



Long Dark Gaps in the Ly β Forest at $z < 6$: Evidence of Ultra-late Reionization from XQR-30 Spectra

Yongda Zhu¹, George D. Becker¹, Sarah E. I. Bosman², Laura C. Keating³, Valentina D’Odorico^{4,5,6}, Rebecca L. Davies^{7,8}, Holly M. Christenson¹, Eduardo Bañados², Fuyan Bian⁹, Manuela Bischetti⁴, Huanqing Chen¹⁰, Frederick B. Davies², Anna-Christina Eilers^{11,17}, Xiaohui Fan¹², Prakash Gaikwad², Bradley Greig^{8,13}, Martin G. Haehnelt¹⁴, Girish Kulkarni¹⁵, Samuel Lai¹⁶, Andrea Pallottini⁵, Yuxiang Qin^{8,13}, Emma V. Ryan-Weber^{7,8}, Fabian Walter², Feige Wang^{12,17}, and Jinyi Yang^{12,18}

¹ Department of Physics & Astronomy, University of California, Riverside, CA 92521, USA; yzhu144@ucr.edu

² Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

³ Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

⁴ INAF-Osservatorio Astronomico di Trieste, Via Tiepolo 11, I-34143 Trieste, Italy

⁵ Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

⁶ IFPU-Institute for Fundamental Physics of the Universe, via Beirut 2, I-34151 Trieste, Italy

⁷ Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia

⁸ ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia

⁹ European Southern Observatory, Alonso de Córdova 3107, Casilla 19001, Vitacura, Santiago 19, Chile

¹⁰ Department of Astronomy & Astrophysics, University of Chicago, Chicago, IL 60637, USA

¹¹ MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

¹² Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

¹³ School of Physics, University of Melbourne, Parkville, VIC 3010, Australia

¹⁴ Kavli Institute for Cosmology and Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK

¹⁵ Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India

¹⁶ Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

Received 2022 February 25; revised 2022 April 15; accepted 2022 May 9; published 2022 June 17

Abstract

We present a new investigation of the intergalactic medium near reionization using dark gaps in the Ly β forest. With its lower optical depth, Ly β offers a potentially more sensitive probe to any remaining neutral gas compared to the commonly used Ly α line. We identify dark gaps in the Ly β forest using spectra of 42 QSOs at $z_{\text{em}} > 5.5$, including new data from the XQR-30 VLT Large Programme. Approximately 40% of these QSO spectra exhibit dark gaps longer than $10 h^{-1}$ Mpc at $z \simeq 5.8$. By comparing the results to predictions from simulations, we find that the data are broadly consistent both with models where fluctuations in the Ly α forest are caused solely by ionizing ultraviolet background fluctuations and with models that include large neutral hydrogen patches at $z < 6$ due to a late end to reionization. Of particular interest is a very long ($L = 28 h^{-1}$ Mpc) and dark ($\tau_{\text{eff}} \gtrsim 6$) gap persisting down to $z \simeq 5.5$ in the Ly β forest of the $z = 5.85$ QSO PSO J025–11. This gap may support late reionization models with a volume-weighted average neutral hydrogen fraction of $\langle x_{\text{HI}} \rangle \gtrsim 5\%$ by $z = 5.6$. Finally, we infer constraints on $\langle x_{\text{HI}} \rangle$ over $5.5 \lesssim z \lesssim 6.0$ based on the observed Ly β dark gap length distribution and a conservative relationship between gap length and neutral fraction derived from simulations. We find $\langle x_{\text{HI}} \rangle \leq 0.05, 0.17$, and 0.29 at $z \simeq 5.55, 5.75$, and 5.95 , respectively. These constraints are consistent with models where reionization ends significantly later than $z = 6$.

