only 25 per cent of haloes host emitters), which we do not do in this paper. McQuinn et al. (2007) also ran simulations with different minimum source halo mass cut-offs, either significantly higher ($M_{min} = 4 \times 10^{10} \, M_{\odot}$) or significantly lower ($M_{min} = 10^8 \, M_{\odot}$) than the one we have here ($M_{min} = 2.2 \times 10^9 \, M_{\odot}$). It is possible, therefore, that our results are more relevant than these previous works, if the low-mass sources missing in our current simulation are in fact strongly suppressed during the late stages of reionization due to Jeans-mass filtering. This can only be resolved conclusively by detailed simulations which actually follow the complicated radiative feedback effects on low-mass haloes, which is well beyond the scope of this paper. We have run several tests in order to try to understand these differences, following suggestions by the referee. One possibility we investigated was that our H II regions size distribution is more strongly peaked and narrower than the one found in the above works, which, if it were the case might have explained some of the clustering differences. We found, however, that the H II region size distributions derived from our simulations is in fair agreement with the one from the McQuinn et al. (2007) simulations, and hence this offers no plausible explanation of the differences. We also compared the Lyα damping wing optical depth distributions for sources of different mass at a range of reionization stages, as derived by Mesinger & Furlanetto (2008b) (their figs 2 and 3) and again found no significant differences between our results and theirs. We conclude that more detailed and direct comparisons will be required in order to evaluate and understand any differences between our results.

6 SUMMARY AND CONCLUSIONS

We considered the effects which the reionizing IGM has on the observations of high-redshift Lyα sources. For this we utilized detailed structure formation and radiative transfer simulations, which allowed us to evaluate many features which can only be studied by detailed simulations, as well as to quantify better a number of previously proposed effects. We followed the full reionization history self-consistently and accounted for the actual source distribution, neutral fraction, density and velocity fields.

We find that the density, neutral fraction and velocity fields are all highly anisotropic, which results in large variations in the IGM transmission and source visibility among different LOS. The velocity effects, both gas infall and source peculiar velocity are most important for massive, luminous sources. The most luminous sources are found in highest peaks of the density field, which at late times are significantly overdense out to $\sim$10 comoving Mpc (cMpc) and are surrounded by infall extending to $\sim$20 cMpc. The infall of gas blueshifts it in frequency space and results in significant absorption on the red side of the line centre, while the peculiar velocity of the source itself can either alleviate or exacerbate this effect, depending on the halo and infalling velocity alignment.

The spherically averaged local density enhancement and gas infall have been modelled analytically in approximate ways (Barkana 2004), and thus can be incorporated in semi-analytical models (e.g. Dijkstra et al. 2007). However, such models are unable to account for the strong intrinsic anisotropies of the neutral fraction, density and velocity fields. The analytical and semi-analytical models typically assume spherical symmetry and full ionization inside the H II regions, both of which assumptions are quite unrealistic.

The Lyα lines we derive are generally asymmetric and vary hugely from LOS to LOS. The luminous sources form at the highest density peaks and as a consequence their line centres are always highly absorbed even though their proximity regions are very highly ionized, with typical neutral fractions $x_{HI} \sim 10^{-5}$ to 10^{-6}. The luminous sources also more affected by infall and exhibit more pronounced proximity region with higher transmission of the blue wing of the line.

High-redshift sources are strongly clustered around the high peaks of the density field. The central source contributes the majority of the ionizing flux only in its immediate vicinity, within 1–2 cMpc. Beyond that distance the ionizing flux is dominated by the fainter sources clustered around it. This dominance is particularly strong at late times, when both many more sources form and the ionized regions become larger, resulting in the fainter sources contributing up to two orders of magnitude more photons than the central source.

Compared to single-source ionized bubbles, the larger H II regions from clustered sources diminish the effects from the damping wing of the line. Nevertheless, these remain significant until fairly late (ionized mass fraction $x_m = 0.3$–0.7, which for the simulation considered here corresponds to redshifts $z \sim 9$–8). Interestingly, the average damping wing effect is similar for luminous and typical sources, even though naively one might expect that damping could be weaker for the former, since they are typically in the middle of large bubbles, away from the neutral patches, unlike the fainter sources, which are more evenly distributed.

