Becker, George D. and Bolton, James S. and Madau, Piero and Pettini, Max and Ryan-Weber, Emma V. and Venemans, Bram P. (2015) Evidence of patchy hydrogen reionization from an extreme Lyα trough below redshift six. Monthly Notices of the Royal Astronomical Society, 447 (4). pp. 3402-3419. ISSN 1365-2966

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **447,** 3402–3419 (2015) doi:10.1093/mnras/stu2646

Evidence of patchy hydrogen reionization from an extreme Lyα trough below redshift six

George D. Becker,[1]* James S. Bolton,[2] Piero Madau,[3] Max Pettini,[1]
Emma V. Ryan-Weber[4] and Bram P. Venemans[5]

[1] *Kavli Institute for Cosmology and Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[2] *School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, UK*
[3] *Department of Astronomy and Astrophysics, University of California, 1156 High Street, Santa Cruz, CA 95064, USA*
[4] *Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia*
[5] *Max-Planck Institute for Astronomy, Knigstuhl 17, D-69117 Heidelberg, Germany*

Accepted 2014 December 12. Received 2014 December 12; in original form 2014 July 17

ABSTRACT

We report the discovery of an extremely long ($\sim$110 Mpc h^{-1}) and dark ($\tau_{\rm eff} \gtrsim 7$) Lyα trough extending down to $z \simeq 5.5$ towards the $z_{\rm em} \simeq 6.0$ quasar ULAS J0148+0600. We use these new data in combination with Lyα forest measurements from 42 quasars at $4.5 \le z_{\rm em} \le 6.4$ to conduct an updated analysis of the line-of-sight variance in the intergalactic Lyα opacity over $4 \le z \le 6$. We find that the scatter in transmission among lines of sight near $z \sim 6$ significantly exceeds theoretical expectations for either a uniform ultraviolet background (UVB) or simple fluctuating UVB models in which the mean free path to ionizing photons is spatially invariant. The data, particularly near $z \simeq 5.6$–5.8, instead require fluctuations in the volume-weighted hydrogen neutral fraction that are a factor of 3 or more beyond those expected from density variations alone. We argue that these fluctuations are most likely driven by large-scale variations in the mean free path, consistent with expectations for the final stages of inhomogeneous hydrogen reionization. Even by $z \simeq 5.6$, however, a large fraction of the data are consistent with a uniform UVB, and by $z \sim 5$ the data are fully consistent with opacity fluctuations arising solely from the density field. This suggests that while reionization may be ongoing at $z \sim 6$, it has fully completed by $z \sim 5$.

Key words: intergalactic medium – quasars: absorption lines – cosmology: observations – dark ages, reionization, first stars – large-scale structure of Universe.

1 INTRODUCTION

Determining how and when the intergalactic medium (IGM) became reionized is currently one of the key goals of extragalactic astronomy. Within roughly one billion years of the big bang, ultraviolet photons from the first luminous objects ionized nearly every atom in the IGM. The details of this process reflect the nature of the first stars, galaxies, and active galactic nuclei (AGN), as well as the characteristics of large-scale structure, and therefore continue to be the subject of considerable observational and theoretical effort.

Some of the most fundamental constraints on when reionization ended come from the evolution of intergalactic Lyα opacity near $z \sim 6$, as measured in the spectra of high-redshift quasars (e.g. Becker et al. 2001; Djorgovski et al. 2001; Fan et al. 2002, 2006; White et al. 2003; Songaila 2004) and gamma-ray bursts (e.g. Chornock et al. 2013, 2014). The largest data set to date was provided by Fan et al. (2006), who measured the opacity in the Lyα forest towards a sample of 19 $z \sim 6$ quasars. The fact that transmitted flux is observed in the Lyα forest up to $z \sim 6$ suggests that reionization had largely ended by that point, at least in a volume-averaged sense. Fan et al. noted a rapid increase in the mean Lyα opacity at $z > 5.7$, however, which suggests a decline in the intensity of the ultraviolet background (UVB) near 1 Ryd (see also Bolton & Haehnelt 2007; Calverley et al. 2011; Wyithe & Bolton 2011). They also noted a large sightline-to-sightline scatter (see also Songaila 2004), which they interpreted as evidence of large (factor of $\gtrsim$4) fluctuations in the UVB near $z \sim 6$. Further evidence for a decline in the UVB from $z \sim 5$ to 6 is also potentially seen in the changing ionization state of metal-enriched absorbers over this interval (Becker, Rauch & Sargent 2009; Ryan-Weber et al. 2009; Becker et al. 2011b; Simcoe et al. 2011; D'Odorico et al. 2013; Keating et al. 2014).

The inferred rapid evolution in the UVB over $5 \lesssim z \lesssim 6$ stands in stark contrast to its nearly constant value over $2 < z < 5$ (e.g. Bolton et al. 2005; Faucher-Giguère et al. 2008; Becker & Bolton 2013).

---

* E-mail: gdb@ast.cam.ac.uk

Published by Oxford University Press on behalf of the Royal Astronomical Society

It is unclear, however, whether a rapidly evolving UVB necessarily indicates a recent end to reionization. As pointed out by McQuinn et al. (2011), a modest increase in the global ionizing emissivity may produce a large increase in the mean free path to ionizing photons, leading to a strong increase in the UVB. Such an evolution may be driven by the increase in the star formation rate density from $z \sim 6$ to 5 (e.g. Bouwens et al. 2007) even if reionization ended significantly earlier.

On the other hand, Lidz et al. (2007) and subsequently Mesinger (2010) have pointed out that existing measurements of the intergalactic Lyα opacity do not firmly rule out the final stages of reionization occurring at $z \lesssim 6$. The spatially inhomogeneous nature of reionization and the limited number of quasar sightlines available at $z > 5$ may conspire together such that isolated, neutral patches in the IGM remain as yet undetected at $z \sim 5$–6.

In this context, the case for recent (or ongoing) reionization at $z \sim 6$ would be significantly clarified by determining whether the observed scatter in Lyα opacity at $z \sim 6$ is truly driven by fluctuations in the UVB, as proposed by Fan et al. (2006). The claim of large UVB fluctuations was queried by Lidz, Oh & Furlanetto (2006), who argued that significant sightline-to-sightline variations in opacity are expected due to large-scale density fluctuations alone. Lidz et al. used analytic and numerical arguments to demonstrate that the scatter should rise sharply as the mean opacity increases, leading to variations at $z \sim 6$ on ∼40–50 Mpc h^{-1} scales that are comparable to the Fan et al. (2006) measurements. If correct, this would significantly weaken the direct evidence that the evolution in the UVB near $z \sim 6$ is related to patchy reionization. Furthermore, Bolton & Haehnelt (2007) and Mesinger & Furlanetto (2009) demonstrated that even in the presence of a fluctuating UVB with a spatially invariant mean free path, the impact of the resulting ionization fluctuations on the effective optical depth is modest. The largest fluctuations in the UVB typically occur in overdense regions which are already optically thick to Lyα photons. Any observational evidence for scatter in the Lyα forest opacity in excess of that expected from density fluctuations or simple fluctuating UVB models alone would therefore be indicative of variations in the mean free path and spatial inhomogeneity in the IGM neutral fraction, which are potential hallmarks of reionization.

In this paper we provide a new analysis of the intergalactic Lyα opacity over $4 \lesssim z \lesssim 6$. Our work is largely motivated by deep Very Large Telescope (VLT)/X-Shooter observations of a single $z_{\rm em} \sim 6$ quasar, ULAS J0148+0600 ($z_{\rm em} = 5.98$), which was discovered in the UKIDSS Large Area Survey (Lawrence et al. 2007). As we demonstrate below, this object shows an extremely dark ($\tau_{\rm eff} \gtrsim 7$) and extended ($\Delta l \simeq 110$ Mpc h^{-1}) Lyα trough. Most remarkably, the trough extends down to $z \simeq 5.5$, where other lines of sight show high levels of transmitted flux. We add Lyα opacity measurements from ULAS J0148+0600 and six other $z_{\rm em} > 5.7$ quasars to the Fan et al. (2006) sample, along with 16 quasars over $4.5 \le z_{\rm em} \le 5.4$ observed at moderate-to-high resolution to provide a lower redshift sample for comparison. We compare measurements from this expanded sample to predictions from simple IGM Lyα transmission models based on numerical simulations to determine whether fluctuations in the UVB are present. Our hydrodynamical simulations include a suite of large boxes ($l_{\rm box} = 25$–100 Mpc h^{-1}) in order to allow us both to evaluate the expected scatter in Lyα opacity from large-scale structure alone, as well as couple simple fluctuating UVB models directly to the density field.

We introduce the new data in Section 2 and our numerical simulations in Section 3. In Section 4 we compare the Lyα opacity measurements to predictions for a uniform UVB and for simple UVB models that assume the ionizing opacity is characterized by a single mean free path. We argue that fluctuations in the mean free path must be present, and discuss the implications for the end of reionization in Section 5. Our results are summarized in Section 6. Convergence tests for our models are presented in Appendix . We quote comoving distances generally assuming ($\Omega_{\rm m}$, Ω_Λ, h) = (0.308, 0.692, 0.678), consistent with recent results from the *Planck* satellite (Planck Collaboration XVI 2014). Cosmological parameters are further discussed in Section 3.

2 DATA

2.1 Quasar spectra

This paper builds upon the sample of 19 $z_{\rm em} > 5.7$ quasars analysed by Fan et al. (2006) in two respects. First, we add a further seven objects at $z_{\rm em} > 5.8$. Notably, this new sample includes the $z_{\rm em} = 5.98$ quasar ULAS J0148+0600,[1] whose Lyα trough is the primary motivation for this work. Spectra for these objects are presented in Fig. 1. We also add 16 quasars spanning $4.5 \le z_{\rm em} \le 5.4$, primarily to provide a lower redshift baseline for evaluating the evolution of the Lyα forest at $z > 5$. Similar data at these redshifts were obtained by Songaila (2004). The present sample allows us to evaluate the evolution in Lyα opacity, including its scatter between lines of sight, in a self-consistent manner over the entire redshift range $3.8 < z < 6.3$. All spectra in this study were obtained at moderate or high spectral resolution with Keck/High Resolution Echelle Spectrometer (HIRES), Keck/Echellette Spectrograph and Imager (ESI), Magellan/Magellan Inamori Kyocera Echelle (MIKE), or VLT/X-Shooter, and thus are suited to the same type of analysis applied by Fan et al. (2006) to their Keck/ESI data. A summary of the spectra is presented in Table 1.

Reduction and continuum fitting procedures for all but two of our objects have been presented elsewhere (see Table 1). For the $z_{\rm em} > 5.7$ objects, the continuum over the Lyα forest was generally estimated using a power law normalized in regions relatively free of emission lines over ∼1285–1350 Å in the rest frame, and out to ∼1450 Å when possible. A low-order spline fit was generally used for lower redshift quasars, although the spline was typically placed near the power-law estimate. Uncertainties in the Lyα opacity measurements related to continuum fitting are discussed below.

New observations for ULAS J0148+0600 and ULAS J1319+0959 (Mortlock et al. 2009) were obtained with the X-Shooter spectrograph on the VLT (D'Odorico et al. 2006). Each object was observed for 10 h using 0.7 and 0.6 arcsec slits in the visible (VIS) and near-infrared (NIR) arms, respectively. The spectra were flat-fielded, sky-subtracted using the method described by Kelson (2003), optimally extracted (Horne 1986) using 10 km s^{-1} bins, and corrected for telluric absorption using a suite of custom routines (see Becker et al. 2012 for more details). These data will be described more fully in an upcoming work (Codoreanu et al., in preparation). For ULAS J0148+0600 we adopt a redshift of $z_{\rm em} = 5.98 \pm 0.01$ based on the peak of the Mg II emission line. For ULAS J1319+0959 we adopt $z_{\rm em} = 6.133$ based on [C II] 158 μm measurements from Wang et al. (2013).

As discussed below, ULAS J0148+0600 displays an extremely dark absorption trough in the Lyα forest. Since estimates of the mean opacity in such regions are sensitive to flux zero-point

---

[1] For coordinates and magnitudes, see Bañados et al. (2014).

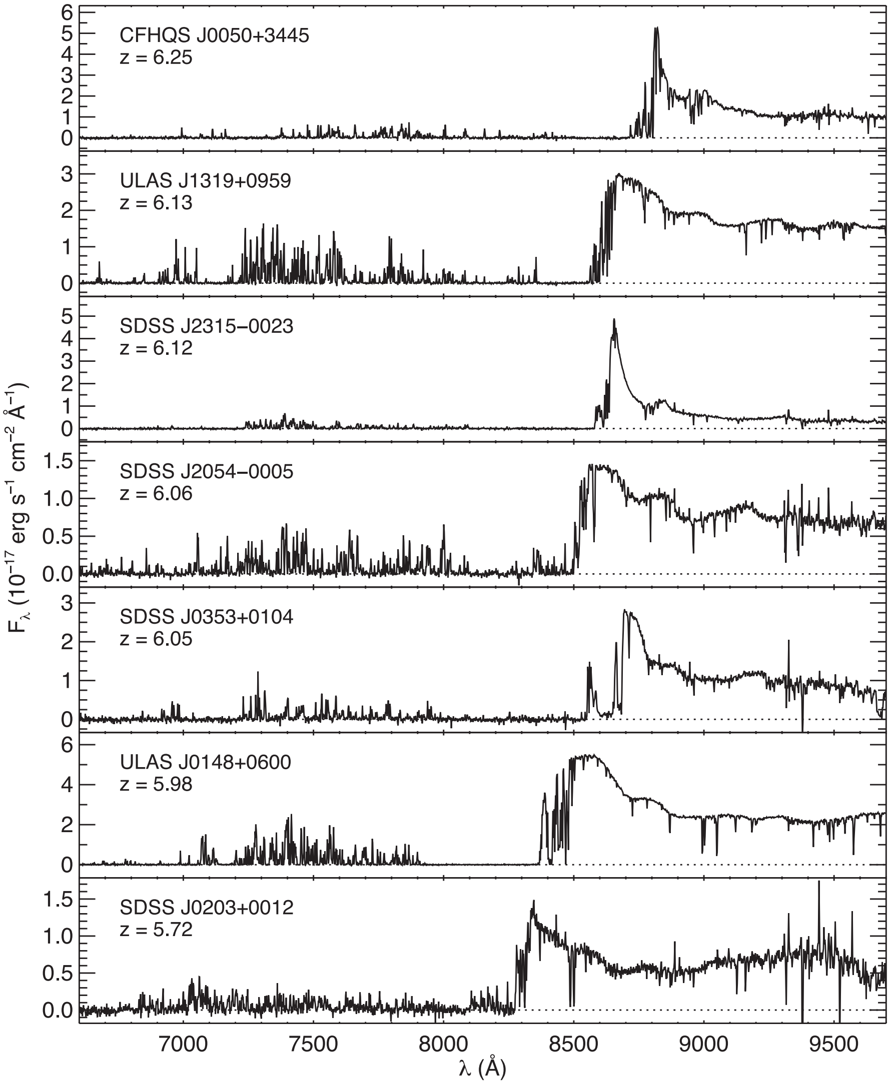

Figure 1. Spectra of $z \sim 6$ quasars analysed in this work that are in addition to those in the Fan et al. (2006) sample. ULAS J1319+0950 and ULAS J0148+0600 were observed with VLT/X-Shooter, while the remainder were observed with Keck/ESI (see Table 1). Approximate fluxing is based on published z'-band magnitudes. The spectra have been binned for display. Note that the Lyα forest flux for SDSS J2315−0023 appears depressed because the y-axis has been scaled to accommodate the strong Lyα emission line.

uncertainties, we adopted a reduction strategy intended to minimize such errors. Individual exposures were combined using an inverse variance weighting scheme, where the variance in each two-dimensional reduced frame was estimated from the measured scatter about the sky model in regions not covered by the object trace, rather than derived formally from the sky model and detector characteristics. This avoids biases when combining multiple exposures due to random errors in the sky estimate, which can be problematic when the sky background is relatively low. We checked our combined one-dimensional X-shooter spectra for evidence of zero-point errors blueward of the quasar's Lyman limit, where there should be no flux, and found the errors to be negligible.

