



**HAL**  
open science

# Constraining the second half of reionization with the Ly $\beta$ forest

Laura C. Keating, Girish Kulkarni, Martin G. Haehnelt, Jonathan Chardin,  
Laurence Voutquenne-Nazabadioko

► **To cite this version:**

Laura C. Keating, Girish Kulkarni, Martin G. Haehnelt, Jonathan Chardin, Laurence Voutquenne-Nazabadioko. Constraining the second half of reionization with the Ly  $\beta$  forest. Monthly Notices of the Royal Astronomical Society, 2020, 497 (1), pp.906-915. 10.1093/mnras/staa1909 . hal-02459822

**HAL Id: hal-02459822**

**<https://hal.science/hal-02459822v1>**

Submitted on 24 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Constraining the second half of reionization with the Ly $\beta$ forest

Laura C. Keating<sup>1</sup>,<sup>\*</sup> Girish Kulkarni<sup>2</sup>,<sup>\*</sup> Martin G. Haehnelt<sup>3,4</sup> Jonathan Chardin<sup>5</sup>  
and Dominique Aubert<sup>5</sup>

<sup>1</sup>Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St George Street, Toronto, ON M5S 3H8, Canada

<sup>2</sup>Department of Theoretical Physics, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India

<sup>3</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

<sup>4</sup>Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

<sup>5</sup>Observatoire Astronomique de Strasbourg, 11 rue de l'Université, F-67000 Strasbourg, France

Accepted 2020 June 12. Received 2020 June 12; in original form 2019 December 11

## ABSTRACT

We present an analysis of the evolution of the Lyman-series forest into the epoch of reionization using cosmological radiative transfer simulations in a scenario where reionization ends late. We explore models with different mid-points of reionization and gas temperatures. We find that once the simulations have been calibrated to match the mean flux of the observed Lyman- $\alpha$  (Ly  $\alpha$ ) forest at  $4 < z < 6$ , they also naturally reproduce the distribution of effective optical depths of the Lyman- $\beta$  (Ly  $\beta$ ) forest in this redshift range. We note that the tail of the largest optical depths that is most challenging to match corresponds to the long absorption trough of ULAS J0148+0600, which we have previously shown to be rare in our simulations. We consider the evolution of the Lyman-series forest out to higher redshifts, and show that future observations of the Ly  $\beta$  forest at  $z > 6$  will discriminate between different reionization histories. The evolution of the Ly  $\alpha$  and Ly  $\gamma$  forests are less promising as a tool for pushing studies of reionization to higher redshifts due to the stronger saturation and foreground contamination, respectively.

**Key words:** methods: numerical – galaxies: high-redshift – intergalactic medium – quasars: absorption lines – dark ages, reionization, first stars.

## 1 INTRODUCTION

The discovery of the first quasars towards redshift 6 (Fan et al. 2001) provided a new probe of the high-redshift Universe. By analysing the Lyman- $\alpha$  (Ly  $\alpha$ ) forest in the spectra of these quasars, it quickly became clear that the intergalactic medium (IGM) was evolving rapidly in this redshift range and that we were perhaps witnessing the end of the reionization epoch (Fan et al. 2006). As the sample of high-quality spectra of quasars above redshift 5 has increased, so also has the quality of the constraints we can place on the properties of the IGM and on the timing of reionization (see, e.g. Becker, Bolton & Lidz 2015a, for a review of some of the different methods that have been used). Transmission of the Ly  $\alpha$  forest, however, saturates at relatively low HI fractions ( $f_{\text{HI}} \sim 10^{-4}$ ), making it increasingly difficult to use the Ly  $\alpha$  forest as a probe of the IGM at redshifts pushing into the epoch of reionization.

As well as the Ly  $\alpha$  forest, further information can be obtained from the spectra of these quasars by instead analysing the next line in the Lyman series, Lyman- $\beta$  (Ly  $\beta$ ). Ly  $\beta$  absorption occurs at a shorter wavelength than Ly  $\alpha$  ( $\lambda_{\text{Ly}\alpha} = 1215.67 \text{ \AA}$  versus  $\lambda_{\text{Ly}\beta} = 1025.72 \text{ \AA}$ ) and it further has a lower oscillator strength ( $f_{\text{Ly}\alpha} = 0.4164$  versus  $f_{\text{Ly}\beta} = 0.0791$ ). At a given point in the IGM, the optical depth to Ly  $\beta$  absorption is therefore related to Ly  $\alpha$  absorption by

$$\tau_{\text{Ly}\beta} = \frac{f_{\text{Ly}\beta}}{f_{\text{Ly}\alpha}} \frac{\lambda_{\text{Ly}\beta}}{\lambda_{\text{Ly}\alpha}} \tau_{\text{Ly}\alpha} \approx 0.16 \tau_{\text{Ly}\alpha}. \quad (1)$$

In practice, the conversion factor of the *effective* optical depth ( $\tau_{\text{eff}} = -\ln \langle F \rangle$ , where  $\langle F \rangle$  is the mean flux measured along some interval) can also depend on the density structure of the IGM (Oh & Furlanetto 2005; Fan et al. 2006) as well as its temperature (Furlanetto & Oh 2009). The lower optical depth of Ly  $\beta$  allows for transmission even when the HI fraction of the IGM is already high enough to saturate Ly  $\alpha$  transmission, and can potentially be used to push absorption-line constraints on the ionization state of the IGM to higher redshifts. Indeed, Barnett et al. (2017) reported a possible observation of an (unresolved) Ly  $\beta$  transmission spike at  $z = 6.85$ , a redshift much higher than the point where transmission from the Ly  $\alpha$  forest has become saturated. Interpretation of the Ly  $\beta$  forest is complicated, however, by the fact that the Ly  $\beta$  forest is observed behind a foreground of Ly  $\alpha$  forest absorption at a lower redshift :

$$z_{\text{FGLy}\alpha} = \frac{\lambda_{\text{Ly}\beta}}{\lambda_{\text{Ly}\alpha}} (1 + z_{\text{abs}}) - 1, \quad (2)$$

where  $z_{\text{abs}}$  is the redshift of the Ly  $\beta$  absorption. This separation in redshift space is large enough that the density fields can be considered to be uncorrelated (Dijkstra, Lidz & Hui 2004), and the total observed optical depth at an individual pixel can be forward modelled in simulations by calculating  $\tau_{\beta}^{\text{obs}}(z_{\text{abs}}) = \tau_{\beta}(z_{\text{abs}}) + \tau_{\alpha}(z_{\text{FGLy}\alpha})$ , where the Ly  $\beta$  optical depth at the redshift of interest ( $z_{\text{abs}}$ ) and the foreground Ly  $\alpha$  optical depth at  $z_{\text{FGLy}\alpha}$  are calculated for different density fields. Another possibility is to observe absorption of the next Lyman-series transition, Lyman- $\gamma$  (Ly  $\gamma$ ), which occurs at a wavelength  $\lambda_{\text{Ly}\gamma} = 972.54 \text{ \AA}$  and has an oscillator strength  $f_{\text{Ly}\gamma} = 0.029$ . This results in an optical depth  $\tau_{\text{Ly}\gamma} = 0.056 \tau_{\text{Ly}\alpha}$ . In practice, however, observing and interpreting the Ly  $\gamma$  forest will

\* E-mail: lkeating@cita.utoronto.ca

be complicated due to the presence of foregrounds from both the Ly  $\alpha$  and Ly  $\beta$  forests at lower redshift.

The evolution of the effective optical depth of the Ly  $\beta$  forest with redshift has long been used as a probe of the high-redshift IGM (Lidz et al. 2002; Songaila 2004; Fan et al. 2006). Eilers et al. (2019) recently presented a new compilation of effective optical depths of the Ly  $\beta$  forest along 19 different sightlines in the redshift range  $5.5 \lesssim z \lesssim 6.1$ . They demonstrated that it was difficult to simultaneously reproduce the observed ratio of co-spatial Ly  $\alpha$  and Ly  $\beta$  optical depths in a variety of models that modelled the IGM assuming a uniform ionizing background, UV fluctuations from a varying mean-free path (Davies & Furlanetto 2016; D’Aloisio et al. 2018), and temperature fluctuations from inhomogeneous reionization (D’Aloisio, McQuinn & Trac 2015; Keating, Puchwein & Haehnelt 2018). They found that the most successful model was one in which the temperature–density relation of the IGM was inverted, which is somewhat challenging to explain theoretically at these redshifts. If the gas was reionized early, one might expect the gas to have already cooled into the usual temperature–density relation (Hui & Gnedin 1997) or, if it has been recently ionized, it may instead be close to isothermal.