Unified Astronomy Thesaurus concepts: Reionization (1383); Intergalactic medium (813); Quasar absorption line spectroscopy (1317); High-redshift galaxies (734)

Supporting material: data behind figure, figure set

1. Introduction

Determining when and how reionization occurred is essential for understanding the intergalactic medium (IGM) physics and galaxy formation in the early universe (e.g., Muñoz et al. 2022). Recent observations have made significant progress on establishing the timing of reionization and largely point to a midpoint of $z \sim 7$ – 8 . These observations include the electron

optical depth to the cosmic microwave background (CMB) photons (Planck Collaboration et al. 2020; see also de Belsunce et al. 2021), the Ly α damping wing in $z \gtrsim 7$ QSO spectra (e.g., Bañados et al. 2018; Davies et al. 2018b; Wang et al. 2020; Yang et al. 2020b; Greig et al. 2022), the decline in observed Ly α emission from $z > 6$ galaxies (e.g., Mason et al. 2018, 2019; Hoag et al. 2019; Hu et al. 2019, and references therein; but see Jung et al. 2020; Wold et al. 2022), and the IGM thermal state measurements at $z > 5$ (e.g., Boera et al. 2019; Gaikwad et al. 2021).

The observations noted above are generally consistent with reionization ending by $z \sim 6$, a scenario supported by existing measurement of the fraction of dark pixels in the Lyman series forests (e.g., McGreer et al. 2011, 2015). Other observations, however, suggest a significantly later end of reionization. The

¹⁷ NASA Hubble Fellow.

¹⁸ Strittmatter Fellow.



large-scale fluctuations in the measured Ly α effective optical depth, $\tau_{\text{eff}} = -\ln\langle F \rangle$, where F is the continuum-normalized transmission flux (e.g., Fan et al. 2006; Becker et al. 2015; Eilers et al. 2018; Yang et al. 2020a; Bosman et al. 2022), together with long troughs extending to or below $z \simeq 5.5$ in the Ly α forest (e.g., Becker et al. 2015; Zhu et al. 2021), potentially indicate the existence of large neutral IGM islands (e.g., Kulkarni et al. 2019; Keating et al. 2020b; Nasir & D’Aloisio 2020). The underdensities around long dark gaps traced by Ly α emitters (LAEs) are also consistent with a late reionization model wherein reionization ends at $z < 6$ (Becker et al. 2018; Kashino et al. 2020; Christenson et al. 2021).

These Ly α forest and LAE results are potentially consistent with an alternative scenario wherein the IGM is ionized by $z = 6$ but retains large-scale fluctuations in the ionizing UV background down to lower redshifts (Davies et al. 2018a). On the other hand, recent measurements of the mean free path of ionizing photons measured at $z = 5.1$ and 6.0 (Becker et al. 2021) are difficult to reconcile with models wherein reionization ends at $z > 6$ and may instead prefer models wherein the IGM is still 20% neutral or more at $z = 6$ (Becker et al. 2021; Cain et al. 2021; Davies et al. 2021). In addition, a reionization ending at $z < 6$ is consistent with models that reproduce a variety of observations (e.g., Weinberger et al. 2019; Choudhury et al. 2021; Qin et al. 2021).

One way of searching for signatures of late ($z_{\text{end}} < 6$) reionization in the Ly α forest is by investigating dark gaps, i.e., contiguous regions of strong absorption (e.g., Songaila & Cowie 2002; Furlanetto et al. 2004; Paschos & Norman 2005; Fan et al. 2006; Gallerani et al. 2008; Gnedin et al. 2017; Nasir & D’Aloisio 2020). In Zhu et al. (2021, hereafter Paper I), we explored long dark gaps in the Ly α forest and found that a fully ionized IGM with a homogeneous ultraviolet background (UVB) is strongly ruled out down to $z \simeq 5.3$. In contrast, late reionization models and a model wherein reionization ends by $z \sim 6$ but retains large-scale UVB fluctuations are consistent with the observations. Predictions for the Ly α dark gap statistics are similar between the two types of models. This is largely because realistic late reionization models also include UVB fluctuations, which are often associated with the neutral islands. Ly α also tends to saturate at a relatively low ($x_{\text{HI}} \sim 10^{-3}$) neutral fraction, limiting its sensitivity to neutral gas.