2.2 Lyα opacity measurements

Following Fan et al. (2006), we measure the mean opacity of the IGM to Lyα in discreet regions along the lines of sight towards individual objects. We quantify the opacity in terms of an effective optical depth, which is conventionally defined as $\tau_{\rm eff} = -\log \langle F \rangle$, where F is the continuum-normalized flux. Since our sample spans a broad redshift range, we measure $\tau_{\rm eff}$ in bins of fixed comoving length (50 Mpc h^{-1}), rather than fixed redshift intervals. This length scale, however, roughly matches the $\Delta z = 0.15$ bins used by Fan et al. (2006) over $z \sim 5$–6.

Our Lyα flux measurements for all 23 objects are given in Table 2. Error estimates do not include continuum errors, which are instead incorporated into the modelling (see Section 4). In order to avoid contamination from the quasar proximity region or from associated Lyβ or O VI absorption, we generally restrict our measurements to the region between rest-frame wavelengths 1041 and 1176 Å. This also minimizes uncertainties in the continuum related to the blue wing of the Lyα emission line. For four of the six $z_{\rm em} > 5.9$ objects, however, we choose the maximum wavelength to lie just blueward of the apparent enhanced transmission in the proximity zone, as done by Fan et al. (2006). Exceptions to this are SDSS J0353+0104, which is a broad absorption line (BAL), and SDSS J2054−0005, for which edge of the region of enhanced flux is unclear. In these cases we use a maximum rest-frame wavelength of 1176 Å.

Table 1. List of QSOs analysed in this work that are in addition to those in the Fan et al. (2006) sample.

QSO	$z_{\rm em}$	Instrument	Ref.[a]
CFHQS J0050+3445	6.25	ESI	5
ULAS J1319+0950	6.13	X-Shooter	7
SDSS J2315−0023	6.12	ESI	5
SDSS J2054−0005	6.06	ESI	5
SDSS J0353+0104	6.05	ESI	5
ULAS J0148+0600	5.98	X-Shooter	7
SDSS J0203+0012	5.72	ESI	5
SDSS J0231−0728	5.42	X-Shooter	6
SDSS J1659+2709	5.32	HIRES	3
SDSS J1208+0010	5.27	X-Shooter	6
SDSS J0915+4244	5.20	HIRES	2
SDSS J1204−0021	5.09	HIRES	2
SDSS J0040−0915	4.98	MIKE	3
SDSS J0011+1446	4.95	HIRES	3
SDSS J2225−0014	4.89	MIKE	4
SDSS J1616+0501	4.88	MIKE	4
BR 1202−0725	4.70	HIRES	1
SDSS J2147−0838	4.60	MIKE	3
BR 0353−3820	4.59	MIKE	3
BR 1033−0327	4.52	MIKE	7
BR 0006−6208	4.52	MIKE	7
BR 0714−6449	4.49	MIKE	3
BR 0418−5723	4.48	MIKE	3

[a] Spectrum references: 1 – Lu et al. (1996); 2 – Becker et al. (2006); 3 – Becker et al. (2011a); 4 – Calverley et al. (2011); 5 – Becker et al. (2011b); 6 – Becker et al. (2012); 7 – this work.

Where no transmitted flux is formally detected, we adopt a lower limit on $\tau_{\rm eff}$ assuming a mean transmitted flux equal to twice the formal uncertainty. In these cases, we also searched the spectra for individual transmission peaks whose flux may have been smaller than the formal uncertainty for the total 50 Mpc h^{-1} region. A peak was considered significant if it had at least four adjacent pixels that exceeded the 1σ error estimate, and if the combined significance of the flux in these pixels was $\geq 5\sigma$. The identified peaks are shown in Fig. 2. In regions where one or more peaks were detected, we adopt an upper limit on $\tau_{\rm eff}$ assuming that the total flux in that 50 Mpc h^{-1} region is equal to the 2σ lower limit on the flux in those peaks alone (Table 3).

Following Fan et al. (2006), we made no attempt to correct for contamination from intervening metal lines or damped Lyα systems (DLAs). Metal lines, at least at $z < 5$, generally account for only a few per cent of the absorption in the Lyα forest (Schaye et al. 2003; Kirkman et al. 2005; Kim et al. 2007; Becker et al. 2011a), and so are not expected to strongly affect $\tau_{\rm eff}$ measurements at $z \gtrsim 4$ or add significantly to their scatter on 50 Mpc h^{-1} scales. DLAs are potentially more problematic, as a single system will increase $\tau_{\rm eff}$ in a 50 Mpc h^{-1} region by ∼0.4. They become increasingly difficult to identify at $z > 5$, however, where the high levels of absorption in the Lyα forest mean that DLAs must often be identified via their associated metal lines, coverage of which varies between lines of sight and is often incomplete. We tested the impact of DLAs on $\tau_{\rm eff}$ at $z < 5$ by repeating our measurements after masking DLAs visible in the Lyα forest. This naturally lowered $\tau_{\rm eff}$ in some regions, although the difference was not large enough to affect the interpretation of the data presented below. We detect no metal lines in the spectrum of ULAS J0148+0600 over the redshift range spanned by its Lyα trough. A more detailed inventory of metals along this line of sight will be presented by Codoreanu et al.

Table 2. Formal continuum-normalized mean Lyα transmitted flux measurements in 50 Mpc h^{-1} regions. The mean redshift in each section is given by $z_{\rm abs}$.

QSO	$z_{\rm em}$	$z_{\rm abs}$	$\langle F/F_{\rm C}\rangle$
CFHQS J0050+3445	6.25	6.074	-0.00761 ± 0.00196
		5.902	0.01149 ± 0.00263
		5.737	0.01578 ± 0.00209
		5.577	0.03836 ± 0.00187
		5.423	0.10284 ± 0.00178
ULAS 1319+0950	6.13	5.948	0.00820 ± 0.00075
		5.781	0.01404 ± 0.00056
		5.620	0.02544 ± 0.00062
		5.464	0.07504 ± 0.00067
SDSS J2315−0023	6.12	5.965	-0.00913 ± 0.00273
		5.797	-0.00094 ± 0.00307
		5.635	0.03103 ± 0.00197
		5.479	0.02652 ± 0.00206
		5.328	0.05167 ± 0.00150
		5.183	0.07156 ± 0.00136
SDSS J2054−0005	6.06	5.747	0.00375 ± 0.00280
		5.586	0.09962 ± 0.00253
		5.432	0.07989 ± 0.00227
		5.283	0.09959 ± 0.00190
		5.139	0.16812 ± 0.00196
SDSS J0353+0104	6.05	5.737	-0.00561 ± 0.00312
		5.577	0.02299 ± 0.00281
		5.423	0.04055 ± 0.00270
		5.274	0.06021 ± 0.00210
		5.130	0.06568 ± 0.00249
ULAS J0148+0600	5.98	5.796	0.00041 ± 0.00037
		5.634	-0.00013 ± 0.00051
		5.478	0.04605 ± 0.00043
		5.327	0.06339 ± 0.00043
		5.182	0.13982 ± 0.00040
SDSS 0203+0012	5.72	5.423	0.04361 ± 0.00674
		5.275	0.06883 ± 0.00505
		5.131	0.11174 ± 0.00586
		4.992	0.07700 ± 0.00758
		4.858	0.14021 ± 0.00519
SDSS J0231−0728	5.42	5.138	0.23508 ± 0.00062
		4.999	0.16729 ± 0.00069
		4.864	0.17614 ± 0.00063
		4.734	0.19190 ± 0.00071
		4.608	0.19291 ± 0.00074
SDSS J1659+2709	5.32	5.043	0.12661 ± 0.00058
		4.907	0.16273 ± 0.00059
		4.776	0.15605 ± 0.00059
		4.648	0.25978 ± 0.00068
		4.525	0.31753 ± 0.00059
SDSS J1208+0010	5.27	4.996	0.25503 ± 0.00107
		4.861	0.12560 ± 0.00083
		4.731	0.22512 ± 0.00100
		4.605	0.30364 ± 0.00101
		4.484	0.29008 ± 0.00105
SDSS J0915+4244	5.20	4.929	0.17196 ± 0.00120
		4.797	0.10740 ± 0.00103
		4.669	0.18277 ± 0.00124
		4.545	0.16960 ± 0.00132
		4.425	0.37799 ± 0.00109
SDSS J1204−0021	5.09	4.824	0.19396 ± 0.00147
		4.696	0.18355 ± 0.00157
		4.571	0.35492 ± 0.00172
		4.450	0.28457 ± 0.00161
		4.333	0.32831 ± 0.00153
SDSS J0040−0915	4.98	4.720	0.15083 ± 0.00085
		4.594	0.28959 ± 0.00088

Table 2 – *continued*

QSO	$z_{\rm em}$	$z_{\rm abs}$	$\langle F/F_{\rm C}\rangle$
		4.473	0.28650 ± 0.00085
		4.355	0.35385 ± 0.00095
		4.242	0.31313 ± 0.00096
SDSS J0011+1446	4.95	4.691	0.29507 ± 0.00045
		4.567	0.23270 ± 0.00040
		4.446	0.42480 ± 0.00037
		4.329	0.38612 ± 0.00036
		4.216	0.34395 ± 0.00036
SDSS J2225−0014	4.89	4.634	0.27845 ± 0.00190
		4.511	0.29665 ± 0.00181
		4.393	0.22136 ± 0.00196
		4.278	0.29759 ± 0.00206
		4.166	0.34961 ± 0.00235
SDSS J1616+0501	4.88	4.625	0.20429 ± 0.00184
		4.502	0.20725 ± 0.00185
		4.384	0.45577 ± 0.00221
		4.269	0.33738 ± 0.00222
		4.158	0.29428 ± 0.00254
BR 1202−0725	4.70	4.453	0.27003 ± 0.00090
		4.336	0.32692 ± 0.00205
		4.223	0.40195 ± 0.00216
		4.113	0.39365 ± 0.00197
		4.007	0.47038 ± 0.00191
SDSS J2147−0838	4.60	4.358	0.39491 ± 0.00075
		4.244	0.42531 ± 0.00087
		4.134	0.35357 ± 0.00090
		4.027	0.34943 ± 0.00091
		3.923	0.38447 ± 0.00097
BR 0353−3820	4.59	4.349	0.45173 ± 0.00042
		4.235	0.28621 ± 0.00043
		4.125	0.46655 ± 0.00051
		4.018	0.27414 ± 0.00045
		3.915	0.45430 ± 0.00053
BR 1033−0327	4.52	4.282	0.38280 ± 0.00171
		4.170	0.27687 ± 0.00179
		4.062	0.32834 ± 0.00184
		3.958	0.45227 ± 0.00200
		3.856	0.51813 ± 0.00205
BR 0006−6208	4.52	4.282	0.46323 ± 0.00244
		4.170	0.38739 ± 0.00287
		4.062	0.42680 ± 0.00301
		3.958	0.46304 ± 0.00323
		3.856	0.41029 ± 0.00329
BR 0714−6449	4.49	4.253	0.47062 ± 0.00093
		4.143	0.29638 ± 0.00104
		4.035	0.41984 ± 0.00111
		3.932	0.40853 ± 0.00116
		3.831	0.48802 ± 0.00127
BR 0418−5723	4.48	4.244	0.29407 ± 0.00070
		4.134	0.41240 ± 0.00075
		4.027	0.38504 ± 0.00073
		3.923	0.45380 ± 0.00075
		3.822	0.47881 ± 0.00079

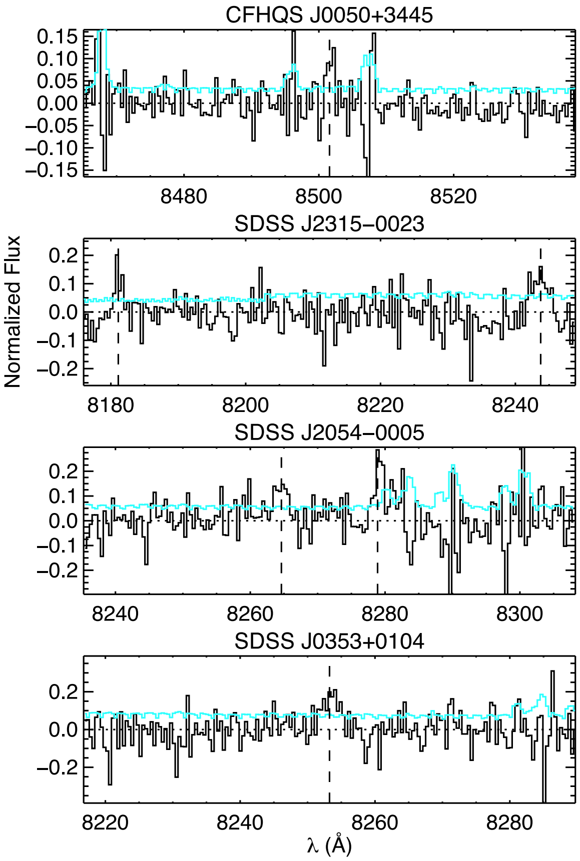

Figure 2. Probable Lyα transmission peaks in regions of the forest near $z \sim 6$ without a formal mean transmitted flux detection. The dark histograms in each panel show the flux, while the light histograms show the 1σ uncertainty. The locations of probable transmission peaks are indicated by vertical dashed lines.

Table 3. Lower limits on the continuum-normalized Lyα flux for $50\,{\rm Mpc}\,h^{-1}$ regions that do not have a formal 2σ detection yet show individually significant transmission peaks.

QSO	$z_{\rm em}$	$z_{\rm abs}$	$\langle F/F_{\rm C}\rangle$
CFHQS J0050+3445	6.25	6.074	>0.0005
SDSS J2315−0023	6.12	5.797	>0.0021
SDSS J2054−0005	6.06	5.747	>0.0010
SDSS J0353+0104	6.05	5.737	>0.0018

(in preparation). Here we note, however, that a detailed search revealed no metal systems at $z > 5.5$ traced by C IV, Si IV, or Mg II doublets, or by groups of low-ionization lines such as O I, C II, Si II, Fe II, and Al II. We are therefore reasonably confident that classical metal-enriched DLAs are not affecting this line of sight,[2] and so are not responsible for the most extreme values of $\tau_{\rm eff}$ in our data. Future, more detailed analysis of the Lyα forest at these redshifts, however, may require a more comprehensive treatment of both metal lines and optically thick absorbers.

2.2.1 *The ULAS J0148+0600 Lyα trough*

The Lyα forest towards ULAS J0148+0600 displays an unusually long ($\sim$110 Mpc h^{-1}) Lyα trough over $5.523 \le z \le 5.879$ (Fig. 3).

