Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **514,** 55–76 (2022) https://doi.org/10.1093/mnras/stac1046

Hydrogen reionization ends by $z = 5.3$: Lyman-α optical depth measured by the XQR-30 sample

Sarah E. I. Bosman [1★] Frederick B. Davies [1] George D. Becker [2] Laura C. Keating [3] Rebecca L. Davies,[4,5] Yongda Zhu [2] Anna-Christina Eilers [6†] Valentina D'Odorico [7,8] Fuyan Bian,[9] Manuela Bischetti,[7,10] Stefano V. Cristiani [7] Xiaohui Fan,[11] Emanuele P. Farina [12] Martin G. Haehnelt,[13,14] Joseph F. Hennawi [15,16] Girish Kulkarni [17] Andrei Mesinger [8] Romain A. Meyer,[1] Masafusa Onoue,[1] Andrea Pallottini [7] Yuxiang Qin [18,5] Emma Ryan-Weber [4,5] Jan-Torge Schindler [1,16] Fabian Walter [1] Feige Wang [11†] and Jinyi Yang [11]

[1]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[2]*Department of Physics & Astronomy, University of California, Riverside, CA 92521, USA*
[3]*Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, Potsdam D-14482, Germany*
[4]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia*
[5]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*
[6]*MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Ave., Cambridge, MA 02139, USA*
[7]*INAF - Osservatorio Astronomico di Trieste, Via Tiepolo 11, I-34143 Trieste, Italy*
[8]*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*
[9]*European Southern Observatory, Alonso de Córdova 3107, Casilla 19001, Vitacura, Santiago 19, Chile*
[10]*INAF - Osservatorio Astronomico di Roma, Via Frascati 33, I–00078 Monte Porzio Catone, Italy*
[11]*Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA*
[12]*Max Planck Institut für Astrophysik, Karl–Schwarzschild–Straße 1, D-85748 Garching bei München, Germany*
[13]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[14]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[15]*Department of Physics, Broida Hall, University of California, Santa Barbara Santa Barbara, CA 93106-9530, USA*
[16]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[17]*Department of Theoretical Physics, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India*
[18]*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*

Accepted 2022 April 7. Received 2022 March 22; in original form 2021 August 11

ABSTRACT

The presence of excess scatter in the Ly-α forest at $z \sim 5.5$, together with the existence of sporadic extended opaque Gunn-Peterson troughs, has started to provide robust evidence for a late end of hydrogen reionization. However, low data quality and systematic uncertainties complicate the use of Ly-α transmission as a precision probe of reionization's end stages. In this paper, we assemble a sample of 67 quasar sightlines at $z > 5.5$ with high signal-to-noise ratios of >10 per ≤ 15 km s^{-1} spectral pixel, relying largely on the new XQR-30 quasar sample. XQR-30 is a large program on VLT/X-Shooter which obtained deep (SNR $>$ 20 per pixel) spectra of 30 quasars at $z > 5.7$. We carefully account for systematics in continuum reconstruction, instrumentation, and contamination by damped Ly-α systems. We present improved measurements of the mean Ly-α transmission over $4.9 < z < 6.1$. Using all known systematics in a forward modelling analysis, we find excellent agreement between the observed Ly-α transmission distributions and the homogeneous-UVB simulations Sherwood and Nyx up to $z \leq 5.2$ ($<1\sigma$), and mild tension ($\sim 2.5\sigma$) at $z = 5.3$. Homogeneous UVB models are ruled out by excess Ly-α transmission scatter at $z \geq 5.4$ with high confidence ($>3.5\sigma$). Our results indicate that reionization-related fluctuations, whether in the UVB, residual neutral hydrogen fraction, and/or IGM temperature, persist in the intergalactic medium until at least $z = 5.3$ ($t = 1.1$ Gyr after the big bang). This is further evidence for a late end to reionization.

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

1 INTRODUCTION

The epoch of reionization, during which the bulk of intergalactic hydrogen became ionized, is of great interest for both astrophysicists

★ E-mail: bosman@mpia.de
† NASA Hubble Fellow

© 2022 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

and cosmologists. The timing and morphology of the transition relate to the properties of the first galaxies and other potential reionizing sources, holding crucial information on the large-scale properties of the intergalactic medium (IGM) as well as galaxy formation and evolution at early cosmic times (see e.g. Dayal & Ferrara 2018). Next-generation 21-cm experiments aim to directly detect the signature of neutral gas in the first stages of reionization at $z \gtrsim 10$ within the coming decade (DeBoer et al. 2017; Trott & Pober 2019). Meanwhile, the end stages of reionization at $z \lesssim 7$ are already being probed through quasar absorption in the Lyman-α (Ly-α) and Lyman-β (Ly-β) hydrogen transitions (e.g. Fan et al. 2002; Mesinger & Haiman 2004; Mortlock et al. 2011; Bosman & Becker 2015; Greig et al. 2017; Davies et al. 2018b; Eilers et al. 2019; Wang et al. 2020).

Over 400 quasars are now known at $z > 5.7$, corresponding to the first billion years after the big bang (Bosman 2020). The first observational constraints on the end of reionization originated from detections of Gunn–Peterson (GP; Gunn & Peterson 1965) troughs at $z > 6$: total absorption of quasar continuum emission by neutral hydrogen in the IGM (Fan et al. 2000, 2006). Saturation of Ly-α absorption occurs in the presence of IGM gas with a hydrogen neutral fraction $\gtrsim$0.01 per cent, with dependence on the density and temperature of the gas. The interpretation of GP troughs for reionization is complex. Measurements of Ly-α transmission towards quasars have revealed that saturation occurs sporadically down to $z \sim 5.6$, and also on very large contiguous scales $\gtrsim$100 cMpc h^{-1} (Becker et al. 2015). Observed differences in Ly-α optical depth between sightlines at fixed redshift far exceed expectations from cosmic density fluctuations alone (Becker et al. 2015; Bosman et al. 2018; Eilers, Davies & Hennawi 2018; Yang et al. 2020), implying a more protracted or 'patchy' end of reionization than was unforeseen by standard models (but see Lidz, Oh & Furlanetto 2006; Mesinger 2010).

Determining the nature of these $z < 6$ optical depth fluctuations is currently a major goal of reionization theory. The existence of late-persisting GP troughs and the observed optical depth scatter at $z \sim 5.8$ can be matched by a late end of reionization in which some voids with hydrogen neutral fractions >10 per cent persist down to $z \sim 5.6$ (Kulkarni et al. 2019; Keating et al. 2020; Nasir & D'Aloisio 2020). Roughly half of the cosmic volume would then be occupied by neutral gas at $z \sim 7$, with important consequences such as e.g. facilitating the observation of the 21-cm signal (Raste et al. 2021; Šoltinský et al. 2021). In addition, scatter in the Ly-α optical depth at $z \sim 5.8$ can also arise from a short and fluctuating photon mean free path, which alters the propagation of ionizing photons through the IGM (Davies & Furlanetto 2016; D'Aloisio et al. 2018). Recent observations have suggested a shorter-than-expected ionizing mean free path at $z \sim 6$ (Becker et al. 2021; see also Bosman 2021). While not explicitly requiring a late end to reionization, a short mean free path at $z \sim 6$ poses tight requirements on the ionizing power of early galaxies (Cain et al. 2021; Davies et al. 2021). Further models have explored the importance of additional sources of scatter, such as relic IGM temperature fluctuations (D'Aloisio, McQuinn & Trac 2015; Keating, Puchwein & Haehnelt 2018) or a potential significant role of quasars (Chardin, Puchwein & Haehnelt 2017; Meiksin 2020). Meanwhile, observations of Ly-α transmission at $z \lesssim 5$ are fully consistent with IGM models including only the effects of density fluctuations in a homogeneous (i.e. permeated) ultra-violet background (UVB) (Becker et al. 2013, 2015; Rollinde et al. 2013). The transition between these two regimes across $5.0 \leq z \leq 5.8$ therefore holds crucial clues to the changes in IGM properties as reionization finishes.

Furthermore, the first measurements of the Ly-α optical depth distribution at $z \sim 5.8$ with large quasar samples have lead to the first results from semi-numerical models of reionization's patchy end stages. The Bayesian inference enabled by these semi-numerical models allowed us to statistically constrain the end of reionization to $z < 5.6$ (Choudhury, Paranjape & Bosman 2021; Qin et al. 2021), as well as disfavor a strong evolution of the ionizing escape fraction in reionizing galaxies (Qin et al. 2021).

The advent of expensive and specifically tuned simulations, as well as sensitive inference models of reionization, necessitate that measurements of Ly-α optical depth have a firm grasp on possible observational biases. To capture cosmic variance, studies require very large samples of quasars – but until now, this has come at the expense of data homogeneity and potential instrumental and reduction biases which are not known accurately (Bosman et al. 2018). There has been some tension between results from different groups (Bosman et al. 2018; Eilers et al. 2018; Yang et al. 2020) which can be largely attributed to differing choices of methods for reconstructing the underlying quasar emission (Bosman et al. 2021, thereafter B21). Out of necessity (insufficient data) and low relative importance compared to sample size, a rigorous quantitative examination of those biases and uncertainties has been neglected until now. Indeed, the existence of excess scatter in Ly-α optical depth at $z \sim 5.7$ is established very robustly even with the most pessimistic assumptions on measurement errors (Becker et al. 2015). The rigour and precision required for quantitative inference and comparison to new models, however, requires a higher level of attention to observational biases and uncertainties. The XQR-30 sample (D'Odorico et al. in preparation) consisting of 30 new high-SNR spectra of $z \gtrsim 5.8$ quasars, enables such a careful analysis for the first time without sacrificing sample size.

In this paper, we use the XQR-30 sample together with archival spectra of equal quality to significantly refine measurements of Ly-α optical depth at $5.0 \leq z \leq 6.0$. The observational data is described in Section 2. Restricting the analysis to high-quality data enables the suitable treatment of a slew of systematics and rigorous error estimation, which we describe in Section 3. We present the new distributions in Section 4. Finally, we compare our measurements with expectations from a homogeneously-ionized Universe in Section 5. The comparison to models both tests whether our analysis has accounted for all significant systematics at $z \sim 5$, where no reionization-related fluctuations are expected in Ly-α transmission, and quantifies the point of transition beyond which these fluctuations are detected. We summarize our results in Section 6.

Throughout the paper we assume a Planck Collaboration (2020) cosmology with $H_0 = 67.74$ and $\Omega_m = 0.3089$. Wavelengths always refer to the rest-frame unless explicitly stated. Comoving and proper distances are always labelled explicitly (e.g. cMpc).

2 DATA

2.1 XQR-30

We primarily use data from the XQR-30 program (1103.A-0817(A)), which is ongoingly building a legacy sample of high-resolution spectra of 30 quasars at $z \gtrsim 5.8$ with the X-Shooter instrument (Vernet et al. 2011) on the *Very Large Telescope*. An example spectrum from the program is shown in Fig. 1. The XQR-30 quasars were selected to have the highest apparent luminosities at $z > 5.7$. Observations were carried out using the 0.9 and 0.6 arcsec slits in the visible and near-infrared arms of X-Shooter, respectively. We use 25 quasars from the XQR-30 sample which do not show strong broad absorption lines (BALs) precluding the modelling of the intrinsic continuum. We however retain the BAL quasars ATLAS J2211-3206, PSO J239-

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

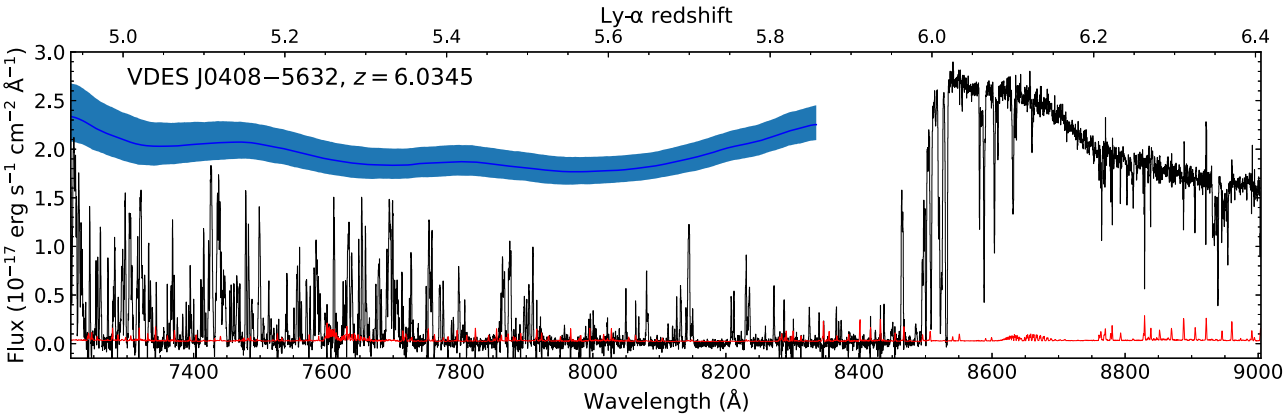

Figure 1. X-Shooter spectrum of the Ly-α transmission region in the XQR-30 quasar VDES J0408-5632 at $z = 6.0345$. The flux uncertainty is shown in red and the PCA-reconstructed continuum and its 1σ uncertainties are shown in blue. The PCA reconstruction is plotted over the wavelength range 1126 Å $< \lambda <$ 1185 Å which we use in the mean flux measurement. The pixel scale is 10 km s^{-1} and the SNR of the Ly-α region (reconstruction divided by uncertainty) is SNR = 86. The exposure time was 13.5 h.

Table 1. XQR-30 quasars with X-Shooter spectra included in this work. References correspond to (Discovery, Redshift determination). The full list of references is given in the caption of Table 3.

Quasar ID	z_{qso}	SNR per pixel	Refs.
PSO J323+12	6.5872	35.9	(1,27)
PSO J231-20	6.5869	42.3	(1,27)
VDES J0224-4711	6.5223	24.4	(3,39)
PSO J1212+0505	6.4386	55.8	(1,4)
DELS J1535+1943	6.3932	22.6	(5,–)
ATLAS J2211-3206	6.3394	37.5	(6/7,4)
PSO J060+24	6.192	49.7	(8,–)
PSO J065-26	6.1871	77.9	(8,27)
PSO J359-06	6.1722	68.8	(9,40)
PSO J217-07	6.1663	33.3	(8,8)
PSO J217-16	6.1498	73.0	(8,4)
PSO J239-07	6.1102	56.3	(8,40)
SDSS J0842+1218	6.0754	83.2	(11/12,27)
ATLAS J158-14	6.0685	60.3	(6,40)
VDES J0408-5632	6.0345	86.6	(3,3)
ATLAS J029-36	6.021	57.1	(14,13)
SDSS J2310+1855	6.0031	113.4	(15,16)
PSO J007+04	6.0015	54.4	(12/17,27)
PSO J029-29	5.984	65.6	(8,8)
PSO J108+08	5.9485	104.8	(8,8)
PSO J183-12	5.917	61.8	(17,–)
PSO J025-11	5.844	50.6	(8,–)
PSO J242-12	5.837	22.9	(8,–)
PSO J065+01	5.833	25.1	(2,–)
PSO J308-27	5.7985	53.2	(8,2)

07, and PSO J239-07, whose BAL features are well-resolved and confined to highly ionized absorption (Bischetti et al. 2022). All XQR-30 spectra have signal-to-noise ratios (SNRs) larger than 20 per 10 km s^{-1} pixel measured over 1165 Å $< \lambda <$ 1170 Å (Table 1). We use the reconstructed continua described in Section 3.2 to calculate the SNR over the range most relevant to our study. The X-Shooter instrument has a resolution of ∼34 km s^{-1} in the visible (5500 Å $< \lambda <$ 10200Å) and ∼37 km s^{-1} in the infrared (10 200 Å $< \lambda <$ 24 800 Å), although better-than-average seeing during observations means the effective resolution is slightly higher. Observations are first flat-fielded and sky-subtracted following the method of Kelson (2003), then the spectra extracted optimally (Horne 1986) separately for the visible and infrared arms of the instrument. Our reductions routines are described in more detail in Becker, Rauch & Sargent (2009); further details, including comparisons with the publicly available FESOREX (Freudling et al. 2013) and PYPEIT (Prochaska et al. 2020a) pipelines for X-Shooter, will be presented in D'Odorico et al. (in preparation). The optical and infrared arms are then stitched together over the 10 110 Å $< \lambda_{obs} <$ 10 130 Å spectral window, by rescaling the infrared spectrum to match the observed mean flux in the optical arm after two rounds of sigma-clipping and discarding of all pixels with SNR<2. The spectrum is then interpolated over the overlap window. This somewhat aggressive procedure is adopted to minimize the risk of creating an artificial 'step' in the spectrum between the arms, to which the continuum-fitting method may be non-linearly sensitive (c.f. Section 3.2).

The quasar spectra used int his work are listed in Tables 1, 2 and 3.

2.2 Other X-Shooter spectra

We supplement the XQR-30 quasars with 26 archival X-Shooter spectra of equal SNR $>$ 10 per 10 km s^{-1} pixel from the literature (Table 2), including three quasars at $5.5 < z < 5.7$ to better sample the Ly-α transmission at $z \leq 5.3$. The spectra were reduced in an identical manner to the XQR-30 quasars, except six of them which had already been reduced with PYPEIT. PYPEIT is an open-source PYTHON package designed to automate the reduction of spectroscopic data for (currently) 28 different spectrographs (Prochaska et al. 2020b). Similarly to our custom reduction pipeline, PYPEIT performs joint extraction of objects and a model of sky emission in each observed frame. We conducted a comparative analysis on a sub-sample of quasars reduced via both methods, which showed only a negligible (<0.5 per cent) effect on the large-scale Ly-α transmission.