However, the models presented in Eilers et al. (2019) all make the assumption that the IGM is completely ionized above redshift 5.5. Kulkarni et al. (2019) used cosmological radiative transfer simulations to show that the broad distribution of Ly  $\alpha$  forest opacities along different lines of sight (Becker et al. 2015b; Bosman et al. 2018; Eilers, Davies & Hennawi 2018) can be explained by a model where reionization ends late, with islands of neutral hydrogen persisting down to  $z \sim 5.3$ . Such a scenario was also proposed by Lidz et al. (2007) and Mesinger (2010). Keating et al. (2019) further showed that a late-end reionization model also explained the observed long absorption trough in the spectrum of ULAS J0148+0666 (Becker et al. 2015b), as well as its anticorrelation with Ly  $\alpha$  emitters (Becker et al. 2018). This result was subsequently demonstrated using a seminumeric method in Nasir & D’Aloisio (2019).

In this paper, we present a study of the Ly  $\alpha$ , Ly  $\beta$ , and Ly  $\gamma$  forests in late-end reionization models, using cosmological radiative transfer simulations carefully calibrated to match the properties of the observed Ly  $\alpha$  forest. In Section 2, we describe the simulations we analyse here. We compare these simulations to the existing observations of Eilers et al. (2019) in Section 3, and make predictions for future observations in Section 4. Finally, in Section 5, we make our conclusions.

## 2 COSMOLOGICAL RADIATIVE TRANSFER SIMULATIONS OF THE INTERGALACTIC MEDIUM

We model the ionization state of the IGM by post-processing cosmological hydrodynamic simulations with the radiative transfer code ATON (Aubert & Teyssier 2008, 2010). The underlying hydrodynamic simulation was performed with the Tree-PM SPH code P-GADGET-3 (last described in Springel 2005). The simulation volume has box size  $160 \text{ Mpc } h^{-1}$  and was run with  $2048^3$  gas particles, resulting in gas particle masses  $m_{\text{gas}} = 6.4 \times 10^6 M_{\odot} h^{-1}$ . We used a gravitational softening length  $l_{\text{soft}} = 3.1 \text{ kpc } h^{-1}$ . The initial conditions were taken from a simulation performed as part of the Sherwood simulation suite (Bolton et al. 2017) and were generated at  $z = 99$ . The cosmological parameters used were  $\Omega_{\text{m}} = 0.308$ ,  $\Omega_{\Lambda} = 0.692$ ,  $h = 0.678$ ,  $\Omega_{\text{b}} = 0.0482$ ,  $\sigma_8 = 0.829$ , and  $n_{\text{s}} = 0.961$  (Planck Collaboration XVI 2014). The simulation uses a simplified and computationally efficient prescription for star formation, which turns any gas with density

**Table 1.** Details of the simulations presented in this paper.

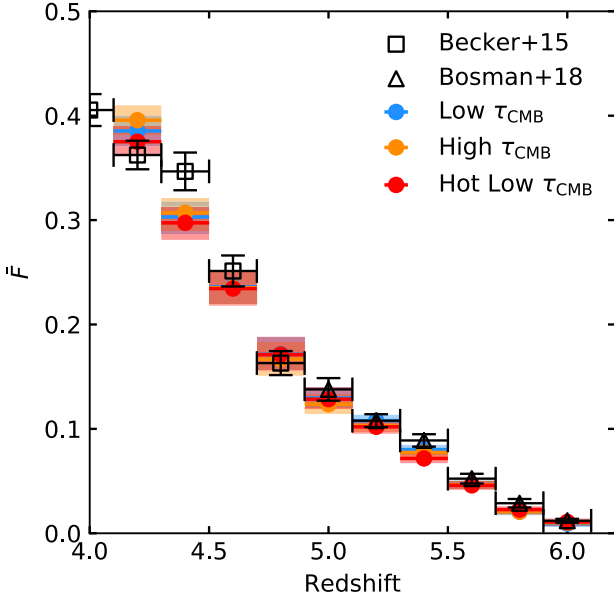
Label	$E_{\gamma}$ (eV)	$z_{50}$	$z_{99.9}$	$\tau_{\text{CMB}}$
Low $\tau_{\text{CMB}}$	17.1	6.7	5.2	0.051
High $\tau_{\text{CMB}}$	17.1	8.4	5.3	0.071
Hot low $\tau_{\text{CMB}}$	18.6	6.7	5.2	0.051

*Notes.* For each simulation, we give the photon energy ( $E_{\gamma}$ ), the redshift at which 50 per cent of the gas is ionized ( $z_{50}$ ), the redshift at which 99.9 per cent of the gas is ionized ( $z_{99.9}$ ), and the optical depth to Thomson scattering of CMB photons ( $\tau_{\text{CMB}}$ ).

1000 times the mean density of the Universe and temperature  $T < 10^5 \text{ K}$  into collisionless star particles. Removing this cold and dense gas from the simulation reduces the time required to perform the simulation, but does not have a significant effect on the low-density gas making up the IGM (Viel, Haehnelt & Springel 2004). As we perform the radiative transfer in post-processing, we still wish to account for the pressure smoothing of the density field and so include the effect of photoheating in the hydrodynamical simulation by imposing a uniform UV background (Haardt & Madau 2012).

We use outputs from this hydrodynamical simulation every 40 Myr starting at  $z = 19.5$  and down to  $z = 4$  as the input density fields for our radiative transfer calculation. As ATON requires Cartesian grids as inputs, we map the SPH particles on to  $2048^3$  grids (with cellsize  $\Delta x = 78.125 \text{ kpc } h^{-1}$ ) using a cubic spline interpolation scheme. To assign sources, we follow the method outlined in Chardin et al. (2015). Using the halo catalogues generated from the simulation, we assign sources to all haloes with  $M_{\text{halo}} > 10^9 M_{\odot} h^{-1}$ . At each redshift, we assume a total volume emissivity (described in more detail below). This total emissivity is divided up among all the sources and weighted by the mass of the host halo, such that the ionizing emissivity of each source is directly proportional to the mass of the host halo (see also Iliev et al. 2006).

ATON solves the radiative transfer equation using a moment-based method with the M1 closure relation for the Eddington tensor (Levermore 1984; Gnedin & Abel 2001; Aubert & Teyssier 2008). ATON makes use of GPU acceleration, which significantly reduces the time required to perform a simulation and allows us to use the full speed of light in the simulations. We perform the radiative transfer using a single frequency band, with appropriate photon energies and photoionization cross-sections calculated using the optically thick grey approximation (see, e.g. Pawlik & Schaye 2011) and assume that the input spectrum is a blackbody. We present three different radiative transfer simulations here, described in Table 1. We present models with two different ionization histories (the low- $\tau_{\text{CMB}}$  and high- $\tau_{\text{CMB}}$  models) that use the same ionizing photon energy, and models with different ionizing photon energies but similar ionization histories (the low- $\tau_{\text{CMB}}$  and hot low- $\tau_{\text{CMB}}$  models). We use two different input blackbody temperatures:  $T = 30\,000$  (as in Keating et al. 2019) and  $40\,000 \text{ K}$  (as in Gaikwad et al. 2020; Puchwein et al., in preparation). For the  $T = 30\,000 \text{ K}$  blackbody, this results in a photon energy  $E_{\gamma} = 17.1 \text{ eV}$  and a photoionization cross-section  $\sigma_{\gamma} = 3.9 \times 10^{-18} \text{ cm}^2$ . For the  $T = 40\,000 \text{ K}$  blackbody, this results in a photon energy  $E_{\gamma} = 18.6 \text{ eV}$  and a photoionization cross-section  $\sigma_{\gamma} = 3.4 \times 10^{-18} \text{ cm}^2$ . As in Keating et al. (2019), we further account for a rising temperature at  $z < 5.2$  by increasing the photon energy linearly with the cosmic scale factor in the simulations, such that it reaches  $23.8 \text{ eV}$  by  $z = 4$ . This is intended to mimic the effect of the photoheating of the IGM during He II reionization and helps us to match the mean flux of the Ly  $\alpha$  forest approaching redshift 4.



**Figure 1.** The evolution of the mean flux with redshift measured in our radiative transfer simulations and compared with observations from Becker et al. (2015b) and Bosman et al. (2018).

We choose the ionizing emissivity in the simulations by aiming to match the observed mean flux of the Ly  $\alpha$  forest (Becker et al. 2015b; Bosman et al. 2018). We measure this mean flux from spectra constructed along sightlines taken from a lightcone extracted from the simulation on the fly, to allow us to capture the rapid evolution in the IGM towards the end of reionization. A comparison between the mean flux measured in our simulations and the observations is shown in Fig. 1. Due to the short time required to perform a simulation with ATON, we can run many models and modify the emissivity as required until we reach good agreement with the observations. The starting point for our low- $\tau_{\text{CMB}}$  and hot low- $\tau_{\text{CMB}}$  models is the input emissivity of the Puchwein et al. (2019) synthesis model of the UV background. As demonstrated in Kulkarni et al. (2019), this produces an ionization history that is in good agreement with the measurements of the CMB optical depth to Thomson scattering (Planck Collaboration VI 2018). We further present here a model that uses the emissivity model of the Haardt & Madau (2012) UV background as a starting point for our high- $\tau_{\text{CMB}}$  model, which results in more ionizing photons at high redshift and shifts the mid-point of reionization to a higher redshift (and hence increases the  $\tau_{\text{CMB}}$ ).

