

# A tale of seven narrow spikes and a long trough: constraining the timing of the percolation of H II bubbles at the tail end of reionization with ULAS J1120+0641

Jonathan Chardin,<sup>1,2★</sup> Martin G. Haehnelt,<sup>1</sup> Sarah E. I. Bosman<sup>1</sup> and Ewald Puchwein<sup>1</sup>

<sup>1</sup>Kavli Institute for Cosmology and Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

<sup>2</sup>Université de Strasbourg, CNRS, Observatoire Astronomique de Strasbourg, UMR 7550, F-67000 Strasbourg, France

Accepted 2017 September 8. Received 2017 September 8; in original form 2017 July 13

## ABSTRACT

High signal-to-noise observations of the Ly $\alpha$  forest transmissivity in the  $z = 7.085$  quasi-stellar object (QSO) ULAS J1120+0641 show seven narrow transmission spikes followed by a long  $240 \text{ cMpc } h^{-1}$  trough. Here, we use radiative transfer simulations of cosmic reionization previously calibrated to match a wider range of Ly $\alpha$  forest data to show that the occurrence of seven transmission spikes in the narrow redshift range  $z = 5.85\text{--}6.1$  is very sensitive to the exact timing of reionization. Occurrence of the spikes requires the most underdense regions of the intergalactic medium to be already fully ionized. The rapid onset of a long trough at  $z = 6.12$  requires a strong decrease of the photoionization rate  $\Gamma$  at  $z \gtrsim 6.1$  in this line of sight, consistent with the end of percolation at this redshift. The narrow range of reionization histories that we previously found to be consistent with a wider range of Ly $\alpha$  forest data have a reasonable probability of showing seven spikes and the mock absorption spectra provide an excellent match to the spikes and the trough in the observed spectrum of ULAS J1120+0641. Larger samples of high signal-to-noise searches for rare Ly $\alpha$  transmission spikes at  $z > 5.8$  should therefore provide important further insights into the exact timing of the percolation of H II bubbles at the tail end of reionization.

**Key words:** methods: numerical – Galaxy: evolution – quasars: general – diffuse radiation – cosmology: theory.

## 1 INTRODUCTION

The Lyman alpha (Ly $\alpha$ ) forest is the main probe of the physical state of the intergalactic medium (IGM) after hydrogen reionization (see Becker, Bolton & Lidz 2015a; McQuinn 2016 for recent reviews). The rapid increase of the Ly $\alpha$  optical depth at  $z > 5$  suggests that at  $z \sim 6$  we are seeing the last stages of the hydrogen reionization process (Fan et al. 2006; Becker et al. 2013, 2015b). The rapid decrease of the Ly $\alpha$  emitters space density at  $z > 7$  (Bolton & Haehnelt 2013; Dijkstra et al. 2014) and the low value of the Thomson optical depth suggested by recent Planck data (Planck Collaboration XIII 2016; Planck Collaboration XLVII 2016) would indicate a somewhat smaller duration and a later start of the reionization process than suggested by the *Wilkinson Microwave Anisotropy Probe* results.

Currently Ly $\alpha$  forest studies allow us to probe the redshift evolution of the ionization state of the IGM up to redshift  $z \sim 6$ . However, because the scattering cross-section of the Ly $\alpha$  transition is very high, it becomes rather challenging to use the forest to put constraints on the mean neutral hydrogen fraction  $x_{\text{HI}}$  at  $z \gtrsim 5.7$ . When the neutral hydrogen fraction at mean density exceeds values of  $x_{\text{HI}} \gtrsim 10^{-4}$  at  $z \sim 6$ , this results in average Ly $\alpha$  optical depths of a few and an almost completely absorbed Ly $\alpha$  forest. Moreover, the apparent decline of the space density of bright QSOs as we enter the reionization period allows us to probe the evolution of the reionization process only with rather small samples with only a handful of objects detected at  $z > 6.5$  (Mortlock et al. 2011; Venemans et al. 2013, 2015; Matsuoka et al. 2016, 2017; Reed et al. 2017; Tang et al. 2017).

Recently, Barnett et al. (2017, B17 hereafter) obtained a very high signal-to-noise (S/N) spectrum of the QSO with the highest redshift known, ULAS J1120+0641 at  $z = 7.084$  (Mortlock et al. 2011). B17 used *Hubble Space Telescope* imaging to flux-calibrate and combined additional data with their previous spectroscopic

\* E-mail: jonathan.chardin@astro.unistra.fr

observations with the X-shooter instrument at the Very Large Telescope. They reported the detection of seven narrow transmission spikes in the redshift range  $5.858 < z < 6.122$  followed by the longest Gunn–Peterson (GP) trough (Gunn & Peterson 1965) of complete absorption known in a QSO spectrum. The long trough extends from  $z = 6.122$  up to  $z = 7.04$  at the edge of the QSO near-zone and corresponds to a length of  $240 \text{ cMpc } h^{-1}$ .

B17 then used the Sherwood simulation (Bolton et al. 2017) to constrain the redshift evolution of the Ly  $\alpha$  optical depth at  $z > 5.7$  by trying to reproduce the  $240 \text{ cMpc } h^{-1}$  long absorption trough in mock absorption spectra produced from the simulation. They concluded that the occurrence of this long GP trough provides only weak additional constraints on the evolution of the neutral hydrogen fraction ( $x_{\text{HI}} \geq 10^{-4}$ ) much weaker than the constraints inferred from the red damping wing by Mortlock et al. 2011 ( $x_{\text{HI}} > 0.1$ ), but also see Bolton et al. (2011) and Bosman & Becker (2015). B17 used their analysis to provide an updated power-law parametrization of the effective optical depth  $\tau_{\text{eff}}(z)$  constrained by the overall transmitted flux in the redshift range  $5.858 < z < 6.122$ . They, however, did not try to reproduce the observed transmission spikes in detail.

Here, we will model both the observed transmission spikes and the long trough in the spectrum of ULAS J1120+0641 with radiative transfer simulations that were originally introduced in Chardin et al. (2015). Our main goal is to model the transmission spikes identified by B17 in the observed spectrum of ULAS J1120+0641 (see also Gallerani, Choudhury & Ferrara 2006 for some modelling of the dependence of high-redshift transmission spikes on the reionization history based on a semi-analytical model of QSO absorption spectra). In Chardin et al. (2015), we have presented a set of radiative transfer simulations carefully calibrated with Ly  $\alpha$  forest data. We have shown that these simulations reproduce the observed evolution of  $\tau_{\text{eff}}(z)$  and the inferred hydrogen photoionization rate  $\Gamma$  very well. These simulations also reproduce the evolution of the average mean free path  $\lambda_{\text{mfp}}^{912}$  as measured by Worseck et al. (2014) in the post-reionization period and are consistent with the latest measurement of the Thompson optical depth inferred by Planck (Planck Collaboration XLVII 2016). The simulations follow self-consistently the growth and percolation of H II bubbles and are well suited to put constraints on the exact timing of the end of reionization by comparing the occurrence of transmission spikes at the tail end of reionization in simulated spectra with that in observed QSO absorption spectra.

This paper is organized as follows. In Section 2, we introduce our numerical simulations and the different steps to construct Ly  $\alpha$  spectra that mimic the observed spectrum of ULAS J1120+0641. Section 3 presents our analysis of the mock absorption spectra and the comparison with the observed spectrum. Finally, in Section 4 we discuss our results and their implications for future observations to constrain in more details the evolution of the ionization history at the tail end of reionization before giving our conclusions in Section 5. Throughout the paper, we use the following cosmology (as derived from the 2014 Planck temperature power spectrum data alone; Planck Collaboration XVI 2014):  $\Omega_{\text{m}} = 0.3175$ ,  $\Omega_{\Lambda} = 0.6825$ ,  $\Omega_{\text{b}} = 0.048$ ,  $h = 0.6711$ ,  $\sigma_8 = 0.8344$  and  $n_s = 0.9624$ .