Given its lower optical depth,¹⁹ Ly β should be a more sensitive probe of neutral gas in the $z \lesssim 6$ IGM. As a result, ultra-late reionization models wherein neutral islands persist down to $z < 5.5$ may produce more long Ly β dark gaps than can be produced by UVB fluctuations alone. Based on this feature, we can potentially use dark gaps in the Ly β forest to place stronger constraints on the timing of reionization and distinguish the late reionization models from the pure fluctuating UVB models. As presented in Nasir & D’Aloisio (2020), distributions of dark gaps in the Ly β forest differ between these models most strongly on scales of $L \gtrsim 10 h^{-1} \text{ Mpc}$. We are therefore particularly interested in these long dark gaps.

In this work, we use 42 high-quality QSO spectra that allow us to search for dark gaps in the Ly β forest over the redshift

range $5.5 < z < 6.0$. The sample includes high-quality X-Shooter spectra from the XQR-30 VLT large program²⁰ (V. D’Odorico et al. 2022, in preparation). In addition to comparing the results to model predictions, we also constrain $\langle x_{\text{HI}} \rangle$ based on a conservative relationship between dark gap length and neutral fraction derived from simulations.

This paper is organized as follows. In Section 2 we describe the data and results from the observations. Section 3 compares our results to model predictions, discusses the implications for reionization, and infers constraints on x_{HI} . Finally, we conclude the findings in Section 4. Throughout this paper, we assume a Λ CDM cosmology with $\Omega_{\text{m}} = 0.308$, $\Omega_{\Lambda} = 0.692$, and $h = 0.678$ (Planck Collaboration et al. 2014). Distances are quoted in comoving units unless otherwise noted.

2. The Data and Results

2.1. QSO Spectra

Our sample includes 42 out of the 43 spectra of QSOs at $5.77 \lesssim z_{\text{em}} \lesssim 6.31$ that were used for Paper I. The exception is PSO J004+17, whose spectrum has lower signal-to-noise ratio (S/N) that does not meet the requirement of our flux threshold for the Ly β forest (Section 2.3). The spectra are taken with the Echelle Spectrograph and Imager (ESI) on Keck (Sheinis et al. 2002) and the X-Shooter spectrograph on the Very Large Telescope (VLT; Vernet et al. 2011). Among these, 19 X-Shooter spectra are from the XQR-30 VLT large program (V. D’Odorico et al. 2022, in preparation). The details of the data reduction procedures are given in Paper I and Becker et al. (2019). We note that the targets are selected without foreknowledge of dark gaps in the Ly β forest. Figure Set 1 displays the spectra, continuum fits, and dark gaps detected in the Ly β forest for each QSO (see details below). An example is given in Figure 1.

2.2. Continuum Fitting

We use principal component analysis (PCA) to predict the unabsorbed QSO continuum and normalize the transmission in the Ly β forest. We follow a similar method to that described in Paper I to fit and predict the continuum. Briefly, we implement and apply the log-PCA method of Davies et al. (2018c) in the Ly α and Ly β forest portion of the spectrum following Bosman et al. (2022). The continuity between the Ly α forest and the Ly β forest continuum is broken intentionally to correct for the effect of overlapping Ly α absorption in the Ly β forest in the PCA training sample. We fit the red-side (rest-frame wavelength $\lambda_0 > 1230 \text{ \AA}$) continuum up to 2000 \AA in the rest frame for X-Shooter spectra with NIR observations. The ESI spectra are fit using an optical-only PCA, which is presented in Bosman et al. (2021). The Ly β dark gap detection is not very sensitive to the continuum, and we also test that using a power-law continuum does not significantly change the dark gap results in this work.