We present the resulting emissivity evolution for the three models in the top left-hand panel of Fig. 2. All models require a rather rapidly declining emissivity towards the end of reionization. Note that this should be considered as an ‘effective’ ionizing emissivity required to match the observed mean flux, as our simulations do not properly resolve the interstellar medium and escape of ionizing photons from the host galaxies. As shown here, using a different ionizing spectrum requires a small change in the shape of the emissivity evolution (due to the temperature dependence of the recombination rate, which changes the H I fraction in the ionized IGM and hence the mean flux). As in Keating et al. (2019), we find that matching the mean flux of the Ly  $\alpha$  forest at  $z \lesssim 5.5$  requires a emissivity that increases with decreasing redshift. This could perhaps be explained by an increasing contribution of active galactic nuclei (AGN) to the UVB towards lower redshift, raising the total emissivity of ionizing photons. This motivates our choice to increase the photon energy in the simulation

as the harder photons emitted from AGN will ionize He II and heat the gas. A more detailed account of the evolution relative contributions of galaxies and AGN with redshift would require more careful modelling of multiple populations of sources, and likely a larger volume simulation than presented here. For these reasons, and the reasons discussed above, we do not make any strong conclusions on the redshift evolution of our emissivity, and just note that using this redshift evolution results in synthetic spectra that match the observed mean flux quite well over the redshift range considered here.

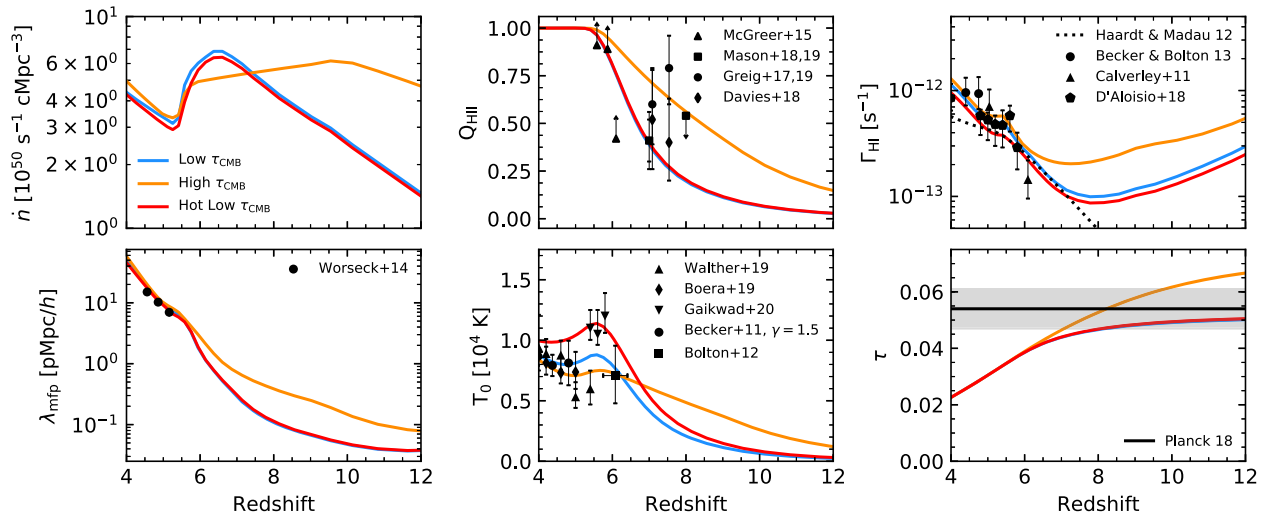
As well as matching the mean flux constraints, we also compare these models with a variety of measured properties of the IGM above redshift 4 (Fig. 2). The redshift evolution of the H I fraction in our models roughly spans the allowed range from the Ly  $\alpha$  absorption constraints (both from Ly  $\alpha$  forest statistics and Ly  $\alpha$  emission from galaxies). As expected, the low- $\tau_{\text{CMB}}$  and hot low- $\tau_{\text{CMB}}$  models come in at the lower end of this range, with a mid-point of reionization of  $z = 6.7$  in both models. These models also have optical depths to Thomson scattering in good agreement with the results from the *Planck* satellite. In contrast, the high- $\tau_{\text{CMB}}$  model has a higher mid-point of reionization ( $z = 8.4$ ) and it is disfavoured by the CMB results (Planck Collaboration VI 2018), although it is just about allowed by the Ly  $\alpha$  absorption/attenuation constraints and we therefore include it to explore the sensitivity of the Ly  $\alpha$  and Ly  $\beta$  forests to different reionization histories.

The temperature of the gas in these models depends both on the reionization history and the energy of ionizing photons we assume in the simulation. We show the evolution of the volume-weighted mean temperature at mean density in the bottom middle panel of Fig. 2. The models with a lower photon energy (low  $\tau_{\text{CMB}}$  and high  $\tau_{\text{CMB}}$ ) are in reasonable agreement with a range of temperature measurements of the high-redshift IGM, although they do fall below the recent measurements of the IGM temperature from the widths of transmission spikes in high-resolution quasar spectra (Gaikwad et al. 2020). These transmission spikes probe low-density regions, and seem to prefer higher temperatures than those inferred from analyses of the flux power spectrum (e.g. Boera et al. 2019; Walther et al. 2019). We remain agnostic as to the actual temperature of the IGM and simply also present a hotter model (hot low  $\tau_{\text{CMB}}$ ). We also note that the resolution of these simulations is insufficient to capture the widths of ionization fronts required to accurately model the IGM temperature (D’Aloisio et al. 2019), so the photon energies we require to match the observed temperature constraints are probably overestimates. We find that all three models are in good agreement with estimates of the H I photoionization rate and mean-free path at 912 Å.

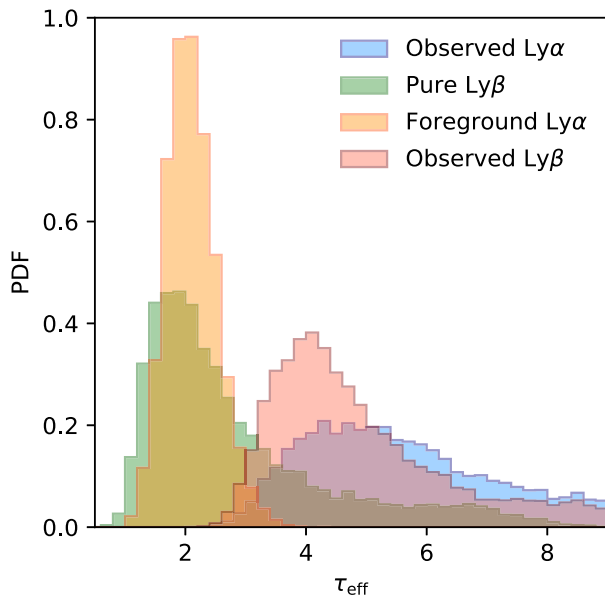
### 3 JOINT ANALYSIS OF THE LY $\alpha$ AND LY $\beta$ FORESTS IN A LATE-END REIONIZATION MODEL

We next examine the evolution of the Ly  $\alpha$  and Ly  $\beta$  forests in these simulations. We construct synthetic spectra along lines of sight taken through the lightcone we extracted from the simulations on the fly. The optical depth is calculated using the analytic approximation to a Voigt profile presented in Tepper-García (2006). Following Eilers et al. (2019), we measure the effective optical depth along skewers with length 40 Mpc. As our simulations contain significant amounts of neutral hydrogen in the redshift range where we are comparing the models against observations, we calculate the Ly  $\beta$  optical depth directly to properly account for the effect of the weaker damping wings (e.g. Malloy & Lidz 2015), rather than simply rescaling by the oscillator strengths and wavelengths as described in Section 1





**Figure 2.** Top left-hand panel: evolution of the volume ionizing emissivity ( $\dot{n}$ ) with redshift in the three simulations we present. Top middle panel: evolution of the ionized hydrogen fraction ( $Q_{\text{HII}}$ ) with redshift compared to constraints from the dark pixel fraction (McGreer, Mesinger & D’Odorico 2015), the fraction of continuum-selected galaxies showing Ly  $\alpha$  emission (Mason et al. 2018, 2019), and analyses of the strength of the damping wings observed in  $z > 7$  quasars (Greig et al. 2017; Davies et al. 2018; Greig, Mesinger & Bañados 2019). Top right-hand panel: evolution of the photoionization rate ( $\Gamma_{\text{HI}}$ ) with redshift in ionized regions in the simulations. Also shown is the Haardt & Madau (2012) model for the UV background used in the underlying hydrodynamic simulation. We compare with observational constraints from Becker & Bolton (2013), Calverley et al. (2011), and D’Aloisio et al. (2018). Bottom left-hand panel: evolution of the mean free path ( $\lambda_{\text{mfp}}$ ) at 912 Å with redshift in the simulations compared to the observations of Worseck et al. (2014). Bottom middle panel: evolution of the volume-weighted mean temperature at the mean density of the Universe ( $T_0$ ) with redshift in the simulations. Also shown are measurements from Walther et al. (2019), Boera et al. (2019), Gaikwad et al. (2020), Becker et al. (2011), and Bolton et al. (2012). Bottom right-hand panel: evolution of the optical depth to Thomson scattering of CMB photons ( $\tau_{\text{CMB}}$ ) with redshift. Also shown is the constraint from Planck Collaboration VI (2018).