## 2 METHODOLOGY

Here, we briefly describe the radiative transfer simulations and our methodology to construct mock Ly  $\alpha$  forest spectra that mimic the observed spectrum of ULAS J1120+0641.

**Table 1.** Summary of the different simulations studied in this work.

Model name	Box size	Particle number
512-20-good_h (Fiducial)	20	$512^3$
512-20-good_l (Late)	20	$512^3$
512-20-early (Early)	20	$512^3$
512-40 (Low-res)	40	$512^3$
512-10 (High-res)	10	$512^3$

### 2.1 The radiative transfer simulations

Our reionization simulations are described in more detail in our recent studies (Chardin et al. 2015; Chardin, Puchwein & Haehnelt 2017) and have been performed in two steps. First, hydrodynamic simulations were run and the radiative transfer calculations were then performed by post-processing the hydrodynamic simulations for a range of ionizing source models.

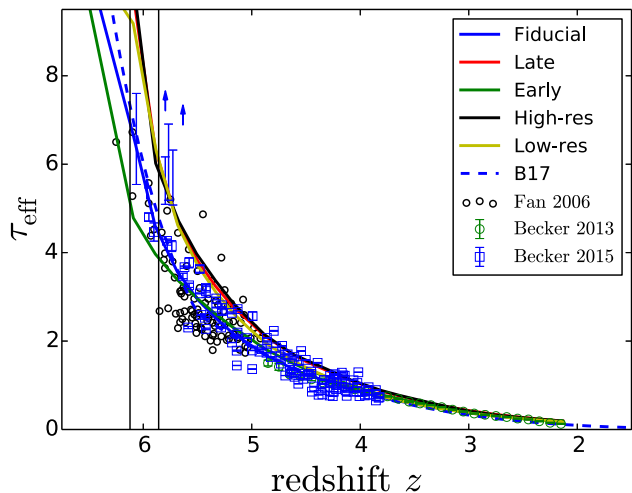
In our simulations, the evolution of dark matter and the hydrodynamics of the gas were performed with the RAMSES code (Teyssier 2002). All the simulations presented in this work use a coarse grid with  $512^3$  cells. The UV background of Haardt & Madau (2012) was used in the RAMSES simulations to model the evolution of the thermal and ionization state of the IGM in the hydrodynamic simulation which results in good agreement with observations from the Ly  $\alpha$  forest in the post-reionization Universe (Puchwein et al. 2015). As discussed in detail in Chardin et al. (2015), the temperature of the hydrosimulation was then also used in the following radiative transfer calculations that were performed in post-processing with the ATON code (Aubert & Teyssier 2008).

ATON is a radiative transfer code utilizing a moment-based description of the radiative transfer equation ported on GPU architecture to accelerate the calculations. We assumed all ionizing photons to have an energy of 20.27 eV in all simulations. Because of our monofrequency treatment, high-energy ionizing photons are not modelled properly. Therefore, we use the temperature from the RAMSES simulation to be consistent with the observed evolution in the post-reionization IGM. Note that we will also discuss simulations later where we have changed this temperature. Radiative transfer ATON outputs were created for snapshots of the RAMSES simulations separated by 40 Myr.

Our ionizing source modelling is assuming the dark matter haloes identified in the RAMSES simulations as the sources of ionizing photons and they are assumed to emit continuously. The ionizing luminosity is assumed to scale with the mass of dark matter haloes in a similar way as in Iliev et al. (2006, but see also Chardin, Aubert & Ocvirk 2012; Chardin, Aubert & Ocvirk 2014). The normalization varies with redshift in such a way that the integrated comoving ionizing emissivity is similar to that of the Haardt & Madau (2012) UV background model, but somewhat modified in order to get a better match with the observations of the hydrogen photoionization rates. In Chardin et al. (2015), we have shown that our calibrations of the ionizing emissivity obtained in this way are also consistent with the redshift evolution of the luminosity function of high-redshift galaxies.

### 2.2 Set of models

We will study here different models for the reionization history. All these different simulations have already been discussed in Chardin et al. (2015). Table 1 summarizes some of the properties of the different models.



**Figure 1.** Evolution of the effective Ly  $\alpha$  optical depth  $\tau_{\text{eff}(z)}$  in our different models. The two black vertical lines show the redshift range where the seven spikes have been detected in ULAS J1120+0641. The blue dashed line shows the evolution from B17 which matches the evolution in our fiducial model very well.

Our reference model is the 512-20-good\_h model of Chardin et al. (2015) where 512 stands for the number of cells in one dimension and 20 for the box size in  $\text{cMpc } h^{-1}$ . We also consider two others models here with a slightly later/earlier reionization history compared to the reference model. The model where reionization occurs slightly later is the 512-20-good\_1 and the model where it occurs slightly earlier is the 512-20-early model from Chardin et al. (2015).

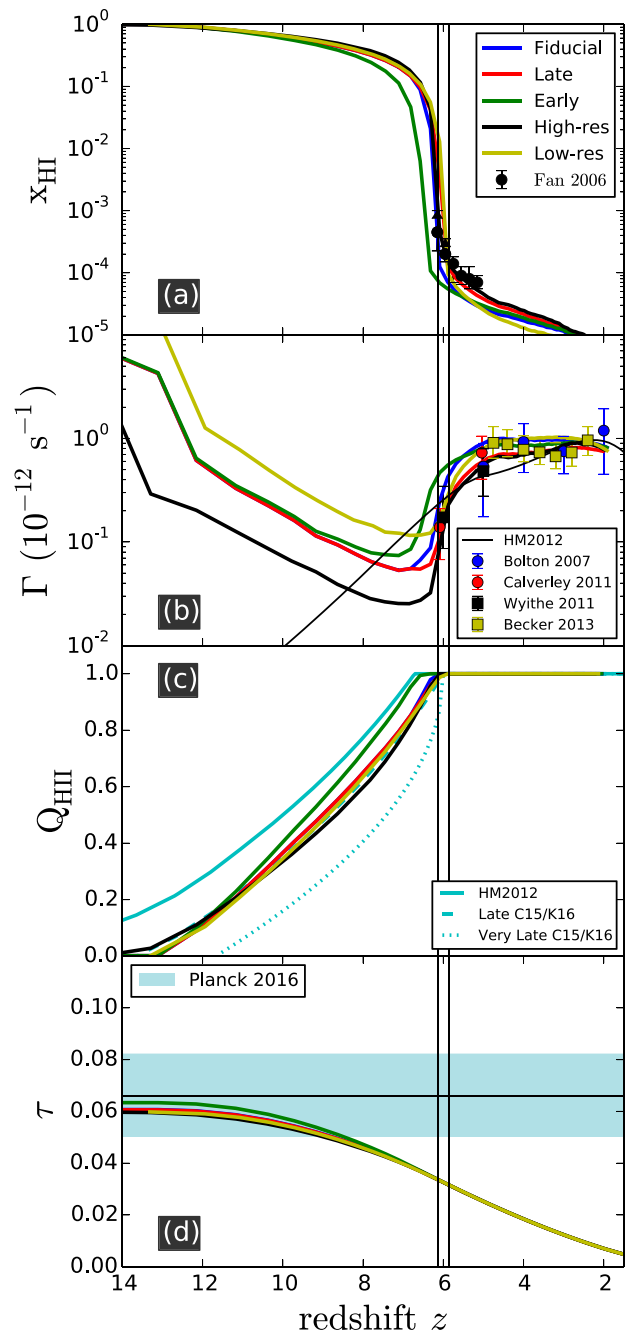
Finally, we also consider two further simulations to study the impact of the resolution of the simulations. For these, we used the models 512-40 and 512-10 of late reionization, also described in Chardin et al. (2015) with  $512^3$  grid cells as in the other models, but with a box size of 40 and 10  $\text{cMpc } h^{-1}$ , respectively. This allows us to test the effect of raising/lowering the mass resolution by a factor of 8 compared to our fiducial model with a box size of 20  $\text{cMpc } h^{-1}$ .