---

[2] Our O I column density detection limit over this redshift range is $N_{\rm O\,I} \simeq 10^{13.5}$ cm^{-2}, which corresponds to a DLA ($N_{\rm H\,I} \ge 10^{20.3}$ cm^{-2}) metallicity of [O/H] $\lesssim -3.5$. This is a factor of 5 in metallicity below the most metal-poor DLAs reported in the literature (e.g. Cooke et al. 2011, and references therein).

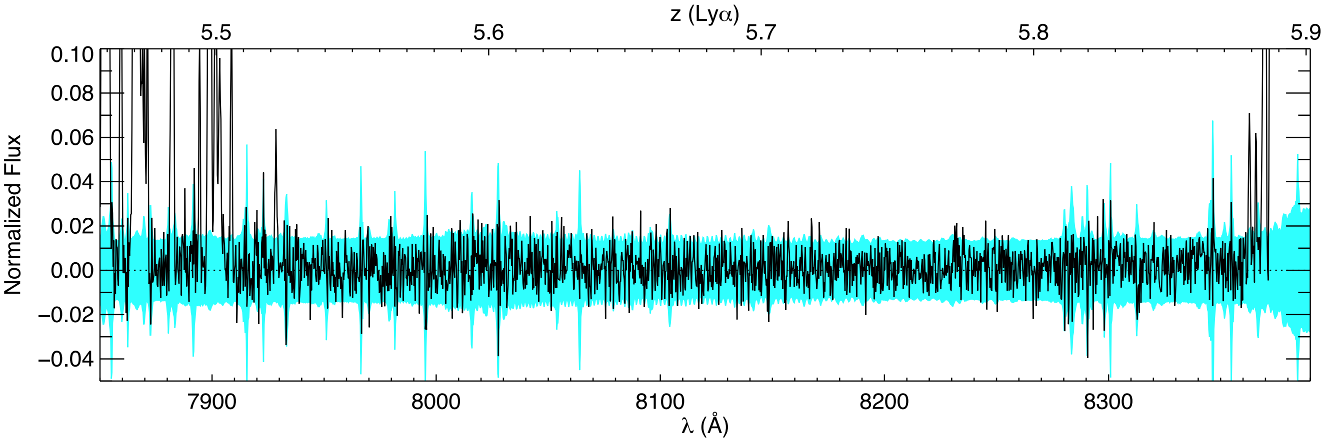

Figure 3. The Lyα trough towards ULAS J0148+0600. The flux, binned in 10 km s^{-1} pixels, is shown by the dark histogram. The shaded region shows the $\pm 2\sigma$ uncertainty interval. Note the scale on the vertical axis. Lyα redshifts are shown at the top of the plot. No significant Lyα transmission peaks appear over $5.523 \le z \le 5.879$.

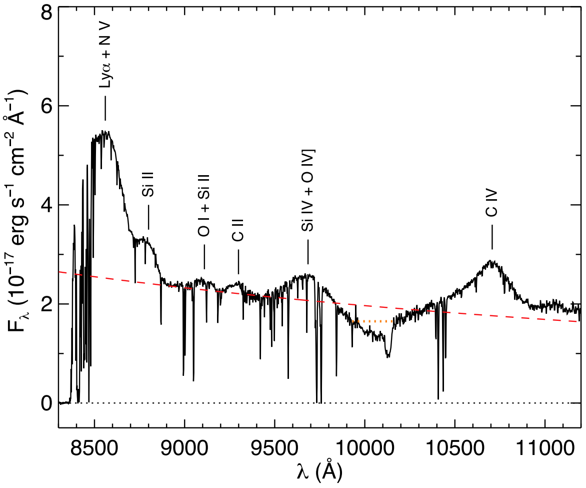

Figure 4. X-Shooter spectrum of ULAS J0148+0600 redward of the Lyα forest. Prominent emission lines are marked. The dashed line is an estimate of the underlying power-law continuum. The C IV BAL is visible as a depression below this estimate between the Si IV and C IV emission lines. A mini-BAL component at $z \simeq 5.54$ is visible near 10 130 Å.

This is roughly twice as long as the longest troughs reported previously (White et al. 2003; Fan et al. 2006), and extending down to significantly lower redshifts. As discussed in more detail below, the depth of the trough towards ULAS J0148+0600 is in significant contrast with other lines of sight at the same redshifts. For the 50 Mpc h^{-1} regions centred at $z = 5.63$ and 5.80 we find 2σ lower limit of $\tau_{\rm eff} \ge 6.9$ and 7.2, respectively. For the complete trough we measure $\tau_{\rm eff} \ge 7.4$.

As noted above, we find no intervening metal absorbers over this redshift interval that would suggest the presence of DLAs. ULAS J0148+0600 does show a mild, broad depression in its spectrum between the Si IV and C IV emission lines ($9840 \lesssim \lambda \lesssim 10\,320$ Å) compared to a power-law estimate of the continuum, however (Fig. 4). BALs are therefore a potential concern, since a C IV BAL over this interval could indicate Lyα and/or N V absorption in the Lyα forest. The clearest indication of broad C IV absorption is the narrow mini-BAL feature over $10\,110 \lesssim \lambda \lesssim 10\,150$ Å. It is unclear, however, the extent to which the remainder of the depression in Fig. 4 is due to broad absorption. A study of BAL quasars in the Sloan Digital Sky Survey (SDSS) by Allen et al. (2011) found no compelling examples of such wide, shallow BALs (Hewett, private communication). The lack of distinct features (apart from the mini-BAL) between the Si IV and C IV emission lines instead suggests that most of the depression is an intrinsic feature of the quasar spectrum, rather than a BAL. No broad absorption in Si IV is seen, which further suggests that most of the depression is not a C IV BAL, although the gas could have either a high ionization state or a low column density. We suspect that the genuine broad C IV absorption is restricted to at most a modest depression over $9930 \lesssim \lambda \lesssim 10170$ Å, indicated by the dotted line in Fig. 4. To be conservative, however, we assume that the entire depression below the power-law estimate is due to a C IV BAL, and model the impact of corresponding N V absorption. We modify our continuum estimate for ULAS J0148+0600 by including an estimate of the N V broad absorption that assumes the C IV and N V BAL profiles are similar in velocity structure and amplitude (e.g. Baskin, Laor & Hamann 2013). This has a relatively minor affect, decreasing $\tau_{\rm eff}$ in the 50 Mpc h^{-1} region centred at $z = 5.63$ by 0.3 and by <0.1 elsewhere. The flux measurements given in Table 2 and the $\tau_{\rm eff}$ lower limits quoted above take this estimate of the N V absorption into account. The lack of Si IV suggests that any BAL would be weak in Lyα (Hamann et al., in preparation). The strongest potential C IV absorption, moreover, occurs at $z < 5.54$, which for Lyα falls blueward of the trough. We repeat our modelling procedure for O VI λ1032, 1038 broad absorption that may be present in the Lyβ forest, however.

Although no Lyα transmission peaks are detected in the ULAS J0148+0600 trough, we can set an upper limit on the Lyα opacity using the fact that transmission is seen in the Lyβ forest over the same redshifts (Fig. 5). We use hydrodynamical simulations (see Section 3) to model the transmission in the Lyβ forest over the ~80 Mpc h^{-1} interval from $z = 5.62$ to 5.88 ($6790 \lesssim \lambda \lesssim 7060$), where the upper limit in redshift corresponds to the start of the Lyα trough, and the lower limit is set by the onset of Lyγ absorption for a quasar at redshift $z_{\rm em} = 5.98$. Over this interval we measure a total effective optical depth of $\tau_{\rm eff}^{\rm tot} = 5.17 \pm 0.05$. We modelled this absorption by superposing simulated Lyβ forest spectra at $z_{\rm trough} = 5.620$ or 5.831 on to foreground Lyα absorption at $z_{\rm fg} = (1 + z_{\rm trough})(\lambda_\alpha/\lambda_\beta) - 1$. The Ly$\beta$ and Lyα spectra were drawn randomly, with the optical depths in the foreground Lyα sample collectively scaled to reproduce the mean Lyα opacity at

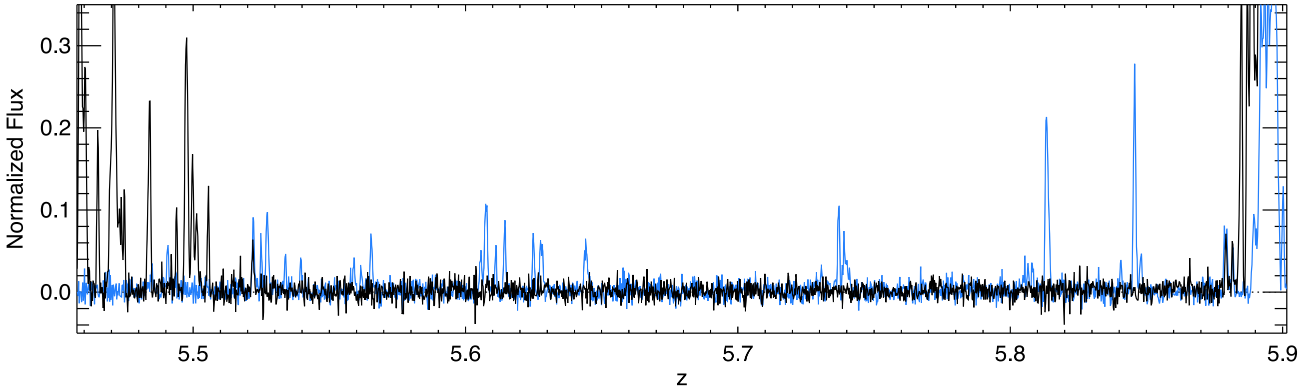

Figure 5. The Lyα trough (black) towards ULAS J0148+0600 overlaid with the Lyβ forest (blue) at the same redshifts. Lyγ absorption also occurs in the Lyβ forest at $z \leq 5.63$.

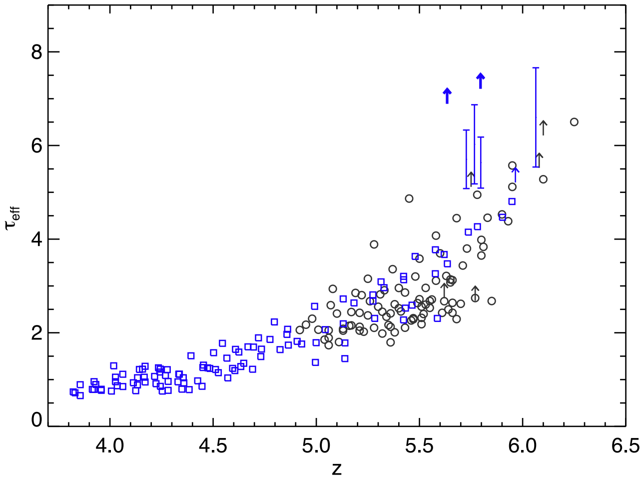

Figure 6. Lyα $\tau_{\rm eff}$ measurements. Black circles and arrows (lower limits) are from Fan et al. (2006). Blue squares, error bars, and arrows are from this work. The points at $z \simeq 5.63$ and 5.80 with $\tau_{\rm eff} \gtrsim 7$ are from the line of sight towards ULAS J0148+0600.

$z_{\rm fg}$ measured by Becker et al. (2013). For each trial, we then scaled the Lyβ optical depths such that the combined opacity matched our measured value in the Lyβ forest, and then calculated the corresponding Lyα opacity at $z = z_{\rm trough}$. In principle this procedure can be used to set both lower and upper bounds of $\tau_{\rm eff}$ for Lyα; however, we find that the conversion from $\tau_{\rm eff}^{\beta}$ to $\tau_{\rm eff}^{\alpha}$ is not converged for our simulations (see Appendix A5), in the sense that $\tau_{\rm eff}^{\alpha}$ is probably too high for a given $\tau_{\rm eff}^{\beta}$ for even our highest resolution simulation. A lower limit on $\tau_{\rm eff}^{\alpha}$ in the trough set by this procedure would therefore not be reliable, although an upper limit will be conservative. For the ∼80 Mpc h^{-1} stretch containing only Lyβ and foreground Lyα absorption, our measurement of $\tau_{\rm eff}^{\rm tot}$ implies a 95 per cent upper limit of $\tau_{\rm eff}^{\alpha} \leq 12.3$. Here we have interpolated the results from adopting $z_{\rm trough} = 5.620$ and 5.831 for our simulations on to the mean redshift of the Lyβ trough ($z \simeq 5.75$).

The combined Lyα $\tau_{\rm eff}$ data are shown in Fig. 6. As pointed out by Fan et al. (2006), $\tau_{\rm eff}$ exhibits both a strong overall increase with redshift and an enhanced scatter at $z > 5$. Our new measurements support this trend, with the ULAS J0148+0600 trough providing the starkest demonstration that lines of sight with very strong absorption exist at the same redshift as lines of sight where the absorption is far more modest. The primary goal of this paper is to determine whether these sightline-to-sightline variations are predicted by simple models of the UVB, or whether more complicated effects – potentially relating to hydrogen reionization – are needed. We now turn to interpreting the $\tau_{\rm eff}$ measurements within the context of simple models for the evolution of the ionizing UVB. These models jointly consider the large-scale radiation and density fields using the numerical simulations described below.

3 HYDRODYNAMICAL SIMULATIONS

The large-scale distribution of gas in the IGM at $z > 4$ is modelled in this work using a set of 11 cosmological hydrodynamical simulations. These simulations are summarized in Table 4, and were performed using the smoothed particle hydrodynamics (SPH) code GADGET-3, which is an updated version of the publicly available code GADGET-2 last described by Springel (2005).

The fiducial cosmological parameters adopted in the simulations are ($\Omega_{\rm m}$, Ω_Λ, $\Omega_{\rm b}\,h^2$, h, σ_8, $n_{\rm s}$) = (0.26, 0.74, 0.023, 0.72, 0.80, 0.96). These calculations were all started at redshift $z = 99$, with initial conditions generated using the Eisenstein & Hu (1999) transfer function. The gravitational softening length was set to 1/25th the

Table 4. Hydrodynamical simulations used in this work. The columns list the model name, the box size in comoving h^{-1} Mpc, the number of gas and dark matter particles, and the gas particle mass for each simulation. The fiducial cosmological parameters adopted in the simulations are ($\Omega_{\rm m}$, Ω_Λ, $\Omega_{\rm b}\,h^2$, h, σ_8, $n_{\rm s}$) = (0.26, 0.74, 0.023, 0.72, 0.80, 0.96), and the fiducial UVB is from Haardt & Madau (2001). The two exceptions are the Planck model, which adopts an alternative set of cosmological parameters, and the Dz12_g1.0 model, which assumes an alternative thermal history (see text for further details).

Model	L (Mpc h^{-1})	Particles	$M_{\rm gas}$ (M$_\odot$ h^{-1})
100–1024	100	2×1024^3	1.15×10^7
100–512	100	2×512^3	9.18×10^7
100–256	100	2×256^3	7.34×10^8
50–1024	50	2×1024^3	1.43×10^6
50–512	50	2×512^3	1.15×10^7
50–256	50	2×256^3	9.18×10^7
25–1024	25	2×1024^3	1.79×10^5
25–512	25	2×512^3	1.43×10^6
25–256	25	2×256^3	1.15×10^7
Planck	100	2×1024^3	1.25×10^7
Dz12_g1.0	100	2×1024^3	1.15×10^7

mean linear interparticle separation. Star formation was modelled using an approach designed to optimize Lyα forest simulations, where all gas particles with overdensity $\Delta = \rho/\langle\rho\rangle > 10^3$ and temperature $T < 10^5$ K are converted into collisionless star particles. The photoionization and heating of the IGM were included using a spatially uniform UVB, applied assuming the gas in the simulations is optically thin (Haardt & Madau 2001). The fiducial thermal history in this work corresponds to model C15 described in Becker et al. (2011a); see also Appendix A1.