2.3 ESI spectra

Finally, we also complement our sample with 16 archival spectra of $z \gtrsim 5.7$ quasars taken by the ESI instrument (Sheinis et al. 2002) on the *Keck Telescope* (Table 3). The spectral resolution of ESI is lower than that of X-Shooter, at ∼60 km s^{-1}, and ESI's wavelength coverage only includes the optical up to $\lambda <$ 10 500 Å. While we strive to reduce systematics arising from instrument and data reduction by minimizing the number of different instruments, we include ESI

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

Table 2. Quasars with literature and archival X-Shooter spectra included in this work. References correspond to (Discovery, Redshift determination). The full list of references is given in the caption of Table 3.

Quasar ID	z_{qso}	SNR per pixel	Refs.
PSO J036+03	6.5405	61.4	(18,27)
PSO J011+09	6.4695	14.5	(1,40)
PSO J159-02	6.386	22.9	(8,–)
SDSS J0100+2802	6.3269	560.5	(20,27)
ATLAS J025-33	6.318	127.3	(14,13)
SDSS J1030+0524	6.309	69.6	(21,22)
VDES J0330-4025	6.239	17.0	(3,10)
PSO J308-21	6.2355	24.4	(8,27)
VIK J2318-3029	6.1456	16.5	(7,27)
ULAS J1319+0950	6.1347	81.7	(23,27)
CFHQS J1509-1749	6.1225	43.0	(24,4)
CFHQS J2100-1715	6.0807	12.4	(25,27)
ULAS J1207+0630	6.0366	29.2	(12,4)
SDSS J1306+0356	6.033	65.3	(21,27)
PSO J340-18	5.999	29.9	(17,13)
ULAS J0148+0600	5.998	152.0	(12,13)
SDSS J0818+1722	5.997	132.1	(19,13)
VIK J0046-2837	5.9926	15.0	(28,29)
PSO J056-16	5.9676	32.0	(8,40)
PSO J004+17	5.8166	15.9	(8,40)
SDSS J0836+0054	5.804	73.8	(21,–)
SDSS J0927+2001	5.7722	53.8	(19,26)
PSO J215-16	5.7321	30.2	(31,31)
J1335-0328	5.693	35.0	(32,13)
J0108+0711	5.577	20.0	(32,13)
J2207-0416	5.529	16.9	(9,13)

spectra with SNR $>$ 10 per 15 km s^{-1} pixel since they constitute the largest collection of deep, publicly available observations of $z > 5.7$ quasar spectroscopy with a single spectrograph besides X-Shooter. The ESI spectra were reduced using the same methods and algorithms as our X-Shooter pipeline, applying optimal spectral extraction after flat-fielding and sky subtraction. All but three of the ESI spectra we employ were also included in the 'GOLD' sample of Bosman et al. (2018), where their reduction is further described. The three new spectra were reduced in an identical manner, but were not included in Bosman et al. (2018) due to the availability of deeper MMTRCS (Schmidt, Weymann & Foltz 1989) or HIRES spectroscopy (Vogt et al. 1994). Here we prefer the slightly shallower ESI spectra in order to preserve instrumental consistency and reduce possible instrumentation systematics. In a preliminary study (B21), we analysed the impact of ESI's lesser resolution and wavelength coverage on systematics arising from quasar continuum reconstruction in the context of Ly-α transmission. We found that while continuum uncertainties were increased by $\sim$50 per cent compared to using spectra with X-Shooter's wavelength coverage, no systematic biases arose. Six of our X-Shooter spectra were also observed to SNR $>$ 10 depth by ESI, enabling an empirical test of potential biases linked to instrumentation which we present in Section 3.3.

3 METHODS

The effective Ly-α optical depth $\tau_{\rm eff}$ is defined using the mean transmitted flux fraction in the Ly-α forest,

$$\tau_{\rm eff} = -\ln \left\langle \frac{F(\lambda)}{F_{\rm cont}(\lambda)} \right\rangle , \quad (1)$$

Table 3. Quasars with archival ESI spectra included in this work. References relate to (Discovery, Redshift determination). (–) This paper; (1) Mazzucchelli et al. (2017); (2) D'Odorico et al. (in preparation); (3) Reed et al. (2017); (4) Decarli et al. (2018); (5) Wang et al. (2019); (6) Chehade et al. (2018); (7) Farina et al. (2019); (8) Bañados et al. (2016); (9) Wang et al. (2016); (10) Eilers et al. (2020); (11) De Rosa et al. (2011); (12) Jiang et al. (2015); (13) Becker et al. (2019); (14) Carnall et al. (2015); (15) Jiang et al. (2016); (16) Wang et al. (2013); (17) Bañados et al. (2014); (18) Venemans et al. (2015); (19) Fan et al. (2006); (20) Wu et al. (2015); (21) Fan et al. (2001); (22) Jiang et al. (2007); (23) Mortlock et al. (2009); (24) Willott et al. (2007); (25) Willott et al. (2010); (26) Wang et al. (2010); (27) Venemans et al. (2020); (28) Venemans et al. (2018); (29) Schindler et al. (2020); (30) Kurk et al. (2007); (31) Morganson et al. (2012); (32) Yang et al. (2017); (33) Fan et al. (2003); (34) Willott, Bergeron & Omont (2015); (35) Shen et al. (2019); (36) Fan et al. (2004); (37) Jiang et al. (2008); (38) Fan et al. (2000); (39) Wang et al. (2021); (40) Eilers et al. (2021). For quasars without discovery papers, we reference the first paper which showcased or used a spectrum featuring broad emission lines.

Quasar ID	z_{qso}	SNR per pixel	References
SDSS J1148+5251	6.4189	118.8	(33,34)
CFHQS J0050+3445	6.251	28.6	(25,35)
SDSS J1623+3112	6.254	16.4	(36,35)
SDSS J1250+3130	6.138	41.2	(19,35)
SDSS J2315-0023	6.124	14.6	(37,13)
SDSS J1602+4228	6.083	24.1	(36,35)
SDSS J1630+4012	6.066	10.3	(33,35)
SDSS J0353+0104	6.057	15.4	(37,35)
SDSS J2054-0005	6.0389	22.6	(37,27)
SDSS J1137+3549	6.009	23.2	(19,35)
SDSS J1411+1217	5.904	42.1	(36,30)
SDSS J1335+3533	5.9012	10.3	(19,26)
SDSS J0005-0006	5.847	18.4	(36,13)
SDSS J0840+5624	5.8441	34.9	(19,26)
SDSS J0002+2550	5.818	119.0	(36,35)
SDSS J1044-0125	5.7846	64.9	(38,27)

where F is the observed flux, $F_{\rm cont}$ is the reconstructed intrinsic quasar continuum, and $\langle\rangle$ is the mean over a fixed interval, traditionally taken to be 50 cMpc h^{-1} (see Section 3.1). The usable range of observed wavelengths is limited by the quasar's effect on its environment on one hand and overlap with Ly-β absorption on the other. To exclude the effect of the background quasars (the so-called 'proximity zone'; Cen & Haiman 2000; Carilli et al. 2010; Eilers et al. 2017) we restrict ourselves to $\lambda < 1185$ Å, beyond which no effect on Ly-α transmission is seen even in the deepest spectral stacks (<0.5 per cent Ly-α flux increas; Bosman et al. 2018). No quasars are known to have proximity zones extending beyond 1185 Å: the longest $z >$ five proximity zone, in quasar SDSS J0100+2802, only extends to $\gtrsim$1189 Å. In fact, we note that our proximity zone cut may be overly conservative, since no effect is seen in deep stacks even at $\lambda < 1195$ Å at $z > 6.1$ and the more conservative cut reduces the probed volume at $z > 6.0$ by $\sim$30 per cent for our sample.

To exclude contamination by the overlapping Ly-β forest at low wavelengths, the redshift of the background quasar must be known precisely to determine its location with respect to the foreground IGM. When possible, we adopt the systemic redshifts of the quasar host galaxies, determined through the identification of sub-mm emission lines (these redshifts can roughly be identified in Tables 1, 2, and 3 by having five significant digits). Redshifts may also be obtained from our rest-UV spectra directly using the quasar broad emission lines, but these features are often blue-shifted from the quasar host galaxies and from each other (Meyer, Bosman & Ellis 2019b; Onoue et al. 2020; Schindler et al. 2020) with a large

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

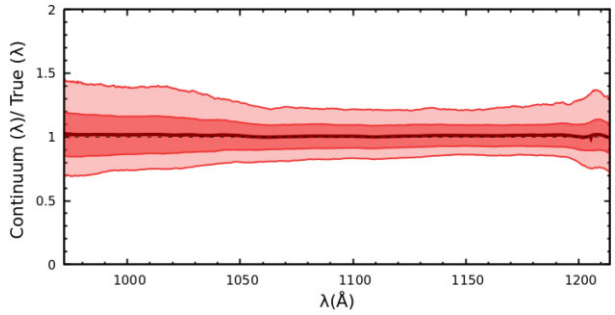

Figure 2. Residuals in the PCA reconstruction of the $\lambda < 1220$ Å blue-side continua of 4597 eBOSS quasars at $2.7 < z < 3.5$, which were not used for training the PCA. No significant wavelength-dependent biases are seen. The average uncertainty over the $1026 < \lambda < 1185$ Å range, used in this paper, is $-7.9/+7.8$ per cent.

scatter $\sim$750 km s^{-1}. An alternative method, which we use here, is to adopt the redshift of the first Ly-α absorber in front of the quasar (Worseck et al. 2014). This method for locating the onset of the IGM has been shown to have relatively little offsets and scatter with respect to sub-mm emission lines, $\Delta v = 180 \pm 180$ km s^{-1} (Becker et al. 2021). We employ it here for cases where fits to the Mg II broad emission line are complicated by absorption, as indicated in Tables 1, 2, and 3.

To err on the side of caution, we round up the Ly-β wavelength of 1025.7 Å and only use wavelengths $\lambda > 1026$ Å. While the presence of the O VI broad emission line renders the continuum prediction slightly more uncertain over the 1026 Å $< \lambda \lesssim 1050$ Å wavelength range, this is carefully quantified and propagated to all our measurements and model comparisons (see Fig. 2 and Section 3.2). We note that even if we use on occasion wavelengths contaminated by Ly-β absorption due to chance redshift errors, the corresponding Ly-β-absorbing gas would be located inside the quasar's proximity zone, and the Ly-β absorption should therefore be relatively small (although difficult to quantify in a model-independent manner).

The data reduction procedure in principle automatically rejects outlier pixels (e.g. cosmic rays) when a large number of exposures are stacked. Nevertheless, we exclude a few ($\lesssim$0.05 per cent of total) anomalous pixels which are flagged if their SNR at the unabsorbed continuum level is <2 per pixel (since an average SNR$\geq$10 is enforced for all our observations) or if pixels have negative flux at $>3\sigma$ significance. Such sigma-clipping can by definition only induce a bias $\ll$0.1 per cent, while it cleans up features which are clearly reduction glitches.

3.1 Redshift or length intervals?

The traditional way of quantifying Ly-α optical depth fluctuations, motivated by efficiency when dealing with small sample sizes and by ease of comparison to theoretical models, has been to divide Ly-α transmission spectra in intervals of constant length (Becker et al. 2015; Bosman et al. 2018; Eilers et al. 2019; Yang et al. 2020). In this approach, the average transmission beyond a quasar's proximity zone is calculated over consecutive bins of fixed length (usually $\Delta L = 50$ cMpc h^{-1}) with variable starting and ending points, and these measurements are then assigned to a redshift interval depending on the mid-point of each bin. We reproduce this approach for the purposes of comparison with the literature, but in our fiducial results we modify it for the purposes of comparison with theoretical models for two main reasons. First, the fixed length definition makes it possible for the same quasar sightline to contribute to the optical depth distribution in a single redshift bin more than once. This is a source of unwanted covariance, since the IGM optical depth is known to be correlated on scales up to 100 cMpc h^{-1} (Becker et al. 2015). Secondly, the definition implies that up to half of the pixels contributing to an optical depth measurement at a given redshift may be located outside of the redshift bin's bounds. The result is artificial scatter in measured optical depth, especially since Ly-α optical depth evolves very quickly at $z > 5$ (B21; see also Worseck et al. 2016). To circumvent these issues, we instead directly measure the opacity in bins with fixed starting and ending points of constant length in redshift space. We divide the spectra in bins of $\Delta z = 0.1$ centred at $z = 5.0, 5.1, ..., 6.0$ corresponding to comoving lengths of $l = 36.0, 35.2, ..., 29.3$ cMpc h^{-1}. We retain measurements if >50 per cent of the corresponding wavelength range is usable. In practice, $\sim$30 per cent of sightlines are truncated by more than 10 per cent; we propagate the resulting uncertainties throughout the analysis.

For the purposes of future comparisons of the data with models, sometimes binning in equal length intervals, with fixed endpoints in redshift, might be preferable. We show the resulting distributions for $\Delta L = 50$ cMpc h^{-1} in Appendix A. Full measurements for $\Delta L = 30$, 50, 100 cMpc h^{-1} and $\Delta z = 0.05$, 0.1 are also available as online material.

3.2 Continuum reconstruction

We employ Principal Component Analysis (PCA) to reconstruct $F_{\rm cont}(\lambda)$ based on the observed quasar continuum at $\lambda > 1280$ Å. Quasar continuum PCA models use a training set of low-z quasar spectra to find optimal linear decompositions of the 'known' red side ($\lambda > 1280$ Å) and the 'unknown' blue side of the spectrum ($\lambda <$ 1220 Å), then determines an optimal mapping between the linear coefficients of the two sides' decompositions (Francis et al. 1992; Yip et al. 2004; McDonald et al. 2005; Suzuki et al. 2005; Pâris et al. 2011; Ďurovčíková et al. 2020). In B21, we conducted a rigorous comparison of the precision and accuracy of six reconstruction techniques used in the literature by using a large sample of 'blind' tests with spectra where the true continuum was known. We found that two PCA methods outperformed both the more traditionally employed power-law extrapolation (e.g. Bosman et al. 2018) and 'stacking of neighbours' methods, both in prediction accuracy and in lack of wavelength-dependent reconstruction residuals. Here, we use a further improved version of the most accurate PCA method identified in B21, the log-PCA approach of Davies et al. (2018c) (see also Davies et al. 2018b).

Our PCA consists of 15 red-side components and 10 blue-side components. Training was performed on 4597 quasars at $2.7 < z < 3.5$ with SNR > 7 from the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013) and the SDSS-IV Extended BOSS (eBOSS; Dawson et al. 2016). Intrinsic continua were obtained automatically using a modified version of the method of Dall'Aglio, Wisotzki & Worseck (2008), originally based on the procedures outlined in Young et al. (1979) and Carswell et al. (1982). The automatically fitted continua are re-normalized to ensure they match the observed mean Ly-α transmission at $z \sim 3$ measured from high-resolution spectra (Faucher-Giguère et al. 2008; Becker et al. 2013), as they would otherwise be biased by the low spectral resolution of the SDSS spectrograph (see discussion in Dall'Aglio, Wisotzki & Worseck 2009).

Testing is performed by using an independent set of 4597 quasars from eBOSS. The asymmetric 1σ and 2σ bounds are measured

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

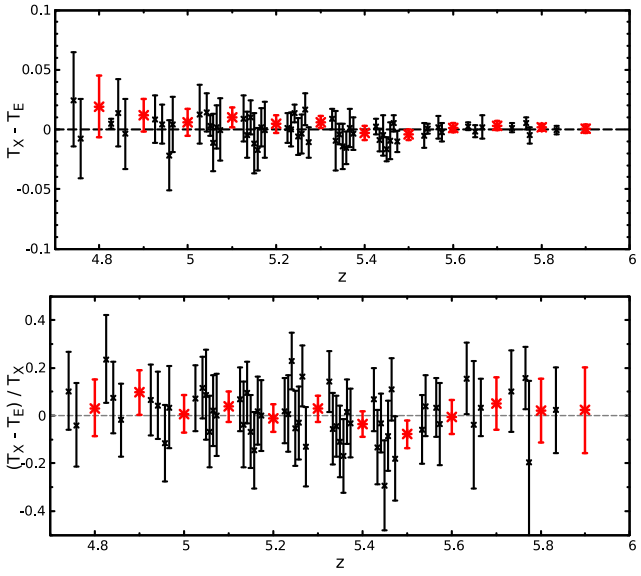

Figure 3. Differences in Ly-α transmission measured with ESI (T_E) and X-Shooter (T_X) spectra of the same six quasars. *Top:* Absolute difference in transmitted flux. The error bars account for observational uncertainties as well as continuum reconstruction uncertainties and biases. Individual measurements are shown in black with the averages in red. *Bottom:* Same as top panel, but showing the fractional differences. The diagnostics show no evidence for instrument biases beyond the ones already accounted for in the measurement uncertainties. In both plots, some random scatter along the *x*-axis (redshift) has been added to improve legibility.

by finding the central 68th and 95th percentile intervals of the prediction error in the testing sample at each wavelength. Fig. 2 shows the wavelength-dependent 1σ and 2σ continuum reconstruction uncertainties, Continuum(λ)/True(λ). No features are visible at any rest-frame wavelength in the residuals, indicating that blue-side emission lines can be reconstructed without bias. The standard deviation is PCA/True $-1 = 0.8^{+7.8}_{-7.9}$ per cent, i.e. the method predicts the underlying continuum within 8 per cent: a large improvement compared to power-law extrapolation methods (>13 per cent) and a slight improvement over the best PCA in B21 (9 per cent). For the ESI spectra covering a shorter red-side wavelength range, we use the 'optical-only' PCA developed in B21 with PCA/True $-1 = 1.0$ per cent$^{+11.2}_{-11.3}$. The lower accuracy is unsurprising since fewer features are available to the PCA modelling. However, no significant wavelength-dependent biases are present.