Fig. 1 shows the evolution of the effective optical depth  $\tau_{\text{eff}}$  in our different simulations. The blue dashed line shows the evolution inferred by B17, parametrized as

$$\tau_{\text{eff}}(z) \begin{cases} 0.85 \left(\frac{1+z}{5}\right)^{4.3} & \text{for } z \leq 5.5, \\ 2.63 \left(\frac{1+z}{6.5}\right)^{11.2} & \text{for } z > 5.5. \end{cases} \quad (1)$$

Interestingly, our fiducial model almost exactly matches the evolution of  $\tau_{\text{eff}}$  inferred by B17. We therefore expect our mock absorption spectra constructed for this model to be in good agreement with the observed spectrum.

In Fig. 2, we show the redshift evolution of a number of volume averaged quantities in our different models such as the mean neutral hydrogen fraction, the mean photoionization rate  $\Gamma$  and the Thompson optical depth. We can see that our simulations bracket well a range of observational constraints. In panel (c), we compare the redshift evolution of the volume filling factor of ionized region for our five different reionization histories with those presented in Choudhury et al. (2015) and Kulkarni et al. (2016). Note that all reionization histories are rather similar to the ‘late’ model in Choudhury et al. (2015) and Kulkarni et al. (2016) that these



**Figure 2.** (a) redshift evolution of the average neutral fraction in our different models. (b) redshift evolution of the average photoionization rate ( $\Gamma_{12} = \Gamma/10^{12}$ ). (c) redshift evolution of the volume filling factor  $Q_{\text{HII}}$  of H II regions. The three different cyan lines show the evolution of  $Q_{\text{HII}}$  in the three models presented in Choudhury et al. (2015) and Kulkarni et al. (2016). (d) redshift evolution of the Thompson optical depth.

authors have shown also to be consistent with the rapid evolution of Ly  $\alpha$  emitters at  $z > 5.5$ .

Here, we will use the occurrence of transmission spikes followed by a long fully absorbed GP trough starting at  $z \approx 6.12$  to constrain the timing of reionization in the line of sight to ULAS J1120+0641.

### 2.3 Creating synthetic Ly $\alpha$ forest spectra from the simulations

For this study, we have generated mock Ly  $\alpha$  forest spectra that mimic the observed spectrum of ULAS J1120+0641 from the five different radiative transfer simulations of cosmic reionization described above. As already discussed, seven statistically significant transmission spikes have been reported for the redshift range  $5.858 < z < 6.122$  by B17. This redshift interval corresponds to a comoving length of 112.9 cMpc for our cosmology, equivalent to  $75.77 \text{ cMpc } h^{-1}$ . This is roughly four times our fiducial box size of  $20 \text{ cMpc } h^{-1}$ . B17 also reported the longest fully absorbed GP trough ever found in a QSO spectrum with a length of  $240 \text{ cMpc } h^{-1}$  from  $z = 6.122$  up to  $z = 7.04$  at the edge of the quasar near-zone. For the slightly different cosmological parameter adopted in our simulations, this corresponds to a comoving length of  $234 \text{ cMpc } h^{-1}$ . The part of the spectrum of interest for our analysis has therefore a total length of  $\sim 310 \text{ cMpc } h^{-1}$ .

In our radiative transfer simulation with ATON, outputs were generated every 40 Myr from redshift  $z \sim 100$  up to redshift  $z \sim 2$ . We use outputs that fall inside the redshift range corresponding to the location of the observed spikes and trough to construct spectra that cover a total comoving length of  $\sim 310 \text{ cMpc } h^{-1}$ . In practice, we concatenate random lines of sight through our simulation outputs to construct the skewers. For each simulation, we calculated 1000 such skewers that mimic the observed spectrum.

In Fig. 3, we show the normalized transmitted flux  $F = e^{-\tau}$  along the same line of sight for the fiducial, late and early reionization histories for our simulations with box size of  $20 \text{ cMpc } h^{-1}$ . The three redshift ranges covered by the skewers for which the spectra are shown are similar to those where spikes occur in the observed spectrum of ULAS J1120+0641. We also show the underlying neutral hydrogen  $n_{\text{HI}}$  density field in the slice where the line of sight is calculated, the value of the photoionization rate  $\Gamma$  along the line of sight, as well as the over density and the temperature. For ease of comparison with the absorption spectrum, the temperature, over-density and photoionization rate along the line of sight are plotted in velocity rather than real space. We can see that the locations where transmitted flux is reported are consistent from one model to another for the same line of sight. However, the amount of transmitted flux is different in the different models. As the underlying density and temperature field is the same in the three models, the difference in transmitted flux is here only due to differences in the photoionization rate values along the skewer due to the different reionization histories. The transmission spikes all coincide with the most underdense regions in our simulations. For all three different reionization histories, percolation is already complete and the spatial fluctuations in the photoionization rate are small.

To construct mock spectra directly comparable to the observed spectrum, we transformed our optical depth in each pixel to a flux in  $\text{erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ . We use the normalization at  $1280 \text{ \AA}$  in the observed spectrum and assume a power law for the normalization as a function of observed wavelength such that

$$f_{\lambda} = \left[ \frac{\lambda}{1280 \times (1+z)} \right]^{-0.5} f_{1280} \quad (2)$$

with  $f_{1280} = 5.5 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ . Then we transform our normalized transmitted flux  $F = e^{-\tau}$  in an observed flux as  $F_{\lambda} = f_{\lambda} \times F$ .

In order to mimic the observed spectrum, we smooth the mock spectra with a Gaussian kernel with FWHM =  $34 \text{ km s}^{-1}$ . We then rebin the spectrum with a pixel size of  $10 \text{ km s}^{-1}$

and add random Gaussian noise to the spectra with a dispersion given by the ULAS J1120+0641 error spectrum with  $\sigma = 0.125 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ .

In Fig. 4, we show one of our mock spectra constructed from our fiducial simulation. Gratifyingly, there are several statistically significant transmission spikes followed by a very long almost fully absorbed GP trough very similar to what is found in the observed spectrum.

### 2.4 Identifying statistically significant transmission spikes

In this subsection, we describe our search for statistically significant transmission spikes in our mock absorption spectra, which follows closely that presented by B17 for the observed spectrum of ULAS J1120+0641. For the search, we used a Gaussian-shaped matched filter (see Bolton et al. 2004) with a range of widths. As in B17, we test different width for the Gaussian-shaped matched filter with the narrowest profile having  $\sigma = 15 \text{ km s}^{-1}$  and the others having  $\sigma$  increasing by factors of  $\sqrt{2}$  ( $\sigma = 21, 30, \dots, 120 \text{ km s}^{-1}$ ).

For each pixel, we then record the S/N of the matched filter search and keep the most significant detections from the search with the templates with different widths. In Fig. 5, we show the corresponding fit with the matched filter together with the mock absorption spectrum for the part of the spectrum where transmission spikes occur in the observed spectrum of ULAS J1120+0641. We also show the S/N we computed assuming a noise level equal to the value of the averaged observed noise spectrum along the wavelength range considered. In the particular example shown, we have identified several transmission spikes similar to those in the observed spectrum with an S/N  $> 5$  (the detection threshold used in B17).

Fig. 6 shows a zoom-in on the seven spikes identified in Fig. 5. This figure can be directly compared to fig. 5 in B17 that shows a zoom of the seven spikes found in the observed spectrum. The spikes in our simulated spectra and those in the observed spectrum look very similar. As we will discuss in more detail later, the spikes in our simulated spectra appear to have, however, a somewhat smaller height and larger width than those in the observed spectrum.