A total of nine simulations were performed to test the impact of box size and resolution on our results (although the two models we use most extensively in this work are the 100–1024 and 25–1024 simulations listed in Table 4). These span a range of box sizes and gas particles masses, from 25 to 100 Mpc h^{-1} and 1.79×10^5 to $7.34 \times 10^8\,\mathrm{M}_\odot\,h^{-1}$. Note, however, that these models (particularly 100–1024) employ a rather low mass resolution relative to that required for fully resolving the low-density Lyα forest at $z > 5$ (cf. Bolton & Becker 2009, who recommend $L \geq 40$ Mpc h^{-1} and $M_{\mathrm{gas}} \leq 2 \times 10^5\,\mathrm{M}_\odot\,h^{-1}$).

In this work, however, our goal is to examine spatial fluctuations in the Lyα forest opacity and UVB on large scales. The typical scales are difficult to capture correctly in smaller ($\sim$10 Mpc h^{-1}) boxes with high mass resolution; the mean free path to Lyman limit photons at $z = 5$ is $\sim$60 Mpc h^{-1} (comoving; e.g. Prochaska, Worseck & O'Meara 2009; Songaila & Cowie 2010; Worseck et al. 2014). Since computational constraints mean we are unable to perform simulations in boxes with $L \sim 100$ Mpc h^{-1} at the mass resolution needed to fully resolve the low-density IGM, a compromise must then be made on this numerical requirement. We have, however, verified that this choice will not alter the main conclusions of this study. This is examined in further detail in Appendix A1, where we present a series of convergence tests with box size and mass resolution.

In addition to the nine simulations used to test box size and mass resolution convergence, we also perform two further simulations in which the cosmological parameters and IGM thermal history are varied. These models are used to test the impact of these assumptions on our results. The Planck simulation adopts $(\Omega_{\mathrm{m}}, \Omega_\Lambda, \Omega_{\mathrm{b}}\,h^2, h, \sigma_8, n_{\mathrm{s}}) = (0.308, 0.692, 0.0222, 0.678, 0.829, 0.961)$, consistent with the recent results from *Planck* (Planck Collaboration XVI 2014). The Dz12_g1.0 model adopts an alternative IGM heating history which reionizes earlier ($z_{\mathrm{r}} = 12$, cf. $z_{\mathrm{r}} = 9$ for our fiducial model), and heats the gas in the low-density IGM to higher temperatures. Further details and tests using these models maybe found in Appendices A2 and A3.

Finally, we extract synthetic Lyα forest spectra from the output of the hydrodynamical simulations using a standard approach (e.g. Theuns et al. 1998) under the assumption of a spatially uniform H I photoionization rate, $\Gamma_{\mathrm{H\,I}}$. As we now discuss in the next section, these spectra will also be generated using a model for spatial fluctuations in the ionization rate which is applied in post-processing.

4 UV BACKGROUND MODELS

4.1 Uniform UVB

We begin by considering models with a uniform ionizing background, where the scatter in τ_{eff} between lines of sight is driven entirely by variations in the density field. Lidz et al. (2006) found that such a model could potentially accommodate much of the observed scatter in τ_{eff} without invoking additional factors such as fluctuations in the UVB related to the end of reionization. Our first task is therefore to reassess this conclusion in light of the additional data presented herein.

We calculate the expected scatter in τ_{eff} at each simulation redshift by fixing the volume-averaged neutral fraction, $\langle f_{\mathrm{H\,I}}\rangle$, assuming a uniform UVB, and measuring the mean flux along randomly drawn 50 Mpc h^{-1} sections of Lyα forest. We use 100–1024 simulation for our fiducial estimates in order to include the maximum amount of large-scale structure. Trials with the other simulations in Table 4 show relatively little dependence on box size, but decreased scatter in τ_{eff} towards higher mass resolution (see Appendix A1). This trend is driven by the fact that the low-density regions, which dominate the transmission at these redshifts, become better resolved with decreasing particle mass (see Bolton & Becker 2009). Our choice of the 100 Mpc h^{-1} box is therefore conservative in determining whether the observed scatter can be reproduced with a uniform UVB.

Our nominal $\langle f_{\mathrm{H\,I}}\rangle$ evolution is given by

$$\langle f_{\mathrm{H\,I}}\rangle(z) = (1.3 \times 10^{-5}) \left(\frac{1+z}{5.6}\right)^{\eta}, \qquad (1)$$

with $\eta = 2.9$ (5.0) for $z \leq 4.6$ ($z > 4.6$). The evolution at $z \leq 4.6$ is chosen to reproduce the mean opacity measurements of Becker et al. (2013) (although we note that the precise neutral fraction will depend on the simulation parameters, primarily mass resolution). The evolution in $\langle f_{\mathrm{H\,I}}\rangle$ at $z > 4.6$ is chosen such that the lower bound in τ_{eff} roughly traces the lower envelope of the observed values over $4.6 < z < 5.8$. At these redshifts, the majority of τ_{eff} measurements tend to cluster along a relatively narrow locus bounded by this envelope, while outlying points tend to scatter towards higher values. While our choice of $\langle f_{\mathrm{H\,I}}\rangle$ evolution at $z > 4.6$ is not a fit, anchoring the τ_{eff} distribution near this lower boundary provides one way of determining whether all of the data, at least up to $z \simeq 5.8$, can be reproduced using a uniform UVB.

Our uniform UVB model is compared to the τ_{eff} measurements in Fig. 7. As noted above, we do not include continuum uncertainties in the τ_{eff} values measured from the data, but instead incorporate these effects directly into our models. The dark shaded region in

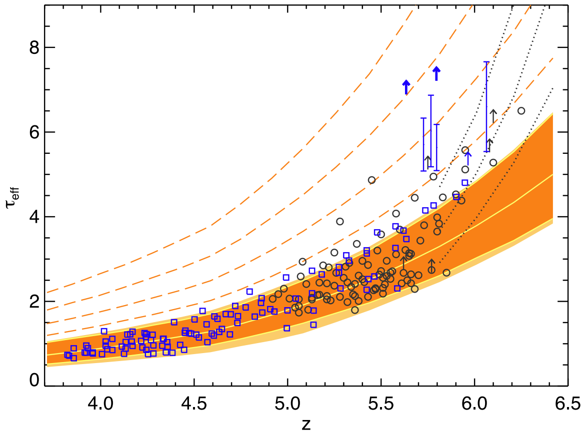

Figure 7. Predicted distribution of Lyα τ_{eff} for our uniform UVB model. Data points are as in Fig. 6. The dark shaded region spans the two-sided 95 per cent range in τ_{eff} for the evolution of the hydrogen neutral fraction given by equation (1). The light shaded bands on either side of this region show the additional scatter due to random continuum errors (see text). The dashed lines give the one-sided 95 per cent upper limit in τ_{eff} for 50 Mpc h^{-1} regions when the neutral fraction is increased by 0.15, 0.3, 0.45, and 0.6 dex (bottom to top). The dotted lines give the two-sided 95 per cent interval in τ_{eff} for a separate evolution in the neutral fraction (see text).

Fig. 7 shows the predicted range in $\tau_{\rm eff}$ without continuum errors, while the outer, lighter shaded regions include random continuum errors with an rms amplitude linearly interpolated between 5, 10, and 20 per cent at $z = 3$, 4, and 5, respectively, and a constant 20 per cent at $z > 5$. At $z < 5$, the scatter in the data is well reproduced by the simulations, suggesting that the UVB near 1 Ryd is reasonably uniform at these redshifts. This is not surprising given that the mean free path to hydrogen ionizing photons at $z < 5$ is long with respect to the typical separation between star-storming galaxies (e.g. Prochaska et al. 2009; Songaila & Cowie 2010; Worseck et al. 2014). Even up to $z \lesssim 5.3$ the scatter in $\tau_{\rm eff}$ outside that expected for a uniform UVB is minimal.

Over $5.3 < z < 5.8$ the scatter in the uniform UVB model still spans a large fraction of the data; however, an increasing number of points fall above the upper model bound with increasing redshifts. The most extreme scatter occurs near $z \simeq 5.6$ where, in order to span the collection of points with $\tau_{\rm eff} \simeq 2.5$, the 97.5 per cent upper limit for the uniform UVB model is $\tau_{\rm eff} \le 3.7$. In contrast, the 50 Mpc h^{-1} section at $z = 5.63$ towards ULAS J0148+0600 has $\tau_{\rm eff} > 6.9$. Several other points, although not as extreme as the ULAS J0148+0600 values, also lie significantly above the upper bound in $\tau_{\rm eff}$, even down to $z \simeq 5.3$. This strongly suggests that the Lyα forest along these lines of sight is inconsistent with a uniform UVB model that is required to fit the observed lower envelope in $\tau_{\rm eff}$ at $z < 5.8$. We emphasize that a simple rescaling of $\langle f_{\rm H\,I} \rangle$ is unable to produce a reasonable fit to all the data at these redshifts. For example, if we increase $\langle f_{\rm H\,I} \rangle$ at $z \simeq 5.6$ by 0.45 dex verses the nominal value in equation (1), the two-sided 95 per cent range in $\tau_{\rm eff}$ becomes $4.2 \le \tau_{\rm eff} \le 7.1$. Although this would accommodate the ULAS J0148+0600 value, the large majority of points near this redshift would then fall below the lower bound.

Before proceeding further, we note that although the $\tau_{\rm eff}$ values at $z > 5.8$ are markedly higher than the data at $5.3 < z < 5.8$, they do not on their own necessarily require an inhomogeneous UVB. The dotted lines in Fig. 7 are for a case where the neutral fraction evolves as $\langle f_{\rm H\,I} \rangle \propto (1 + z)^{15}$ at $z > 5.8$. This is a somewhat arbitrary choice, but the bounds in $\tau_{\rm eff}$ span the existing measurements and lower limits at these redshifts. Thus, while larger samples may ultimately require a non-uniform UVB at $z > 5.8$, the present Lyα data do not currently demand it. We note, however, that if the UVB contains significant fluctuations over $5.3 < z < 5.8$, perhaps due to variations in the mean free path (see below), then it is unlikely to be uniform at higher redshifts. Scatter in the UVB may also be required to account for the range in Lyβ opacities at $z \gtrsim 6$ measured by Fan et al. (2006).

We can use our uniform UVB model to estimate the minimum amplitude of UVB fluctuations required to explain the strongest outliers in $\tau_{\rm eff}$. The dashed lines in Fig. 7 show the one-sided 95 per cent upper limit in $\tau_{\rm eff}$ expected when $\langle f_{\rm H\,I} \rangle$ is increased with respect to the nominal value (equation 1) by 0.15, 0.3, 0.45, and 0.6 dex (bottom to top). The majority of points can be accommodated by a factor of 2 increase in $\langle f_{\rm H\,I} \rangle$; however, the two points for ULAS J0148+0600 at $z = 5.63$ and 5.80 require an increase in $\langle f_{\rm H\,I} \rangle$ by a factor of $\gtrsim$3. Note that the lower bound on $\tau_{\rm eff}$ will also increase when $\langle f_{\rm H\,I} \rangle$ is increased, and so a higher $\langle f_{\rm H\,I} \rangle$ will not accommodate the lowest $\tau_{\rm eff}$ points. In Fig. 8 we replaced the 50 Mpc h^{-1} points for ULAS J0148+0600 with our lower limit for the complete ∼110 Mpc h^{-1} trough, while the dashed lines show the expected upper limits for 100 Mpc h^{-1} regions. The complete trough similarly requires a factor of $\gtrsim$3 increase in $\langle f_{\rm H\,I} \rangle$ from our nominal values. Hence, it appears very likely that significant fluctuations in $\langle f_{\rm H\,I} \rangle$ in excess of those expected from density fluctuations alone must be present at $z \sim 5.6$–5.8.

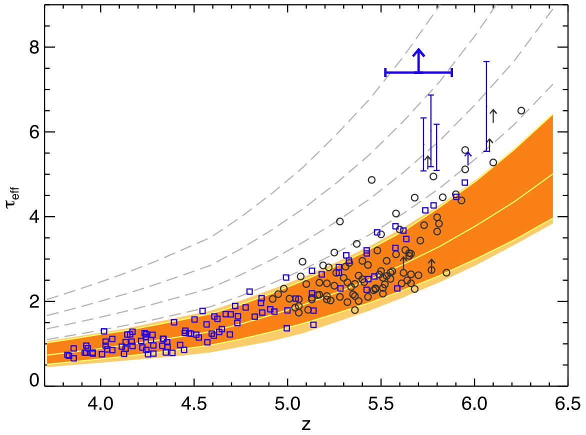

Figure 8. Same as Fig. 7, except that the 50 Mpc h^{-1} regions at $z = 5.63$ and 5.80 towards ULAS J0148+0600 have been replaced by the lower limit for the complete Lyα trough (large arrow). The dashed lines give the one-sided 95 per cent upper limit in $\tau_{\rm eff}$ for 100 Mpc h^{-1} regions when the neutral fraction is increased by 0.15, 0.3, 0.45, and 0.6 dex (bottom to top) versus the nominal value at each redshift.

4.2 Galaxy UVB model

Some level of spatial variation in the UVB is expected simply due to the fact that the photons are emitted by discrete sources. When the mean free path of ionizing photons is sufficiently long, however, the radiation field at any location will tend to reflect the contributions from a large number of sources, and so the amplitude of the fluctuations will be small unless there is a nearby, bright source such as a quasar. On the other hand, recent measurements have shown that the mean free path decreases steeply with redshift over $2 \lesssim z \lesssim 6$ (Prochaska et al. 2009; Songaila & Cowie 2010; O'Meara et al. 2013; Rudie et al. 2013; Worseck et al. 2014). It is therefore possible that the fluctuations in the UVB at $z > 5$ we infer from the Lyα forest are a natural consequence of the shortening of the mean free path.

In this section we model the expected distribution in $\tau_{\rm eff}$ for simple UVB models where star-forming galaxies are assumed to provide the majority of ionizing photons at $z > 5$ (e.g. Haardt & Madau 2012; Robertson et al. 2013). AGN are not included, as their contribution to the UVB near 1 Ryd is believed to be small at these redshifts (e.g. Cowie, Barger & Trouille 2009). To construct the UVB we first populate our simulation with sources and then calculate the intensity of the radiation field as a function of position assuming a spatially uniform mean free path. Although this is clearly an oversimplification, it gives us a first-order method for coupling the radiation field to the density field, which is essential for determining how fluctuations in the UVB affect the transmitted flux statistics. Below we assess whether this 'vanilla' UVB model can reproduce the observed variations in $\tau_{\rm eff}$ along different lines of sight.

We model the UVB within our 100–1024 simulation following an approach similar to that of Bolton & Viel (2011) and Viel et al. (2013). First we assign star-forming galaxies to dark matter haloes using an abundance matching scheme (e.g. equation 1 in Trenti et al. 2010). The haloes in the simulation volume were identified using a friends-of-friends halo finder, assuming a linking length of 0.2 and a minimum of 32 particles per halo. We assume a galaxy duty cycle of unity in our calculations, but have verified that a lower values for the duty cycle of 0.5 and 0.1 have little effect on our results (see also Appendix A4.)