In the rest of the paper, we always correct for the residual wavelength-dependent mean bias (<1 per cent) to our reconstructions of $F_{\rm cont}(\lambda)$ and forward-model the full wavelength-dependent uncertainties into all measurements and model comparisons. We refer the reader to B21 for further details of the PCA training and testing procedures. Figures showing all PCA fits and blue-side predictions are shown in Zhu et al. (2021) and the PCA fits for XQR-30 spectra will be made public with the first XQR-30 data release (D'Odorico et al. in preparation).

3.3 Instrumental effects

To empirically check whether our data reduction and continuum reconstruction methods have accounted for all differences between ESI and X-Shooter spectra, we compare optical depth measurements for six quasars which have deep spectra with both X-Shooter and ESI. Fig. 3 shows the difference between the Ly-α transmission observed with X-Shooter, $T_X = F_{\rm X\text{-}Shooter}/F_{\rm cont,\ X\text{-}Shooter}$, and with ESI, $T_E = F_{\rm ESI}/F_{\rm cont,\ ESI}$. The continua were reconstructed using the two PCAs discussed in Section 3.2. The top and bottom panels show the absolute and fractional difference between T_X and T_E, respectively. No statistically significant bias is detected at any redshift. Across all observations, the average fractional bias between the instruments is 1.3 per cent with an observed scatter of 2.3 per cent. Since the effect is very sub-dominant compared to continuum uncertainties (∼11 per cent for ESI spectra) and we did not detect a statistically significant bias, we disregard instrumental differences between ESI and X-Shooter spectra beyond what is already included in the reduction pipelines.

3.4 DLA exclusion

Damped Ly-α absorbers (DLAs), named after their prominent Ly-α damping wings, are intervening systems along quasar sightlines with hydrogen column densities $N_{\rm H\,I} \geq 10^{20.3}$ cm^{-2} (Wolfe, Gawiser & Prochaska 2005; Rafelski et al. 2012). DLAs near quasars at $z \gtrsim 6$ can completely absorb Ly-α transmission over intervals $\Delta v = 2000$ km s^{-1}, with significant suppression of the transmission over $\Delta v \gtrsim 5000$ km s^{-1} (D'Odorico et al. 2018; Bañados et al. 2019; Davies 2020). Since reionization models typically do not include the effect of DLAs, we strive to remove them from our nominal measurements.

The detection of $z \gtrsim 5$ DLAs relies on the identification of associated low-ionization metal absorption lines, since their Ly-α absorption may not contrast against the highly-opaque IGM. DLA metallicities at $z \gtrsim 5$ are very diverse, and some can be highly sub-solar (Bañados et al. 2019), such that even relatively weak metal absorption might indicate a DLA. The identification of intervening metal absorbers in the XQR-30 sample will be described in detail in Davies et al. (in preparation). For the other quasars, we used where relevant the published lists of intervening metal systems of Cooper et al. (2019), D'Odorico et al. (2018), Meyer et al. (2019a), and Becker et al. (2019). We supplemented the literature where necessary by conducting our own metal search, following closely the standard procedure described in Bosman et al. (2017). Pairs of absorption lines corresponding to the same ion or frequently co-occurring ions (C IV, Mg II, Fe II, O I + C II) are searched for automatically before being confirmed manually. Due to the high SNR of the X-Shooter spectra, we expect to be >90 per cent complete to absorption corresponding to $\log N_{\rm Mg\,II}/{\rm cm}^{-2} \gtrsim 13$. The metal identification in the ESI spectra similarly relies on literature studies which employed infrared spectra of the objects.

We adopt the following criteria: we mask the central $\Delta v = 3000$ km s^{-1} for systems with metal column densities $\log N_{\rm C\,II}/{\rm cm}^{-2} > 13$, $\log N_{\rm O\,I}/{\rm cm}^{-2} > 13$, or $\log N_{\rm Si\,II}/{\rm cm}^{-2} > 12.5$, measured through the $\lambda = 1334.53$, 1302.16, and 1526 Å transitions, respectively. When none of these ions are accessible, we also exclude the central $\Delta v = 3000$ km s^{-1} for systems with $\log N_{\rm Mg\,II}/{\rm cm}^{-2} > 13$ based on the high rates of co-occurrence of the Mg II 2796.35, 2803.53 Å doublet (Cooper et al. 2019). We exclude a larger window of $\Delta v = 5000$ km s^{-1} around intervening systems with $\log N_{\rm OI,\ C\,II,\ Si\,II,\ Mg\,II} > 14$ cm^{-2} due to the likely presence of extended damping wings.

We do not exclude systems based on the presence of highly-ionized ions alone (e.g. C IV, Si IV) since the corresponding gas is likely highly ionized (Cooper et al. 2019). Finally, we exclude $\Delta v = 5000$ km s^{-1} around the suspected location of strong O VI associated absorption (from systems detected from strong associated

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

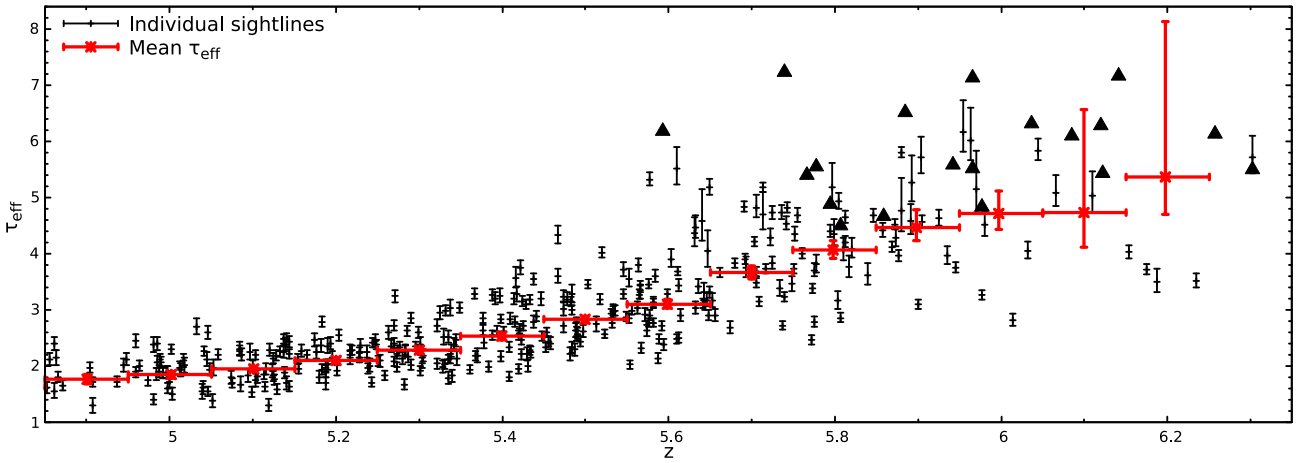

Figure 4. Mean Ly-α flux measured along 67 quasar sightlines at $4.9 < z < 6.2$, measured in consecutive 50 cMpc h^{-1} bins along each sightline (black). Non-detections are shown with upwards pointing triangles. The mean fluxes measured in intervals of $\Delta z = 0.1$ are shown with red points. Uncertainties correspond to the 16th and 84th percentile contours of a bootstrap resampling in each redshift interval. Non-detections are shown at the 2σ limit. The observed scatter between sightlines increases drastically above $z \gtrsim 5.4$.

C IV absorption), which overlaps with the Ly-α forest in quasars J1411+1217 and J1623+3112.

4 RESULTS

Fig. 4 shows the mean Ly-α transmission measured in intervals of 50 cMpc h^{-1}. The average transmission evolves smoothly over $5 < z < 6$, but an increase in scatter between measurements at equal redshift becomes clear at $z \geq 5.4$. By $z = 5.6$, the sampling of the distribution is visibly limited. The number of fully opaque Gunn–Peterson troughs with non-detections (2σ) at $\tau > 6$ increases sharply, with the first occurrence found at $z \sim 5.6$.

4.1 Mean transmission across $4.8 < z < 6.2$

We calculate the mean transmission in bins of $\Delta z = 0.1$ and give the results in Table 4. We do not weight the measurement: all pixels corresponding to Ly-α transmission inside a given redshift interval contribute equally (after the masking of bad regions as described above). The uncertainties are calculated via bootstrap re-sampling in each redshift bin. We quote the 16 per cent and 84 per cent percentiles of the bootstrap results. The observational uncertainties, taking into account only uncertainties in individual measurements, are a factor 5–10 smaller than the bootstrap uncertainties at all redshifts. The uncertainties are therefore dominated by the intrinsic width of the Ly-α transmission distribution. The mean Ly-α transmission over the range $4.8 < z < 5.7$ is empirically well-described by a linear decline of the form,

$$F_{Ly-\alpha}(1+z) = a \times (1+z) + b. \quad (2)$$

We fit this functional form to our observations using least-squares regression, and obtain best-fitting parameters $a = -0.191, b = 1.307$. Both parameters are constrained to better than 0.1 per cent. We show the resulting curve in Fig. 5. An empirical parametric description of effective Ly-α optical depth evolution with redshift, used for instance by Becker et al. (2013), is an power-law function with a constant offset of the form,

$$\tau_{\rm eff}(1+z) = \tau_0 \left(\frac{1+z}{1+z_0} \right)^{\beta} + C. \quad (3)$$

Table 4. Mean Ly-α flux transmission at $4.75 < z < 6.25$, measured in $\Delta z = 0.1$ intervals centred on the redshift given in the first column. Uncertainties correspond to the 16th and 84th percentiles from bootstrap resampling. The measurement uncertainties on their own are a factor 5–10 smaller than the bootstrap uncertainties quoted here. $N_{\rm los}$ sightlines contribute to each measurement.

z	$<F_{\rm Ly-\alpha}>$	-1σ	$+1\sigma$	$N_{\rm los}$
4.8	0.194	−0.015	+0.018	15
4.9	0.171	−0.014	+0.014	17
5.0	0.1581	−0.0089	+0.0082	37
5.1	0.1428	−0.0054	+0.0068	48
5.2	0.1222	−0.0054	+0.0046	55
5.3	0.1031	−0.0050	+0.0056	58
5.4	0.0801	−0.0048	+0.0061	64
5.5	0.0591	−0.0035	+0.0039	64
5.6	0.0447	−0.0036	+0.0033	59
5.7	0.0256	−0.0029	+0.0031	51
5.8	0.0172	−0.0028	+0.0022	45
5.9	0.0114	−0.0030	+0.0029	28
6.0	0.0089	−0.0029	+0.0033	19
6.1	0.0088	−0.0074	+0.0082	10
6.2	0.0047	−0.0044	+0.0045	8

Setting $z_0 = 4.8$, we run a least-squares regression and find best-fitting parameters $\tau_0 = 0.30 \pm 0.08$, $\beta = 13.7 \pm 1.5$, and $C = 1.35 \pm 0.12$. We fit this form to the mean optical depth over $4.8 < z < 5.9$ and show the resulting best-fitting model in Fig. 6. We sample the covariance matrix of the three parameters and calculate the upper and lower envelopes encompassing 68 per cent of the variance about the best fit, which are shown by the orange shaded region. The evolution of $\tau_{\rm eff}$ with redshift is much steeper at $z > 4.8$ than over $2 < z < 5$, where Becker et al. (2013) found a best-fitting $\beta = 2.90$.

Our measurements are in fair agreement with past literature, as shown in Fig. 5. The quasars used in this work have considerable overlap (∼30−50 per cent) with the ones employed by Becker et al. (2015), Eilers et al. (2018), Bosman et al. (2018), and Yang et al. (2020), such that differences are unlikely to be due to cosmic variance alone. Systematic differences in continuum reconstruction methods are a known cause of bias: as shown in B21, the tension

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

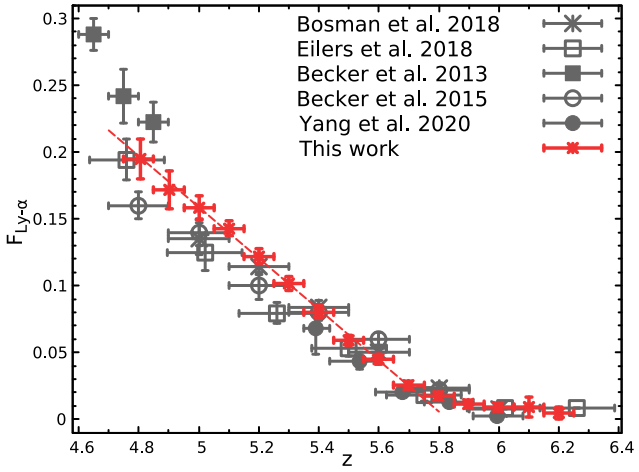

Figure 5. Average Ly-α transmission evolution with redshift. Uncertainties are obtained via bootstrap resampling in this work as well as in the literature comparison samples of Becker et al. (2013, 2015), Bosman et al. (2018), and Eilers et al. (2018), and should therefore encompass cosmic variance as long as the underlying optical depth distributions are well-sampled. Differences at $z \gtrsim 5.5$ are due to under-sampling of cosmic variance, as well as systematic biases in older work. At $z > 5.4$, cosmic variance is 5–10 times larger than measurement uncertainties. The red dashed line shows the optimal linear fit to the data over $4.8 < z < 5.7$ (equation 2, see text).

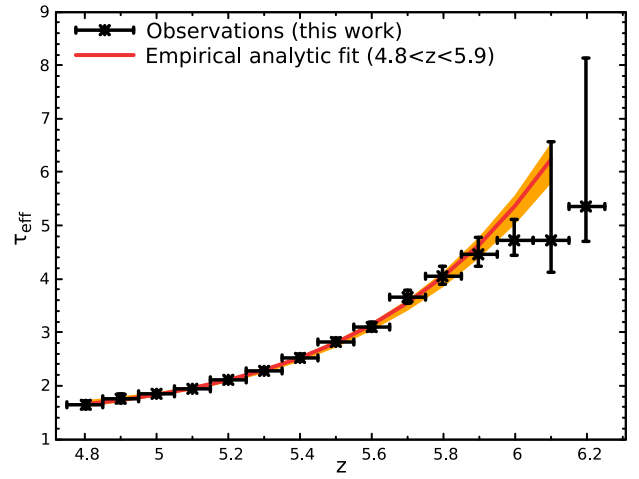

Figure 6. Evolution of $\tau_{\rm eff}$ with redshift measured in our sample across $4.8 < z < 6.2$ in steps of $\Delta z = 0.1$ (black). The red line shows the best-fitting power-law model and its 68 per cent uncertainty envelope (equation 3, see text). Observational uncertainties are obtained via bootstrap resampling.

between the measurements of Bosman et al. (2018) and Eilers et al. (2018) can be explained almost entirely by the different continuum reconstruction methods employed by the two studies. Bosman et al. (2018) employed power-law extrapolation, while Eilers et al. (2018) used a linear PCA originating in a small number of hand-fitted continua in Pâris et al. (2011). Both methods were found to introduce non-trivial wavelength-dependent biases which are virtually absent from more recent log-space PCA and neural-network-mapped PCAs (e.g. Davies et al. 2018c; Ďurovčíková et al. 2020; see B21 for details). Such biases depend sensitively on the redshifts of the background quasars and corrections unfortunately cannot be applied *post hoc*. Becker et al. (2015), Bosman et al. (2018), and Yang et al. (2020) all employed power-law reconstructions, and therefore carry similar biases; this may explain why our results are offset from all three studies in the same direction at $z < 5.4$ (where power-law-induced uncertainties and biases are the largest). We also note that Eilers et al. (2018), Bosman et al. (2018), and Becker et al. (2015) had substantial overlap in quasar sightlines, and should therefore be affected by cosmic variance in a coherent way compared to our sample. Yang et al. (2020) calculated mean optical depths by using a weighted spectral stack, without providing measurements of scatter between sightlines. In order to provide a better comparison with this study, we re-calculate the optical depth values from Yang et al. (2020)'s sample by using their published list of measured optical depths in each quasar sightline, and we estimate the cosmic variance uncertainties via bootstrap resampling.

We find a very smooth, linear evolution of the mean Ly-α transmitted flux across $4.8 < z < 5.7$, with no sudden steepening, in contrast with several past studies. Matching the results of Becker et al. (2013) at $z < 4.8$ still seems to require a faster steepening at $z \sim 4.7$ (Fig. 5). However, both their measurements and ours are at the (opposite) edges of their redshifts of validity at $z \sim 4.8$. In our study, measurements at $z < 5$ rely on $N < 20$ objects, and are the most sensitive to errors in background quasar redshift via contamination by Ly-β absorption. Conversely, the measurements of Becker et al. (2013) at $z > 4.5$ use the Ly-α forest at the shortest separations from the background quasars, where the continuum is under-predicted by power-law extrapolations due to the large width of the broad Ly-α emission line (B21); the measurements of Becker et al. (2013) are also based on spectra with much lower SNR than our sample. Properly sampling the overlapping region would therefore require a sample of deep quasar spectra of intermediate redshifts, i.e. $5 < z_{\rm qso} < 5.5$.