Both the mean IGM transmission and the typical photoionization rates we find are high compared to observations at $z \sim 6$, indicating that our adopted source efficiencies are also high. The mean IGM transmissivity decreases only slowly towards the higher redshifts and the GP transparency occurs significantly after the actual overlap epoch. For the simulation considered here overlap (defined as 1 per cent neutral fraction) occurs at $z = 6.6$, while average neutral fraction of 10^{-4} is reached only by $z = 6$ and even then relatively small fraction (few to 10 per cent) of the flux is transmitted. By overlap the spectra start showing significant transmission gaps in the mean IGM (i.e. away from the proximity region of a luminous source).

We find that for a given number density of sources (e.g. as determined by observations) the clustering of these sources depends only weakly on the IGM absorption during reionization. As a

consequence, the reionization patchiness has little effect on the observed Lyα source clustering, which implies that source clustering is not a good indicator for reionization patchiness.

Our derived luminosity function assuming constant mass-to-light ratio provides an excellent match to the shape of the observed luminosity function at $z = 6.6$ with faint-end slope of $\alpha = -1.5$. The resulting mass-to-light ratio implies that the majority of sources responsible for reionization are too faint to be observed by the current surveys.

REFERENCES

Barkana R., 2004, MNRAS, 347, 59
Barkana R., Loeb A., 2004a, ApJ, 601, 64
Barkana R., Loeb A., 2004b, ApJ, 609, 474
Bolton J. S., Haehnelt M. G., 2007a, MNRAS, 374, 493
Bolton J. S., Haehnelt M. G., 2007b, 382, 325
Bunker A., Stanway E., Ellis R., McMahon R., Eyles L., Lacy M., 2006, New Astron. Rev., 50, 94
Cen R., Haiman Z., 2000, ApJ, 542, L75
Cen R., McDonald P., 2002, ApJ, 570, 457
Dijkstra M., Lidz A., Wyithe J. S. B., 2007, MNRAS, 377, 1175
Doré O., Holder G., Alvarez M. A., Iliev I. T., Mellema G., Pen U.-L., Shapiro P. R., 2007, Phys. Rev. D, 76, 043002
Fan X. et al., 2002, AJ, 123, 1247
Fan X. et al., 2004, AJ, 128, 515
Fan X., Carilli C. L., Keating B., 2006a, ARA&A, 44, 415
Fan X. et al., 2006b, AJ, 132, 117
Furlanetto S. R., Hernquist L., Zaldarriaga M., 2004a, MNRAS, 354, 695
Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004b, ApJ, 613, 1
Furlanetto S. R., Oh S. P., Briggs F. H., 2006a, Phys. Rep., 433, 181
Furlanetto S. R., Zaldarriaga M., Hernquist L., 2006b, MNRAS, 365, 1012
Gallerani S., Ferrara A., Fan X., Choudhury T. R., 2008, MNRAS, 386, 359
Gnedin N. Y., Jaffe A. H., 2001, ApJ, 551, 3
Gruzinov A., Hu W., 1998, ApJ, 508, 435
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Haiman Z., Cen R., 2005, ApJ, 623, 627
Haiman Z., Holder G. P., 2003, ApJ, 595, 1
Holder G. P., Iliev I. T., Mellema G., 2007, ApJ, 663, L1
Hu E. M., Cowie L. L., McMahon R. G., Capak P., Iwamuro F., Kneib J.-P., Maihara T., Motohara K., 2002, ApJ, 568, L75
Hu E. M., Cowie L. L., Capak P., Kakazu Y., 2005, in Williams P., Shu C.-G., Menard B., eds, IAU Colloq. 199, Probing Galaxies through Quasar Absorption Lines. Cambridge Univ. Press, Cambridge, p. 363
Hu W., 2000, ApJ, 529, 12
Iliev I. T., Shapiro P. R., Ferrara A., Martel H., 2002, ApJ, 572, L123