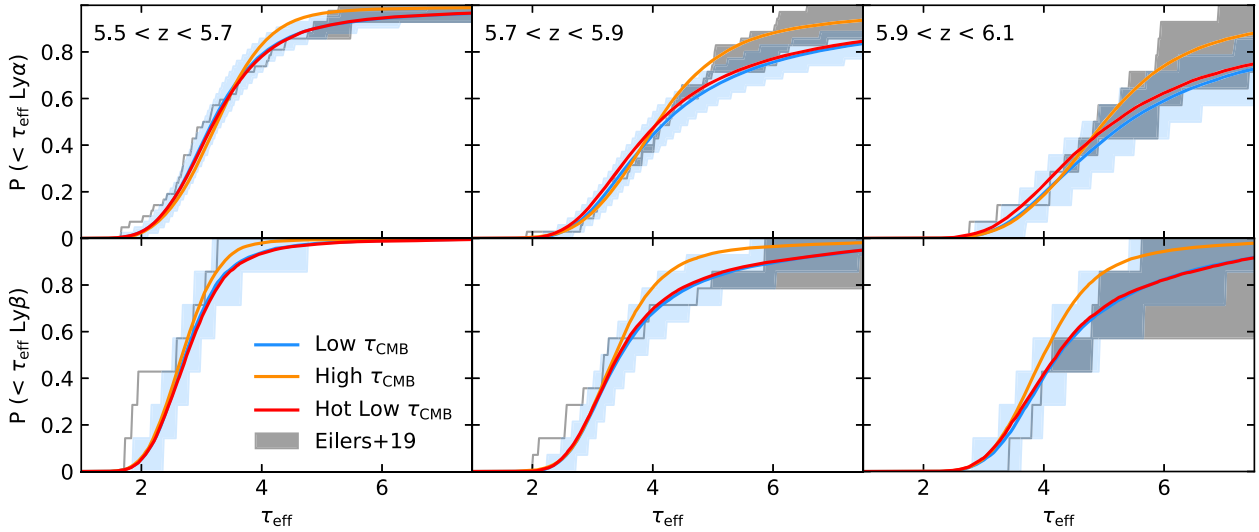


**Figure 3.** Distribution of effective optical depths in the low- $\tau_{\text{CMB}}$  model in the redshift range  $5.9 < z < 6.1$ . The foreground Ly  $\alpha$  comes from the redshift range  $4.8 \lesssim z \lesssim 5$ . We show the observed Ly  $\alpha$  effective optical depths (blue), the pure Ly  $\beta$  effective optical depths (green), the foreground Ly  $\alpha$  effective optical depths (orange), and the observed Ly  $\beta$  effective optical depths (red).

and Eilers et al. (2019). As discussed in Section 1, when comparing the Ly  $\beta$  forest in our simulations against the observations, we must also account for the contribution of the foreground Ly  $\alpha$  absorption. Following Eilers et al. (2019), we show how this changes the observed Ly  $\beta$  effective optical depth in Fig. 3 for the low- $\tau_{\text{CMB}}$  model (but the results are similar for our high- $\tau_{\text{CMB}}$  and hot low- $\tau_{\text{CMB}}$  models).

As expected, accounting for the foreground Ly  $\alpha$  absorption shifts the observed Ly  $\beta$  effective optical depths to higher values. We note that one difference between our simulations and the equivalent figure shown in Eilers et al. (2019) is the long tail towards high effective optical depths for the pure Ly  $\beta$  case (and hence also the observed Ly  $\beta$  case) that is present in our simulations. This is due to the late-end reionization models we present here, with the darkest sightlines in our models corresponding to regions that contain residual islands of neutral hydrogen below redshift 6.

We compare the cumulative distribution functions (CDFs) of Ly  $\alpha$  and Ly  $\beta$  effective optical depths calculated in our simulations to the Eilers et al. (2019) measurements, shown in the top panels of Fig. 4. Looking at the Ly  $\alpha$  distributions, the first thing to note is that although the simulations were calibrated to match the Bosman et al. (2018) measurements of the mean flux in this redshift range, the agreement with the Eilers et al. (2019) data is also quite good. This may suggest that different groups are now converging on a consensus value for the mean flux at high redshift. Secondly, as discussed in Eilers et al. (2019), accounting for noise in the mock observations can decrease the width of the distributions. We have not added noise to our spectra here, but instead show ‘optimistic’ and ‘pessimistic’ cases for the observations as in Nasir & D’Aloisio (2019). The optimistic case shows the case where upper limits are treated as detections. The pessimistic case instead takes the observed value for the mean flux (even if is below twice the estimate of the noise level). In cases where the observed flux is negative, we assume the sightlines are completely dark. We note that for most models, our distributions fall below the CDFs calculated from the upper limits, with long tails towards high effective optical depths. Therefore, adding the noise would bring the models into closer agreement with the CDFs constructed treating upper limits as detections. Comparing the differences between the Ly  $\alpha$  CDFs of the models, we find that although the three simulations we present have nearly identical mean flux, their effective optical



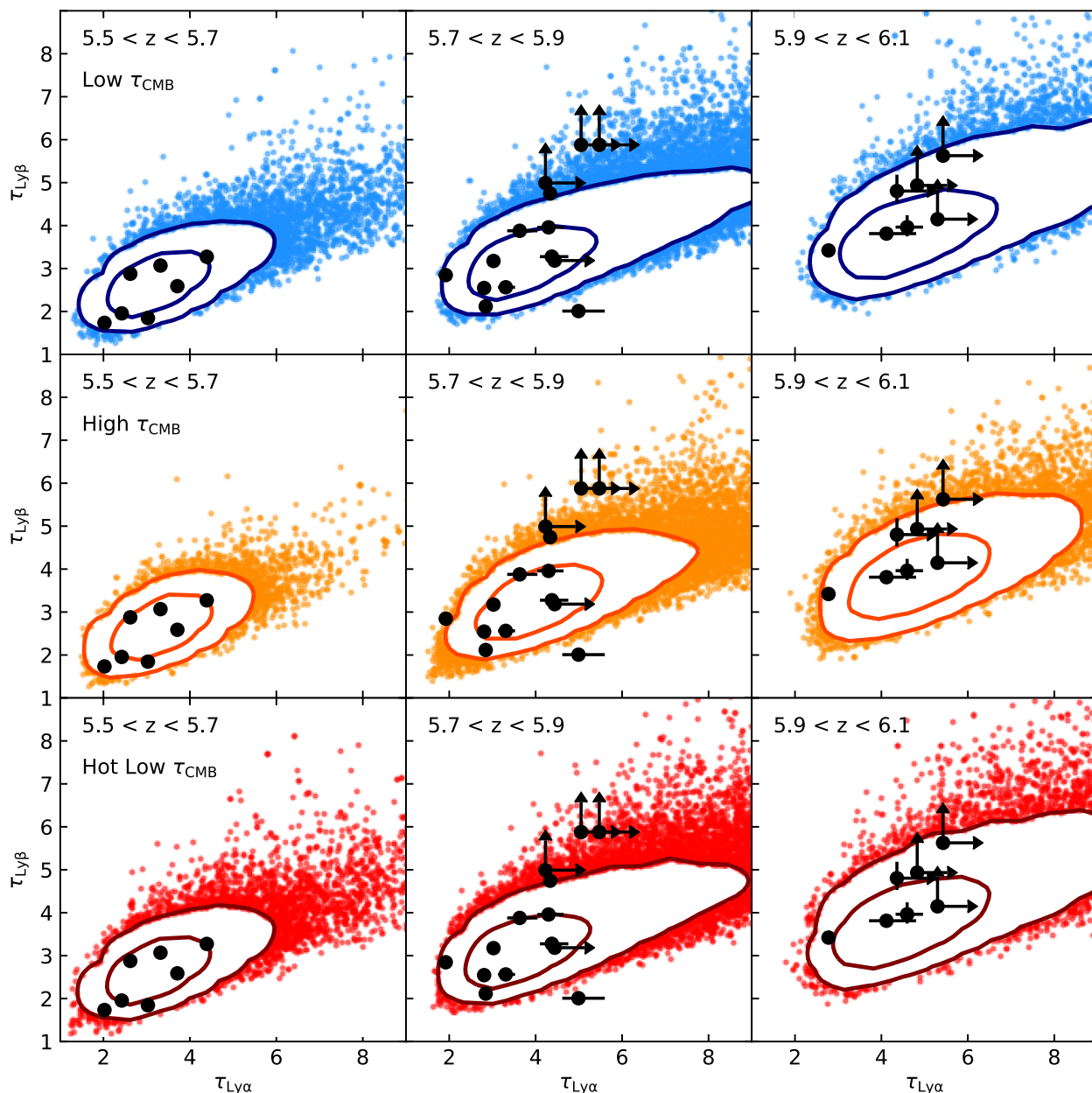
**Figure 4.** Cumulative distribution of the effective optical depth of the Ly  $\alpha$  forest (top panels) and Ly  $\beta$  forest (bottom panels) measured along skewers of length 40 Mpc. We show three different redshift intervals:  $5.5 < z < 5.7$  (left-hand panels),  $5.7 < z < 5.9$  (middle panels), and  $5.9 < z < 6.1$  (right-hand panels). The coloured lines show the median CDFs for each model (blue for the low- $\tau_{\text{CMB}}$  model, orange for the high- $\tau_{\text{CMB}}$  model, and red for the hot low- $\tau_{\text{CMB}}$  model). The blue shaded region accounts for the effect of cosmic variance in the low- $\tau_{\text{CMB}}$  model: The shaded region spans the 15th and 85th percentile range for 1000 CDFs calculated using the same number of sightlines as the Eilers et al. (2019) compilation. The strength of this effect is similar for all models. The grey shaded regions show the cumulative distributions calculated from the Eilers et al. (2019) sample. Upper limits are taken into account by showing both the case where the mean flux is equal to twice the noise (upper boundary), or assuming that the detected mean flux is correct (lower boundary). In cases where the observed mean flux is negative, the sightlines are assumed to have  $\tau_{\text{eff}} > 8$ .