Parameters for the non-ionizing ($\lambda_{\rm rest} \sim 1500$ Å) UV luminosity function are determined by interpolating fits from Bouwens et al. (2014). For our fiducial models we integrate down to $M_{\rm AB} \le -18$, and assign luminosities to dark matter haloes by randomly sampling the luminosity function. Although this magnitude limit neglects contributions from fainter sources, we find that the impact on the $\tau_{\rm eff}$ distribution is converged at this limit (see Appendix A4). We assume a galaxy spectral energy distribution (SED) that is flat at $\lambda > 912$ Å, follows a power law $L_\nu \propto \nu^{-\alpha}$ at $\lambda < 912$ Å, and has a break at $\lambda = 912$ Å of $A_{912} = L_\nu(1500)/L_\nu(912)$. For our fiducial model we adopt $\alpha = 2$ and $A_{912} = 6.0$. The amplitude of the UVB is then multiplied by a scaling factor, $f_{\rm ion}$, which is chosen as described below. This factor nominally represents the escape fraction of ionizing photons; however, in our models it is degenerate with α and A_{912}, where for young stellar populations the latter may be a factor of 2 smaller than what we assume (e.g. Eldridge & Stanway 2012). It is also degenerate with any contribution from galaxies fainter than $M_{\rm AB} = -18$. Based on the luminosity functions of Bouwens et al. (2014), these fainter galaxies may increase the total emissivity by a factor of ∼2–3 over $4 < z < 6$. Hence, $f_{\rm ion}$ may be up to a factor of ∼6 larger than the true luminosity-weighted mean escape fraction at these redshifts. For an escape fraction $f_{\rm esc} \le 1$, therefore, $f_{\rm ion} \lesssim 6$ represents a reasonable upper limit for this parameter. We note that it is obviously simplistic to assume that all galaxies have the same SED and escape fraction; if the escape fraction increases with luminosity, for example, we would expect larger fluctuations in the UVB. We find, however, that even a model with contributions solely from galaxies with $M_{\rm AB} \le -21$ produces only a modest increase in the predicted scatter in $\tau_{\rm eff}$ (see Appendix A4).

For our fiducial mean free path evolution we adopt the fit from Worseck et al. (2014). Using measurements over $2.4 \le z \le 5.2$ based on composite quasar spectra (Prochaska et al. 2009; Fumagalli et al. 2013; O'Meara et al. 2013; Worseck et al. 2014) they find

$$\lambda^{912}_{\rm mfp}(z) = 130 \left(\frac{1+z}{5} \right)^{-4.4} h^{-1}\,{\rm Mpc}, \qquad (2)$$

where we have converted their values into comoving units. Equation (2) is broadly similar to, albeit somewhat steeper than, the evolution in $\lambda^{912}_{\rm mfp}$ found by Songaila & Cowie (2010) out to $z \sim 6$. These values are only taken as a reference point; however, a range of value in $\lambda^{912}_{\rm mfp}$ are explored below. At each spatial position along our simulated lines of sight we then compute the specific intensity of the ionizing background between 1 and 4 Ryd by summing over the contribution from each galaxy as

$$J(\boldsymbol{r}, \nu) = \frac{1}{4\pi} \sum_{i=1}^{N} \frac{L_i(\boldsymbol{r}_i, \nu)}{4\pi |\boldsymbol{r}_i - \boldsymbol{r}|^2} {\rm e}^{-\frac{|\boldsymbol{r}_i - \boldsymbol{r}|}{\lambda^{912}_{\rm mfp}} \left(\frac{\nu}{\nu_{912}} \right)^{-3(\beta-1)}}. \qquad (3)$$

Here, ν_{912} is the frequency at the H I ionizing edge, and β is the slope of the H I column density distribution, which sets the dependence of mean free path on frequency. We adopt $\beta = 1.3$ (e.g. Songaila & Cowie 2010; Becker & Bolton 2013). The evolution in equation (2) roughly agrees with that derived by Songaila & Cowie (2010). In practice we perform the sum in equation (2) for all sources with $|\boldsymbol{r}_i - \boldsymbol{r}| \le L_{\rm box}/2$ (within the periodic box), and add a contribution from larger distances assuming a spatially uniform distribution of sources. Note that we neglect the redshifting of ionizing photons, though this should have a relatively minor impact at $z > 4$ (e.g. Becker & Bolton 2013). The H I photoionization rate is then computed as

$$\Gamma(\boldsymbol{r}) = 4\pi \int_{\nu_{912}}^{4\nu_{912}} \frac{{\rm d}\nu}{h\nu} J(\boldsymbol{r}, \nu) \sigma_{\rm H\,I}(\nu), \qquad (4)$$

where $\sigma_{\rm H\,I}(\nu)$ is the photoionization cross-section.

We focus our analysis on $z \simeq 5.6$, where the measured variation in $\tau_{\rm eff}$ is largest. The density field and UVB at $z = 5.62$ for two values of $\lambda^{912}_{\rm mfp}$ are shown for a slice through our simulation box in Fig. 9. As expected, the ionization rate correlates with the density, although for $\lambda^{912}_{\rm mfp} = 38\,{\rm Mpc}\,h^{-1}$ (the nominal value given by equation 2), the UVB in low-density regions is still relatively uniform. The mean intensity of the UVB scales as $\langle J \rangle \propto \lambda^{912}_{\rm mfp}$, but decreasing $\lambda^{912}_{\rm mfp}$ has the largest impact in low-density regions, which are least populated by ionizing sources. This effect has the potential, at least, to increase the scatter in $\tau_{\rm eff}$ between lines of sight.

The predictions for our galaxy UVB model are compared to the data in Fig. 10. In each panel we plot the observed cumulative probability distribution function, $P(\le \tau_{\rm eff})$, over $5.5 < z < 5.7$. Note that, for simplicity, we construct $P(\le \tau_{\rm eff})$ treating lower limits as measurements, although we do not include the two lower limits from the Fan et al. (2006) data that fall below $\tau_{\rm eff} = 3$. In cases where we have both lower and upper limits on $\tau_{\rm eff}$ we adopt their midpoint when constructing $P(\le \tau_{\rm eff})$. We then overplot the expected $P(\le \tau_{\rm eff})$ for $\lambda^{912}_{\rm mfp} = 38$, 24, 15, and $9.5\,{\rm Mpc}\,h^{-1}$, which are factors of 1.0, 0.63, 0.40, and 0.25 times the nominal value expected from equation (2). The model distributions include a 20 per cent rms uncertainty in the continuum placement, meant to mimic the effect of random continuum errors in the data. The solid lines show the model predictions when $f_{\rm ion}$ is tuned such that $P(\le \tau_{\rm eff})$ roughly matches the lower end of the observed distribution. The $\lambda^{912}_{\rm mfp} = 38\,{\rm Mpc}\,h^{-1}$ case uses $f_{\rm ion} = 0.8$, which is reasonable given the model parameters (see above). The $\lambda^{912}_{\rm mfp} = 9.5\,{\rm Mpc}\,h^{-1}$ case, however, uses $f_{\rm ion} = 4.0$, which is close to the expected upper limit of ∼6 for this parameter. Shorter mean free paths are therefore probably not realistic for these models. The $\lambda^{912}_{\rm mfp} = 38\,{\rm Mpc}\,h^{-1}$ case produces nearly the same $P(\le \tau_{\rm eff})$ as a uniform UVB model, which is also plotted in the upper left-hand panel. This reflects the fact that the radiation field in the voids, which dominate the transmission at $z > 5$, is relatively uniform for large value of $\lambda^{912}_{\rm mfp}$ (e.g. Fig. 9). Fluctuations in $\tau_{\rm eff}$ therefore remain dominated by variations in the density field (see also Bolton & Haehnelt 2007; Mesinger & Furlanetto 2009). Both the uniform UVB and $\lambda^{912}_{\rm mfp} = 38\,{\rm Mpc}\,h^{-1}$ models strongly underpredict the number of high-$\tau_{\rm eff}$ lines of sight.

The general agreement with observations does, in some sense, improve towards smaller values of $\lambda^{912}_{\rm mfp}$ (and correspondingly higher emissivities). The model $P(\le \tau_{\rm eff})$ for $\lambda^{912}_{\rm mfp} = 9.5\,{\rm Mpc}\,h^{-1}$, $f_{\rm ion} = 4.0$ (lower right-hand panel, solid line) has the broadest distribution and roughly matches most of the data. Even in this case, however, the probability of observing the highest $\tau_{\rm eff}$ value is essentially zero. We also emphasize that this model requires an ionizing emissivity of $\sim 5 \times 10^{51}$ photons ${\rm s}^{-1}\,{\rm Mpc}^{-3}$, which is a factor of 5 higher than the most recent estimate at $z \simeq 4.8$ (Becker & Bolton 2013). The dashed line in the lower right-hand panel is for $\lambda^{912}_{\rm mfp} = 9.5\,{\rm Mpc}\,h^{-1}$, $f_{\rm ion} = 2.3$. This is the only combination of parameters for which $P(\le \tau_{\rm eff})$ is non-negligible for both the highest and lowest $\tau_{\rm eff}$ values in the data ($P(\le 2.2) = 0.005$, $P(\le 6.9) = 0.994$). An Anderson–Darling test rejects the hypothesis that the data were drawn from this distribution at >99.99 per cent confidence. The remaining models are ruled out on the grounds

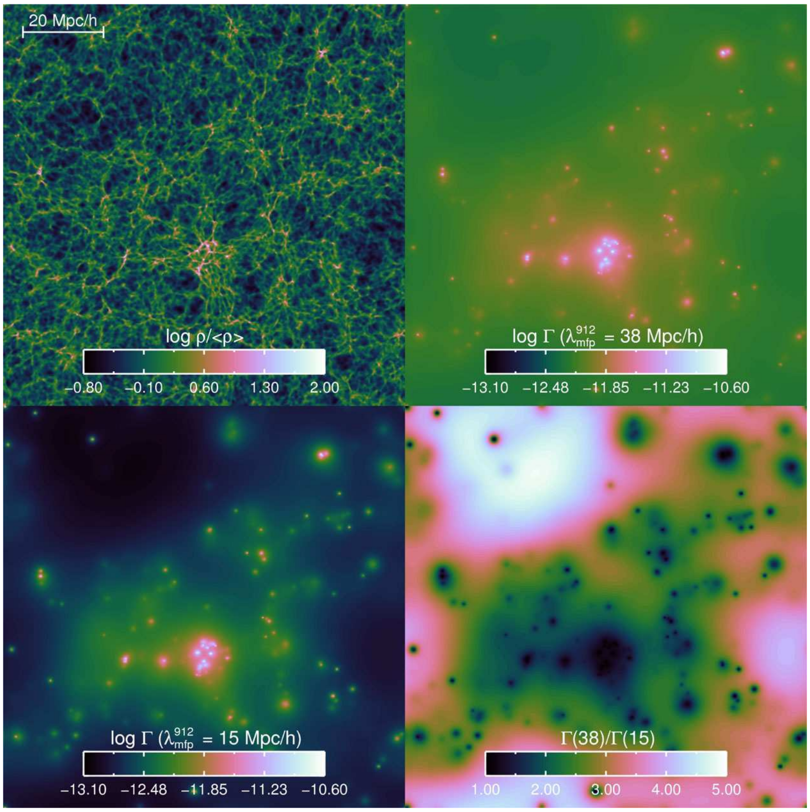

Figure 9. A slice through our 100–1024 simulation at $z = 5.62$. The top left-hand panel shows the density field integrated over a 200 kpc h^{-1} thick region. The top right-hand and bottom left-hand panels show the logarithm of the H I ionization rate, where Γ is in units of s^{-1}, for $\lambda_{\rm mfp}^{912} = 38$ and 13 Mpc h^{-1}, respectively. The bottom right-hand panel shows the ratio of the two ionization rates. Each panel is 100 Mpc h^{-1} on a side.

that the predicted probabilities of observing the extreme values in the data are too small to be meaningfully calculated.

For the three cases with $\lambda_{\rm mfp}^{912} < 38$ Mpc h^{-1} in Fig. 10 we also show the predicted $P(\le \tau_{\rm eff})$ when $f_{\rm ion}$ is fixed to the value used for the 38 Mpc h^{-1} case (dotted lines). As expected, $P(\le \tau_{\rm eff})$ shifts towards higher values of $\tau_{\rm eff}$, yet a single $\lambda_{\rm mfp}^{912}$ is again unable to match the full observed $\tau_{\rm eff}$ distribution. For this value of $f_{\rm ion}$, the lowest observed $\tau_{\rm eff}$ values only appear in the $\lambda_{\rm mfp}^{912} = 38$ Mpc h^{-1} case, while the highest observed value is only predicted to occur with significant frequency when $\lambda_{\rm mfp}^{912} \le 15$ Mpc h^{-1}. Hence, for a given emissivity, fluctuations in $\lambda_{\rm mfp}^{912}$ by factors of $\gtrsim 2.5$ appear necessary to bracket the observed $P(\le \tau_{\rm eff})$.

In summary, the failure of either a uniform UVB model or our simple galaxy UVB model to reproduce the full distribution of $\tau_{\rm eff}$ values, particularly near $z \simeq 5.6$, suggests that more complicated ionization-driven fluctuations in the volume-averaged neutral fraction are present at these redshifts. Although variations in gas temperature could technically produce variations in Lyα opacity, the high $\tau_{\rm eff}$ values towards ULAS J0148+0600 would require those regions of the IGM to be roughly a factor of 5 *colder* than average, a scenario that is physically implausible in an ionized IGM. We therefore conclude that substantial ($\gtrsim 0.5$ dex), large-scale[3] (possibly $l \gtrsim 50$ Mpc h^{-1}) fluctuations in the neutral fraction must be present throughout at least part of the IGM at these redshifts. Given that the ionizing emissivity from galaxies is likely to be comparatively uniform on these scales, we expect that the $\tau_{\rm eff}$ fluctuations are driven primarily by fluctuations in the mean free path. We now turn towards examining the $\tau_{\rm eff}$ data in more detail, and hence determining how these fluctuations may be evolving with redshift.

5 REDSHIFT EVOLUTION OF THE LYα OPACITY: EVIDENCE FOR PATCHY REIONIZATION

In the previous sections we have argued that the Lyα $\tau_{\rm eff}$ distribution at $z \simeq 5.6$–5.8 is inconsistent with either line-of-sight density variations alone or a spatially fluctuating UVB with a fixed mean free path. We thus argue that spatial variations in the mean free path must be present at these redshifts. As seen in Fig. 7, however, the scatter in the observed $\tau_{\rm eff}$ diminishes rapidly with redshift, until at $z \lesssim 5$ it becomes consistent with that expected from fluctuations in the IGM density field alone. We now investigate in more detail how the $\tau_{\rm eff}$ distribution evolves with redshift. As we demonstrate below, simple models of the UVB, while they are unable to fully describe the $\tau_{\rm eff}$ data at $z > 5$, can nevertheless provide insight into how the IGM is evolving at these redshifts.