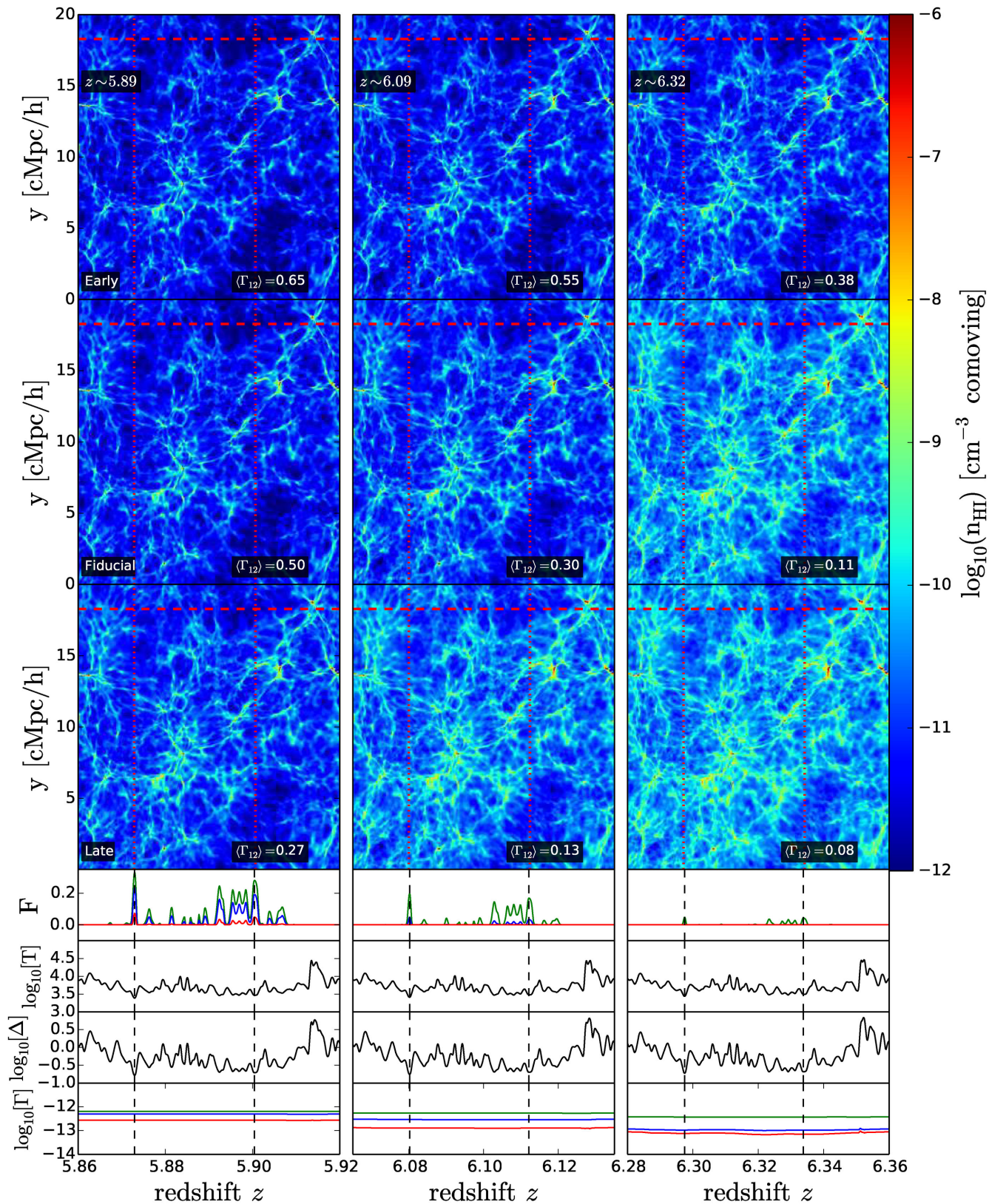
## 3 RESULTS

In this section, we present our main results. First, we look at the incidence rate of transmission spikes in the different simulations and the total amount of transmitted flux. We then investigate whether we find a GP trough as long as in the observed spectrum.

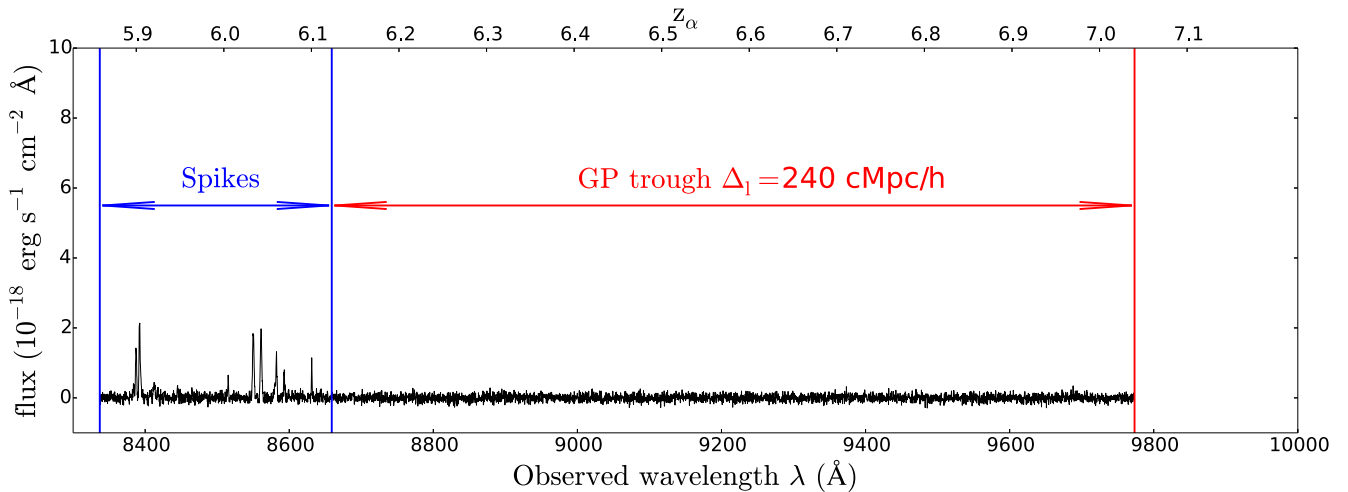
### 3.1 Incidence rate of transmission spikes

In Fig. 7, we show the normalized probability distribution function (PDF) of the number of spikes in the different simulations calculated from 1000 mock spectra for random lines of sight through each simulation. The search for statistically significant transmission spikes is performed over the wavelength range corresponding to the observed detection of spikes in ULAS J1120+0641. We count a transmission spike as statistically significant if at least one pixel has S/N  $> 5$  as in the observational study. The black vertical dashed line shows the value of seven transmission spikes identified by B17 in the observed spectrum of ULAS J1120+0641.

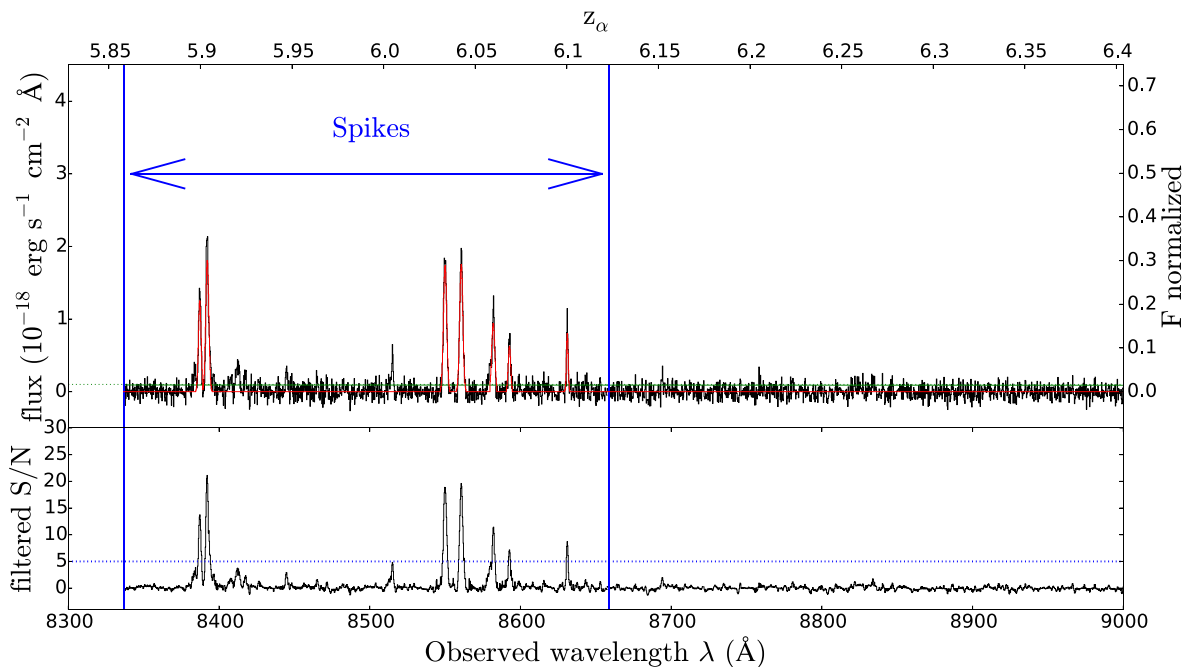
The first thing to note is that the incidence rate of statistically significant transmission spikes is extremely sensitive to the reionization history. Recall that our reionization histories are actually all rather