4.2 Optical depth distributions at $5.0 < z < 6.1$

We show the distributions of Ly-α optical depth at $5.0 < z < 6.1$ for comparison with previous studies in Fig. 7. We first use the traditional method of measuring the optical depth in constant 50 cMpc h^{-1} intervals whose centres are then binned by redshift, since this definition was employed by all past studies. We will highlight the biases created by this definition later in Fig. 8.

Non-detections of transmission over a given interval (defined at the 2σ level) give rise to lower limits on optical depth. Limits can either be represented as corresponding to flux equal to twice the measurement uncertainty (i.e. just below the detection threshold; e.g. Becker et al. 2015) or as corresponding to infinite optical depth. Following Bosman et al. (2018), we display the cumulative distribution functions (CDFs) showing both bounds. The lower CDF assumes that all non-detections are infinitely opaque, while the upper CDF assumes all non-detections correspond to flux just below the detection limit.

Owing to the much higher SNR of our sample, the number of non-detections over 50 cMpc h^{-1} is highly reduced at $5.5 < z < 5.9$ compared to Bosman et al. (2018), which used a sample of size comparable to ours ($N_{\rm qso} = 64$ compared to our $N_{\rm qso} = 67$) but with widely varying SNR. Our study only employs spectra which are able to probe optical depths up to (at least) $\tau = 4.5$ in $\Delta z = 0.1$ bins. At $5.5 < z < 5.7$, the only non-detection is in quasar PSO J025-11 (1/67 sightlines) while ∼10 per cent of sightlines were undetected in Bosman et al. (2018). The quasar J0148+0600 (the longest trough from Becker et al. 2015) was formerly the most opaque at this redshift, but a slight update to the quasar's redshift shifts the exact start and end of the measurement bin in our study such that transmission is detected ($\tau = 5.33$) in a bin centred at $z =$

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

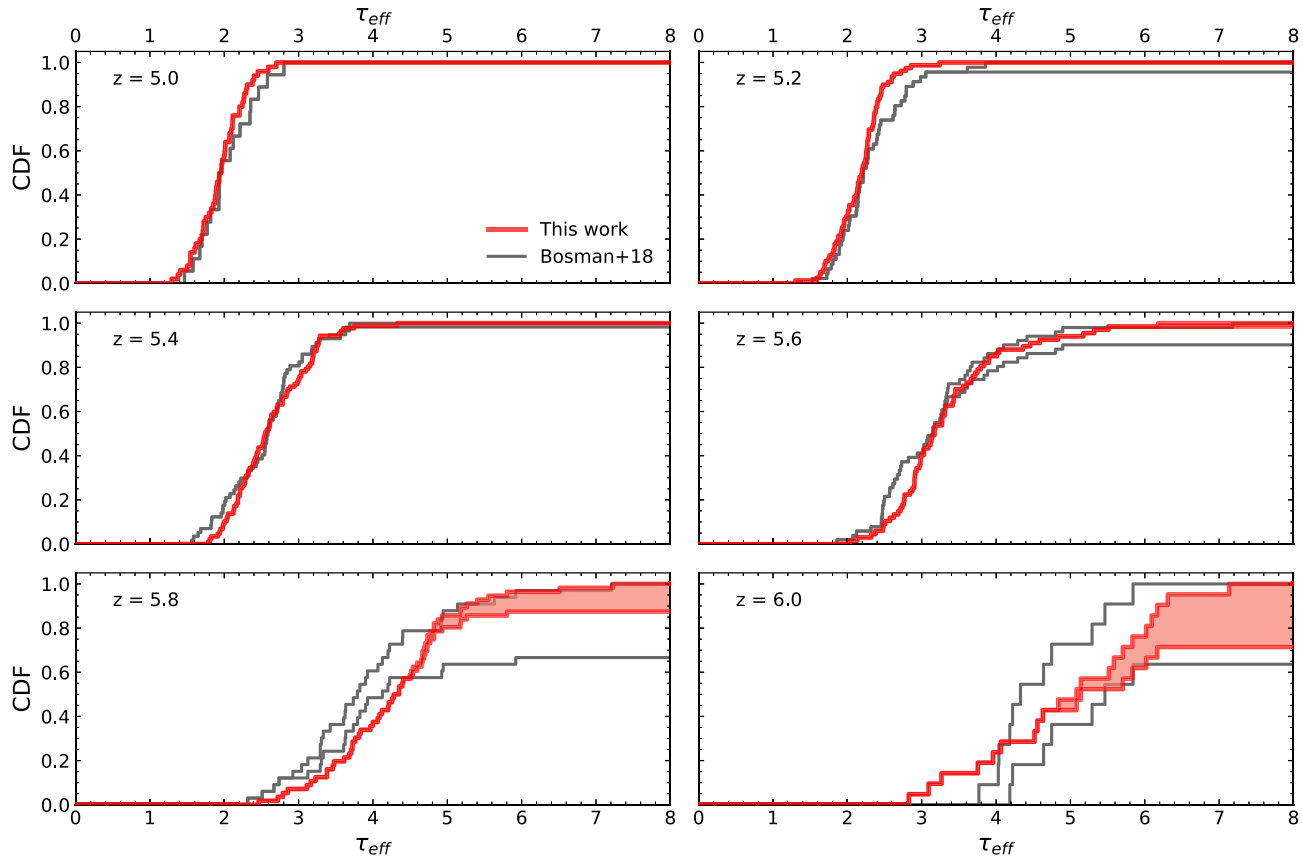

Figure 7. CDFs of Ly-α optical depth (red) computed in 50 cMpc h^{-1} intervals, compared to results from Bosman et al. (2018; black). The lower and upper CDFs correspond to assumptions that non-detections are infinitely opaque, or just below the detection threshold, respectively. While the sample sizes are comparable ($N = 67$ and $N = 64$, respectively), the highly improved data quality and new correction of systematic biases linked to instrumentation and continuum reconstruction result in overall smoother distributions and fewer non-detections.

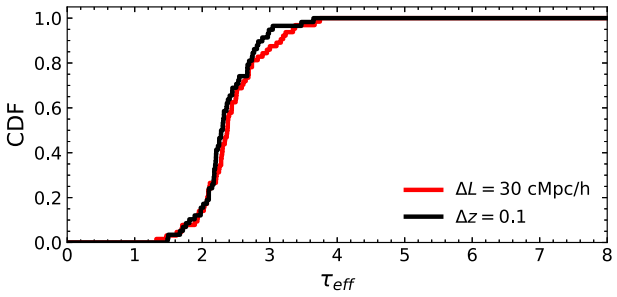

Figure 8. Comparison of the optical depth distribution measured over the same interval, $5.25 < z < 5.35$, using bins of constant length with boundaries that vary (red) or constant Δz boundaries (black). The length of 30 cMpc h^{-1} corresponds to the redshift interval $\Delta z = 0.1$; the only difference between the distributions is the definition of binning and not the lengths over which the optical depth is intrinsically computed. The excess of highly opaque sightlines in the red curve originates from length bins whose centres lie near the top end of the range ($z \sim 5.35$), artificially increasing the scatter by including contributions from pixels outside the nominal redshift range. To avoid this bias, we adopt binning in constant redshift for the purposes of inference.

5.577. This issue highlights one of the problems with the classical definition of optical depth binning: results depend non-trivially on the assumed redshifts of the background quasars (while our binning explicitly does not). At $5.7 < z < 5.9$, the number of non-detections is reduced from ∼35 per cent (in Bosman et al. 2018) to 12.5 per cent (7/56 sightlines have $\tau > 4.5$). In contrast, our increased sensitivity does not reduce the number of sightlines with non-detections at $5.9 < z < 6.1$, where ∼30 per cent of sightlines (6/21) remain fully absorbed. A large fraction of $z \gtrsim 5.9$ sightlines are therefore more opaque than $\tau = 4.5$, a limit which is unlikely to be exceeded for large samples of quasars with current instrumentation. Significant advances in sensitivity, which may be required to detect residual transmission in the bulk of quasars at $z > 6$, could be brought by the next-generation Extremely Large Telescope (Gilmozzi & Spyromilio 2007) or the *Thirty Meter Telescope* (Sanders 2013).

To determine the lowest redshift at which the optical depth distribution is in agreement with fluctuations from density alone, we use redshift bins with $\Delta z = 0.1$. We choose this binning size in order to resolve the fast evolution in the mean optical depth (Fig. 5). In Section 3.1 we highlighted some potential biases inherent to the classical definition of binning optical depth measurements of constant length. In Fig. 8 we demonstrate these biases, which become more pronounced as the redshift intervals are shortened. The optical depth distribution at $z = 5.3$ is artificially broadened by the inclusion of transmission outside the nominal redshift range, as shown in the red curve. This effect is non-negligible when the binning length becomes comparable to the redshift interval, when the evolution in the mean is rapid (as shown in Fig. 5), or when using sightlines near the end of their usable wavelength ranges (i.e. the lowest redshift bins). We adopt binning in constant redshift intervals for the purposes of inference in order to avoid this bias, and show the resulting CDFs

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

in $\Delta z = 0.1$ in Fig. 9. The new definition also avoids any covariance of sightlines in distributions at fixed redshift. The distributions will still be covariant between redshifts, since opaque sightlines show coherence over scale $\Delta z > 0.1$.

4.3 $z \geq 5.9$ transmissive sightlines

In Fig. 10 we show the most transmissive sightlines at $z = 5.9$ and $z = 6.1$ which stand out from the distributions in Fig. 9. The XQR-30 quasar J1535+1943 is the most transmissive at $z = 5.9$ with $\tau_{\rm eff} = 2.50$, showing Ly-α transmission over the entire $5.85 < z < 5.95$ interval (top panel). The same quasar is the second most transmissive at $z = 6.1$, with two strong transmission spikes at $z \sim 6.07$ resulting in $\tau_{\rm eff} = 3.79$; this suggests elevated transmission over scales $\gtrsim$100 cMpc h^{-1}. At $z = 6.1$, the X-Shooter archival quasar PSO J011+09 is the most transmissive by far, with three very strong transmission spikes resulting in $\tau_{\rm eff} = 2.59$ (bottom panel). The transmission is affected by increased uncertainties due to corrections for telluric absorption; however, excluding the regions affected by increased uncertainties actually further lowers the measured optical depths. J1535+1943 and PSO J011+09 display flux transmission larger than the mean at $z = 5.9$ and $z = 6.1$ by factors of 4.3 and 4.9, respectively (corresponding to optical depths 50 per cent smaller than the mean). The discovery of such rare transmissive sightlines is only possible by employing large samples of quasars to sample cosmic variance: at $z = 5.9$, only 1/51 sightlines has an optical depth $\tau_{\rm eff} < 3$. Characterizing the extrema of the optical depth distribution at fixed redshift is crucial in order to design models of UVB fluctuations which reproduce the full variety of environments at the end of reionization. The quasar J1535+1943 was not included in any previous measurements of optical depth; its addition to our sample raises the average transmitted flux by $\sim$10 per cent. While this change is comfortably included within our quoted bootstrap uncertainties, it may account for some of the systematic disagreements between our study and past work which did not include this quasar (Fig. 5).

In addition, Ly-α transmission spikes can been used to measure the thermal state of the IGM (e.g. Gaikwad et al. 2020) and to pose constraints on reionization history through their statistical distribution (Barnett et al. 2017; Chardin et al. 2018). The identification of strong transmission spikes at $z > 5.8$ therefore opens up complementary analyses, which will explored in a separate paper (Gaikwad et al., in preparation).

5 COMPARISON WITH HOMOGENEOUS-UVB MODELS

The existence of completely opaque troughs with $\tau > 5$ at $z \sim 5.8$ rules out models of reionization with a homogeneous UVB and IGM temperature-density relation (Becker et al. 2015; Bosman et al. 2018). Even in the absence of large opaque troughs, the observed scatter in Ly-α optical depth at fixed redshift suggests excess fluctuations at even later times (Fig. 4). Determining the redshift evolution of these fluctuations can quantify the transition redshift beyond which the IGM no longer retains reionization-related structure from the point of view of Ly-α absorption.

We compare our results to predictions from two different homogenous-UVB simulations, Sherwood (Bolton et al. 2017) and Nyx (Almgren et al. 2013). In both models, scatter between sightlines results solely from fluctuations in the density field within a constant, fully permeated UVB. Compared to Nyx, the Sherwood simulation is run with smaller boxes but provides finer redshift sampling ever $\Delta z = 0.1$ in redshift, while in Nyx the optical depth distribution must be extrapolated from three snapshots at $z = 5.0$, $z = 5.5$, and $z = 6.0$. The two simulation suites also employ different models of the UVB with different base rescalings of the ionizing intensity. Neither simulation resolves the gas densities corresponding to DLAs nor Lyman-limit systems. We give more details of the simulation suites below.

The Sherwood simulation suite was designed to reproduce Ly-α transmission post-reionization, at $2 < z < 5$, where it is in remarkable agreement with observations (Bolton et al. 2017). Sherwood employs the hydrodynamics code P-GADGET 3 (Springel 2005) and a uniform Haardt & Madau (2012) UVB. The gas particle masses are $M_{\rm gas} = 9.97 \times 10^4$ M$_\odot$ and the box includes 2×2048^3 particles. We use the simulated boxes which are 40 cMpc h^{-1} on the side; we prefer those boxes over the lower-resolution 80 cMpc h^{-1} runs of Sherwood since they resolve the Ly-α transmission and thus provide a closer comparison to the Nyx simulation. Snapshots were taken every $\Delta z = 0.1$ from $z = 5.0$ to $z = 6.0$. We draw 5000 lines of sight through the simulation box with lengths corresponding to $\Delta z = 0.1$ at each redshift.

Nyx is an Eulerian grid cosmological hydrodynamical simulation code which is optimized for simulations of the Lyα forest (Lukić et al. 2015). We use the Nyx simulation described in Davies et al. (2018a), 100 cMpc h^{-1} on a side with 4096^3 dark matter particles and 4096^3 baryon grid cells, sufficient box size and resolution for converged Lyα forest statistics at $z \lesssim 6$ (Oñorbe et al. 2017). Snapshots at $z = 5.0$, 5.5, and 6.0 were used to simulate the Lyα forest at $z = 5.0$–5.2, 5.3–5.7, and 5.8, respectively. At redshifts not equal to the snapshot redshift, we re-scaled the physical gas densities by $(1 + z)^3$ to account for cosmological expansion, effectively ignoring the impact of structure formation over these intervals of cosmic time. We draw 40 000 lines of sight through each snapshot starting from random positions within the volume towards a random direction along the grid axes. While the simulation was originally run with the Haardt & Madau (2012) UVB for heating and cooling rates, here we initially construct Lyα forest skewers assuming a fixed photo-ionization rate $\Gamma_{\rm H\,I} = 10^{-12.1}$ s^{-1} comparable to observational estimates at $z \sim 5$ (Becker & Bolton 2013).

We post-process sightlines drawn from simulations in the following way. First, we shorten sightlines to the length corresponding to $\Delta z = 0.1$ and project them on to a wavelength array with constant velocity sampling. We then randomly assign each simulated sightline to a real observation in the same redshift interval, and interpolate the simulated flux on to the observed wavelength array (including any masking of bad regions). We add random noise sampled from the corresponding observed error array by drawing from a Gaussian with width of the 1σ uncertainty at each pixel. Finally, we multiply the sightline by a wavelength-dependent continuum error drawn from a normal distribution with scale of the observed 1σ bound of the continuum uncertainty. This assumes that the continuum uncertainty is fully covariant, while formally we would need to draw from the full PCA posterior (see e.g. Davies et al. 2018b). However, since we care not about the details of wavelength-dependence, our approach is both more computationally efficient and conservative. Shifting the continuum reconstruction at all wavelengths by the same standard deviation introduces a more coherent shift than selecting a random draw with the same PCA likelihood. This procedure will tend to introduce 'pessimistic' continuum-reconstruction scatter into the post-processed simulations, in the sense that it slightly lower the evidence for fluctuations (which is conservative for our purposes). Note that we do not need to convolve our simulated sightlines to match the observed instrumental resolutions, since the convolution operation explicitly conserves the total flux.