2.3. Dark Gap Detection

Similar to the definition of a dark gap in the Ly α forest in Paper I, we define a dark gap in the Ly β forest to be a continuous spectral region in which all pixels binned to $1 h^{-1} \text{ Mpc}$ have an observed normalized flux $F = F_{\text{obs}}/F_c < 0.02$, where F_{obs} is the

¹⁹ Ly β absorption has a shorter wavelength ($\lambda_{\text{Ly}\beta} = 1025.72 \text{ \AA}$) and a lower oscillator strength ($f_{\text{Ly}\beta} = 0.0791$) compared to those of Ly α absorption ($\lambda_{\text{Ly}\alpha} = 1215.67 \text{ \AA}$, $f_{\text{Ly}\alpha} = 0.4164$). The ratio of optical depth is given by $\tau_{\text{Ly}\beta}/\tau_{\text{Ly}\alpha} = (f_{\text{Ly}\beta}\lambda_{\text{Ly}\beta})/(f_{\text{Ly}\alpha}\lambda_{\text{Ly}\alpha}) \simeq 0.16$.

²⁰ <https://xqr30.inaf.it>

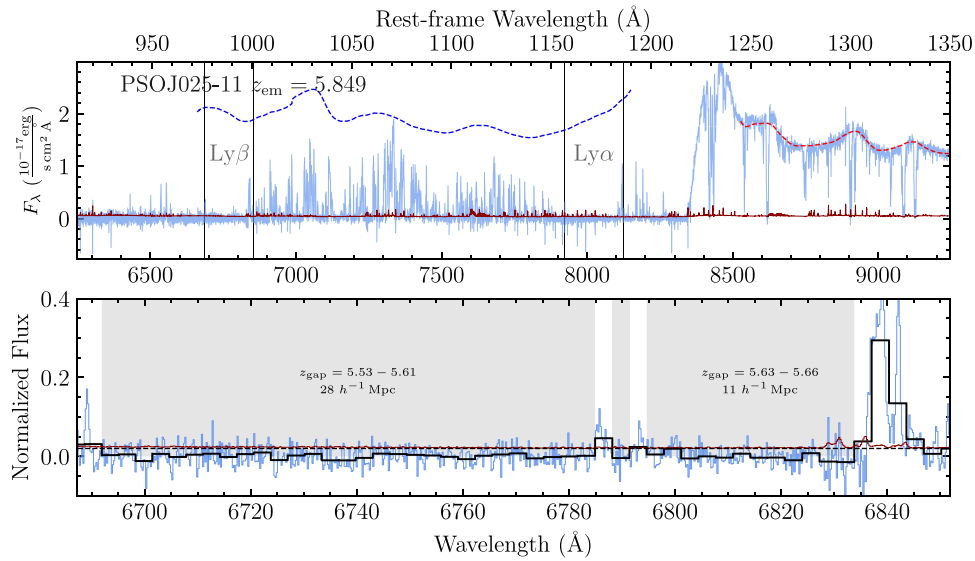


Figure 1. Top panel: spectrum and continuum fits for the $z_{\text{em}} = 5.849$ QSO PSO J025–11. The light-blue and dark-red lines represent flux and flux error in the original binning, respectively. Dashed curves redward and blueward of the $\text{Ly}\alpha$ peak show the best-fitting and predicted QSO continuum in absence of absorption, respectively. The fitting and prediction are based on PCA. The continuity between the $\text{Ly}\alpha$ forest and the $\text{Ly}\beta$ forest continuum is broken intentionally near 1020 Å in the rest frame. We label the wavelength range over which we search for dark gaps in the $\text{Ly}\beta$ forest and its corresponding $\text{Ly}\alpha$ forest in redshift. Bottom panel: $\text{Ly}\beta$ forest and dark gaps detected. The dashed black line labels the flux threshold of 0.02. The thick black line displays the flux binned to $1 h^{-1}$ Mpc. Light-blue and dark-red lines show the flux and flux error in the original binning, respectively. Dark gaps detected are shaded with gray. We also label the redshift range and length of each long ($L \geq 10 h^{-1}$ Mpc) dark gap.