**Figure 3.** Example of a line of sight (shown with the red dashed horizontal lines) through the volume of our three  $20 \text{ cMpc } h^{-1}$  simulations at three different redshifts. For each model, we show the transmitted flux  $F$ , as well as the temperature  $T$ , the overdensity  $\Delta$  and the photoionization rate  $\Gamma$  in velocity space along the line of sight. The background shows the neutral hydrogen comoving density  $n_{\text{HI}}$  in the different models in a slice of  $39.0625 \text{ kpc } h^{-1}$  thickness.



**Figure 4.** Example of a synthetic  $\text{Ly}\alpha$  forest spectrum corresponding to the wavelength range of the observed spectrum of ULAS J1120+0641 presented by B17. The simulated spectrum has been first convolved to the instrumental resolution with a Gaussian of  $\text{FWHM} = 34 \text{ km s}^{-1}$  and then rebinned to a pixel size of  $10 \text{ km s}^{-1}$  as in the data presented in B17. Random Gaussian noise has been added to every pixel with a standard deviation corresponding to the average value of the observed spectrum presented in B17. The spectrum is divided into two sections that correspond to the wavelength range of the observed spikes in ULAS J1120+0641 and to that of the long GP trough of  $240 \text{ cMpc h}^{-1}$  length.



**Figure 5.** Top: Zoom-in of one of our synthetic spectra of the wavelength region of the observed spectrum of ULAS J1120+0641 showing transmission spikes. The original smoothed plus rebinned plus noise spectrum are shown in black. The horizontal green solid line shows the assumed noise level corresponding to the average noise value in the observed spectrum of ULAS J1120+0641 as presented by B17. The red solid curves correspond to the results of Gaussian line fitting with the best  $\chi^2$  values. Bottom: the corresponding filtered  $S/N$  ratio.

similar and were all chosen to be consistent with the observed overall evolution of the  $\text{Ly}\alpha$  optical depth in the observed high-redshift QSO spectra. Interestingly, the model that appears to best match the data is the ‘Early’ reionization history and not our fiducial reionization history that matches best the evolution of  $\tau_{\text{eff}}$  inferred from the observed spectrum in B17 (see equation 1 and Fig. 1), but this is of course of limited significance for only one line of sight. As expected, the larger  $\tau_{\text{eff}}$  is in the simulation, the smaller the num-

ber of spikes detected in the skewer. Our fiducial model is certainly still consistent with the observed number of spikes, but the peak of the PDF is at somewhat lower values at  $\sim 2\text{--}3$  spikes. It is important to note here again that we are using the temperature from the hydrosimulations and reionization occurs rather early with the (homogeneous) HM2012 UV background employed in the hydrosimulations. The temperatures in the underdense regions are thus rather low (about 5000 K). As discussed by D’Aloisio,

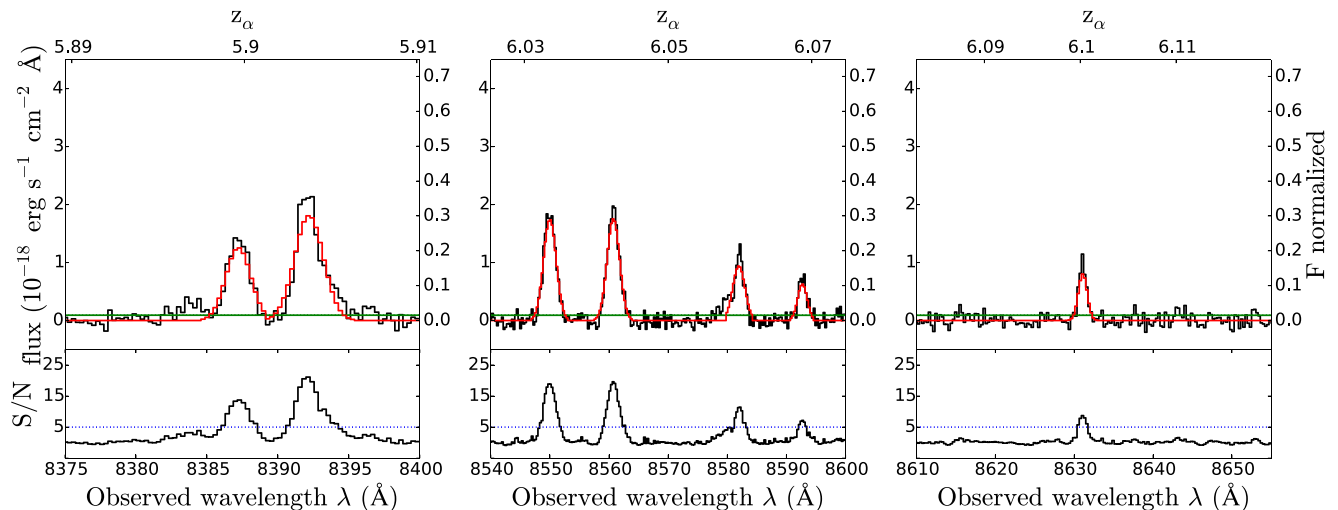


Figure 6. Zoom-in on the different transmission spikes from Fig. 5.

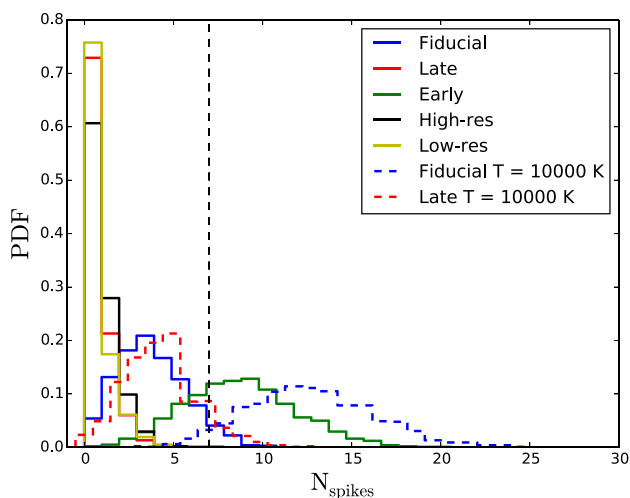


Figure 7. The PDF of the number of identified transmission spikes obtained from a sample of 1000 lines of sight covering the wavelength range where transmission spikes occur in the observed spectrum of ULAS J1120+0641. The black vertical dashed line shows the value (seven transmission spikes) found in the study of B17.