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

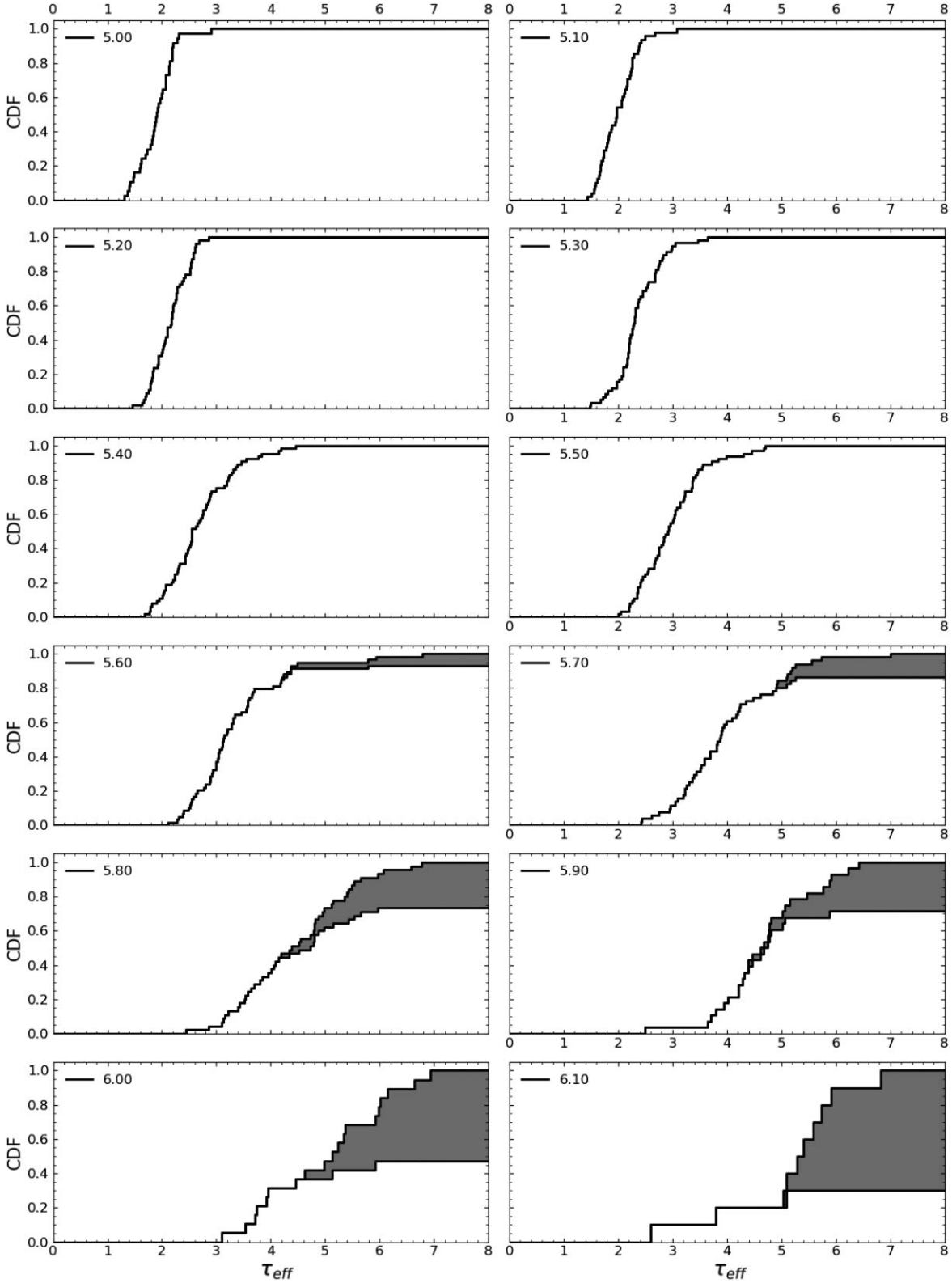

Figure 9. CDFs of Ly-α optical depth in bins of constant $\Delta z = 0.1$. The distributions begin to appear 'elongated' around $z \sim 5.4$. The fraction of non-detections (as seen by the difference between the curves at the right edges of the CDFs) increases sharply above $z = 5.7$, but some transparent sightlines remain even at $z = 6.1$. The most transparent sightlines at $z = 5.9$ and $z = 6.1$ are shown in Fig. 10.

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

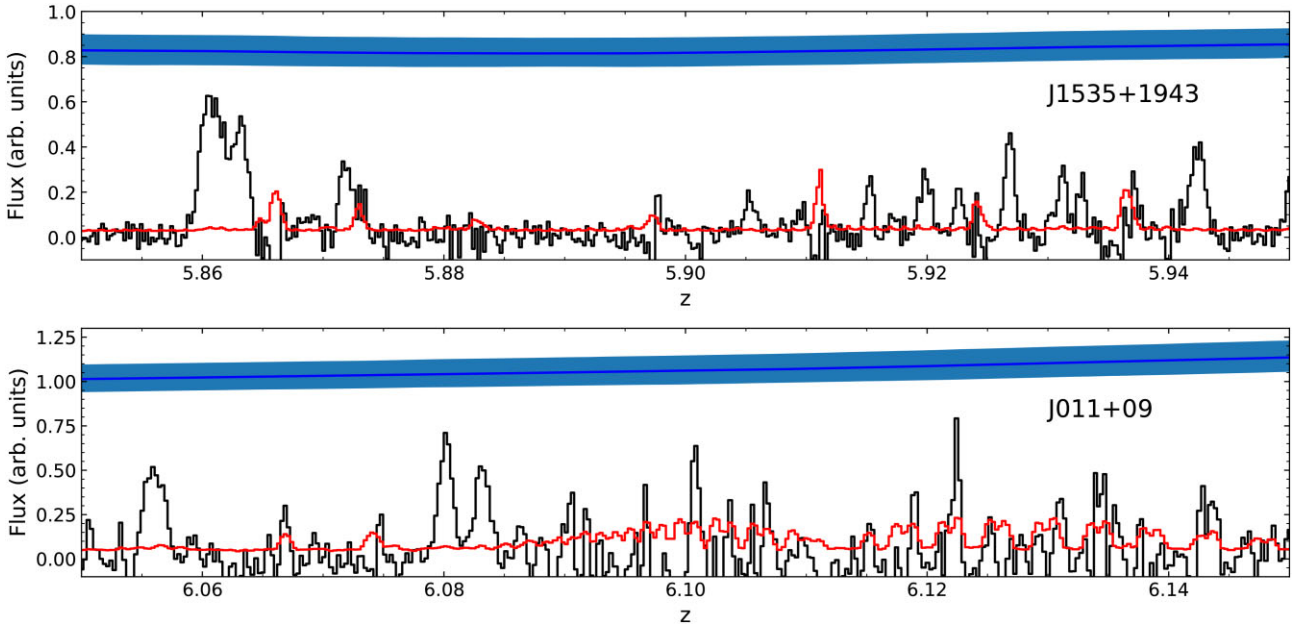

Figure 10. Black: Transmitted Ly-α flux in the most transmissive sightlines at $z = 5.9$ (J1535+1943, top) and at $z = 6.1$ (PSO J011+09, bottom). Red: Observational uncertainties. Blue: PCA continuum reconstruction and its $\pm 1\sigma$ uncertainties. The $6.10 < z < 6.14$ range is affected by increased uncertainties due to sky emission corrections, but excluding this region slightly raises the overall mean transmitted flux over the $6.05 < z < 6.15$ interval.

5.1 Maximum likelihood analysis

In order to calculate the likelihood of our observations given a model, optical depths at each simulated pixel must first be rescaled in bulk. The optical depth rescaling is expressed as a multiplicative factor on optical depth, $\tau_{\rm rescaled} = A\tau_{\rm sim}$ where the rescaling factor A is different at each redshift.[1] Rescaling corresponds to adjusting the ionizing background intensity in the simulations, i.e. for Sherwood it reflects a deviation from Haardt & Madau (2012) in the average ionizing emissivity which can be a factor of a few. Optical depth rescaling factors are usually chosen to match the observed mean fluxes at each redshift, $\langle e^{A\tau_{\rm sim}} \rangle = \langle F \rangle_{\rm obs}$, but this may lead to bias when large sightline scatter is present. A few highly transmissive sightlines will lead to a very low average τ, which might make it difficult to match opaque sightlines. However, the highly transmissive sightlines also carry uncertainties. Therefore, a rescaling to a slightly lower average flux than observed leads to a better agreement between models and observations, because both opaque and transmissive sightlines can be produced by random noise. Motivated by this observation, we choose the rescaling factor to *maximize the likelihood of the observations* instead of matching the mean flux explicitly. We note our rescaling factors still give rise to mean simulated fluxes consistent with observed mean fluxes at $<1\sigma$ at all redshifts where the simulations are a good fit to the data (see below).

We determine the likelihood of the observations by combining the likelihoods of each individual measurement τ_n made in sightline S_n,

$$\mathcal{L}_{\rm data} = \prod_{n=1..N} p(\tau_n|S_n)$$

where N is the total number of sightlines contributing to the distribution at a given redshift. The probability $p(\tau_n|S_n)$ is obtained by post-processing all simulated sightlines with the observational properties of sightline S_n: wavelength sampling and masking, random flux uncertainties, and a random continuum uncertainty. The resulting distribution of predicted τ given S_n is then used to build a kernel density estimator (KDE). To obtain a smooth KDE from the Sherwood simulation with a relatively small number of sightlines, we over-sample each sightline six times.[2] The KDE is then evaluated at the observed value τ_n to produce $p(\tau_n|S_n)$. The process is repeated for each observation S_n to obtain $\mathcal{L}_{\rm data}$ via equation (2).

For the purposes of comparison with models, we always assume that flux non-detections correspond to intrinsic flux just below the detection threshold (i.e. the upper CDF bounds in Fig. 9). We thus ensure that simulations are given the 'best possible chance' at reproducing the optical depth scatter in the observations, since homogeneous-UVB models are known to always under-estimate (and never over-estimate) Ly-α optical depth scatter. Non-detections of mean flux only occur at $z \geq 5.6$ therefore this definition is equivalent to using the measured values of flux at all redshifts where the models are a good description of the observations. We pick the rescaling factor to maximize $\mathcal{L}_{\rm data}$ by sampling A in steps of 0.0025.

Table 5 gives the optimally rescaled photo-ionization rates, $\Gamma = \Gamma_{\rm sim}/A$. The Sherwood simulation required rescaling Γ down by up a factor of 2 at $z \sim 5.0$ ($A = 0.5$) while Nyx's photo-ionization rate was rescaled up by up to a factor 2 at $z \sim 6.0$ ($A = 2$). After rescaling, Γ is in good agreement between the simulations, within ~12 per cent. This remaining difference at $z \leq 5.3$ can most likely be attributed to small differences in cosmological parameters (e.g. $\Omega_{\rm m}$, σ_8) between the simulations, to which Γ is known to be sensitive (Bolton & Haehnelt 2007). Γ shows a discontinuity in Nyx at $z = 5.3$ due to switching from extrapolating the $z = 5.0$ snapshot to the $z = 5.5$ snapshot; the difference between using the two snapshots is about 20 per cent. Any tensions are far below current measurement

[1] Note that this rescaling is applied at each simulated pixel before computing the mean optical depth.

[2] meaning the observational uncertainties are chosen at random six times for each sightline to produce six predicted values of τ.

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

Table 5. Rescaled photo-ionization rates which maximize the likelihood of the observed optical depth distributions for Sherwood and Nyx at each redshift. The rescaling factors themselves are in the range 0.5–2. The discontinuity at $z = 5.3$ in Nyx is to due switching from the $z = 5.0$ to the $z = 5.5$ snapshots (see text); the value in brackets gives the rescaled photo-ionization rate when the $z = 5.0$ snapshot is used instead of $z = 5.5$.

z	$\Gamma_{\rm Sherwood}/10^{-13}$ s^{-1}	$\Gamma_{\rm Nyx}/10^{-13}$ s^{-1}
5.0	7.85	7.34
5.1	7.57	6.92
5.2	7.39	6.62
5.3	7.03	7.32 (6.13)
5.4	6.17	6.41
5.5	5.77	5.56
5.6	4.46	4.47
5.7	4.05	4.02
5.8	3.79	3.94

uncertainties in Γ due to the IGM's thermal state which are a factor of $\sim$2 (e.g. D'Aloisio et al. 2018).

We now calculate the probability of drawing a full dataset with $\mathcal{L}_{\rm data}$ from the simulations. We generate 10 000 fully forward-modelled datasets by post-processing N randomly-selected model sightlines, each assigned to the uncertainties of an observed sightline S_n. All simulated datasets therefore have the same size as the observations. The likelihood is calculated for each simulated dataset in the same manner as the data,[3] giving rise to a distribution of $\{\mathcal{L}_{\rm sim}\}$. We build a KDE on the distribution of simulated likelihoods and evaluate it at $\mathcal{L}_{\rm data}$ to finally obtain the probability of the entire set of observations given the simulation model. These probabilities p formally coincide with the p-values, and we also convert them to standard deviations via stdev $= \sqrt{2}{\rm erf}^{-1}(2p)$ where erf^{-1} is the inverse error function.

5.2 Results

Figs 11 and 12 show the results of the likelihood analysis for the Sherwood and Nyx simulations, respectively. The data likelihood falls within $\pm1\sigma$ expectations at $5.0 \leq z \leq 5.2$ for both Sherwood and Nyx. Forward-modelling introduces some optical depth scatter due to uncertainties, most visible at $z = 5.0$. The extra scatter is expected, and provides a better fit to observations: e.g. at $z = 5.1$ and $z = 5.2$ in Sherwood, the post-processed elongated distribution (red line) provides a better fit to the data than the model without post-processing (green line). The excellent agreement with models at $z \leq 5.2$ implies that the intrinsic physics within the simulations combined with our known observational uncertainties account for all the variance observed in the data. A homogeneous UVB acting on density fluctuations is therefore a sufficient description of Ly-α transmission up to $z = 5.2$.

Conversely, the Ly-α transmission scatter observed at $z \geq 5.4$ is in excess of model predictions at $>3.5\sigma$ in both models. Since we sampled 5000 sightlines from the Sherwood simulation, we are limited in determining the nature of outliers to the $\lesssim 1/5000 \simeq 3.5\sigma$ level. The 40 000 sightlines from the Nyx simulation enable us to push the analysis to $\lesssim 1/40\,000 \simeq 4\sigma$ outliers. We find that the post-processed Sherwood simulations fail to match the observations at the saturation level (3.5σ) at all redshifts $z \geq 5.4$. Nyx similarly fails to match observations at the corresponding 4σ level at $z \geq 5.4$. In both models, the rescaling factor which maximizes the likelihood of observations results in mean simulation fluxes in close agreement with observed values (within 1σ of the values in Table 4) at $z \leq 5.3$; but the mean fluxes are in disagreement at $z \geq 5.4$ where the 'most likely' mean fluxes are closer to the median (Figs 11 and 12). As expected, matching the median transmission increases the likelihood of a extended distribution since both extremely opaque and extremely transmissive sightlines then have reasonable probabilities.

Both the Nyx and Sherwood simulations transition from providing good fits to the data to being in strong tension with observations at $z = 5.3$, where they are in mild tension with the data (excluded at 2.4σ and 2.5σ, respectively). We conclude that extra Ly-α optical depth scatter is present in the observations, and its magnitude is in excess of differences between simulations due to box size and different choices of UVB models. However, the tension is mild and we cannot completely rule out either homogeneous UVB model. Examining the difference between Nyx and Sherwood in more detail, we find that Nyx provides a statistically better fit to the data at all redshifts in the absolute (i.e. the likelihood of the observed dataset is higher in Nyx). We attribute this to Nyx's larger box size, which makes the model more apt to capture density fluctuations on large scales. However, both models are in agreement with the data at $z < 5.3$, in strong tension at $z > 5.3$, and in mild tension at $z = 5.3$.

Redshift $z = 5.4$ is the lowest redshift at which the observed distribution of optical depths is in strong tension with both models ($>3.5/4\sigma$). To evaluate the robustness of the result, we test whether the tension is driven primarily by the most opaque sightlines at $z = 5.4$ or by the extended shape of the entire distribution. We arbitrarily remove the most opaque three sightlines, which have observed $\tau > 4$ – corresponding to 5 per cent of the sample. While none of them show signs of foreground absorption by DLAs in the form of intervening metal absorbers, some DLAs at $z > 5$ may be particularly metal-poor. Even though we find no evidence for such metal-poor DLAs in the other redshift bins, unlucky alignment cannot be completely excluded. We roughly estimate that such DLAs would need metallicities of [X/H] $\lesssim -2.5$ to avoid detection in our spectra; this will be calculated in more detail in future work.

However, we find that even after arbitrarily removing the most opaque three sightlines from the distribution, the observations are still in strong tension with the $z = 5.4$ Sherwood simulation at 3.43σ ($p = 0.0305$ per cent with $A = 0.58$). A similar result is obtained with Nyx, where omitting the most opaque sightlines still results in tension at 3.68σ. We show the corresponding likelihood distributions in Appendix B. We therefore conclude that the whole extended shape of the distribution, and not just a few sightlines, are driving the disagreement between homogeneous UVB models and observations at $z = 5.4$.

5.3 Discussion

Fig. 13 summarizes the results of the likelihood analysis. Both homogeneous-UVB simulations, Sherwood and Nyx, provide an excellent fit to observations at $z \leq 5.2$. Post-processing the simulations slightly broadens the predicted distribution of optical depths in this regime, bringing predictions in agreement with the data. There

[3]However, we do not apply the optimal choice of rescaling factor to each simulated dataset as for the data, since this would be computationally unfeasible. A test of the impact reveals that the wings of the likelihood distribution may shift to higher values by up to $\Delta\mathcal{L} \sim 2$, which is insufficient to quantitatively affect our results.