depth distributions are somewhat different. The low- $\tau_{\text{CMB}}$  and hot low- $\tau_{\text{CMB}}$  models have CDFs that are slightly broader than the high- $\tau_{\text{CMB}}$  model, with tails of more opaque sightlines. This difference becomes more apparent with increasing redshift, and the implications of this will be discussed in Section 4. We also note that as the observations probe a limited number of sightlines, the effect of cosmic variance can play a role, and we estimate the magnitude of this effect for the low- $\tau_{\text{CMB}}$  model in the blue shaded region of Fig. 4.

We next compare our models to the CDFs of Ly  $\beta$  effective optical depth against the observations of Eilers et al. (2019), shown in the bottom panels of Fig. 4. We find that, in contrast to the models presented in Eilers et al. (2019), our models do a reasonable job of matching the observed CDFs. There are however some discrepancies, most notably the most transmissive sightlines at  $5.5 < z < 5.7$  which are not reproduced within the 15th/85th percentile range of our mock CDFs. However, we do find sightlines such as this if we extend this range out to the 5th/95th percentile range. Unlike Nasir & D’Aloisio (2019), who also investigated the evolution of the Ly  $\beta$  forest in a late-end reionization model, we do find that our models produce dark Ly  $\beta$  sightlines in the redshift range  $5.7 < z < 5.9$ . This is likely due to differences in modelling the ionization state of the gas introduced between the seminumeric method they employ and the radiative transfer simulations presented here. Furthermore, unlike Nasir & D’Aloisio (2019), we do not perform any rescaling of the photoionization rate to match the mean flux of the Ly  $\alpha$  forest. This means that the reionization history (and mean-free path) in our models is completely consistent with the mean flux of our synthetic spectra. The simulations presented here also resolve the ionization state of the gas at a factor of 10 higher spatial resolution than in Nasir & D’Aloisio (2019). We do not find any substantial differences between our three models, but again note that the high- $\tau_{\text{CMB}}$  model produces comparatively fewer dark sightlines than the low- $\tau_{\text{CMB}}$  and hot low- $\tau_{\text{CMB}}$  models towards high redshift.

We next investigate the redshift evolution of co-spatial measurements of the Ly  $\alpha$  and Ly  $\beta$  effective optical depths. The results for the three models in the three redshift bins are shown in Fig. 5. As in Fig. 4, we do not find large differences between our three models, although as discussed above, the low- $\tau_{\text{CMB}}$  and hot low- $\tau_{\text{CMB}}$  models predict a higher incidence rate of opaque sightlines. We compare with the data from Eilers et al. (2019). Again, we have not added noise to the mock spectra as this is already accounted for in the error bars and upper limits measured by Eilers et al. (2019). We also do not account for the uncertainties in fitting the continua of these spectra, as this was shown in Eilers et al. (2019) to only have a small effect on the scatter of the distribution of the Ly  $\alpha$  and Ly  $\beta$  effective optical depths. We plot contours enclosing 68 and 98 per cent of our sightlines by calculating a 2D kernel density estimate of the data using a Gaussian kernel with bandwidth equal to 0.2. Outside these contours, we plot the results from individual sightlines. We find that the results are generally in quite good agreement, with all the observations in the  $5.5 < z < 5.7$  and  $5.9 < z < 6.1$  bins lying within or very close to the 95 per cent confidence interval of our models.

As also pointed out in Nasir & D’Aloisio (2019), we find that the largest inconsistency between the models and the observations occurs in the  $5.7 < z < 5.9$  bin. We note however that the most extreme points, with  $\tau_{\text{eff, Ly } \beta} \sim 6$  in that redshift bin, are measurements from the sightline ULAS J0148+0600, which is the sightline that hosts the extreme  $110 \text{ Mpc } h^{-1}$  absorption trough identified in Becker et al. (2015b). We previously showed in Keating et al. (2019) that sightlines as extreme as this were uncommon in our simulation (perhaps due to the limited volume of the simulation box) and this has also been demonstrated in Nasir & D’Aloisio (2019). It is therefore unsurprising that these points lie outside of the 95 per cent confidence interval calculated for our models. We do note though that there are a handful of sightlines in all three models we present that have Ly  $\alpha$  and Ly  $\beta$  effective optical depths high enough to be consistent with the observations. We note that reproducing converged results



**Figure 5.** Relation between the  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  effective optical depths measured in our three models: low  $\tau_{\text{CMB}}$  (top panels), high  $\tau_{\text{CMB}}$  (middle panels), and hot low  $\tau_{\text{CMB}}$  (bottom panels). The contours show the region enclosing 68 and 95 per cent of the points. Outside these contours, the coloured points represent individual sightlines. The three columns show three different redshift intervals:  $6.1 < z < 6.3$  (left-hand panels),  $6.3 < z < 6.5$  (middle panels), and  $6.5 < z < 6.7$  (right-hand panels). The black points are taken from Eilers et al. (2019).

in the ratio of  $\text{Ly}\alpha$  to  $\text{Ly}\beta$  effective optical depths may require hydrodynamical simulations with higher mass resolution than the one we utilize here (see, e.g. the appendices of Becker et al. 2015b and Eilers et al. 2019). The simulation presented here represents our best compromise between mass resolution and a volume large enough to capture the process of reionization.

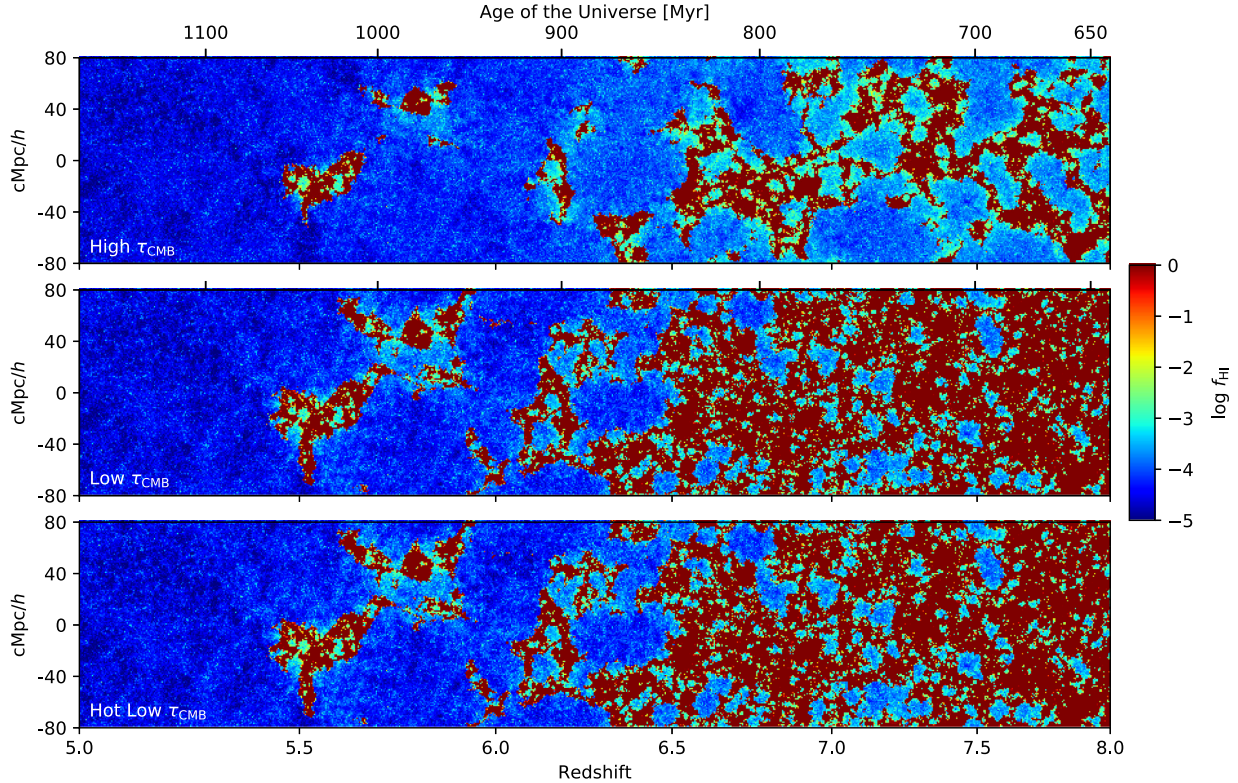
The late-end reionization models we present here are consistent with the observed effective optical depths of the  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  forests above  $z = 5.5$ . The current data do not however differentiate between models with different mid-points of reionization or gas temperatures, as all of our models show similar results. There does not appear to be a need to invoke an inverted temperature–density relation to explain the observations, as in Eilers et al. (2019). However, the most extreme

sightlines have a low incidence rate in the simulations presented here and depend on the details of the modelling (see Nasir & D’Aloisio 2019). Further theoretical work in larger simulation volumes, in tandem with a search for more extreme absorption troughs in high-quality spectra above  $z = 5.5$ , is therefore required to understand whether the incidence rate of these sightlines may require an even later end to reionization.