McQuinn & Trac (2015) and Keating, Puchwein & Haehnelt (2017), temperatures in recently ionized underdense regions may be significantly hotter than this. For a fixed photoionization rate, this will decrease the neutral hydrogen fraction and thus the Ly  $\alpha$  opacity. To test this, we have run two further simulations where we have fixed the temperature of the gas in *ATON* simulations of our fiducial and late reionization histories at 10000 K, independent of the density of the gas. The resultant PDF of the number of spikes is shown as the dashed blue and red curves. As expected with higher temperatures, the transmissivity of the underdense region responsible for the spikes increases and the PDF shifts to larger numbers. For a temperature of 10000 K, the best match lies somewhere between the fiducial and late reionization histories, while even higher temperatures may suggest an even later end of reionization in the line of sight to ULAS J1120+0641. Finally, our lower and higher resolution simulations with a reionization history similar to that of the late reionization model show very similar results to the simulation

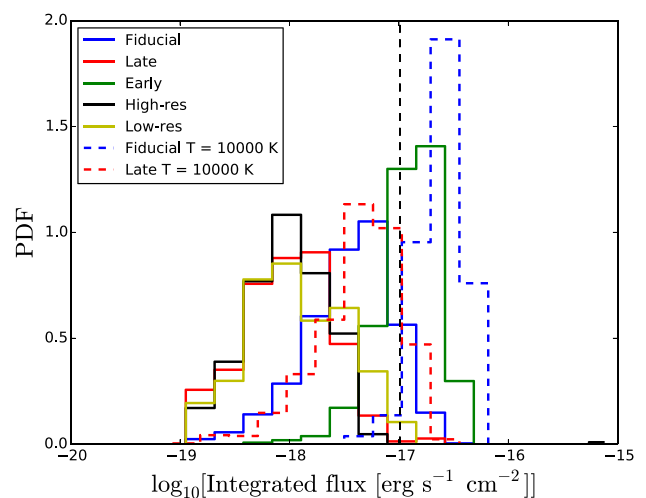
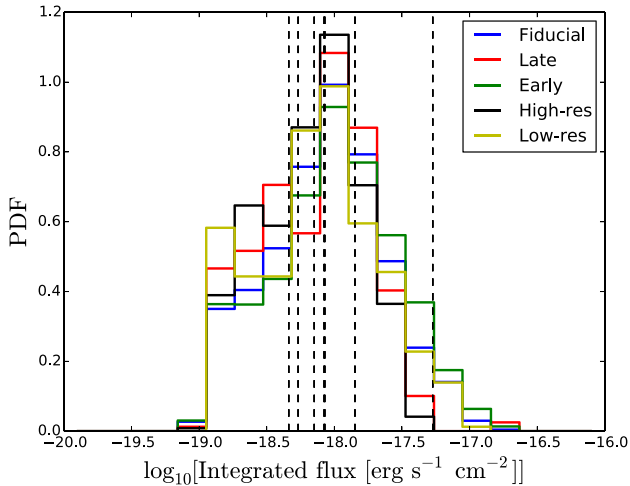


Figure 8. The PDF of the total flux integrated over all the detected transmission spikes from a sample of 1000 lines of sight through the simulations. The black vertical dashed line shows the value found in the study of B17 corresponding to a total transmitted flux of  $F = 1.02 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ .

at our default resolution, demonstrating that the number of significant transmission spikes is rather insensitive to the resolution of the simulation.

### 3.2 Total flux in transmission spikes

In Fig. 8, we show the PDF of the total transmitted flux of all transmission spikes detected in a given spectrum as obtained from 1000 lines of sight through the simulations with the different reionization histories. The vertical black dashed line shows the value of  $(1.02 \pm 0.03) \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$  reported by B17 for the observed spectrum of ULAS J1120+0641. As for the number of transmission spikes, we observe that the early reionization history model is the one that best matches the observed total flux from the real spectrum with the low temperatures of the *RAMSES* simulation. For our fiducial reionization history, the total flux is also consistent with the observed transmitted flux, but the peak of the PDF is located at values somewhat lower than the total flux in the



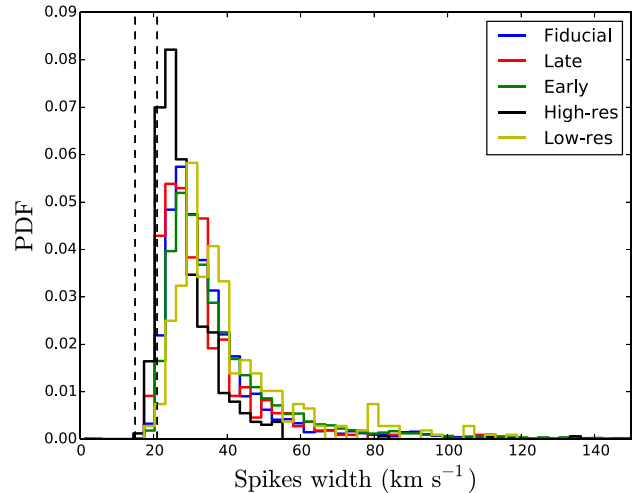
**Figure 9.** The PDF of the flux of individual transmission spikes calculated from 1000 lines of sight through the simulation of each model. The seven black vertical dashed lines show the value found in the study of B17 corresponding to the seven spikes identified in the observed spectrum of ULAS J1120+0641.

observed spectrum. As shown by the dashed blue and red curves for a temperature of 10 000 K in the underdense regions, the best match lies somewhere between the fiducial and late reionization histories, similarly to what happened in case of the number of transmission spikes.

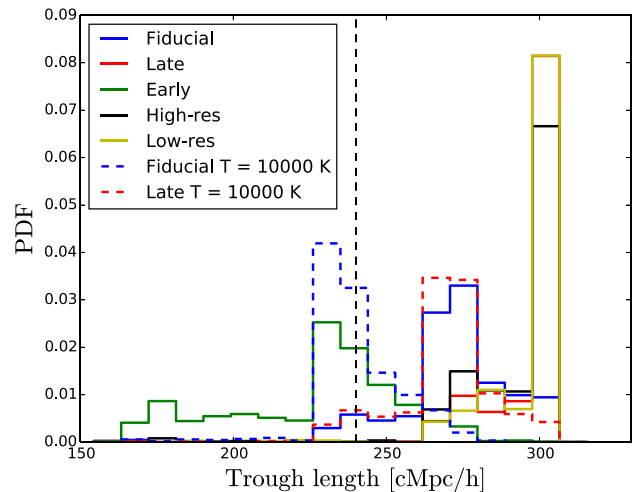
In Fig. 9, we show the PDF of the transmitted flux of individual transmission spikes for the different reionization histories. The vertical black dashed lines correspond to the transmitted flux values in the seven transmission spikes in the observed spectrum. All models appear to be consistent with the values reported for the observed spectrum. This is perhaps not surprising as the minimum S/N for a significant detection corresponds to a lower limit of the flux for a transmission spike to be counted as significant detection that depends only weakly on the width of the transmission spike. Note, however, that the observed spike with the highest transmitted flux dominates the total observed flux.

### 3.3 The width of the transmission spikes

In Fig. 10, we show the PDF of the width of the Ly $\alpha$  line for transmission spikes with S/N > 5 in terms of the  $\sigma$  of the best-fitting Gaussian. We note here that this is somewhat different to what was done by B17 for the observed spectrum, as they give the width of the best-fitting Gaussian of their (rather coarse) matched filter search while we give the value of  $\sigma$  of the best-fitting Gaussian for each transmission spike identified. The values reported for the seven spikes by B17 are 15 km s $^{-1}$  for six of the transmission spikes and 21 km s $^{-1}$  for the remaining spike with the highest transmitted flux. We show these two values in Fig. 10 with the two black vertical dashed lines. At our fiducial resolution, the spikes in our mock spectra are clearly broader by a factor of 1.5–2 than the observed transmission spikes. Looking at the higher and lower resolution simulations of the late reionization history shows, however, that the width of the transmission spikes is not converged and decreases with increasing resolution. This is consistent with the results of Bolton & Becker (2009), who also found that very high resolution (in their case 512 $^3$  for a 10 Mpc  $h^{-1}$  SPH simulation or better)



**Figure 10.** The PDF of the width of the transmission spikes as identified with a Gaussian-shaped matched filter calculated from 1000 lines of sight for each model. The two black vertical dashed lines show the value found in the study of B17 corresponding to the seven spikes observed in the ULAS J1120+0641 quasar spectrum.