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

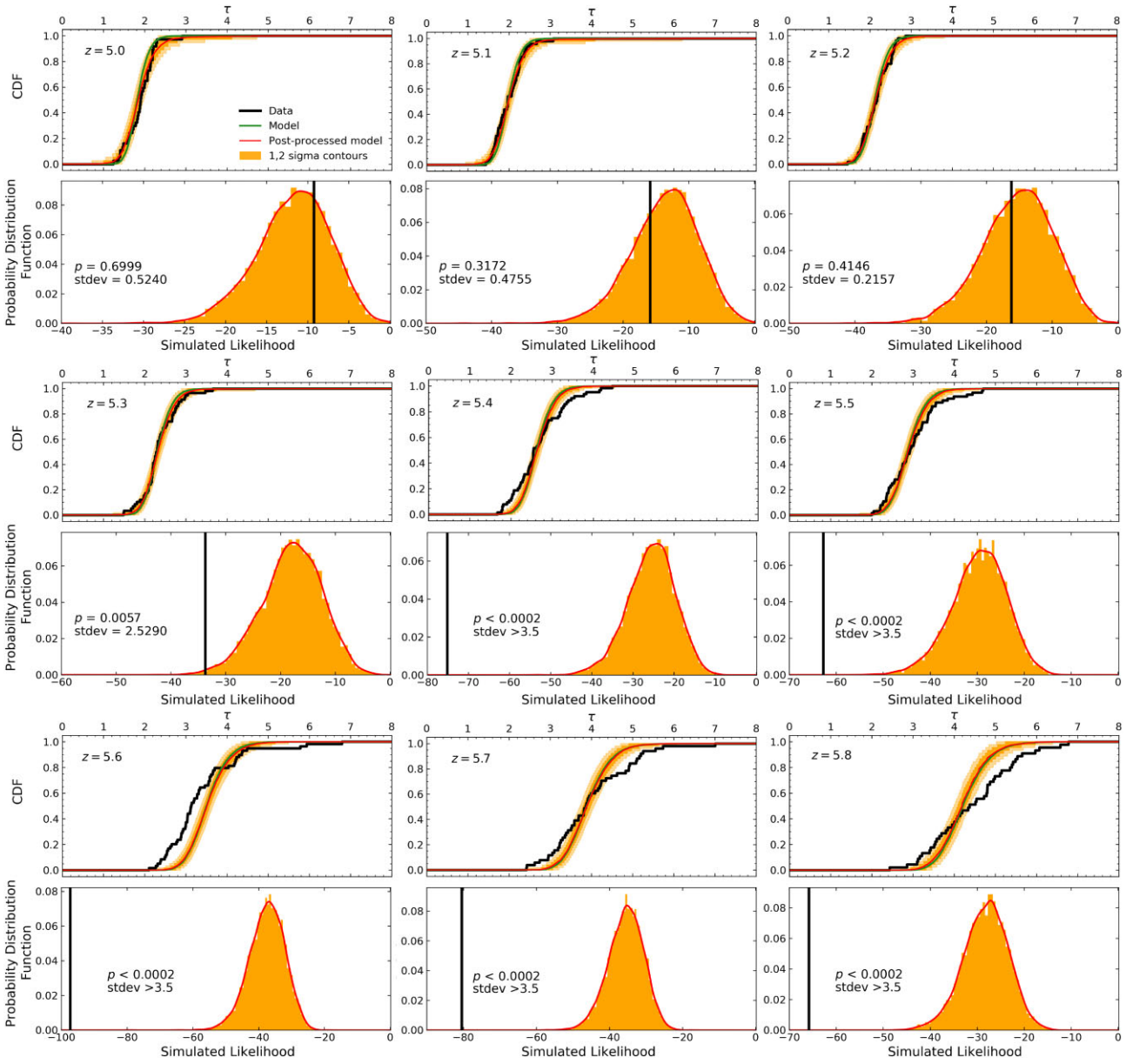

Figure 11. Top panels: Observed Ly-α optical depth distribution (black), compared to the Sherwood simulations without any post-processing (green), and with post-processing (red, orange). The light and dark orange contours show 1σ and 2σ envelopes from bootstrap resampling the post-processed models. Bottom panels: Probability distribution of log-likelihoods for fully forward-modelled datasets (orange). The distributions are used to build KDEs (red) which are evaluated at the location of the likelihood of the observed dataset (thick vertical black lines). The Sherwood model is a great fit to observations at $z \leq 5.2$, but excluded at $>3.5\sigma$ at $z \geq 5.4$.

is no evidence that any extra sources of fluctuations are necessary at $z \leq 5.2$, such as, for example, a spatially varying thermal state of the IGM. In particular, the Sherwood simulation successfully matches Ly-α optical depth over $2 \leq z \leq 5.2$ without any such modifications (Bolton et al. 2017). The tension observed at $z \geq 5.4$ is therefore highly significant, and marks the breakdown of one or more simplifying assumptions in the post-reionization high-z IGM.

A potential caveat to our maximum likelihood analysis is that the statistical power of the homogeneous-UVB simulations may be limited by box size rather than by the number of simulated sightlines. Indeed, the Nyx simulation box only contains $\sim$50 independent volumes of scale comparable to the lengths of observed sightlines ($\sim$30–35 cMpc h^{-1}) while the Sherwood box contains only a few. This is much lower than the 10 000 independent draws necessary to establish statistical significance at the 4σ level. However, the fact that the two simulations result in very similar large-scale optical depth CDFs suggests that the modes of the density field which dominate the large-scale opacity fluctuations are actually much smaller than the total path-length which should thus be much better sampled (see e.g. the appendix of Becker et al. 2015, who found that 50 Mpc h^{-1}-scale fluctuations were extremely similar between 50 and 100 Mpc h^{-1} simulation volumes). Treating the sampling as being limited by the number of independent large-scale modes would therefore be somewhat too conservative. Nevertheless, our analysis is only strictly valid in the context of the specific simulation boxes we used. While our results suggest that density fluctuations on scales

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

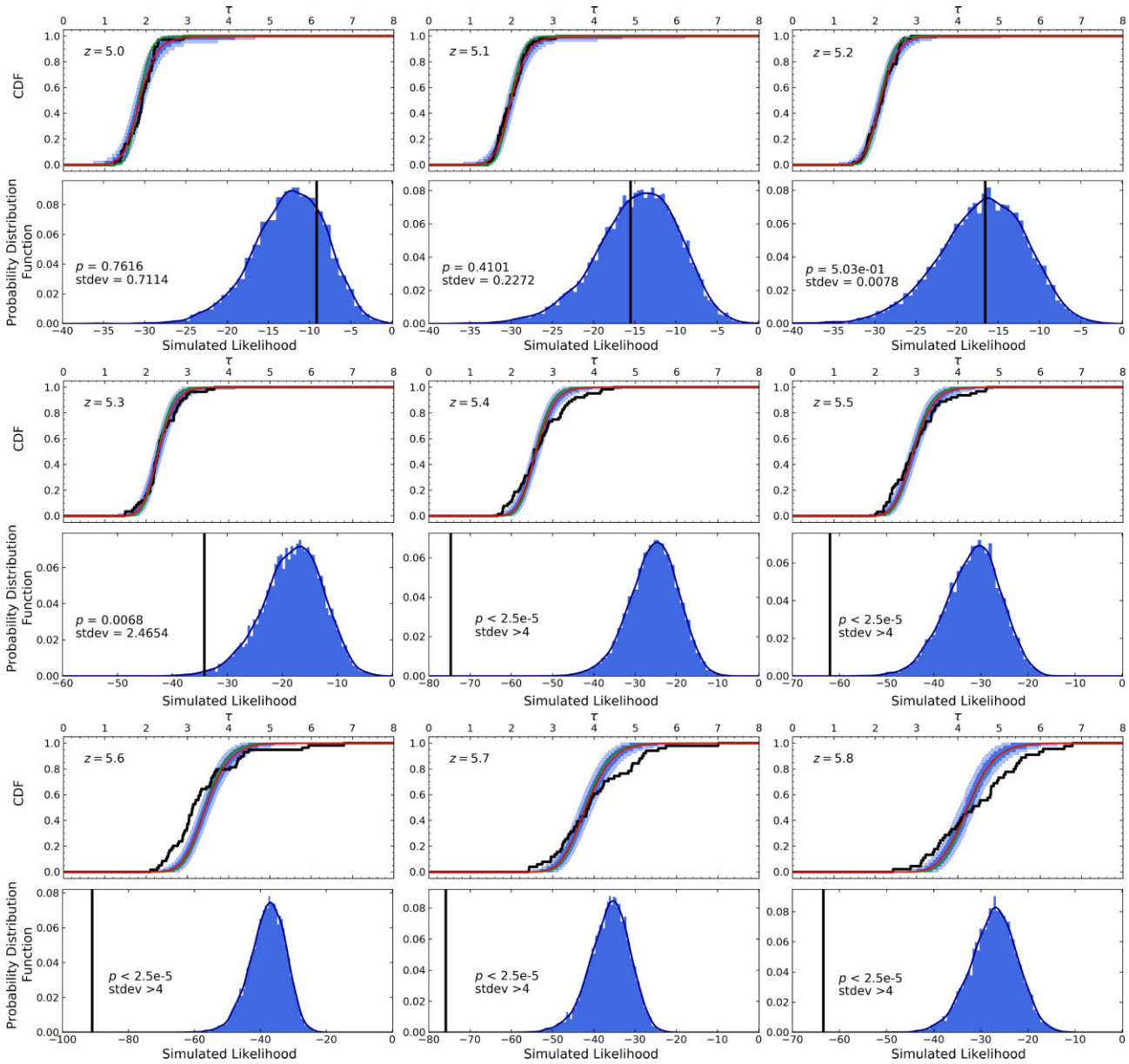

Figure 12. Same as Fig. 11 but for the Nyx simulation (blue). The Nyx model is a great fit to observations at $z \leq 5.2$, is in mild tension with data at $z = 5.3$ (2.5σ) and excluded at $<3.5\sigma$ at $z \geq 5.4$.

larger than $\gtrsim$40 cMpc h^{-1} play a negligible role in determining the Ly-α optical depth at $z < 5.3$, we note that larger simulated volumes are crucial to modelling the reionization process at higher redshifts, especially in models where bright rare sources play a significant role (e.g. Chardin et al. 2017; Meiksin 2020).

The presence of large opaque troughs $\gtrsim$100 cMpc h^{-1} in length in the Ly-α forest down to $z \sim 5.6$ already independently rules out homogeneous ionization at that redshift (Becker et al. 2015; Bosman et al. 2018; Zhu et al. 2021). Opaque troughs persisting at late times have been theorised to arise from patches of significantly neutral gas ($x_{\rm HI} > 10$ per cent; Kulkarni et al. 2019; Keating et al. 2020; Nasir & D'Aloisio 2020; see also Lidz et al. 2006; Mesinger 2010). At the same time, recent measurements have reported a very short mean free path of ionizing photons at $z = 6$, of $\sim$0.75 pMpc (Becker et al. 2021). Evidence therefore points to a late end of reionization, with remnant fluctuations in the UVB and/or IGM temperature persisting for at least 70–80 Myr after the demise of the last highly neutral 'patches' at $z \simeq 5.6$ (see also Cain et al. 2021; Davies et al. 2021). From the point of view of Ly-α transmission homogeneity, hydrogen reionization is not over before $z = 5.3$.

5.3.1 *Hydrogen neutral fraction*

We calculate the volume-averaged $x_{\rm H\,I}$ directly from the 40 000 Nyx skewers at each redshift assuming ionization equilibrium, shown in Fig. 14. Our nominal measurements correspond to re-scalings of the UVB which maximize the likelihood of the $\tau_{\rm eff}$ distribution (Table 6). We also measure upper and lower bounds corresponding to re-scalings that reproduce, respectively, the lower and upper bounds of the mean transmitted flux. At $z = 5.3$, we find a significant difference between the $x_{\rm H\,I}$ values obtained from rescaling the $z = 5.0$ snapshot of Nyx as opposed to the $z = 5.5$ snapshot; we attribute this

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

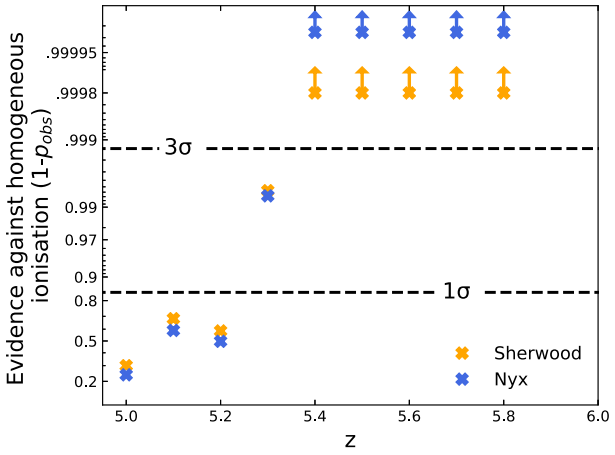

Figure 13. Summary of the likelihood analysis comparing the observed distributions of optical depth to the homogeneous-UVB Sherwood (orange) and Nyx (blue) simulations. By including forward-modelling of all known uncertainties, both models provide an excellent fit to the data at $z \leq 5.2$ but are in strong tension at $z \geq 5.4$.

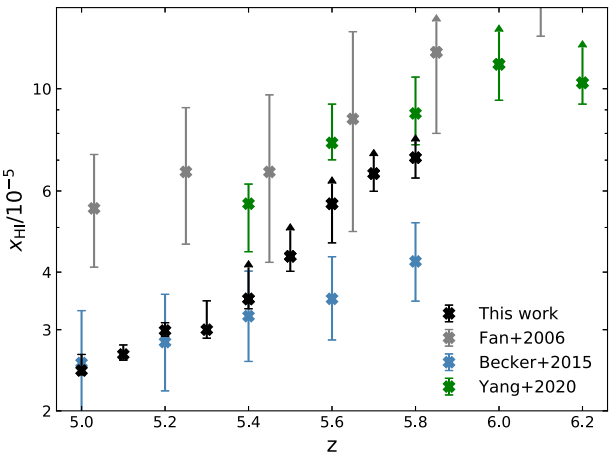

Figure 14. Evolution of the volume-averaged neutral fraction $x_{\rm H\,I}$ with redshift. We only provide lower limits in the regime where the Nyx simulation, which we use in the calculation of $x_{\rm H\,I}$, is excluded by the observations. The model is a fairly poor fit to the data at $z = 5.3$ (2.5σ tension), which may be causing an offset. Our measurements employ samples factors 5–10 larger than past measurements. Tension with the Fan et al. (2006) values (blue) may be the result of a number of observational, systematic, or modelling differences.

difference to evolution in the cosmic structure which neither snapshot captures perfectly. We list the most pessimistic bounds on $x_{\rm H\,I}$ among both snapshot re-scalings. For redshifts $z \geq 5.4$, the strong tension between our data and the maximum-likelihood Nyx $\tau_{\rm eff}$ distribution implies that reionization may not yet be complete (see also Kulkarni et al. 2019; Nasir & D'Aloisio 2020; Choudhury et al. 2021; Qin et al. 2021). The $x_{\rm H\,I}$ estimate from the mean flux is not sensitive to the fraction of fully neutral regions, so we show our measurements at $z \geq 5.4$ as lower limits. Since the homogeneous-UVB simulations are rescaled to maximize the likelihood of 'optimistic' observations where non-detections are treated as flux just below the detection limit (Section 5.1), our $x_{\rm H\,I}$ limits might be too conservative by $\lesssim$5 per cent. This effect goes in the same direction as the lack of inclusion of self-shielding in the models.

The calculation of $x_{\rm H\,I}$ has traditionally assumed an optically thin IGM without self-shielding by dense fluctuations (Fan et al. 2006; Becker et al. 2015; Yang et al. 2020). To estimate the impact of this assumption on our measurements, we post-process a set of skewers with the prescription of Rahmati et al. (2013) and show the results in Table 6. The inclusion of self-shielding results in an increase of $x_{\rm H\,I}$ by ∼25 per cent. Unlike previous works where the effect was comparatively negligible, uncertainties related to the treatment of self-shielding dominate over our statistical uncertainties. Since the Nyx simulations do not resolve dense gas, we cannot provide a physically realistic inclusion of self-shielding at the <5 per cent level required to match the statistical uncertainties. We show the best-fitting values without self-shielding in Fig. 14 in order to compare to past work which universally assumed an optically thin IGM.

Table 6. Volume-averaged hydrogen neutral fraction and $\pm1\sigma$ bounds computed by comparison with the homogeneous-UVB Nyx simulations. We show the values both with and without self-shielding included (s-s). The numbers in brackets have the highest likelihood according our model, but note that above $z \geq 5.4$, homogeneous-UVB simulations are a poor match to data and only enable lower limits on $x_{\rm H\,I}$.

z	$x_{\rm H\,I}/10^{-5}$ (no s-s)	$x_{\rm H\,I}/10^{-5}$ (with s-s)
5.0	$2.446 - 0.051 + 0.205$	$3.020 - 0.058 + 0.230$
5.1	$2.651 - 0.075 + 0.129$	$3.336 - 0.164 + 0.064$
5.2	$2.988 - 0.085 + 0.119$	$3.636 - 0.095 + 0.131$
5.3	$3.000 - 0.125 + 0.466$	$3.598 - 0.145 + 0.566$
5.4	$(3.498) > 3.332$	–
5.5	$(4.328) > 4.016$	–
5.6	$(5.627) > 4.630$	–
5.7	$(6.544) > 5.990$	–
5.8	$(7.087) > 6.401$	–

We are consistent with the inferred $x_{\rm H\,I}$ values of Yang et al. (2020), who employed a homogeneous UVB model up to $z = 5.8$. Our values of the neutral fraction at $5.0 \leq z \leq 5.4$ are a factor ∼2 lower than reported by Fan et al. (2006). This tension (∼2σ) may be due to a number of factors, such as a much smaller sample size than our study, continuum reconstruction systematics, lower SNR, or the very significant differences in the IGM model. We are in good agreement with Becker et al. (2015) up to $z = 5.4$. The measurements of Becker et al. (2015) correspond to the neutral fraction specifically *inside of ionized regions*, which explains the divergence with our lower limits at higher redshifts. The uncertainties of our low-z $x_{\rm HI}$ measurements are very small, reflecting the exquisite precision of the measurement of the mean flux (Fig. 5). We warn that homogeneous UVB models are a fairly poor fit to observations at $z = 5.3$ (2.5σ tension), such that systematic errors in $x_{\rm H\,I}$ may be present in that redshift bin. The use of a -1σ lower bound on $x_{\rm H\,I}$ as a lower limit may therefore be an equally justified choice.