(The complete figure set (42 images) is available.)

observed flux and F_c is the continuum flux. The minimum length of a dark gap is $1 h^{-1}$ Mpc following Paper I. A bin size of $1 h^{-1}$ Mpc enables us to retain significant transmission profiles while reducing the influence of very small peaks on dark gap detection. The precise bin size should have relatively little impact on our results provided that the observations and mock spectra are treated consistently. We have verified that using bin sizes of $0.5 h^{-1}$ Mpc or $1.5 h^{-1}$ Mpc does not change our major conclusions, although the lengths of some dark gaps would change. A flux threshold of 0.02 is used here instead of 0.05, which we used for the $\text{Ly}\alpha$ gaps, because spectra in this subsample have higher S/N levels. In addition, the $\text{Ly}\beta$ forest at the redshifts that we are interested in is less contaminated by sky lines than the $\text{Ly}\alpha$ forest. We have tested that using a threshold of 0.05 will not change our results fundamentally, although the difference between the models (Section 3) may become less significant. In order to reduce false detections caused by foreground $\text{Ly}\alpha$ absorption, we further require that all $\text{Ly}\beta$ dark gaps correspond to $\text{Ly}\alpha$ dark gaps as defined in Paper I over the same redshifts for both the observed and mock (Section 3.1) spectra. That is to say, each $1 h^{-1}$ Mpc bin in the $\text{Ly}\beta$ dark gap also has a normalized flux less than 0.05 in the $\text{Ly}\alpha$ forest at the same redshift.²¹ For reference, we present the relationship between $\text{Ly}\beta$ dark gaps and $\text{Ly}\alpha$ dark gaps in Appendix A.

For each QSO sight line, dark gaps are detected from 976 Å in the rest frame to 11 proper-Mpc blueward of the QSO’s redshift, which corresponds to approximately 1000 Å in the rest frame. The lower wavelength limit ensures that the detection is not affected by the $\text{Ly}\gamma$ absorption. We use the higher limit to

avoid the QSO proximity zone transmission, and the cut is comparable to the choice in, e.g., Bosman et al. (2021). Following Paper I, we also exclude from the statistical analysis an additional $10 h^{-1}$ Mpc buffer zone blueward of the proximity zone cut. This allows the pixel at the red end of each sight line to intersect a possible long ($L \geq 10 h^{-1}$ Mpc) dark gap and hence helps to mitigate potential truncation issues.²²

To avoid the contamination from sky-line subtraction residuals, we mask out $\pm 75 \text{ km s}^{-1}$ intervals of the spectra centered at sharp peaks in the flux error array when searching for dark gaps. The exception is that we do not mask transmission with a $> 3\sigma$ detection. In addition, we make no attempts to correct for the effects of damped $\text{Ly}\alpha$ (DLA) systems or metal-enriched absorbers, although known systems in a subsample of the spectra with a relatively complete identification of metal-enriched systems are discussed in Appendix B. Figure 2 displays an overview of dark gaps detected in the $\text{Ly}\beta$ forest from our sample.

2.3.1. Dark Gaps Statistics

Figure 3 displays the statistical properties of dark gaps detected in the $\text{Ly}\beta$ forest from our sample. We detect 195 dark gaps in total, of which 24 have $L \geq 10 h^{-1}$ Mpc. Panel (a) plots length versus central redshift of these dark gaps. Long dark gaps become less common as redshift decreases. Nevertheless, some long gaps still exist down to $z < 5.6$.

We calculate the cumulative distribution function (CDF) of dark gap length, $P(< L)$. Dark gaps are sorted into two redshift

²¹ Based on our test, whether requiring gaps to be also dark in the $\text{Ly}\alpha$ forest or not only affects a small fraction of gaps and does not change the results in this paper significantly. Although this requirement may not remove all false detection, it partially avoids contamination from random foreground density fluctuations.

²² Without this additional buffer zone, it is possible that the F_{10} (Section 2.3.1) is underestimated near the red end of a sight line, since there can exist otherwise $\geq 10 h^{-1}$ Mpc gaps that are truncated by the edge or peaks in the proximity zone.