Iliev I. T., Scannapieco E., Martel H., Shapiro P. R., 2003, MNRAS, 341, 81
Iliev I. T., Scannapieco E., Shapiro P. R., 2005, ApJ, 624, 491
Iliev I. T., Mellema G., Pen U.-L., Merz H., Shapiro P. R., Alvarez M. A., 2006a, MNRAS, 369, 1625
Iliev I. T. et al., 2006b, MNRAS, 371, 1057
Iliev I. T., Mellema G., Shapiro P. R., Pen U.-L., 2007a, MNRAS, 376, 534
Iliev I. T., Pen U.-L., Bond J. R., Mellema G., Shapiro P. R., 2007b, ApJ, 660, 933
Iliev I. T., Mellema G., Pen U.-L., Bond J. R., Shapiro P. R., 2008a, MNRAS, 384, 863
Iliev I. T., Shapiro P. R., Mellema G., Merz H., Pen U.-L., 2008b, in refereed proceedings of TeraGrid08 (arXiv:0806.2887)
Kashikawa N. et al., 2006a, ApJ, 637, 631
Kashikawa N. et al., 2006b, ApJ, 648, 7
Kodaira K. et al., 2003, PASJ, 55, L17
Kramer R. H., Haiman Z., Oh S. P., 2006, ApJ, 649, 570
Madau P., Meiksin A., Rees M. J., 1997, ApJ, 475, 429
Malhotra S., Rhoads J. E., 2004, ApJ, 617, L5
Martel H., 1991, ApJ, 366, 353
Maselli A., Gallerani S., Ferrara A., Choudhury T. R., 2007, MNRAS, 376, L34
McQuinn M., Furlanetto S. R., Hernquist L., Zahn O., Zaldarriaga M., 2005, ApJ, 630, 643
McQuinn M., Hernquist L., Zaldarriaga M., Dutta S., 2007, MNRAS, 381, 75
Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006a, New Astron., 11, 374
Mellema G., Iliev I. T., Pen U.-L., Shapiro P. R., 2006b, MNRAS, 372, 679
Mellema G., Iliev I. T., Pen U.-L., Shapiro P. R., 2006c, MNRAS, 372, 679
Merz H., Pen U.-L., Trac H., 2005, New Astron., 10, 393
Mesinger A., Furlanetto S. R., 2008a, MNRAS, 385, 1348
Mesinger A., Furlanetto S. R., 2008b, MNRAS, 386, 1990
Mesinger A., Haiman Z., 2004, ApJ, 611, L69
Mesinger A., Haiman Z., 2007, ApJ, 660, 923
Miralda-Escude J., 1998, ApJ, 501, 15
Morales M. F., Bowman J. D., Hewitt J. N., 2006, ApJ, 648, 767
Mortonson M. J., Hu W., 2007, ApJ, 657, 1
Murayama T. et al., 2007, ApJS, 172, 523
Ota K. et al., 2008, ApJ, 677, 12
Rhoads J. E. et al., 2003, AJ, 125, 1006
Santos M. G., Cooray A., Haiman Z., Knox L., Ma C.-P., 2003, ApJ, 598, 756
Santos M. R., 2004, MNRAS, 349, 1137
Shapiro P. R., Iliev I. T., Martel H., Ahn K., Alvarez M. A., 2004, (astro-ph/0409173)
Shapiro P. R., Ahn K., Alvarez M. A., Iliev I. T., Martel H., Ryu D., 2006, ApJ, 646, 681
Shapiro P. R., Iliev I. T., Mellema G., Pen U.-L., Merz H., 2008, in Minchin R., Momjian A., eds, AIP Conf. Proc., The Evolution of Galaxies through the Neutral Hydrogen Window. Am. Inst. Phys., New York, p. 68
Shimasaku K. et al., 2006, PASJ, 58, 313
Shklovskii I. S., 1964, Astron. Zh., 41, 801
Spergel D. N. et al., 2007, ApJS, 170, 377
Stanway E. R. et al., 2004, ApJ, 604, L13
Taniguchi Y. et al., 2005, PASJ, 57, 165
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Totani T., Kawai N., Kosugi G., Aoki K., Yamada T., Iye M., Ohta K., Hattori T., 2006, PASJ, 58, 485
Tozzi P., Madau P., Meiksin A., Rees M. J., 2000, ApJ, 528, 597
Tytler D. et al., 2004, ApJ, 617, 1
Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52
White R. L., Becker R. H., Fan X., Strauss M. A., 2003, AJ, 126, 1
Wyithe J. S. B., Loeb A., 2004, ApJ, 610, 117
Wyithe J. S. B., Loeb A., 2005, ApJ, 625, 1
Wyithe J. S. B., Loeb A., 2007, MNRAS, 374, 960
Wyithe J. S. B., Loeb A., Carilli C., 2005, ApJ, 628, 575
Zaldarriaga M., Furlanetto S. R., Hernquist L., 2004, ApJ, 608, 622

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