The conversion of mean flux measurements to values of the IGM neutral fraction $x_{\rm H\,I}$ is only valid under the assumption of completely homogeneous ionization. This is because, fundamentally, the translation is model-dependent and relies on simulations assuming homogeneous ionization. Models which reproduce the mean flux with late reionization, such as that of Kulkarni et al. (2019), predictably result in significantly higher $x_{\rm H\,I}$ than simulations with homogenous ionization even when they match *the same observed mean flux* (Yang et al. 2020).

5.3.2 *Non-homogeneous UVB models*

We compare our new measurements of Ly-α optical depth distributions with the late-reionization model of Keating et al. (2020; first

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

described in Kulkarni et al. 2019). Their model uses high-resolution cosmological radiative transfer simulations in boxes of 160 cMpc h^{-1} on the side. Similarly to the Sherwood suite, the late-reionization simulation is run with the P-GADGET 3 code and uses the same cosmological initial conditions. The radiative transfer is conducted in post-processing with the ATON code (Aubert & Teyssier 2008, 2010). The simulation employed 2×2048^3 gas and dark matter particles. Lightcones of 50 cMpc h^{-1} are extracted from the simulation on-the-fly, resulting in sightlines with H I fraction and temperature that evolve along the line of sight with redshift. The centre of each such sightline is matched to the mid-point redshift of a measurement from the Bosman et al. (2018) sample, such that all sightlines are at slightly different redshifts. 500 such simulated datasets are generated for each redshift. Fig. 15 shows the resulting 1σ (70 per cent) bounds of the corresponding CDFs.

In order to compare these simulations to our observations, we re-bin the spectra in intervals of 50 cMpc h^{-1} centred on the mid-point redshift of each snapshots. The resulting bins cover redshifts of $5.3245 < z < 5.4754$ for the $z = 5.4$ snapshot, $5.5210 < z < 5.6710$ for the $z = 5.6$ snapshot, etc. In addition to the like-to-like sightline matching detailed above, the predicted Ly-α optical depth distributions from Keating et al. (2020) were also calibrated to the mean Ly-α transmitted flux measurements of Bosman et al. (2018). The late reionization model cannot be trivially re-scaled to our updated mean flux values, because the radiative transfer simulations rescale the emissivity of reionizing sources to match the mean flux and predict Ly-α optical depth fluctuations self-consistently. As such, the model requires time-consuming runs of the simulation to calibrate. We therefore leave a quantitative comparison of the late reionization model with our observations to future work. Fig. 15 shows the excellent qualitative agreement between our new observations and the Keating et al. (2020) model without rescaling to match the new measurements of mean transmitted flux, nor sightline matching. At $z = 6.0$, the late reionization model predicted a significantly larger fraction of sightlines $\tau_{\rm eff} < 4$ than observed in previous work (compare with Fig. 7). At $z = 5.8$, the model also predicted a larger number of highly opaque sightlines, closer to our present measurements than to previous measurements. The agreement with our updated results is therefore excellent despite the lack of specific re-calibration. This is encouraging evidence for a patchy, late end to hydrogen reionization.

6 CONCLUSIONS

We have measured the mean Ly-α optical depth at $4.8 < z < 6.2$ by assembling a sample of 67 high-SNR quasar sightlines, leveraging the new XQR-30 sample of X-Shooter spectra of $z \gtrsim 5.8$ quasars. Our sample represents a ∼3-fold increase in the number of high-quality spectra of Ly-α transmission at the end stages of reionization. We only employ observations taken with two spectrographs, enabling us to rigorously quantify systematics in instrumentation and continuum reconstruction for all our observations. The depth of observations, SNR > 10 per spectral pixel, also enables a more careful removal of possible DLA contaminants than previous studies.

Our measurement of the evolution of the mean Ly-α evolution with redshift is in rough agreement with previous work (Fig. 5). Differences are more likely to originate in previously-uncorrected systematics than in cosmic variance, given our large sample size and overlap with previous studies. We detect no sudden acceleration in the mean flux evolution over $4.8 < z < 5.5$.

We present an extremely transparent sightline with $\tau < 3$ at $z = 5.9$, and 2 rare sightlines with $\tau < 4$ at $z = 6.1$. These rare sightlines correspond to patches of the IGM with factors 5–15 times more transmitted flux than the median. The existence of transparent patches may help constrain future models of reionization, which must be able to generate both sightlines with $\tau \sim 2.5$ and $\tau > 6$ at the same redshift ($z = 5.9$).

Next, we determine the lowest redshift at which excess optical depth scatter in Ly-α emerges, signalling a departure from a uniformly ionized IGM. Using an improved grasp on systematics, we forward-model two simulation models employing homogeneous UVBs, the Sherwood and Nyx simulations. We conduct a maximum-likelihood analysis to obtain the probability of the full observed dataset at each step of $\Delta z = 0.1$. All observational systematics (wavelength masking, observational uncertainties, continuum uncertainties, etc) are included in post-processing of the simulations. These uncertainties result in increased Ly-α optical depth scatter which improves the agreement between models and observations.

We find excellent agreement between the forward-modelled simulations and observations at $5.0 \leq z \leq 5.2$, where the observed data has a high probability of being observed by chance ($<1\sigma$). A homogeneous UVB is in mild tension with observations at $z = 5.3$ (2.5σ) and strongly excluded at $z \geq 5.4$ ($>4\sigma$). To check whether the disagreement at $z = 5.4$ is driven by a few opaque sightlines which may contain DLAs, we arbitrarily remove the three least transmissive sightlines which have $\tau > 4$. Homogeneous UVB models remain excluded at $z = 5.4$ at $>3.5\sigma$ confidence, meaning that the intrinsically large width of the observed distribution, and not just a few sightlines, is driving the tension. Despite differences in the box size, snapshot density, and UVB models between the two suites, our results are consistent between the Sherwood and Nyx simulations.

Since the Sherwood model has been highly successful in modelling the Ly-α forest over a wide range of redshifts ($2 < z < 5.2$), a sudden failure by $z = 5.4$ represents a breakdown of one or more simplifying assumptions. Whether fluctuations in the UVB are present at very late times and/or whether the thermal state of the IGM retains the imprint of recent ionization, it is clear that reionization-related fluctuations persist in the IGM until at least $z = 5.3$.

Finally, we convert our measurements of the mean Ly-α flux to volume-averaged neutral fraction $x_{\rm H\,I}$. We stress that this conversion is model-dependent; here we use the Nyx simulation suite. Since Nyx (and homogeneous-UVB models in general) provides a very poor fit to data at $z \geq 5.4$, only lower limits on $x_{\rm H\,I}$ can be quoted. Our results at $5.0 \leq z \leq 5.3$ are in mild tension with those or Fan et al. (2006; $\sim 2\sigma$), but the vast improvements in data quality, quantity, and understanding of systematics and IGM modelling over the last 15 yr makes it difficult to pinpoint the source of the disagreement.

The XQR-30 sample has qualitatively changed the landscape of the late stages of reionization. Analysis of Ly-α transmission at $z > 5$ has become a precision probe of the post-reionization era, with exciting prospects both on the analysis and theoretical fronts. Through excellent complementarity with upcoming 21-cm probes, IGM transmission studies make it possible to uncover the entire history of reionization from start to end.

ACKNOWLEDGEMENTS

This research has made use of NASA's Astrophysics Data System, and open-source projects including IPYTHON (Perez & Granger 2007), SCIPY (Virtanen et al. 2019), NUMPY (van der Walt, Colbert & Varoquaux 2011), ASTROPY (Astropy Collaboration 2013; Price-Whelan et al. 2018), SCIKIT-LEARN (Pedregosa et al. 2011), and MATPLOTLIB (Hunter 2007).

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

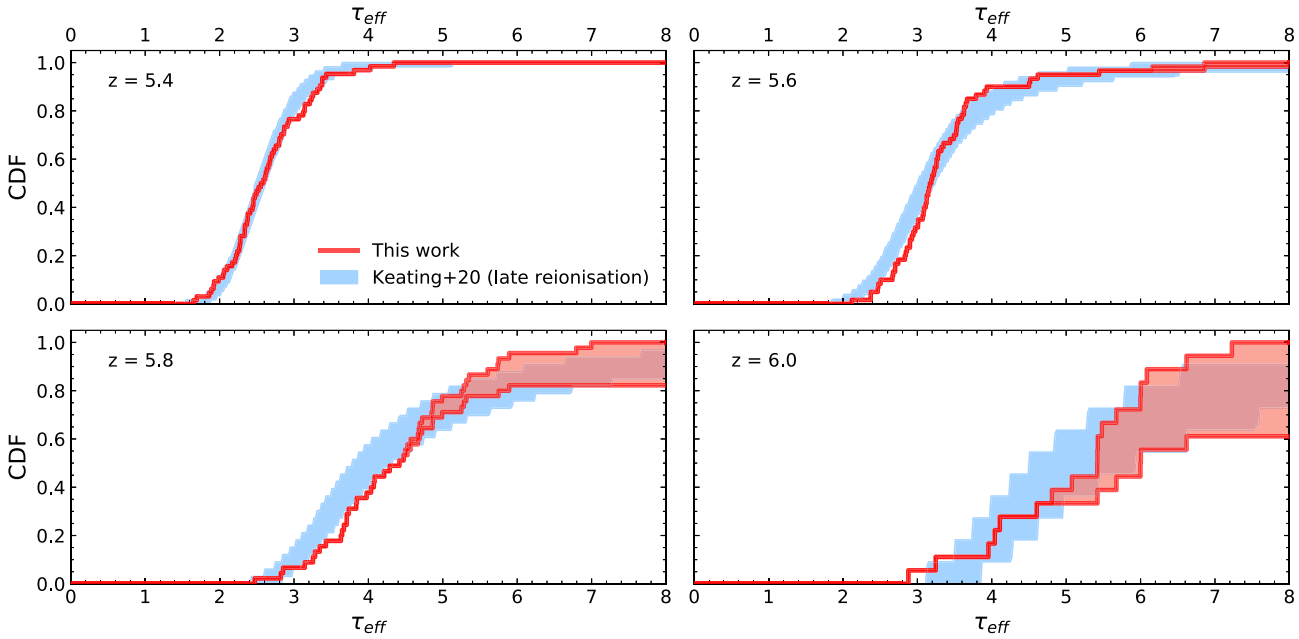

Figure 15. Comparison of the optical depth distributions measured in this work (red) to a model of late reionization (Keating et al. 2020, blue). The contours of the blue distribution correspond to the central 1σ(70 per cent) bounds of the simulated distributions. Late reionization provides an excellent qualitative description of the Ly-α optical depth scatter above $= 5.4$. The late reionization model was not calibrated to reproduce our updated measurements to mean Ly-α transmitted flux, precluding a direct quantitative comparison.

SEIB, RAM, FW, and MO acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 740246 'Cosmic Gas'). AP acknowledges support from the ERC Advanced Grant INTERSTELLAR H2020/740120. AM acknowledges funding from the ERC under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 638809 – AIDA). The results presented here reflect the authors' views; the ERC is not responsible for their use.

LCK was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 885990.

ACE acknowledges support by NASA through the NASA Hubble Fellowship grant #HF2-51434 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. FW acknowledges support by NASA through the NASA Hubble Fellowship grant #*HST*-HF2-51448.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555.

GDB and YZ are supported by the National Science Foundation through grant AST-1751404. JFH acknowledges support from the National Science Foundation under Grant No. 1816006.

RD, ERW, and YQ acknowledge the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

FB acknowledges support from the Australian Research Council through Discovery Projects (award DP190100252) and Chinese Academy of Sciences (CAS) through a China-Chile Joint Research Fund (CCJRF1809) administered by the CAS South America Center for Astronomy (CASSACA).

Based on observations collected at the European Southern Observatory under ESO programme 1103.A-0817(A).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

The Sherwood simulation was performed with super-computer time awarded by the Partnership for Advanced Computing in Europe (PRACE) 8th call. This project also made use of the DiRAC High Performance Computing System (HPCS) and the COSMOS shared memory service at the University of Cambridge. These

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

are operated on behalf of the Science and Technology Facilities Council (STFC) DiRAC HPC facility. This equipment is funded by BIS National E-infrastructure capital grant ST/J005673/1 and STFC grants ST/H008586/1, ST/K00333X/1.

Calculations presented in this paper used resources of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

DATA AVAILABILITY

The quasar spectra used in this analysis will be shared on reasonable request to the corresponding author. The XQR-30 spectra and associated meta-data will further be made public in the upcoming data release of D'Odorico et al. (in preparation).

All measurements of the optical depth generated and used in this work are available in the paper and its online supplementary material.

REFERENCES

Almgren A. S., Bell J. B., Lijewski M. J., Lukić Z., Van Andel E., 2013, ApJ, 765, 39
Astropy Collaboration, 2013, A&A, 558, A33
Aubert D., Teyssier R., 2008, MNRAS, 387, 295
Aubert D., Teyssier R., 2010, ApJ, 724, 244
Bañados E. et al., 2014, AJ, 148, 14
Bañados E. et al., 2016, ApJS, 227, 11
Bañados E. et al., 2019, ApJ, 885, 59
Barnett R., Warren S. J., Becker G. D., Mortlock D. J., Hewett P. C., McMahon R. G., Simpson C., Venemans B. P., 2017, A&A, 601, A16
Becker G. D., Bolton J. S., 2013, MNRAS, 436, 1023
Becker G. D., Rauch M., Sargent W. L. W., 2009, ApJ, 698, 1010
Becker G. D., Hewett P. C., Worseck G., Prochaska J. X., 2013, MNRAS, 430, 2067
Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, MNRAS, 447, 3402
Becker G. D. et al., 2019, ApJ, 883, 163
Becker G. D., D'Aloisio A., Christenson H. M., Zhu Y., Worseck G., Bolton J. S., 2021, MNRAS, 508, 1853
Bischetti M. et al.,2022, preprint (arXiv:2205.00021)
Bolton J. S., Haehnelt M. G., 2007, MNRAS, 382, 325
Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897
Bosman S. E. I., 2021, preprint (arXiv:2108.12446)
Bosman S., 2020, Zenodo Dataset, Zenodo, available at https://doi.org/10.5281/zenodo.3634964
Bosman S. E. I., Becker G. D., 2015, MNRAS, 452, 1105
Bosman S. E. I., Becker G. D., Haehnelt M. G., Hewett P. C., McMahon R. G., Mortlock D. J., Simpson C., Venemans B. P., 2017, MNRAS, 470, 1919
Bosman S. E. I., Fan X., Jiang L., Reed S., Matsuoka Y., Becker G., Haehnelt M., 2018, MNRAS, 479, 1055
Bosman S. E. I., Ďurovčíková D., Davies F. B., Eilers A.-C., 2021, MNRAS, 503, 2077 (B21)
Cain C., D'Aloisio A., Gangolli N., Becker G. D., 2021, ApJ, 917, L37
Carilli C. L. et al., 2010, ApJ, 714, 834
Carnall A. C. et al., 2015, MNRAS, 451, L16
Carswell R. F., Whelan J. A. J., Smith M. G., Boksenberg A., Tytler D., 1982, MNRAS, 198, 91
Cen R., Haiman Z., 2000, ApJ, 542, L75
Chardin J., Puchwein E., Haehnelt M. G., 2017, MNRAS, 465, 3429
Chardin J., Haehnelt M. G., Bosman S. E. I., Puchwein E., 2018, MNRAS, 473, 765
Chehade B. et al., 2018, MNRAS, 478, 1649
Choudhury T. R., Paranjape A., Bosman S. E. I., 2021, MNRAS, 501, 5782
Cooper T. J., Simcoe R. A., Cooksey K. L., Bordoloi R., Miller D. R., Furesz G., Turner M. L., Bañados E., 2019, ApJ, 882, 77
Dall'Aglio A., Wisotzki L., Worseck G., 2008, A&A, 491, 465
Dall'Aglio A., Wisotzki L., Worseck G., 2009, preprint (arXiv:0906.1484)
D'Aloisio A., McQuinn M., Trac H., 2015, ApJ, 813, L38
D'Aloisio A., McQuinn M., Davies F. B., Furlanetto S. R., 2018, MNRAS, 473, 560
D'Odorico V. et al., 2018, ApJ, 863, L29
Davies F. B., 2020, MNRAS, 494, 2937
Davies F. B., Furlanetto S. R., 2016, MNRAS, 460, 1328
Davies F. B., Hennawi J. F., Eilers A.-C., Lukić Z., 2018a, ApJ, 855, 106
Davies F. B. et al., 2018b, ApJ, 864, 142
Davies F. B. et al., 2018c, ApJ, 864, 143
Davies F. B., Bosman S. E. I., Furlanetto S. R., Becker G. D., D'Aloisio A., 2021, ApJ, 918, L35
Dawson K. S. et al., 2013, AJ, 145, 10
Dawson K. S. et al., 2016, AJ, 151, 44
Dayal P., Ferrara A., 2018, Phys. Rep., 780, 1
De Rosa G., Decarli R., Walter F., Fan X., Jiang L., Kurk J., Pasquali A., Rix H. W., 2011, ApJ, 739, 56
DeBoer D. R. et al., 2017, PASP, 129, 045001
Decarli R. et al., 2018, ApJ, 854, 97
Ďurovčíková D., Katz H., Bosman S. E. I., Davies F. B., Devriendt J., Slyz A., 2020, MNRAS, 493, 4256
Eilers A.-C., Davies F. B., Hennawi J. F., Prochaska J. X., Lukić Z., Mazzucchelli C., 2017, ApJ, 840, 24
Eilers A.-C., Davies F. B., Hennawi J. F., 2018, ApJ, 864, 53
Eilers A.-C., Hennawi J. F., Davies F. B., Oñorbe J., 2019, ApJ, 881, 23
Eilers A.-C. et al., 2020, ApJ, 900, 37
Eilers A.-C. et al., 2021, ApJ, 914, 74
Fan X. et al., 2000, AJ, 120, 1167
Fan X. et al., 2001, AJ, 122, 2833
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., 2003, AJ, 125, 1649
Fan X. et al., 2004, AJ, 128, 515
Fan X. et al., 2006, AJ, 132, 117
Farina E. P. et al., 2019, ApJ, 887, 196
Faucher-Giguère C.-A., Prochaska J. X., Lidz A., Hernquist L., Zaldarriaga M., 2008, ApJ, 681, 831
Francis P. J., Hewett P. C., Foltz C. B., Chaffee F. H., 1992, ApJ, 398, 476
Freudling W., Romaniello M., Bramich D. M., Ballester P., Forchi V., García-Dabló C. E., Moehler S., Neeser M. J., 2013, A&A, 559, A96
Gaikwad P. et al., 2020, MNRAS, 494, 5091
Gilmozzi R., Spyromilio J., 2007, Messenger, 127, 11
Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, MNRAS, 466, 4239
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Haardt F., Madau P., 2012, ApJ, 746, 125
Horne K., 1986, PASP, 98, 609
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jiang L., Fan X., Vestergaard M., Kurk J. D., Walter F., Kelly B. C., Strauss M. A., 2007, AJ, 134, 1150
Jiang L. et al., 2008, AJ, 135, 1057
Jiang L., McGreer I. D., Fan X., Bian F., Cai Z., Clément B., Wang R., Fan Z., 2015, AJ, 149, 188
Jiang L. et al., 2016, ApJ, 833, 222
Keating L. C., Puchwein E., Haehnelt M. G., 2018, MNRAS, 477, 5501
Keating L. C., Weinberger L. H., Kulkarni G., Haehnelt M. G., Chardin J., Aubert D., 2020, MNRAS, 491, 1736
Kelson D. D., 2003, PASP, 115, 688
Kulkarni G., Keating L. C., Haehnelt M. G., Bosman S. E. I., Puchwein E., Chardin J., Aubert D., 2019, MNRAS, 485, L24
Kurk J. D. et al., 2007, ApJ, 669, 32
Lidz A., Oh S. P., Furlanetto S. R., 2006, ApJ, 639, L47
Lukić Z., Stark C. W., Nugent P., White M., Meiksin A. A., Almgren A., 2015, MNRAS, 446, 3697
Mazzucchelli C. et al., 2017, ApJ, 849, 91