**Figure 11.** The PDF of the fully absorbed GP trough length in the different models. The values are calculated as the comoving length between the values of  $z_{\min}$  that corresponds to highest redshift where a statistically significant transmission spike is identified and  $z = 7.04$  corresponding to the edge of the proximity zone of ULAS J1120+0641. The black vertical dashed line shows the value found in the study of B17, 240 cMpc  $h^{-1}$ .

is needed to resolve in particular the underdense regions at high redshift. Note, however, that Bolton & Becker (2009) looked at Ly $\alpha$  forest statistics for significantly higher photoionization rates and thus mean flux levels.

### 3.4 The length of totally absorbed GP troughs

In Fig. 11, we show the PDF of the length of the fully absorbed GP trough for all our models. In practice, the trough lengths are calculated by finding the highest redshift of a statistically significant transmission spike in the different mock spectra. Then we calculate



the corresponding comoving length from that redshift up to a redshift  $z = 7.04$  that corresponds to the edge of the observed quasar proximity zone.

Again, we note that it is the simulated spectra with the early reionization history that best match the observed trough length for the low temperatures in underdense regions of the RAMSES simulation. The model with the fiducial reionization history is also producing troughs with the observed lengths although the PDF of the trough length peaks at somewhat larger values. This is consistent with our finding for the incidence rate of statistically significant transmission spike at the lower end of the redshift range probed by the Ly  $\alpha$  forest in the observed spectrum of ULAS J1120+0641. The blue and red dashed curves show the PDF of the absorption trough length for the 10 000 K case for the fiducial and late reionization models. Higher temperatures in the underdense regions lead to a higher number of spikes detected, with some of them found at higher redshift. This results in lower length of GP trough in the spectra.

## 4 DISCUSSION

### 4.1 The Ly $\alpha$ forest of $z > 6.5$ quasars to constrain the reionization history

In the previous sections, we have shown that matching the absorption properties of the Ly  $\alpha$  forest in ULAS J1120+0641 in radiative transfer simulation is very sensitive to the reionization histories assumed in the simulations. It also depends sensitively on the temperature of the IGM in underdense regions. The use of narrow transmission spikes only detectable with high-S/N high-resolution spectra pushes the redshift limit where information about the ionization state of the hydrogen in IGM can be extracted significantly higher. This suggests that obtaining more high-S/N spectra of  $z > 6.5$  QSOs should allow us to put strong constraints on the timing of the overlap of H II bubbles in many lines of sight. Currently, ULAS J1120+0641 is the only QSO with a 30 h high-resolution spectrum. There are already 10 quasars found at  $z \geq 6.5$ . Four were discovered in near-infrared surveys (the quasar at  $z = 7.085$  studied in the current paper was first reported in Mortlock et al. 2011 and three other quasars were reported in Venemans et al. 2013). Three new quasars (Venemans et al. 2015) were discovered recently from the  $3\pi$  Panoramic Survey Telescope and Rapid Response System (Pan-STARRS1 or PS1; Kaiser et al. 2010). Two new faint quasars have been discovered by the Subaru High- $z$  Exploration of Low-Luminosity Quasars survey (Matsuoka et al. 2016, 2017) and one by the Dark Energy Survey (Reed et al. 2017). Moreover, a new high-redshift quasar, PSO J006.1240+39.2219, at  $z = 6.61$  has recently been found (Koptelova et al. 2017; Tang et al. 2017). Re-observing the nine additional quasars at  $z \geq 6.5$  with high S/N can thus be expected to extend our ability to use the Ly  $\alpha$  forest as a probe of the tail end of reionization to significantly higher redshift than previously possible.

### 4.2 Caveats of our modelling

The models presented here all assume that reionization is driven by ‘star-forming’ galaxies. However, motivated by the large observed opacity fluctuations at rather larger scales (Becker et al. 2015b), some authors including us have recently revisited the old idea that QSOs may contribute significantly to the ionizing emissivity at high redshift (Giallongo et al. 2015; Madau & Haardt 2015, but see also Akiyama et al. 2017; Parsa, Dunlop & McLure 2017 for alternative views on the space density of high-redshift AGN found by Giallongo

et al. 2015). As discussed in detail by Chardin et al. (2015, 2017), a significant contribution of QSOs to the ionizing emissivity at  $z \gtrsim 5.5$  could lead to large fluctuations of the photoionization rate well after the percolation of H II bubbles and could explain the broad distribution of the effective optical depth  $\tau_{\text{eff}}$  measured in 50 cMpc  $h^{-1}$  chunks by Becker et al. (2015b). If necessary, modelling a possible QSO contribution to the ionizing emissivity will obviously further complicate the interpretation of the observed transmission spikes. Note, however, that there is currently no evidence for the presence of QSOs close to the line of sight of ULAS J1120+0641 even though QSOs contributing to the ionizing emissivity in the line of sight to ULAS J1120+0641 could have obviously switched off or may not point towards us.

Two alternative explanations have been put forward for the observed large opacity fluctuations on large scales: spatial fluctuations of the temperature–density relation of the IGM (D’Aloisio et al. 2015, but see Keating et al. 2017) and large mean free path fluctuations that could arise when the mean free path of ionizing photons is still shorter than or comparable to the mean distance between ionizing sources (see Davies & Furlanetto 2016; D’Aloisio et al. 2016). Modelling this properly is rather difficult and beyond the scope of this paper but may also further complicate the interpretation of the observed transmission spikes.

## 5 CONCLUSIONS

We have presented here the result of our analysis of mock Ly  $\alpha$  absorption spectra that aim to reproduce the observed high-S/N high-resolution spectrum of the  $z = 7.084$  QSO ULAS J1120+0641 recently presented by B17. The radiative transfer simulation used to construct the mock absorption spectra had been previously carefully calibrated to match the main observational constraints from Ly  $\alpha$  forest data such as the redshift evolution of the effective optical depth, the neutral hydrogen fraction, the inferred average hydrogen photoionization rate and the Thompson optical depth measured from the cosmic microwave background.

Our main results are as follows.

- (i) The occurrence of narrow transmission spikes from highly underdense regions and the onset of a fully absorbed GP trough in our simulations are caused by the rather rapid rise of the photoionization rate following the percolation of H II regions and are very sensitive to the exact timing of the overlap of H II bubbles as well as the temperature of the IGM at low densities.
- (ii) Using our well-calibrated simulations with a narrow range of three reionization histories (early, fiducial and late), we were able to reproduce the occurrence of seven transmission spikes, the observed integrated transmitted flux and the length of the totally absorbed GP trough in the observed spectrum of ULAS J1120+0641 very well.
- (iii) Reionization histories with a later overlap of H II region require the underdense regions to be hotter to be consistent with the observed spectrum.
- (iv) The width of the transmission spikes in our simulated spectra is not fully converged and is about a factor of 1.5–2 larger in spectra produced from our simulations with our fiducial resolution ( $512^3$  grid cell, 20 cMpc  $h^{-1}$  box size).

Our analysis shows that identification of transmission spikes in high-S/N spectra can push the constraints on the ionization state of the IGM to redshifts as large as  $z = 6.1$  and possibly higher, where the average Ly  $\alpha$  optical depth is already very large. This is significantly higher than was previously possible, and constraints on the ionization state of the IGM from Ly  $\alpha$  forest data at these

redshifts were so far limited to the difficult analysis of the near-zones and possible red damping wings in the QSO absorption spectra. Extending the search for and the modelling of transmission spikes at  $z > 5.8$  (still to be obtained) high-S/N absorption spectra of a larger sample of QSOs should provide important insight into the exact timing of the overlap of H II bubbles at the tail end of reionization.

## ACKNOWLEDGEMENTS

We thank James Bolton and George Becker for comments on the manuscript. This work was supported by the ERC Advanced Grant 320596 ‘The Emergence of Structure during the epoch of Reionization’. The RAMSES simulation were performed utilizing the supercomputer COSMOS Shared Memory system at DAMTP, University of Cambridge operated on behalf of the STFC DiRAC HPC Facility. This equipment is funded by BIS National E-infrastructure capital grant ST/J005673/1 and STFC grants ST/H008586/1, ST/K00333X/1. The ATON radiative transfer simulation in this work were performed using the Wilkes GPU cluster at the University of Cambridge High Performance Computing Service (<http://www.hpc.cam.ac.uk/>), provided by Dell Inc., NVIDIA and Mellanox, and part funded by STFC with industrial sponsorship from Rolls Royce and Mitsubishi Heavy Industries.