In Fig. 11 we plot $P(\le \tau_{\rm eff})$ for the data over $3.9 < z < 5.9$ in redshift bins of $\Delta z = 0.2$. For each bin we then overplot $P(\le \tau_{\rm eff})$ for a uniform UVB model with $\langle f_{\rm H\,I} \rangle$ tuned such that that the model matches the data over the maximum possible range in $\tau_{\rm eff}$ starting at the low end. By matching the low-$\tau_{\rm eff}$ end of the model to the low end of the data, these models represent the maximum $\langle f_{\rm H\,I} \rangle$ that

---

[3] Note, however, that it is difficult to quantify the exact scale on which fluctuations in the neutral fraction occur from the 1D line-of-sight data analysed here. Fluctuations which occur on smaller scales in 3D may appear to produce larger scale fluctuations in 1D due to aliasing (e.g. McQuinn et al. 2011).

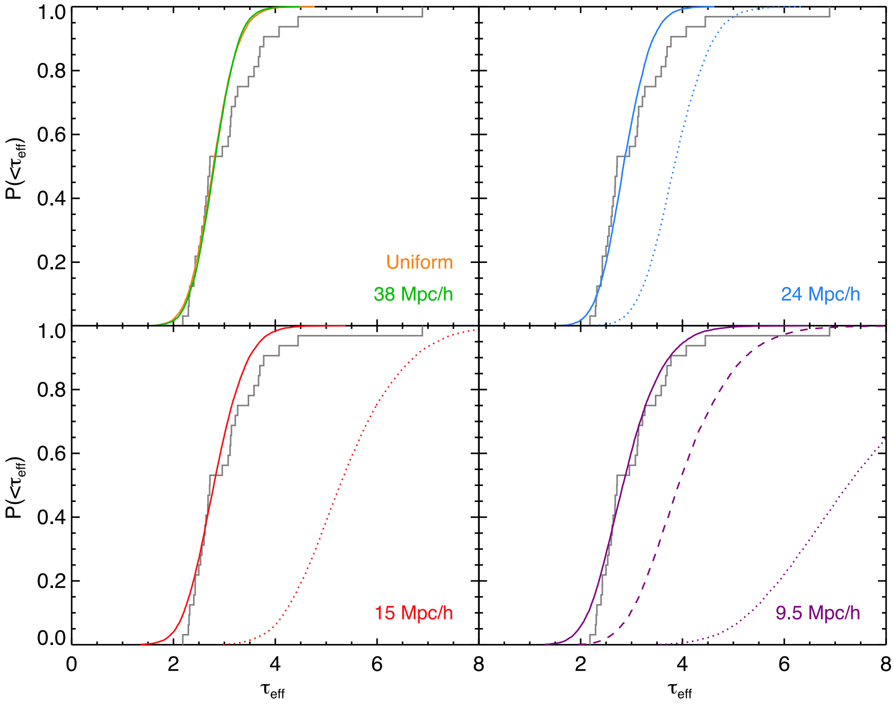

Figure 10. Cumulative probability distribution function of $\tau_{\rm eff}$ near $z = 5.6$. The histogram in each panel shows $P(\le \tau_{\rm eff})$ from the data over $5.5 < z < 5.7$. Continuous lines give the predicted $P(\le \tau_{\rm eff})$ from our galaxy UVB model for the mean free path indicated in the lower right-hand corner of each panel. For the solid lines, $f_{\rm ion}$ has been tuned individually for each $\lambda_{\rm mfp}^{912}$ such that $P(\le \tau_{\rm eff})$ roughly matches the lower end of the data distribution ($f_{\rm ion} = 0.8$, 1.3, 2.3, and 4.0 for $\lambda_{\rm mfp}^{912} = 38$, 24, 15, and 9.5 Mpc h^{-1}, respectively). The dotted lines in the top right-hand and bottom panels use $f_{\rm ion} = 0.8$, as in the $\lambda_{\rm mfp}^{912} = 38$ Mpc h^{-1} case. The dashed line in the lower right-hand panel is for $f_{\rm ion} = 2.3$. In the top left-hand panel, $P(\le \tau_{\rm eff})$ for a uniform UVB is also shown (orange line), although it is nearly indistinguishable from the galaxy UVB case (green line).

can reproduce the most transparent lines of sight. As above, we can then investigate the extent to which these $\langle f_{\rm H\,I}\rangle$ values also predict more opaque regions. Note that this procedure differs from simply matching the global mean observed opacity with simulations, which implicitly assumes a uniform photoionization rate for gas probed by the *entire* $P(\le \tau_{\rm eff})$ distribution, and thus ignores the additional scatter in the $\tau_{\rm eff}$ measurements and potentially underestimates the photoionization rate in the most highly ionized regions (e.g. Bolton & Haehnelt 2007; Mesinger & Furlanetto 2009).

For this comparison we use the 25–1024 simulation (i.e. two simulated lines of sight are used per 50 Mpc h^{-1}region), for which we find $P(\le \tau_{\rm eff})$ to be nearly converged with $\langle f_{\rm H\,I}\rangle$ with respect to box size and mass resolution (see Appendix A1). The model $\tau_{\rm eff}$ distributions are interpolated between simulation output redshifts to match the data. The models also include an rms scatter in the continuum with amplitude linearly interpolated between 10 and 20 per cent between $z = 4$ and 5, and 20 per cent at $z > 5$. The exact amplitude of the continuum scatter is not critical to our analysis. There may also be systematic uncertainties in the continuum placement, however, which we address below. Note again that we construct $P(\le \tau_{\rm eff})$ treating lower limits as measurements.

Over $3.9 < z < 4.9$ the data are well matched by a uniform UVB model over the full range in $\tau_{\rm eff}$. This agrees with the general impression from Fig. 7 that line-of-sight variations in the density field dominate the scatter in $\tau_{\rm eff}$, which is perhaps not surprising given the long mean free paths at these redshifts (equation 2). At $z > 4.9$ the data begin to diverge from the uniform UVB model at the high-$\tau_{\rm eff}$ end. We note, however, that although the divergence increases with redshift, a substantial fraction of the data remain consistent with the uniform UVB model up to at least $z \sim 5.7$. Over $5.5 < z < 5.7$, roughly half of the data follow the expected $P(\le \tau_{\rm eff})$ for a uniform UVB, even while the remaining half follow an extended tail towards higher values. Over $5.7 < z < 5.9$, in contrast, less than 20 per cent of the data appear to be consistent with density-driven fluctuations in $\tau_{\rm eff}$.

The apparent agreement between much of data over $4.9 < z < 5.9$ and the predicted $P(\le \tau_{\rm eff})$ for a uniform UVB suggests that lines of sight matched by the model may trace regions where the H I photoionization rate is reasonably similar, at least in the voids, which dominate the transmission at $z > 5$. The fraction of the $\tau_{\rm eff}$ data that require a somewhat lower photoionization rate, meanwhile, decreases rapidly with decreasing redshift over this interval. This trend is broadly consistent with the final stages of patchy reionization (e.g. Gnedin 2000; Miralda-Escudé, Haehnelt & Rees 2000; Barkana & Loeb 2001). Even once the ionized bubbles in a region of the IGM overlap and the volume-averaged neutral fraction approaches zero, the local mean free path will still evolve rapidly and exhibit a degree of spatial variance as residual patches of neutral hydrogen and/or Lyman limit systems at the edges of H II regions are ionized (e.g. Furlanetto & Oh 2005; Choudhury, Haehnelt & Regan 2009; Alvarez & Abel 2012; Sobacchi & Mesinger 2014). The final stages of reionization will progress until the local mean free path is set by large-scale structure rather than reionization topology. At this point, fluctuations in Lyman limit opacity observed in the existing quasar spectra are primarily driven by variations in density rather than ionization, and the UVB in underdense regions will approach a global value that is relatively uniform. We argue here that the fraction of the data in Fig. 11 consistent with this natural end-point to reionization are small at $z \sim 6$ but approach unity by $z \sim 5$.

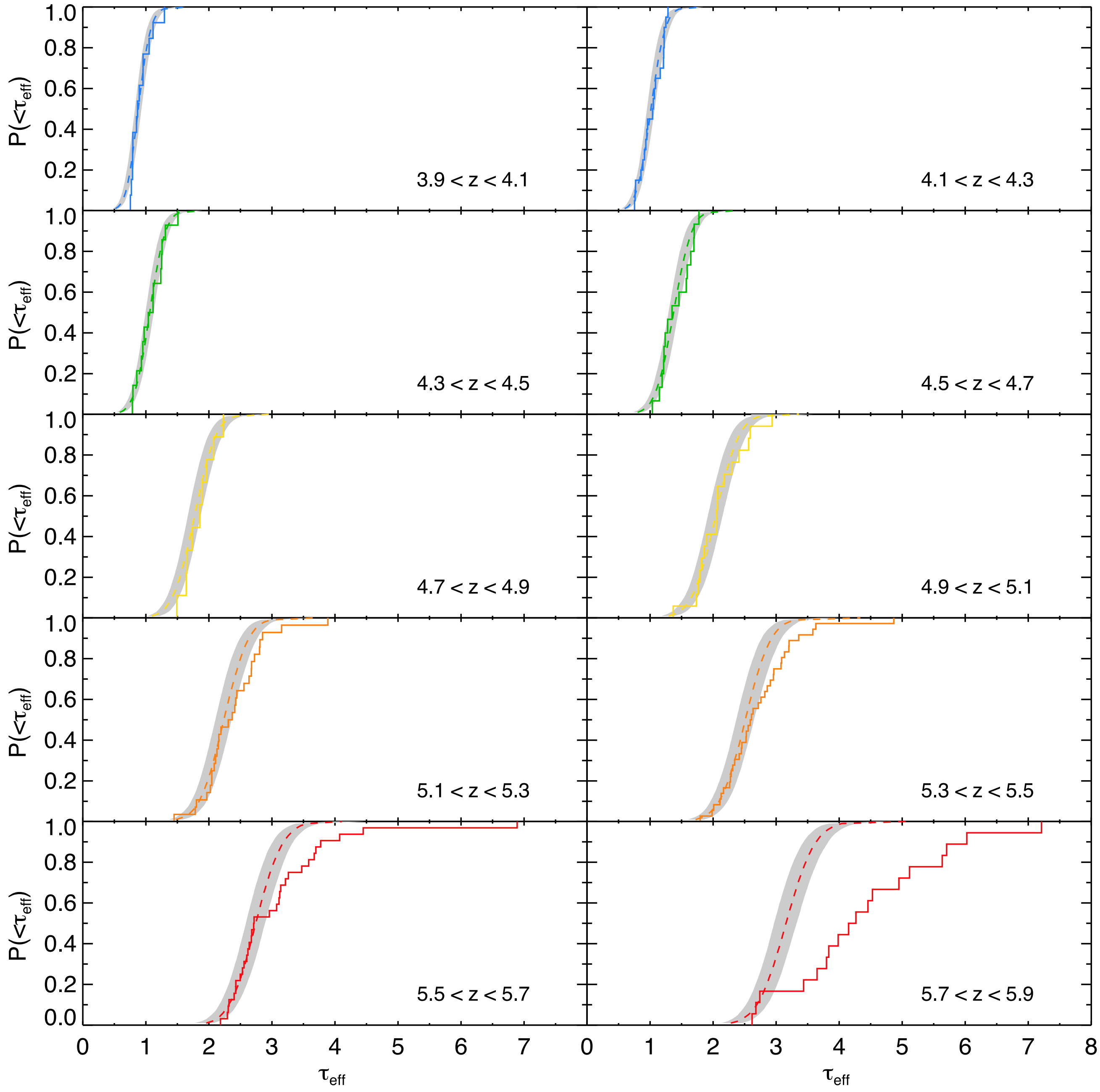

Figure 11. Cumulative probability distribution function of $\tau_{\rm eff}$ in bins of $\Delta z = 0.2$. The histograms in each panel are for the data. Dashed lines give $P(\le \tau_{\rm eff})$ for a uniform UVB (i.e. where the scatter in $\tau_{\rm eff}$ is driven by line-of-sight density fluctuations alone), where $\langle f_{\rm H\,I} \rangle$ has been tuned such that the model $P(\le \tau_{\rm eff})$ matches the lower end of the observed distribution. The shaded regions show the effect of varying $\langle f_{\rm H\,I} \rangle$ by ±10 per cent.

The volume-weighted hydrogen neutral fractions corresponding to the simulated $\tau_{\rm eff}$ distributions in Fig. 11 are shown in Fig. 12. The error bars include possible systematic errors in the quasar continua, which we take to be equal to our adopted random continuum error estimates (5, 10, and 20 per cent at $z = 3$, 4, and 5, respectively, and 20 per cent at $z > 5$). These neutral fractions correspond to regions of the IGM where the line-of-sight variance in $\tau_{\rm eff}$ is consistent with density fluctuations alone. At $z > 5$, since $\langle f_{\rm H\,I} \rangle$ has been tuned to match only the low end of the observed $\tau_{\rm eff}$ distribution, we are implicitly assuming that the matching regions are generally of lower-than-average density. If these regions are actually of higher density, then a higher ionization rate, and hence lower $\langle f_{\rm H\,I} \rangle$, would be required. In this sense, the $\langle f_{\rm H\,I} \rangle$ values at $z > 5$ in Fig. 12 are upper limits. We see, nevertheless, that $\langle f_{\rm H\,I} \rangle$ in these regions evolve gradually with redshift, increasing by only a factor of 2 between $z \sim 5$ and 6. This lends further support to the picture wherein lines of sight that are consistent with the model $P(\le \tau_{\rm eff})$ in Fig. 11 tend to probe regions of the IGM that have transitioned to a state where the mean free path is evolving relatively slowly.

6 SUMMARY

We have presented evidence for ionization-driven fluctuations in the IGM neutral fraction near $z \sim 6$ based on an expanded set of high-redshift quasar spectra. The strongest evidence for fluctuations is at $z \simeq 5.6$–5.8, where the deep Lyα trough towards ULAS J0148+0600

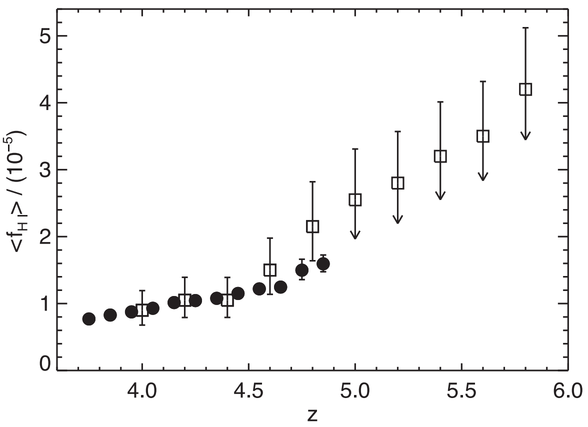

Figure 12. Hydrogen neutral fraction in regions of the IGM where the line-of-sight variance in $\tau_{\rm eff}$ is well described by density fluctuations alone. Open squares show values derived from the $P(\le \tau_{\rm eff})$ fits in Fig. 11. Filled circles show values derived from the mean Lyα opacity measurements of Becker & Bolton (2013).

contrasts strongly with the abundant transmitted flux towards other lines of sight at the same redshift. Using a suite of large ($l_{\rm box} = 25$–$100\,{\rm Mpc}\,h^{-1}$) hydrodynamical simulations, we find that the full distribution of $\tau_{\rm eff}$ values cannot be reproduced with either a uniform UVB or a simple background model that assumes galaxies are the sources of ionizing photons and uses a fixed mean free path.