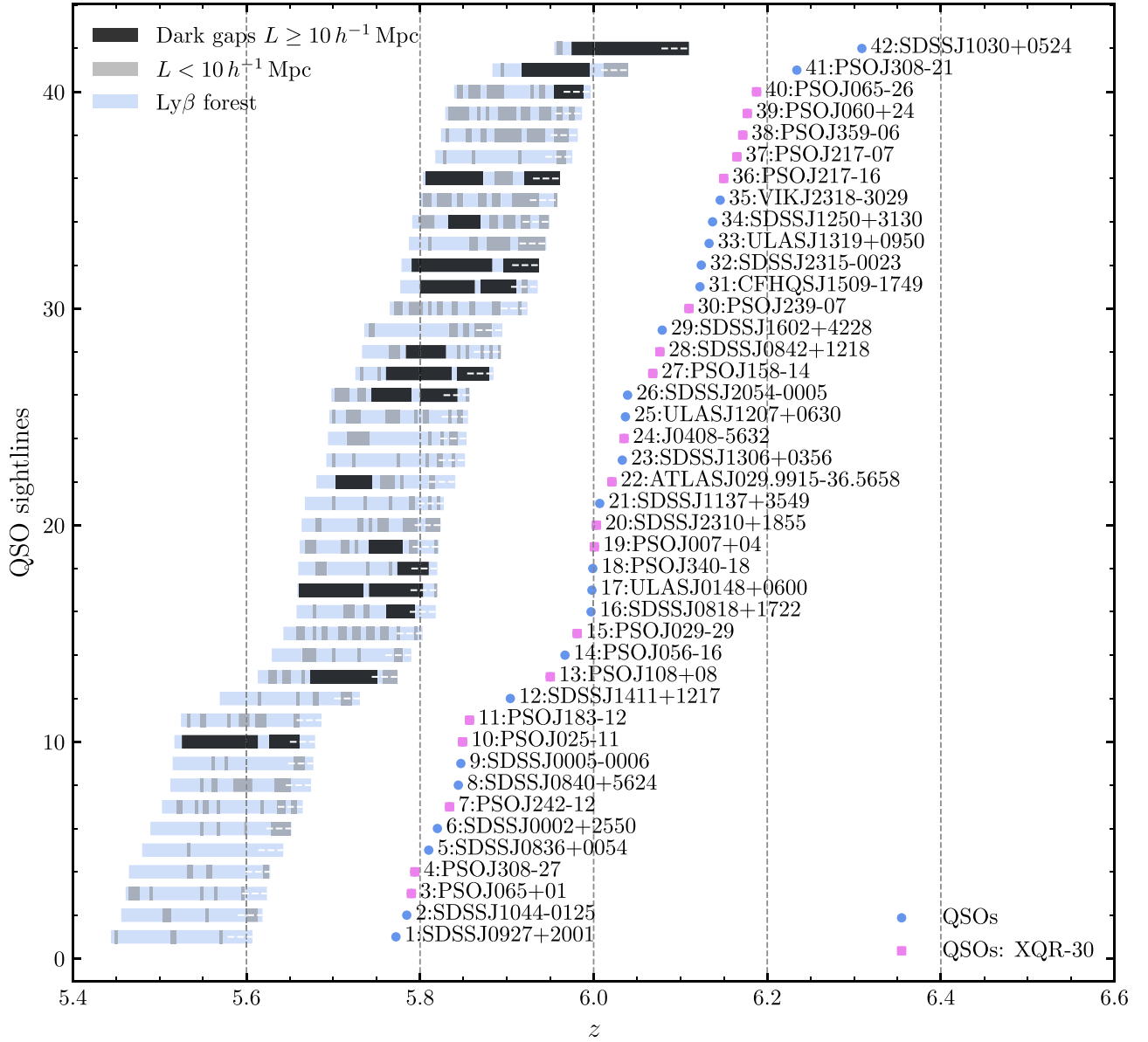


Figure 2. Overview of dark gaps identified in the Ly β forest from our sample of 42 QSO spectra. Black (gray) bars represent dark gaps longer (shorter) than $10 h^{-1}$ Mpc. Pink squares label redshifts of XQR-30 targets, and blue circles mark the redshifts of the rest of the QSOs. Light-blue shaded regions demonstrate the redshift coverage of the Ly β forest. We note that the Ly β forest is truncated at 11 pMpc from the QSO. The Ly β forest shown in this figure includes the $10 h^{-1}$ Mpc buffer zone, labeled with a dashed white line at the red end.