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

McDonald P., Seljak U., Cen R., Bode P., Ostriker J. P., 2005, MNRAS, 360, 1471
Meiksin A., 2020, MNRAS, 491, 4884
Mesinger A., 2010, MNRAS, 407, 1328
Mesinger A., Haiman Z., 2004, ApJ, 611, L69
Meyer R. A., Bosman S. E. I., Kakiichi K., Ellis R. S., 2019a, MNRAS, 483, 19
Meyer R. A., Bosman S. E. I., Ellis R. S., 2019b, MNRAS, 487, 3305
Morganson E. et al., 2012, AJ, 143, 142
Mortlock D. J. et al., 2009, A&A, 505, 97
Mortlock D. J. et al., 2011, Nature, 474, 616
Nasir F., D'Aloisio A., 2020, MNRAS, 494, 3080
Oñorbe J., Hennawi J. F., Lukić Z., Walther M., 2017, ApJ, 847, 63
Onoue M. et al., 2020, ApJ, 898, 105
Pâris I. et al., 2011, A&A, 530, A50
Pedregosa F. et al., 2011, J. Mach. Learn. Research, 12, 2825
Perez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21
Planck Collaboration, 2020, A&A, 641, A6
Price-Whelan A. M. et al., 2018, AJ, 156, 123
Prochaska J., et al., 2020a, J. Open Source Softw., 5, 2308
Prochaska J. X. et al., 2020b, pypeit/PypeIt: Release 1.0.0, Zenodo, available at https://doi.org/10.5281/zenodo.3743493
Qin Y., Mesinger A., Bosman S. E. I., Viel M., 2021, MNRAS, 506, 2390
Rafelski M., Wolfe A. M., Prochaska J. X., Neeleman M., Mendez A. J., 2012, ApJ, 755, 89
Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, MNRAS, 430, 2427
Raste J., Kulkarni G., Keating L. C., Haehnelt M. G., Chardin J., Aubert D., 2021, MNRAS, 507, 4684
Reed S. L. et al., 2017, MNRAS, 468, 4702
Rollinde E., Theuns T., Schaye J., Pâris I., Petitjean P., 2013, MNRAS, 428, 540
Sanders G. H., 2013, J. Astrophys. Astron., 34, 81
Schindler J.-T. et al., 2020, ApJ, 905, 51
Schmidt G. D., Weymann R. J., Foltz C. B., 1989, PASP, 101, 713
Sheinis A. I., Bolte M., Epps H. W., Kibrick R. I., Miller J. S., Radovan M. V., Bigelow B. C., Sutin B. M., 2002, PASP, 114, 851
Shen Y. et al., 2019, ApJ, 873, 35
Šoltinský T. et al., 2021, MNRAS, 506, 5818
Springel V., 2005, MNRAS, 364, 1105
Suzuki N., Tytler D., Kirkman D., O'Meara J. M., Lubin D., 2005, ApJ, 618, 592
Trott C. M., Pober J., 2019, in 2514-3433, The Cosmic 21-cm Revolution. IOP Publishing, p. 8
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Venemans B. P. et al., 2015, MNRAS, 453, 2259
Venemans B. P. et al., 2018, ApJ, 866, 159
Venemans B. P. et al., 2020, ApJ, 904, 130
Vernet J. et al., 2011, A&A, 536, A105
Virtanen P., 2019, scipy/scipy: SciPy 1.2.1. Zenodo, available at https://doi.org/10.5281/zenodo.3533894
Vogt S. S. et al., 1994, in Crawford D. L., Craine E. R., eds, Proc. SPIE Conf. Ser. Vol. 2198, Instrumentation in Astronomy VIII. SPIE, Bellingham, p. 362
Wang F. et al., 2016, ApJ, 819, 24
Wang F. et al., 2019, ApJ, 884, 30
Wang F. et al., 2020, ApJ, 896, 23
Wang F. et al., 2021, ApJ, 908, 53
Wang R. et al., 2010, ApJ, 714, 699
Wang R. et al., 2013, ApJ, 773, 44
Willott C. J. et al., 2007, AJ, 134, 2435
Willott C. J. et al., 2010, AJ, 139, 906
Willott C. J., Bergeron J., Omont A., 2015, ApJ, 801, 123
Wolfe A. M., Gawiser E., Prochaska J. X., 2005, ARA&A, 43, 861
Worseck G. et al., 2014, MNRAS, 445, 1745
Worseck G., Prochaska J. X., Hennawi J. F., McQuinn M., 2016, ApJ, 825, 144
Wu X.-B. et al., 2015, Nature, 518, 512
Yang J. et al., 2017, AJ, 153, 184
Yang J. et al., 2020, ApJ, 904, 26
Yip C. W. et al., 2004, AJ, 128, 2603
Young P. J., Sargent W. L. W., Boksenberg A., Carswell R. F., Whelan J. A. J., 1979, ApJ, 229, 891
Zhu Y. et al., 2021, ApJ, 923, 223

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

online_material.zip

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: ALTERNATIVE DATA BINNING

In this paper, we bin the data in fixed redshift intervals of equal size $\Delta z = 0.1$. Fixed redshift bins to equal comoving length may sometimes be more convenient for model comparison. In Fig. A1 we show a such a sub-division of the data between $4.9 \lesssim z \lesssim 6.2$ in 8 bins with fixed $\Delta L = 50$ cMpc h^{-1}. The mid-points and edges of each bin are given in each panel as well as in Table A1. The average optical depths measured in this manner, as well as the qualitative evolution of the optical depth distribution, are fully consistent with those obtained in the paper's main body. The number of non-detections is slightly lessened due to averaging of the optical depth over a larger interval. Distributions of optical depths on different scales are expected to be sensitive to different optical effects. We make these distributions available as supplemental online material, as well as the distributions with intervals of $\Delta z = 0.05, 0.1, 0.2$ and $\Delta L = 30, 50, 70$ cMpc h^{-1}. Note that intervals $\Delta z \gtrsim 0.1$ are subject to variations of the mean flux $>1\sigma$ between their edges (Fig. 5).

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

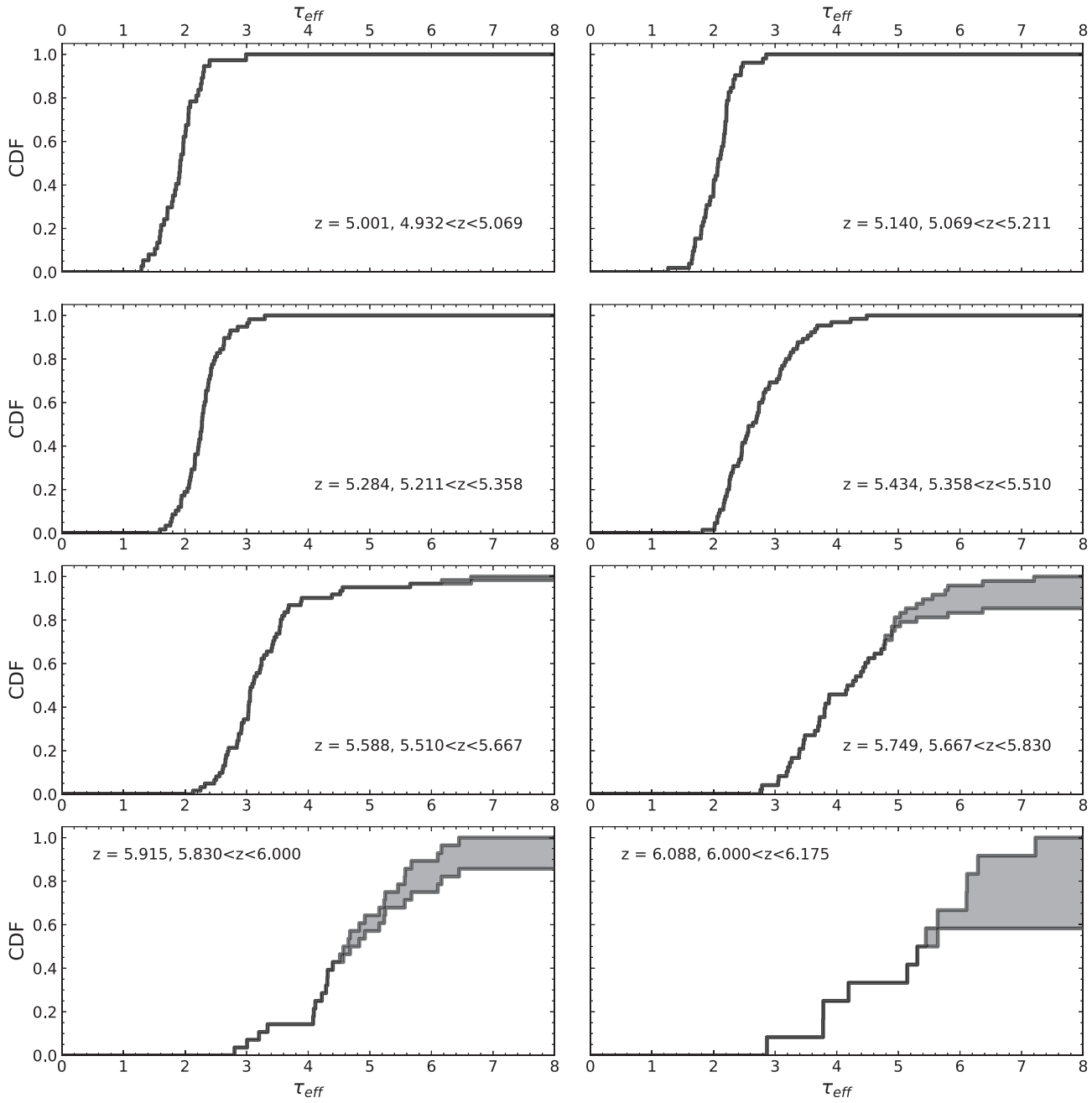

Figure A1. Same as Fig. 9, but with fixed redshift intervals of constant length $\Delta L = 50$ cMpc h^{-1} as indicated in each panel. The qualitative evolution is unchanged: the first opaque troughs emerge at $z \sim 5.6$.

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022

Table A1. Mean Ly-α flux transmission at $4.9 \lesssim z \lesssim 6.2$, measured in $\Delta L = 50$ cMpc h^{-1} bins with fixed redshift bounds $z_{\rm min}$ and $z_{\rm max}$. Uncertainties correspond to the 16th and 84th percentiles from bootstrap resampling. The measurement uncertainties on their own are a factor 5–10 smaller than the bootstrap uncertainties quoted here. $N_{\rm los}$ sightlines contribute to each measurement.

z	$z_{\rm min}$	$z_{\rm max}$	$<F_{\rm Ly\text{-}\alpha}>$	-1σ	$+1\sigma$	$N_{\rm los}$
5.000	4.932	5.069	0.1545	−0.0085	+0.0080	37
5.140	5.069	5.210	0.1329	−0.0056	+0.0054	52
5.284	5.210	5.357	0.1073	−0.0043	+0.0046	58
5.433	5.357	5.509	0.0741	−0.0039	+0.0047	65
5.588	5.509	5.667	0.0458	−0.0032	+0.0031	61
5.748	5.667	5.830	0.0192	−0.0025	+0.0022	48
5.915	5.830	5.999	0.0126	−0.0031	+0.0028	28
6.087	5.999	6.175	0.0091	−0.0051	+0.0051	12

APPENDIX B: $z = 5.4$ DISTRIBUTION WITHOUT MOST OPAQUE SIGHTLINES

To determine whether the tension at $z = 5.4$ is due to outlier sightlines (potential DLAs), we arbitrarily remove the top three most opaque sightlines and re-run the likelihood analysis. The results are shown in Fig. B1. While the tension is reduced compared to including the opaque sightlines, the tension remains above 3σ for both the Sherwood and Nyx simulations. The test therefore indicates that the entire shape of the Ly-α optical depth distribution at $z = 5.4$, and not just a few outliers, drive the tension with homogeneous UVB models.

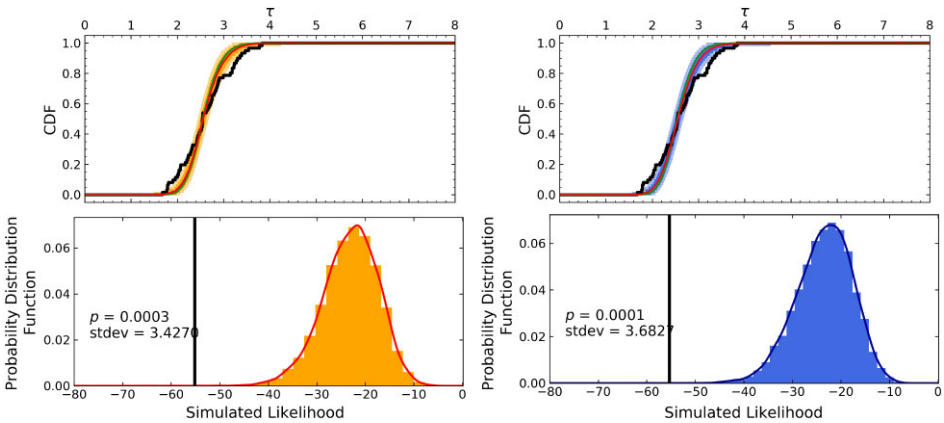

Figure B1. Same as Fig. 11 but distributions at $z = 5.4$ arbitrarily excluding the most opaque three sightlines with $\tau > 4$. Left: Results for Sherwood. Right: Results for Nyx. In both simulations, a tension at $>3\sigma$ remains even after removing the most opaque sightlines.

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

Downloaded from https://academic.oup.com/mnras/article/514/1/55/6598046 by a.venera@unict.it user on 19 July 2022