#### 4 PREDICTIONS FOR FUTURE OBSERVATIONS

The number of high-redshift quasars identified in large surveys continues to grow, with three published quasars with redshifts above 7 to date (Mortlock et al. 2011; Bañados et al. 2018; Wang et al.





**Figure 6.** Lightcones showing the evolution of the HI fraction with redshift for the three models presented here. The top panel shows the high- $\tau_{\text{CMB}}$  model, the middle panel shows the low- $\tau_{\text{CMB}}$  model, and the bottom panel shows the hot low- $\tau_{\text{CMB}}$  model.

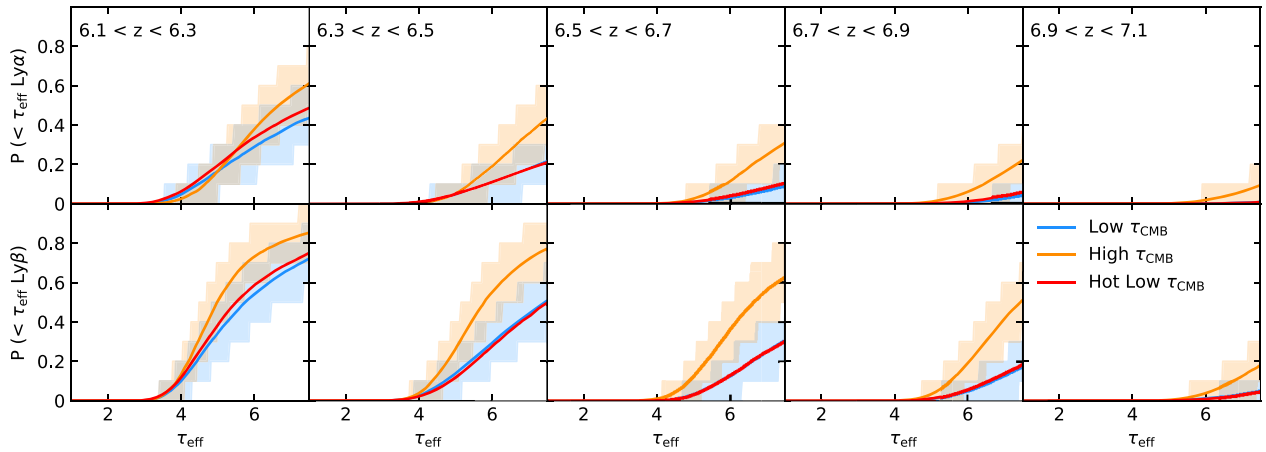
2018) and many more with redshifts above 6.5. Future surveys such as *Euclid* will further increase the number of known quasars above redshift 7 (Euclid Collaboration et al. 2019). It is therefore relevant to make predictions for the effective optical depth distributions beyond redshift 6 that may be measured from future observations of such quasars. As discussed in Section 2, we have modelled here two extremes in reionization history, and emphasize again that the high- $\tau_{\text{CMB}}$  model is already disfavoured by the CMB. It is still interesting however to see what constraints can be obtained from future analyses of absorption-line spectra. We demonstrate again the differences between the reionization histories of our three models in Fig. 6, where we plot the evolution in HI fraction in redshift along a lightcone through the simulation volume for each model. The ionization state of the gas is relatively similar below redshift 6 in all models, where they are constrained by the Ly $\alpha$  forest measurements. At higher redshifts, however, the contrast between both of the low- $\tau_{\text{CMB}}$  and the high- $\tau_{\text{CMB}}$  models is very striking as the differences between the mid-point of reionization in the two models become more apparent. The sizes of ionized bubbles in the high- $\tau_{\text{CMB}}$  model are much larger than in both of the low- $\tau_{\text{CMB}}$  models, due to the higher ionizing emissivity assigned to each halo at early times (top left-hand panel of Fig. 2).

Although there are clearly large reservoirs of ionized gas out to redshift 6.5 in all models, the HI fraction inside the ionized bubbles increases with increasing redshift, as the photoionization rate in the bubbles is lower before they percolate. This means that the incidence of opaque sightlines will increase in all models. Pushing Ly $\alpha$  forest measurements to higher redshifts with current facilities will therefore be challenging, as the observations will become dominated by the noise of the sky background. We nevertheless show the evolution of the Ly $\alpha$  optical depths out to redshift 7.1 in the three models

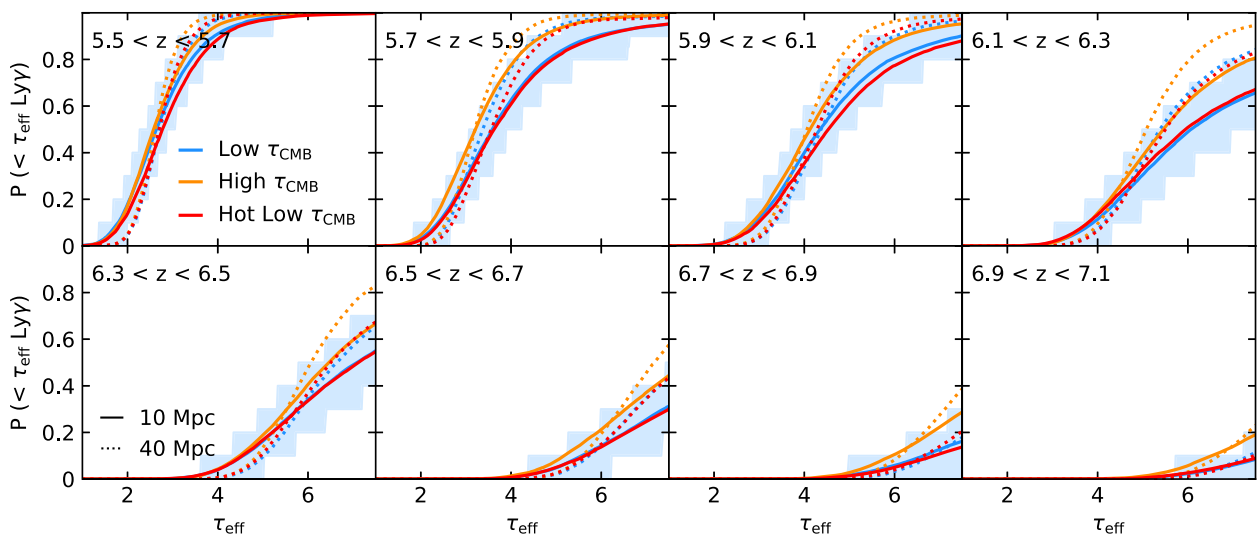
in the top panel of Fig. 7. The evolution of the CDFs for the low- $\tau_{\text{CMB}}$  and hot low- $\tau_{\text{CMB}}$  models are very similar, but the difference between those CDFs and the CDF of the high- $\tau_{\text{CMB}}$  model grows with increasing redshift. Indeed, the high- $\tau_{\text{CMB}}$  model predicts that there should be more than a factor of 2 more sightlines with  $\tau_{\text{eff, Ly}\alpha} \gtrsim 8$  at  $z = 6.7$ . The difference between the models however is close to the scatter that is expected from cosmic variance in measurements from a limited number of sightlines (the 15th/85th percentiles of 1000 CDFs constructed using 10 sightlines is shown for the low- $\tau_{\text{CMB}}$  model in the blue shaded region, and for the high- $\tau_{\text{CMB}}$  model in the yellow shaded region). The situation will be further complicated by the fact that any measurements of such high effective optical depths will likely be upper limits only.