bins according to their redshifts at both ends. For distributions that include dark gaps cut at the red end by the proximity zone limit, we calculate a lower bound on $P(<L)$ by assuming an infinite length for these gaps and an upper bound by assuming that the gap length that would appear in the absence of any proximity effect is the same as the one measured. As shown in Figure 3(b), longer dark gaps become more numerous over $5.75 < z < 6.00$ compared to $5.50 < z < 5.75$. This significant evolution of $P(<L)$ is consistent with the results shown in panel (a).

Following Paper I, we quantify the spatial coverage of large Ly β -opaque regions by calculating the fraction of QSO spectra showing long ($L \geq 10 h^{-1}$ Mpc) dark gaps as a function of redshift, $F_{10}(z)$. Here we use $10 h^{-1}$ Mpc as the threshold because dark gaps longer than this in the late reionization

models (see Section 3) are dominated by those containing neutral islands. Based on our tests, the number of dark gaps predicted by different models differs the most for $L \gtrsim 10 h^{-1}$ Mpc, as also suggested by Nasir & D’Aloisio (2020).

We calculate F_{10} at each redshift and average over $\Delta z = 0.025$ bins. As shown in Figure 3(c), F_{10} has a significant redshift evolution over $5.5 < z < 6.0$. It increases rapidly with redshift over $5.5 < z < 5.8$, from $\sim 10\%$ to $\sim 40\%$, and climbs up to $\sim 80\%$ by $z = 6.0$ after a drop at $z \sim 5.9$. The reason for the drop is unclear, but the limited number of QSO sight lines may produce large statistical fluctuations (Appendix C), as also shown in the model predictions in Section 3. For comparison, we compute F_{01} , the fraction of QSO sight lines exhibiting dark gaps with $L \geq 1 h^{-1}$ Mpc, and find no significant drop near $z = 5.9$ (see Appendix C).