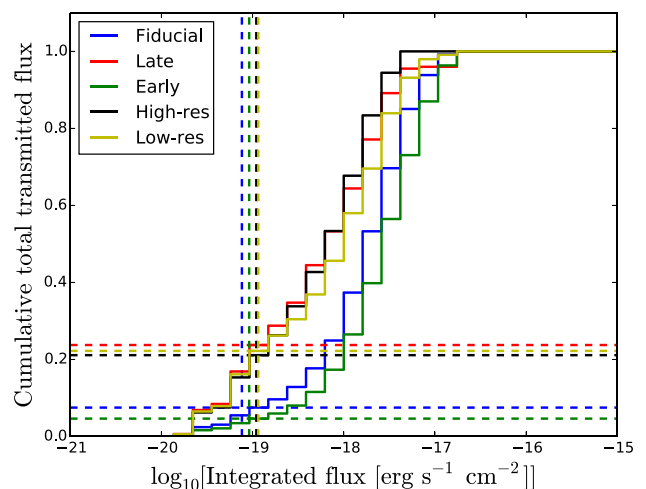
## REFERENCES

- Akiyama M. et al., 2017, preprint (arXiv:1704.05996)  
Aubert D., Teyssier R., 2008, MNRAS, 387, 295  
Barnett R., Warren S. J., Becker G. D., Mortlock D. J., Hewett P. C., McMahon R. G., Simpson C., Venemans B. P., 2017, A&A, 601, A16 (B17)  
Becker G. D., Hewett P. C., Worseck G., Prochaska J. X., 2013, MNRAS, 430, 2067  
Becker G. D., Bolton J. S., Lidz A., 2015a, PASA, 32, e045  
Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015b, MNRAS, 447, 3402  
Bolton J. S., Becker G. D., 2009, MNRAS, 398, L26  
Bolton J. S., Haehnelt M. G., Warren S. J., Hewett P. C., Mortlock D. J., Venemans B. P., McMahon R. G., Simpson C., 2011, MNRAS, 416, L70  
Bolton J. S., Haehnelt M. G., 2013, MNRAS, 429, 1695  
Bolton A. S., Burles S., Schlegel D. J., Eisenstein D. J., Brinkmann J., 2004, AJ, 127, 1860  
Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897  
Bosman S. E. I., Becker G. D., 2015, MNRAS, 452, 1105  
Chardin J., Aubert D., Ocvirk P., 2012, A&A, 548, A9  
Chardin J., Aubert D., Ocvirk P., 2014, A&A, 568, A52  
Chardin J., Haehnelt M. G., Aubert D., Puchwein E., 2015, MNRAS, 453, 2943  
Chardin J., Puchwein E., Haehnelt M. G., 2017, MNRAS, 465, 3429  
Choudhury T. R., Puchwein E., Haehnelt M. G., Bolton J. S., 2015, MNRAS, 452, 261  
D’Aloisio A., McQuinn M., Trac H., 2015, ApJ, 813, L38  
D’Aloisio A., McQuinn M., Davies F. B., Furlanetto S. R., 2016, preprint (arXiv:1611.02711)  
Davies F. B., Furlanetto S. R., 2016, MNRAS, 460, 1328  
Dijkstra M., Wyithe S., Haiman Z., Mesinger A., Pentericci L., 2014, MNRAS, 440, 3309  
Fan X. et al., 2006, AJ, 132, 117  
Gallerani S., Choudhury T. R., Ferrara A., 2006, MNRAS, 370, 1401  
Giallongo E. et al., 2015, A&A, 578, A83  
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633  
Haardt F., Madau P., 2012, ApJ, 746, 125

- Iliev I. T., Mellema G., Pen U., Merz H., Shapiro P. R., Alvarez M. A., 2006, MNRAS, 369, 1625  
Kaiser N. et al., 2010, in Stepp L. M., Gilmozzi R., Hall H. J., eds, Proc. SPIE Conf. Ser. Vol. 7733, Ground-based and Airborne Telescopes III. SPIE, Bellingham, p. 77330E  
Keating L. C., Puchwein E., Haehnelt M. G., 2017, preprint (arXiv:1709.05351)  
Koptelova E., Hwang C.-Y., Yu P.-C., Chen W.-P., Guo J.-K., 2017, Nature Sci. Rep., 7, 41617  
Kulkarni G., Choudhury T. R., Puchwein E., Haehnelt M. G., 2016, MNRAS, 463, 2583  
McQuinn M., 2016, ARA&A, 54, 313  
Madau P., Haardt F., 2015, ApJ, 813, L8  
Matsuoka Y. et al., 2016, ApJ, 828, 26  
Matsuoka Y. et al., 2017, preprint (arXiv:1704.05854)  
Mortlock D. J. et al., 2011, Nature, 474, 616  
Parsa S., Dunlop J. S., McLure R. J., 2017, preprint (arXiv:1704.07750)  
Planck Collaboration XVI, 2014, A&A, 571, A16  
Planck Collaboration XIII, 2016, A&A, 594, A13  
Planck Collaboration XLVII, 2016, A&A, 596, A108  
Puchwein E., Bolton J. S., Haehnelt M. G., Madau P., Becker G. D., Haardt F., 2015, MNRAS, 450, 4081  
Reed S. L. et al., 2017, MNRAS, 468, 4702  
Tang J.-J. et al., 2017, MNRAS, 466, 4568  
Teyssier R., 2002, A&A, 385, 337  
Venemans B. P. et al., 2013, ApJ, 779, 24  
Venemans B. P. et al., 2015, ApJ, 801, L11  
Worseck G. et al., 2014, MNRAS, 445, 1745

## APPENDIX A: THE REMAINING FLUX BELOW DETECTION LIMIT

In Fig. A1, we show the cumulative transmitted flux for all the different models as a function of the flux in individual transmitted spikes. In order to estimate the amount of flux beyond the detection limit of the search performed by B17, we use a signal-to-noise ratio  $S/N = 1$ . To this end, we have rerun a Gaussian-matched filter search



**Figure A1.** The cumulative transmitted flux as a function of flux in individual transmission spikes in the different models, if we assume a signal-to-noise ratio  $S/N = 1$  in order to estimate the flux in transmission that is beyond the current detection limit of the observations by B17. The vertical coloured dashed lines show the detection limit of the flux in individual spikes assuming a signal-to-noise-ratio  $S/N = 5$  (the smallest values found in each simulation in Fig. 9) as in the observations. The corresponding horizontal coloured dashed lines show the percentage of total flux that would be missed by current observations in the different models.

with this value of the S/N. The vertical dashed lines in Fig. A1 show the flux detection limit of the spikes detected with  $S/N = 5$  as in the observations (corresponding to the smallest value of flux in individual spikes found in Fig. 9). The corresponding horizontal dashed lines show the percentage of the total transmitted flux that is beyond the detection limit of Barnett et al. (2017).

About 5 and 7 per cent of the total transmitted flux are missing for the ‘Fiducial’ and ‘Early’ reionization histories, respectively. Larger values of the order of 20 per cent are found for simulations with the ‘Late’ reionization history and for the, ‘Low-res’ and ‘High-res’ simulations. These larger values are not surprising as spikes with large values of transmitted flux are much rarer in these

simulations because reionization ends too late to be consistent with the observations. Therefore, small spikes beyond the detection limit impact the total transmitted flux more significantly in these simulations. Since these models are, however, not fitting the observed number of transmission spikes, we can conclude that our simulations suggest that the majority (about 95 per cent) of the total transmitted flux in ULAS J1120+0641 has been detected by Barnett et al. (2017).

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.