A more promising avenue is to measure the evolution of the effective optical depths of the Ly $\beta$  forest, which we show in the bottom panel of Fig. 7. Again, the difference between the low- $\tau_{\text{CMB}}$  and high- $\tau_{\text{CMB}}$  models increases with increasing redshift. In this case, above  $z = 6.3$ , the difference between the low- $\tau_{\text{CMB}}$  and high- $\tau_{\text{CMB}}$  models is noticeably larger than the 15th/85th percentile scatter predicted from cosmic variance (compare blue and yellow shaded regions). We also find that there is already a significant difference in the models below  $\tau_{\text{eff, Ly}\beta} \sim 5$ , which is perhaps detectable in high signal-to-noise spectra. We therefore suggest that observations of the Ly $\beta$  forest out to high redshift can be used to constrain the evolution of the ionized fraction of the IGM out to  $z \sim 7$ , in the same way that the Ly $\alpha$  forest has been used to fix the ionization history below  $z \sim 6$ . If the mid-point of reionization is close to  $z \sim 7$ , as suggested by the CMB and estimates from quasar damping wings (Davies et al. 2018), then measurements of the effective optical depth of the Ly $\beta$  forest out past  $z = 6.5$  would allow us to constrain the entire second half of reionization from the Ly $\alpha$  and Ly $\beta$  forests alone.





**Figure 7.** Cumulative distribution of the effective optical depth of the  $\text{Ly}\alpha$  forest (top panels) and  $\text{Ly}\beta$  forest (bottom panels) measured along skewers of length 40 Mpc. We show five different redshift intervals. From the left- to right-hand side:  $6.1 < z < 6.3$ ,  $6.3 < z < 6.5$ ,  $6.5 < z < 6.7$ ,  $6.7 < z < 6.9$ , and  $6.9 < z < 7.1$ . The coloured lines show the median CDFs for each model (blue for the low- $\tau_{\text{CMB}}$  model, orange for the high- $\tau_{\text{CMB}}$  model, and red for the hot low- $\tau_{\text{CMB}}$  model). The blue and yellow shaded regions account for the effect of cosmic variance in the low- $\tau_{\text{CMB}}$  and high- $\tau_{\text{CMB}}$  models, respectively: The shaded region spans the 15th and 85th percentile range for 1000 CDFs calculated assuming a future compilation of 10 measurements in each bin. The strength of this effect is similar for the hot low- $\tau_{\text{CMB}}$  model.



**Figure 8.** Cumulative distribution of the effective optical depth of the  $\text{Ly}\gamma$  forest measured along skewers of length 10 (solid lines) and 40 Mpc (dotted lines). The pathlength accessible to observing the  $\text{Ly}\gamma$  forest is not long enough to measure along 40 Mpc skewers, but we show it here for comparison with the  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  CDFs presented elsewhere in this paper. The blue shaded region accounts for the effect of cosmic variance in the low- $\tau_{\text{CMB}}$  model: The shaded region spans the 15th and 85th percentile range for 1000 CDFs calculated assuming a future compilation of 10 measurements in each bin. The strength of this effect is similar for all models.

Further information could, in principle, be obtained from the evolution of the  $\text{Ly}\gamma$  effective optical depth, although now the foregrounds of both the lower redshift  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  absorption must be accounted for. The pathlength that is accessible by higher order Lyman series lines also becomes increasingly short for higher order lines, so measurements along sightlines towards quasars that have small proximity zones will be more favourable (e.g. Fan et al. 2006). We estimate that, for a given sightline, 10 comoving Mpc of the  $\text{Ly}\gamma$  forest could be observable (where we have assumed that  $5000 \text{ km s}^{-1}$  of the  $\text{Ly}\gamma$  forest at the high redshift end should be removed due to contribution from the proximity zone, and  $1000 \text{ km s}^{-1}$  should be removed at the low redshift end to avoid the  $\text{Ly}\delta$  emission peak). In Fig. 8, we present our prediction for the CDFs that would be observed.

Unlike the  $\text{Ly}\beta$  CDFs, we do not see a strong difference between the models with different mid-points of reionization. It seems that this is due to the contamination by the lower redshift  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  forests rather than the shorter pathlength we measure  $\text{Ly}\gamma$  along here. We demonstrate this by showing that we recover very similar results for the CDFs if we imagine we could measure the mean flux along 40 Mpc segments (as we have done elsewhere in this paper). As expected, the 40 Mpc CDFs are somewhat less broad (as the variations due to the density field or residual neutral islands are not felt so strongly), but again there is not a strong difference between the three models we present. We therefore conclude that the best chance of probing deeper into reionization with quasar absorption lines is through observations of the  $\text{Ly}\beta$  forest.

## 5 CONCLUSIONS

We have presented here a study of the evolution of the Ly  $\beta$  forest in cosmological radiative transfer simulations of models with reionization ending at  $z = 5.2\text{--}5.3$ . The models we test have different mid-points of reionization and gas temperatures. We find that all of our models agree reasonably well with the observations of Eilers et al. (2019) over the redshift range  $5.5 < z < 6.1$ , which corresponds to the tail-end of reionization in our models. The main discrepancy between the simulations and observations is in the incidence rate of the sightlines with the highest Ly  $\beta$  optical depths in the redshift bin  $5.7 < z < 5.9$ ; however, our simulations do reproduce several sightlines that are in agreement with observations. We further note that the sightline that these observations are measured along is the long absorption trough of ULAS J0148+0600, which we have already shown to be rare in radiative transfer simulations such as these.

We conclude that the late-end reionization scenario proposed in Kulkarni et al. (2019) to explain the properties of the high redshift Ly  $\alpha$  forest is also in good agreement with observations of the Ly  $\beta$  forest over the same redshift range. We therefore find there is no ‘anomaly’ in the opacity of the Ly  $\alpha$  and Ly  $\beta$  forests at high redshift as has previously been suggested, and that there is no need to invoke an inverted temperature–density relation in the IGM above  $z > 5.5$  to explain the observations. Good agreement between models with reionization ending late and the observational data seems however to require proper modelling of radiative transfer effects, as the seminumerical method of Nasir & D’Aloisio (2019) does not seem to produce as many sightlines that are opaque in Ly  $\beta$  in the redshift range  $5.7 < z < 5.9$  as in the radiative transfer simulations we present here.

We find that the current published effective optical depths do not yet discriminate between our models with different mid-points of reionization. We show that future measurements of the effective optical depth of the Ly  $\beta$  forest pushing out to higher redshifts will differentiate between reionization histories with redshifts of the mid-point of reionization differing by  $\Delta z > 1.5$ . There are already enough  $z > 6.5$  quasars known to measure the evolution of the Ly  $\beta$  forest out towards  $z = 7$ , so all that remains is to take spectra of high enough quality to provide reasonable constraints. An ongoing ESO large programme taking high signal-to-noise spectra of a sample of  $z > 5.8$  quasars with VLT/X-SHOOTER (PI: V. D’Odorico) is a very promising step in this direction.

## ACKNOWLEDGEMENTS

This work was performed using the Cambridge Service for Data Driven Discovery (CSD3), part of which is operated by the University of Cambridge Research Computing on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). The DiRAC component of CSD3 was funded by BEIS capital funding via STFC capital grants ST/P002307/1 and ST/R002452/1 and STFC operations grant ST/R00689X/1. This work further used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility. The equipment was funded by BEIS capital funding via STFC capital grants ST/P002293/1 and ST/R002371/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. This work was supported by the ERC Advanced Grant 320596 ‘The Emergence of Structure During the Epoch of Reionization’.

## DATA AVAILABILITY

The data underlying this paper are available at 10.5281/zenodo.3921822.