These data instead appear to require fluctuations in $\langle f_{\rm H\,I} \rangle$ of at least a factor of 3 on large scales. These variations in $\langle f_{\rm H\,I} \rangle$ must be produced by fluctuations in the ionizing UV radiation field, which we argue are likely to be driven by spatial variation in the local mean free path throughout at least part of the IGM. Our results are broadly consistent with the original conclusions of Fan et al. (2006), although we find that the Fan et al. data alone require only more modest ($\lesssim$0.3 dex) fluctuations. The new data presented here, particularly the deep X-Shooter spectrum of ULAS J0148+0600, are essential for motivating larger variations in $\langle f_{\rm H\,I} \rangle$.

The variations in $\lambda_{\rm mfp}^{912}$ argued for here are consistent with expectations for the final stages of patchy hydrogen reionization (Furlanetto & Oh 2005; Choudhury et al. 2009; Alvarez & Abel 2012; Sobacchi & Mesinger 2014). During this transitional period the IGM can already be highly ionized in a volume-averaged sense, yet the radiation field will be rapidly evolving locally as residual Lyman limit systems and/or remaining diffuse patches of neutral hydrogen are ionized. Based on the observed evolution of the $\tau_{\rm eff}$ distribution, we find that a decreasing fraction of the $\tau_{\rm eff}$ data towards higher redshift ($\lesssim$20 per cent at $z \simeq 5.8$) is consistent with the variance expected from density fluctuations in the IGM alone.

Our analysis uses models of the radiation field that are purposefully simplistic in order to allow us to assess whether fluctuations in the UVB are required to explain the observed $\tau_{\rm eff}$ data, and how these fluctuations may be evolving with redshift. Predictions for $P(\le \tau_{\rm eff})$ from more sophisticated models include both sources and sinks of ionizing photons in large volumes are clearly of interest for developing a more nuanced picture of the IGM at these redshifts. For example, Gnedin & Kaurov (2014) find a long tail towards high values of $\tau_{\rm eff}$ at $z < 6$ in a set of simulations where $\langle f_{\rm H\,I} \rangle$ approaches zero between $z \simeq 6$ and 7. These and other simulations should help to translate the $\tau_{\rm eff}$ data into more detailed constraints on reionization models.

ACKNOWLEDGEMENTS

The authors thank Nick Gnedin, Martin Haehnelt, Fred Hamann, Paul Hewett, and Adam Lidz for helpful conversations, as well as Volker Springel for making GADGET-3 available. This work is based in part on observations made with ESO Telescopes at the La Silla Paranal Observatory under program ID 084.A-0390. Further observations were made at the W.M. Keck Observatory, which is operated as a scientific partnership between the California Institute of Technology and the University of California; it was made possible by the generous support of the W.M. Keck Foundation. This paper also includes data gathered with the 6.5-m Magellan telescopes located at Las Campanas Observatory, Chile. The hydrodynamical simulations used in this work were performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (http://www.hpc.cam.ac.uk/), provided by Dell Inc. using Strategic Research Infrastructure Funding from the Higher Education Funding Council for England. Fig. 9 uses the cube helix colour scheme introduced by Green (2011). GDB has been supported by an Ernest Rutherford Fellowship sponsored by the UK Science and Technology Facilities Council. JSB acknowledges the support of a Royal Society University Research Fellowship. PM acknowledges support by the NSF through grant OIA–1124453 and by NASA through grant NNX12AF87G. BPV acknowledges funding through the ERC grant 'Cosmic Dawn'.

REFERENCES

Allen J. T., Hewett P. C., Maddox N., Richards G. T., Belokurov V., 2011, MNRAS, 410, 860
Alvarez M. A., Abel T., 2012, ApJ, 747, 126
Bañados E. et al., 2014, AJ, 148, 14
Barkana R., Loeb A., 2001, Phys. Rep., 349, 125
Baskin A., Laor A., Hamann F., 2013, MNRAS, 432, 1525
Becker G. D., Bolton J. S., 2013, MNRAS, 436, 1023
Becker R. H. et al., 2001, AJ, 122, 2850
Becker G. D., Sargent W. L. W., Rauch M., Simcoe R. A., 2006, ApJ, 640, 69
Becker G. D., Rauch M., Sargent W. L. W., 2009, ApJ, 698, 1010
Becker G. D., Bolton J. S., Haehnelt M. G., Sargent W. L. W., 2011a, MNRAS, 410, 1096
Becker G. D., Sargent W. L. W., Rauch M., Calverley A. P., 2011b, ApJ, 735, 93
Becker G. D., Sargent W. L. W., Rauch M., Carswell R. F., 2012, ApJ, 744, 91
Becker G. D., Hewett P. C., Worseck G., Prochaska J. X., 2013, MNRAS, 430, 2067
Bolton J. S., Becker G. D., 2009, MNRAS, 398, L26
Bolton J. S., Haehnelt M. G., 2007, MNRAS, 382, 325
Bolton J. S., Viel M., 2011, MNRAS, 414, 241
Bolton J. S., Haehnelt M. G., Viel M., Springel V., 2005, MNRAS, 357, 1178
Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, ApJ, 670, 928
Bouwens R. J. et al., 2014, ApJ, submitted
Calverley A. P., Becker G. D., Haehnelt M. G., Bolton J. S., 2011, MNRAS, 412, 2543
Chornock R., Berger E., Fox D. B., Lunnan R., Drout M. R., fai Fong W., Laskar T., Roth K. C., 2013, ApJ, 774, 26
Chornock R., Berger E., Fox D. B., Fong W., Laskar T., Roth K. C., 2014, ApJ, preprint (arXiv:1405.7400)
Choudhury T. R., Haehnelt M. G., Regan J., 2009, MNRAS, 394, 960
Cooke R., Pettini M., Steidel C. C., Rudie G. C., Nissen P. E., 2011, MNRAS, 417, 1534
Cowie L. L., Barger A. J., Trouille L., 2009, ApJ, 692, 1476
Djorgovski S. G., Castro S., Stern D., Mahabal A. A., 2001, ApJ, 560, L5

D'Odorico S. et al., 2006, Proc. SPIE, 6269, 626933
D'Odorico V. et al., 2013, MNRAS, 435, 1198
Eisenstein D. J., Hu W., 1999, ApJ, 511, 5
Eldridge J. J., Stanway E. R., 2012, MNRAS, 419, 479
Fan X., Narayanan V. K., Strauss M. A., White R. L., Becker R. H., Pentericci L., Rix H.-W., 2002, AJ, 123, 1247
Fan X. et al., 2006, AJ, 132, 117
Faucher-Giguère C.-A., Lidz A., Hernquist L., Zaldarriaga M., 2008, ApJ, 682, L9
Fumagalli M., O'Meara J. M., Prochaska J. X., Worseck G., 2013, ApJ, 775, 78
Furlanetto S. R., Oh S. P., 2005, MNRAS, 363, 1031
Gnedin N. Y., 2000, ApJ, 535, 530
Gnedin N. Y., Kaurov A. A., 2014, ApJ, 793, 30
Green D. A., 2011, Bull. Astron. Soc. India, 39, 289
Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X- rays. Preprint (astro-ph/0106018)
Haardt F., Madau P., 2012, ApJ, 746, 125
Horne K., 1986, PASP, 98, 609
Keating L. C., Haehnelt M. G., Becker G. D., Bolton J. S., 2014, MNRAS, 438, 1820
Kelson D. D., 2003, PASP, 115, 688
Kim T.-S., Bolton J. S., Viel M., Haehnelt M. G., Carswell R. F., 2007, MNRAS, 382, 1657
Kirkman D. et al., 2005, MNRAS, 360, 1373
Lawrence A. et al., 2007, MNRAS, 379, 1599
Lidz A., Oh S. P., Furlanetto S. R., 2006, ApJ, 639, L47
Lidz A., McQuinn M., Zaldarriaga M., Hernquist L., Dutta S., 2007, ApJ, 670, 39
Lu L., Sargent W. L. W., Womble D. S., Barlow T. A., 1996, ApJ, 457, L1
McQuinn M., Hernquist L., Lidz A., Zaldarriaga M., 2011, MNRAS, 415, 977
Mesinger A., 2010, MNRAS, 407, 1328
Mesinger A., Furlanetto S., 2009, MNRAS, 400, 1461
Miralda-Escudé J., Haehnelt M., Rees M. J., 2000, ApJ, 530, 1
Mortlock D. J. et al., 2009, A&A, 505, 97
O'Meara J. M., Prochaska J. X., Worseck G., Chen H.-W., Madau P., 2013, ApJ, 765, 137
Planck Collaboration XVI, 2014, A&A, 571, A16
Prochaska J. X., Worseck G., O'Meara J. M., 2009, ApJ, 705, L113
Robertson B. E. et al., 2013, ApJ, 768, 71
Rudie G. C., Steidel C. C., Shapley A. E., Pettini M., 2013, ApJ, 769, 146
Ryan-Weber E. V., Pettini M., Madau P., Zych B. J., 2009, MNRAS, 395, 1476
Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2003, ApJ, 596, 768
Simcoe R. A. et al., 2011, ApJ, 743, 21
Sobacchi E., Mesinger A., 2014, MNRAS, 440, 1662
Songaila A., 2004, AJ, 127, 2598
Songaila A., Cowie L. L., 2010, ApJ, 721, 1448
Springel V., 2005, MNRAS, 364, 1105
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Trenti M., Stiavelli M., Bouwens R. J., Oesch P., Shull J. M., Illingworth G. D., Bradley L. D., Carollo C. M., 2010, ApJ, 714, L202
Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, Phys. Rev. D, 88, 43502
Wang R. et al., 2013, ApJ, 773, 44
White R. L., Becker R. H., Fan X., Strauss M. A., 2003, AJ, 126, 1
Worseck G. et al., 2014, MNRAS, 445, 1745
Wyithe J. S. B., Bolton J. S., 2011, MNRAS, 412, 1926

APPENDIX A: CONVERGENCE TESTS

In this appendix we address several issues related to numerical convergence. As discussed in Bolton & Becker (2009), the Lyα forest becomes increasingly sensitive to box size and mass resolution towards higher redshifts, since the transmission becomes dominated by rare voids. We therefore focus our tests at $z \simeq 5.6$, which is both near the upper end of the redshift range probed in this paper and the redshift where the largest range in $\tau_{\rm eff}$ values are observed. For our convergence tests we use a suite of nine simulations with box sizes that span 25–100 Mpc h^{-1}, and gas particle masses in the range 1.79×10^5–7.34×10^8 M$_\odot$ h^{-1}. These are listed in Table 4. Except where noted we compute $\tau_{\rm eff}$ over 50 Mpc h^{-1} lines of sight, hence for the 25 Mpc h^{-1} boxes we join two randomly chosen lines of sight per measurement, whereas for the 100 Mpc h^{-1} box we extract two measurements per line of sight.

A1 Numerical convergence

We begin by examining the convergence of our simulated $P(\leq \tau_{\rm eff})$ with box size and mass resolution for a uniform UVB. In Fig. A1 we plot $P(\leq \tau_{\rm eff})$ at $z = 5.62$ where the mean transmitted Lyα flux is fixed to $\langle F \rangle = 0.084$ for all simulations. At fixed $\langle F \rangle$ the simulated $P(\leq \tau_{\rm eff})$ increases marginally with box size

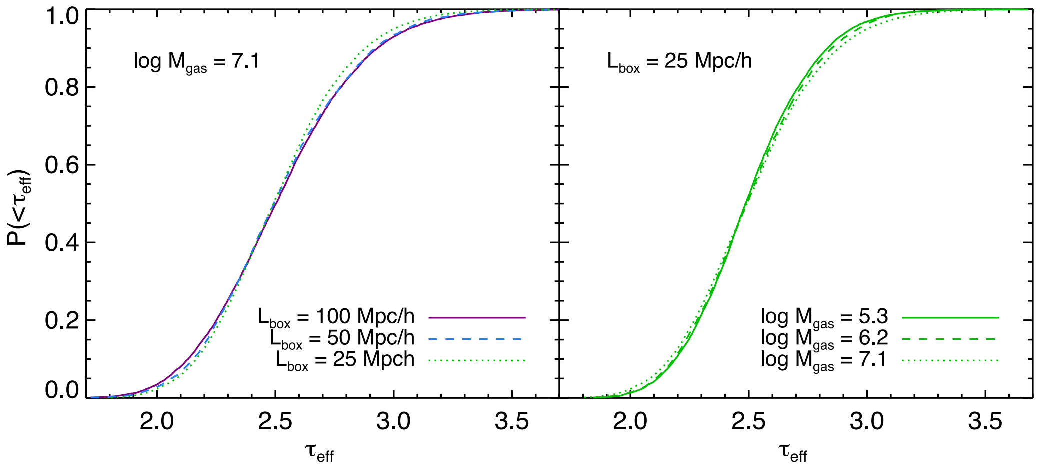

Figure A1. Numerical convergence of $P(\leq \tau_{\rm eff})$ for fixed mean Lyα flux. For each simulation we compute $\tau_{\rm eff}$ values at $z = 5.62$ over 50 Mpc h^{-1} lines of sight after rescaling the photoionization rate such that $\langle F \rangle = 0.084$. The left-hand panel shows the effects of varying the simulation box size while fixing the gas particle mass to $M_{\rm gas} = 1.1 \times 10^7$ M$_\odot$ h^{-1}, while the right-hand panel shows the effects of varying the mass resolution for a fixed box size, $l_{\rm box} = 25$ Mpc h^{-1}.

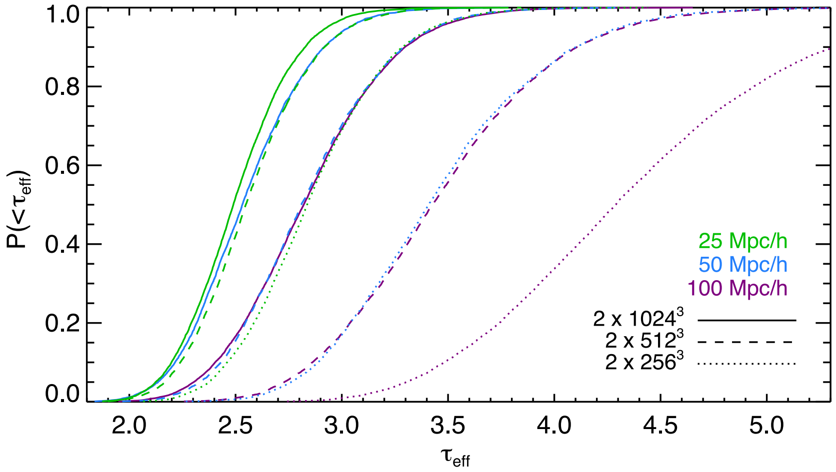

Figure A2. Numerical convergence of $P(\leq \tau_{\rm eff})$ for fixed hydrogen neutral fraction. For each simulation we compute $\tau_{\rm eff}$ values at $z = 5.62$ over 50 Mpc h^{-1} lines of sight after rescaling the photoionization rate such that $\langle f_{\rm H\,I}\rangle = 2.9 \times 10^{-5}$. Green, blue, and purple lines are for $l_{\rm box} = 25$, 50, and 100 Mpc h^{-1}, respectively, while solid, dashed, and dotted lines are for total (dark matter plus gas) particle numbers $n = 2 \times 1024^3$, 2×512^3, and 2×256^3, respectively. Runs with similar mass resolution produce similar $P(\leq \tau_{\rm eff})$ results, with only a weak dependence on box size. Decreasing the mass resolution, however, can produce a significantly more opaque Lyα forest at this redshift.