## REFERENCES

- Aubert D., Teyssier R., 2008, *MNRAS*, 387, 295  
 Aubert D., Teyssier R., 2010, *ApJ*, 724, 244  
 Bañados E. et al., 2018, *Nature*, 553, 473  
 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  
 Becker G. D., Bolton J. S., 2013, *MNRAS*, 436, 1023  
 Becker G. D., Bolton J. S., Haehnelt M. G., Sargent W. L. W., 2011, *MNRAS*, 410, 1096  
 Becker G. D., Bolton J. S., Lidz A., 2015a, *Publ. Astron. Soc. Aust.*, 32, e045  
 Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015b, *MNRAS*, 447, 3402  
 Becker G. D., Davies F. B., Furlanetto S. R., Malkan M. A., Boera E., Douglass C., 2018, *ApJ*, 863, 92  
 Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, *ApJ*, 872, 101  
 Bolton J. S., Becker G. D., Raskutti S., Wyithe J. S. B., Haehnelt M. G., Sargent W. L. W., 2012, *MNRAS*, 419, 2880  
 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., Fan X., Jiang L., Reed S., Matsuoka Y., Becker G., Haehnelt M., 2018, *MNRAS*, 479, 1055  
 Calverley A. P., Becker G. D., Haehnelt M. G., Bolton J. S., 2011, *MNRAS*, 412, 2543  
 Chardin J., Haehnelt M. G., Aubert D., Puchwein E., 2015, *MNRAS*, 453, 2943  
 D’Aloisio A., McQuinn M., Trac H., 2015, *ApJ*, 813, L38  
 D’Aloisio A., McQuinn M., Davies F. B., Furlanetto S. R., 2018, *MNRAS*, 473, 560  
 D’Aloisio A., McQuinn M., Maupin O., Davies F. B., Trac H., Fuller S., Upton Sanderbeck P. R., 2019, *ApJ*, 874, 154  
 Davies F. B., Furlanetto S. R., 2016, *MNRAS*, 460, 1328  
 Davies F. B. et al., 2018, *ApJ*, 864, 142  
 Dijkstra M., Lidz A., Hui L., 2004, *ApJ*, 605, 7  
 Eilers A.-C., Davies F. B., Hennawi J. F., 2018, *ApJ*, 864, 53  
 Eilers A.-C., Hennawi J. F., Davies F. B., Oñorbe J., 2019, *ApJ*, 881, 23  
 Euclid Collaboration et al., 2019, *A&A*, 631, A85  
 Fan X. et al., 2001, *AJ*, 122, 2833  
 Fan X. et al., 2006, *AJ*, 132, 117  
 Furlanetto S. R., Oh S. P., 2009, *ApJ*, 701, 94  
 Gaikwad P. et al., 2020, *MNRAS*, 494, 5091  
 Gnedin N. Y., Abel T., 2001, *New A*, 6, 437  
 Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, *MNRAS*, 466, 4239  
 Greig B., Mesinger A., Bañados E., 2019, *MNRAS*, 484, 5094  
 Haardt F., Madau P., 2012, *ApJ*, 746, 125  
 Hui L., Gnedin N. Y., 1997, *MNRAS*, 292, 27  
 Iliiev I. T., Mellema G., Pen U. L., Merz H., Shapiro P. R., Alvarez M. A., 2006, *MNRAS*, 369, 1625  
 Keating L. C., Puchwein E., Haehnelt M. G., 2018, *MNRAS*, 477, 5501  
 Keating L. C., Weinberger L. H., Kulkarni G., Haehnelt M. G., Chardin J., Aubert D., 2019, *MNRAS*, 491, 1736  
 Kulkarni G., Keating L. C., Haehnelt M. G., Bosman S. E. I., Puchwein E., Chardin J., Aubert D., 2019, *MNRAS*, 485, L24  
 Levermore C. D., 1984, *J. Quant. Spec. Radiat. Transf.*, 31, 149  
 Lidz A., Hui L., Zaldarriaga M., Scoccimarro R., 2002, *ApJ*, 579, 491  
 Lidz A., McQuinn M., Zaldarriaga M., Hernquist L., Dutta S., 2007, *ApJ*, 670, 39  
 McGreer I. D., Mesinger A., D’Odorico V., 2015, *MNRAS*, 447, 499  
 Malloy M., Lidz A., 2015, *ApJ*, 799, 179  
 Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, *ApJ*, 856, 2  
 Mason C. A. et al., 2019, *MNRAS*, 485, 3947  
 Mesinger A., 2010, *MNRAS*, 407, 1328

- Mortlock D. J. et al., 2011, *Nature*, 474, 616  
 Nasir F., D’Aloisio A., 2019, *MNRAS*, 494, 3080  
 Oh S. P., Furlanetto S. R., 2005, *ApJ*, 620, L9  
 Pawlik A. H., Schaye J., 2011, *MNRAS*, 412, 1943  
 Planck Collaboration et al., 2014, *A&A*, 571, A16  
 Planck Collaboration et al., 2018, preprint (arXiv:1807.06209)  
 Puchwein E., Haardt F., Haehnelt M. G., Madau P., 2019, *MNRAS*, 485, 47  
 Songaila A., 2004, *AJ*, 127, 2598  
 Springel V., 2005, *MNRAS*, 364, 1105  
 Tepper-García T., 2006, *MNRAS*, 369, 2025  
 Viel M., Haehnelt M. G., Springel V., 2004, *MNRAS*, 354, 684  
 Walther M., Oñorbe J., Hennawi J. F., Lukić Z., 2019, *ApJ*, 872, 13  
 Wang F. et al., 2018, *ApJ*, 869, L9  
 Worseck G. et al., 2014, *MNRAS*, 445, 1745

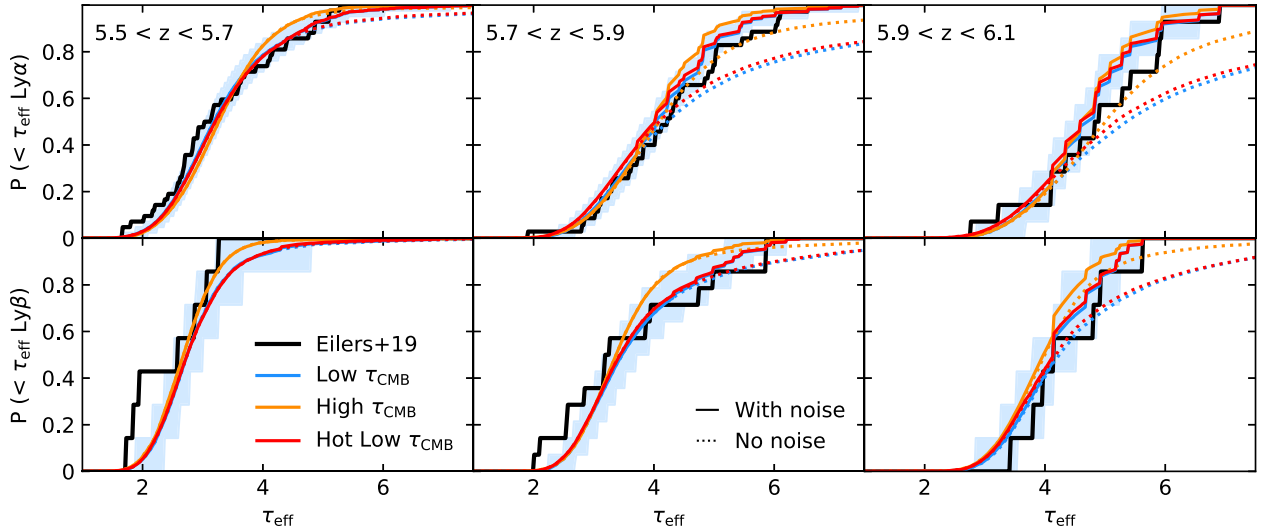
## APPENDIX A: FORWARD MODELLING OF SPECTRAL NOISE

Throughout this paper, we have determined whether the opacity distribution of our simulated spectra are compatible with observations by seeing if they lie within the ‘optimistic’ and ‘pessimistic’ predictions of the data as in Bosman et al. (2018). In this scenario, if the mean flux of a sightline is below the noise, then its mean flux

is assumed to be at most twice the noise. The other extreme is to assume that there is no transmitted flux at all. Rather than comparing the simulations against these limits, one can also forward model the spectral noise as in Eilers et al. (2019). In this appendix, we compare our simulations against observations using the forward modelling technique.

We construct 1000 mock surveys, drawing random sightlines from our simulations and adding noise appropriate to the uncertainties of the Eilers et al. (2019) observations. As with the observed data, we note the simulated sightlines that have a mean flux  $\langle F \rangle < 2\sigma_{\text{obs}}$ , where  $\sigma_{\text{obs}}$  is the uncertainty in the mean flux along a given sightline. In these cases, we assume  $\langle F \rangle = 2\sigma_{\text{obs}}$  for this sightline. Otherwise, we take the mean flux calculated for that sightline.

Our results are shown in Fig. A1. As in Eilers et al. (2019), we find that forward modelling the noise in the simulated data narrows the opacity CDFs. However, we find consistent results with the analysis presented in Section 3. This is in contrast to Eilers et al. (2019), who found that accounting for this spectral noise put their models in tension with the observations. We attribute the difference here to the fact that the opacity CDFs produced in these late reionization models are significantly broader than the UV and temperature fluctuation models before any noise is added.



**Figure A1.** Cumulative distribution of the effective optical depth of the  $\text{Ly}\alpha$  forest (top panels) and  $\text{Ly}\beta$  forest (bottom panels) measured along skewers of length 40 Mpc. We show three different redshift intervals:  $5.5 < z < 5.7$  (left-hand panels),  $5.7 < z < 5.9$  (middle panels), and  $5.9 < z < 6.1$  (right-hand panels). The coloured lines show the median CDFs for each model (blue for low  $\tau_{\text{CMB}}$  model, orange for high  $\tau_{\text{CMB}}$ , and red for hot low  $\tau_{\text{CMB}}$ ). The blue shaded region accounts for the effect of cosmic variance in the low- $\tau_{\text{CMB}}$  model: The shaded region spans the 15th and 85th percentile range for 1000 CDFs calculated using the same number of sightlines as the Eilers et al. (2019) compilation. The strength of this effect is similar for all models. The solid lines are CDFs constructed using sightlines with forward modelled spectral noise. The dashed lines show the case with no noise. The black lines show the cumulative distributions calculated from the Eilers et al. (2019) sample, counting the lower limits on  $\tau_{\text{eff}}$  as detections.

This paper has been typeset from a  $\text{T}_{\text{E}}\text{X}/\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  file prepared by the author.