(left-hand panel), though there is little difference between the 50 and 100 Mpc h^{-1} boxes. $P(\leq\tau_{\rm eff})$ is somewhat narrower for smaller gas particles masses (right-hand panel), consistent with expectations from Bolton & Becker (2009). Our choice of a 100 Mpc h^{-1} box with $M_{\rm gas} = 1.15 \times 10^7\,{\rm M}_{\odot}\,h^{-1}$ is therefore conservative in terms of determining whether the observed scatter in $\tau_{\rm eff}$ can be reproduced using a uniform UVB.

The results are somewhat different if we evaluate $P(\leq \tau_{\rm eff})$ at a fixed hydrogen neutral fraction. We plot $P(\leq \tau_{\rm eff})$ at $z = 5.62$ for our nine numerical convergence runs in Fig. A2, where for each run we have fixed $\langle f_{\rm H\,I}\rangle = 2.9 \times 10^{-5}$. Although the predicted $\tau_{\rm eff}$ distribution shows relatively little dependence on box size, it is strongly sensitive to mass resolution. Runs using smaller gas particle masses generate voids that are more transparent (see discussion in Bolton & Becker 2009). In Fig. A3 we plot $\langle f_{\rm H\,I}\rangle$ at a fixed $\langle F\rangle = 0.084$. This again shows relatively little dependence on box size over 25–100 Mpc h^{-1}, but a strong dependence on mass resolution. The neutral fraction appears to be roughly converged for our 25–1024 run ($M_{\rm gas} = 1.8 \times 10^5\,{\rm M}_{\odot}\,h^{-1}$), which we used to measure the $\langle f_{\rm H\,I}\rangle$ values shown in Fig. 12.

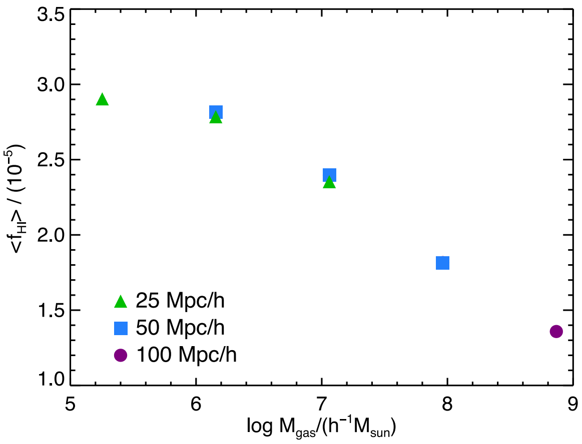

Figure A3. Convergence of the volume-averaged hydrogen neutral fraction with box size and mass resolution. For each simulation, $\langle f_{\rm H\,I}\rangle$ at $z = 5.62$ is computed after rescaling the photoionization rate to produce a fixed mean Lyα flux, $\langle F\rangle = 0.084$.

A2 Cosmology

Our fiducial simulations use a cosmology with ($\Omega_{\rm m}$, Ω_{Λ}, $\Omega_{\rm b}\,h^2$, h, σ_8, $n_{\rm s}$) = (0.26, 0.74, 0.023, 0.72, 0.80, 0.96). To test our sensitivity to cosmological parameters we ran an additional 100 Mpc h^{-1}, 2×1024^3 simulation using ($\Omega_{\rm m}$, Ω_{Λ}, $\Omega_{\rm b}\,h^2$, h, σ_8, $n_{\rm s}$) = (0.308, 0.692, 0.0222, 0.678, 0.829, 0.961), consistent with the recent results from *Planck* (Planck Collaboration XVI 2014). At fixed $\langle F\rangle$ we find a negligible difference in $P(\leq \tau_{\rm eff})$ (Fig. A4); however, the

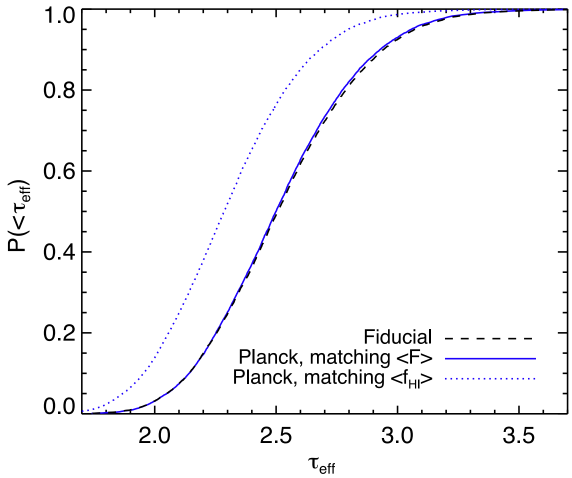

Figure A4. Dependence of $P(\leq \tau_{\rm eff})$ on cosmology. The dashed line shows $P(\leq \tau_{\rm eff})$ for 50 Mpc h^{-1} lines of sight drawn from the 100–1024 run using our fiducial simulation cosmology. The photoionization rate has been tuned such that $\langle F\rangle = 0.084$ and $\langle f_{\rm H\,I}\rangle = 2.4 \times 10^{-5}$. The solid and dotted lines show $P(\leq \tau_{\rm eff})$ for the same box size and mass resolution but using *Planck* cosmology. The solid line shows $P(\leq \tau_{\rm eff})$ with $\langle F\rangle$ matching the fiducial case, while the dotted line shows $P(\leq \tau_{\rm eff})$ when matching in $\langle f_{\rm H\,I}\rangle$.

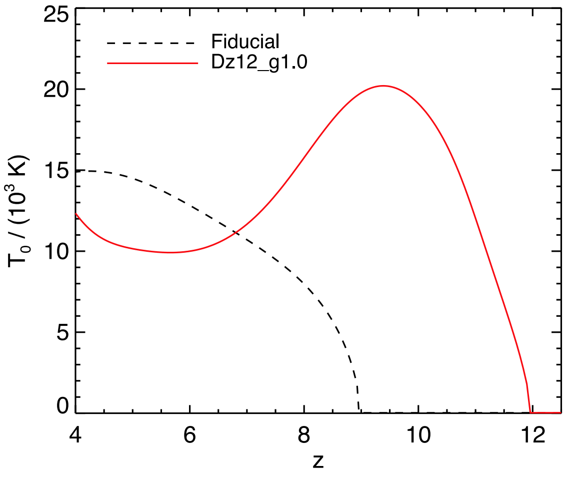

Figure A5. Thermal histories for simulation runs used to probe the impact of Jeans smoothing on $P(\le \tau_{\rm eff})$ for a uniform UVB. The dashed line shows the temperature at the mean density as a function of redshift for our fiducial thermal history, while the solid line is for a case that invokes earlier photoionization heating and a nearly isothermal ($\gamma = 1$) temperature–density relation.

Planck cosmology has a 16 per cent higher neutral fraction. Our results for $\langle f_{\rm H\,I} \rangle$ (Fig. 12) include this correction.

A3 Thermal history

The thermal history of the IGM may also impact $P(\le \tau_{\rm eff})$. Greater heating during hydrogen reionization, for example, can suppress the accretion of mass on to low-mass haloes, leaving more gas in the voids. We tested this affect by running our 100–1024 simulation with two thermal histories, which are shown in Fig. A5. In our fiducial run, the gas is reionized at $z_{\rm r} = 9$ and allowed to heat up gradually. In this run we use a temperature–density relation $T = T_0(\rho/\langle\rho\rangle)^{\gamma-1}$ with $\gamma \simeq 1.4$ at $z \sim 6$. Run Dz12_g1.0, in contrast, reionizes earlier ($z_{\rm r} = 12$), heats the gas more strongly at reionization, and uses $\gamma = 1.0$, which increases the heating in the voids. We find a somewhat broader $P(\le \tau_{\rm eff})$ in this run, although the difference is not large (Fig. A6). For this test we compute $\tau_{\rm eff}$ over 40 Mpc h^{-1} regions in order to facilitate a direct comparison with the results of Lidz et al. (2006). The thermal history for run Dz12_g1.0 is comparable to that used by Lidz et al., and we find a similar $\tau_{\rm eff}$ distribution.

A4 Galaxy UVB parameters

Out fiducial galaxy UVB models presented in Section 4.2 integrate over the ionizing emissivity from galaxies with $M_{\rm AB} \le -18$. In principle this cut-off may cause us to overestimate the scatter in $\tau_{\rm eff}$ since we are neglecting contributions from fainter galaxies that are less biased with respect to the density field. To estimate the magnitude of this effect we calculated our UVB at $z = 5.62$ while varying the upper limit in $M_{\rm AB}$ from −21 to −18, adjusting $f_{\rm ion}$ to achieve the same mean transmitted Lyα flux in each case. The results for $P(\le \tau_{\rm eff})$ are shown in Fig. A7. As expected, models that include only contributions from rarer, brighter galaxies, which we assign to more massive haloes, show a broader range in $\tau_{\rm eff}$. We find, however, that $P(\le \tau_{\rm eff})$ is essentially converged when integrating up to $M_{\rm AB} = -19$. Decreasing the galaxy duty cycle from unity essentially pushes the sources down to lower mass haloes, which we find has little affect on $P(\le \tau_{\rm eff})$. We also find no dependence on the assumed galaxy UV spectral slope.

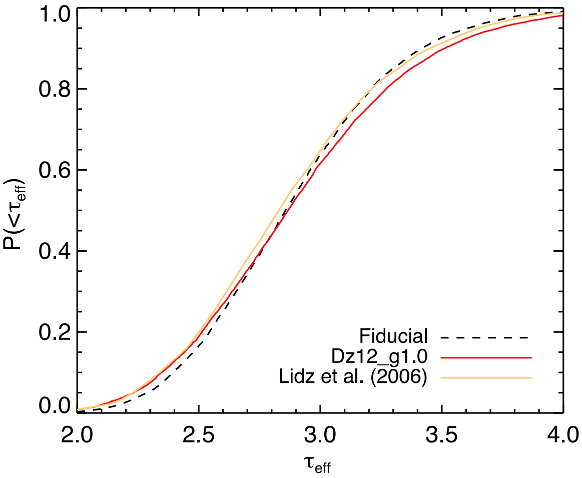

Figure A6. Dependence of $P(\le \tau_{\rm eff})$ on the thermal history of the IGM. The dashed and red solid lines show $P(\le \tau_{\rm eff})$ at $z = 5.62$ for the corresponding thermal histories plotted in Fig. A5. In this case we have computed $\tau_{\rm eff}$ over 40 Mpc h^{-1} lines of sight and rescaled the photoionization rates to produce $\langle F \rangle = 0.06$ in order to compare with $P(\le \tau_{\rm eff})$ from Lidz et al. (2006), shown as a yellow solid line. We note that the Lidz et al. $P(\le \tau_{\rm eff})$ is computed at $z = 5.7$; however, we find that small differences in redshift have relatively little impact provided that $\langle F \rangle$ remains fixed.

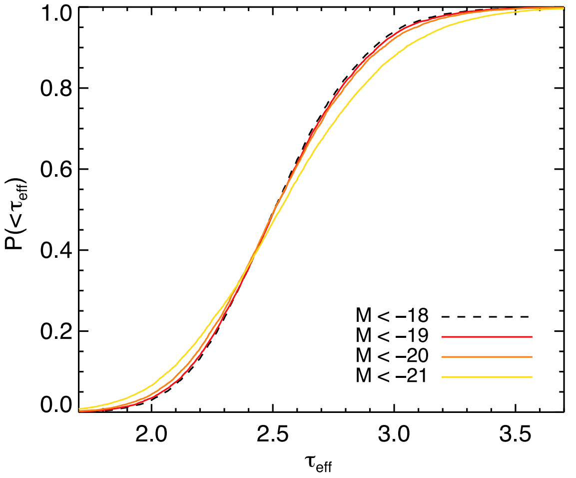

Figure A7. Dependence of $P(\le \tau_{\rm eff})$ in our galaxy UVB models on the limiting absolute magnitude of the sources. In each case we compute $\tau_{\rm eff}$ values at $z = 5.62$ over 50 Mpc h^{-1} lines of sight after scaling the total ionizing emissivity such that $\langle F \rangle = 0.084$. The width of $P(\le \tau_{\rm eff})$ increases as the sources become increasingly rare. At our fiducial cut-off ($M_{\rm AB} \le -18$), however, $P(\le \tau_{\rm eff})$ appears to be well converged.

A5 Lyα/Lyβ ratio

Finally, we examine the dependence of the relationship between Lyα and Lyβ opacity on box size and mass resolution. Since Lyβ effectively probes higher density gas, $\tau^{\beta}_{\rm eff}$ is expected to converge more quickly than $\tau^{\alpha}_{\rm eff}$ in SPH simulations. Moreover, since Lyα and Lyβ probe different density ranges, the predicted $\tau^{\alpha}_{\rm eff}$ at a fixed $\tau^{\beta}_{\rm eff}$ may depend on the simulation parameters. This effect is demonstrated in

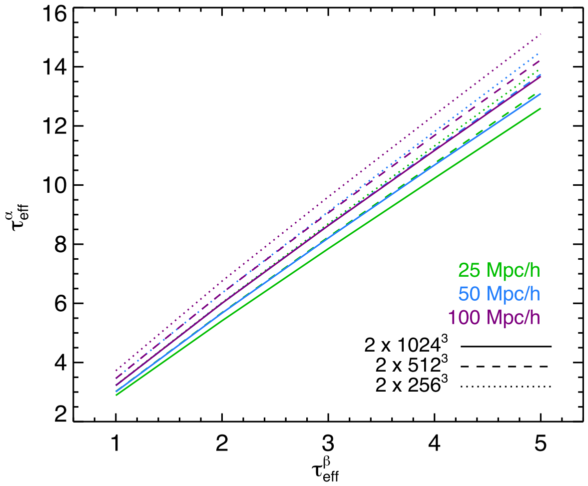

Figure A8. Convergence of relationship between $\tau_{\rm eff}^{\alpha}$ and $\tau_{\rm eff}^{\beta}$ with box size and resolution. Colours and line styles denote different box sizes and particle numbers, as in Fig. A3. For each combination of these parameters we compute $\tau_{\rm eff}^{\alpha}$, averaged over 5000 lines of sight at $z = 5.62$, as a function of $\tau_{\rm eff}^{\beta}$. Different values of $\tau_{\rm eff}^{\beta}$ are achieved by adjusting the photoionization rate and do not include a contribution from foreground Lyα absorption. The relationship between $\tau_{\rm eff}^{\alpha}$ and $\tau_{\rm eff}^{\beta}$, which probe different density ranges, depends relatively little on box size at a fixed mass resolution; however, it is not well converged with mass resolution over the range covered here.

Fig. A8, where we plot $\tau_{\rm eff}^{\alpha}$ as a function of $\tau_{\rm eff}^{\beta}$ for our nine convergence test runs. Box size has relatively little effect; however, $\tau_{\rm eff}^{\alpha}$ is lower in runs with finer mass resolution. This is again due to the fact that the centres of voids are more highly evacuated, and therefore more transparent, in runs that use a smaller gas particle mass. This has a greater impact on Lyα than on Lyβ. We note that we have neglected foreground Lyα absorption in the Lyβ forest for this test. Our upper limit for $\tau_{\rm eff}^{\alpha}$ for the trough in ULAS J0148+0600 based on the Lyβ opacity, for which we used the 25–1024 run, should nevertheless be conservative in terms of numerical convergence.

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