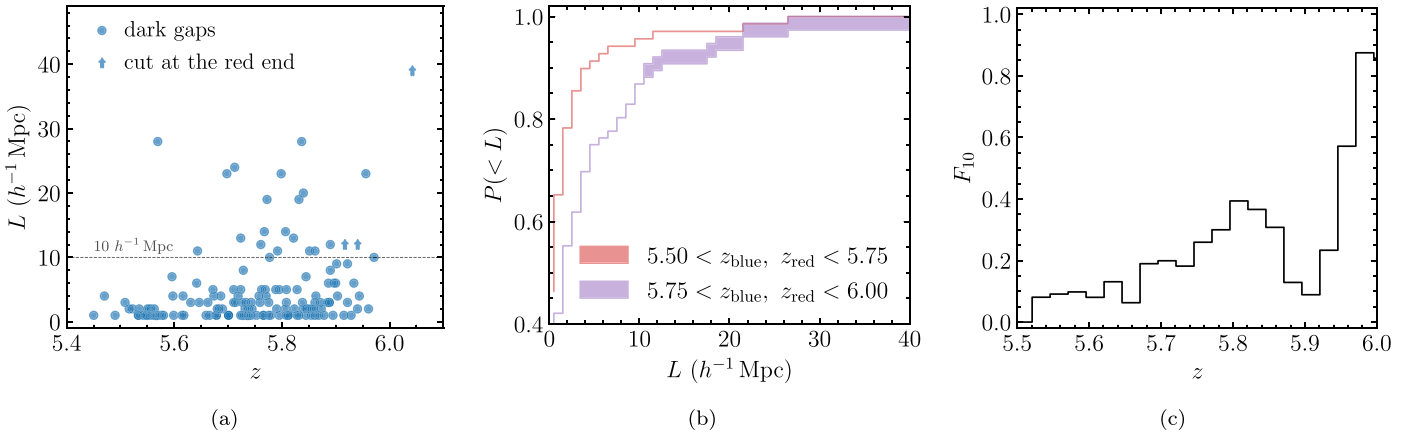


Figure 3. Observed $\text{Ly}\beta$ dark gaps. (a) Gap length vs. central redshift. Dark gaps cut at the red end by the proximity zone cut are labeled with arrows. (b) Cumulative distribution of dark gap length for two redshift intervals. The upper/lower bounds of the shaded region are described in Section 2.3.1. (c) The fraction of QSO spectra showing dark gaps longer than $10 h^{-1} \text{ Mpc}$ as a function of redshift, F_{10} . F_{10} is plotted with a binning of $\Delta z = 0.025$. See text and Appendix C for discussions about the drop of F_{10} near $z = 5.9$.

(The data used to create this figure are available.)

2.3.2. Long Dark Gap toward PSO J025–11

We find a particularly interesting $\text{Ly}\beta$ gap toward the $z = 5.849$ QSO PSO J025–11 (Figure 1). This dark gap is within the redshift interval of a long ($68 h^{-1} \text{ Mpc}$) trough in the $\text{Ly}\alpha$ forest. It spans $5.526 \leq z \leq 5.613$ with a length of $28 h^{-1} \text{ Mpc}$. This is longer and extending to an even lower redshift than the 19 and $23 h^{-1} \text{ Mpc}$ $\text{Ly}\beta$ troughs with $z_{\min} = 5.66$ within the extreme ($110 h^{-1} \text{ Mpc}$) $\text{Ly}\alpha$ trough over $5.523 \leq z \leq 5.879$ toward ULAS J0148+0600 (Becker et al. 2015). This dark gap toward J025–11 contains no apparent transmission peaks, even in the unbinned data. The 2σ lower limit of $\tau_{\text{eff}} \geq 6.067$ measured over the complete trough indicates that it is highly opaque.

There is a possibility that part of the trough may be due to either $\text{Ly}\beta$ or foreground $\text{Ly}\alpha$ absorption from the circumgalactic medium (CGM) around intervening galaxies. In this case, corresponding metal lines may be present. We check for potential CGM absorption using the XQR-30 metal absorber catalog (R. L. Davies et al. 2022, in preparation; see Appendix B for technique details). We find no intervening metal systems within the redshift range of the gap. We note that this line of sight has a DLA near the redshift of the QSO, as evidenced by the damping wing at the blue edge of the $\text{Ly}\alpha$ proximity zone. The $\text{Ly}\beta$ gap described here is at a velocity separation of $>3000 \text{ km s}^{-1}$ from the QSO, however, and is unaffected by the DLA. The XQR-30 catalog does include a CIV system toward J025–11 at $z = 4.5138$, for which $\text{Ly}\alpha$ would fall at the blue end of the $\text{Ly}\beta$ trough. Overall, however, the general lack of metal absorbers associated with this long dark gap may suggest that the gap probes a low-density region. This would favor the association of highly opaque sight lines with galaxy underdensities, as seen in imaging surveys for galaxies surrounding long $\text{Ly}\alpha$ troughs (Becker et al. 2018; Kashino et al. 2020; Christenson et al. 2021).

We examine the possible role of metal-enriched absorbers more broadly in Appendix B, finding little evidence for a strong correlation with long $\text{Ly}\beta$ troughs. We also examined a sample of lower-redshift lines of sight in Appendix D, finding that metal-enriched absorbers in the foreground $\text{Ly}\alpha$ alone are unlikely to create such a long dark gap. We emphasize that this

gap falls in redshift within a long $\text{Ly}\alpha$ trough spanning $5.461 \leq z \leq 5.674$ with $L = 68 h^{-1} \text{ Mpc}$ that does not appear to be affected by metal absorbers (Paper I). This combination of factors gives us confidence that the $L = 28 h^{-1} \text{ Mpc}$ dark gap is genuinely caused by IGM opacity.²³

3. Models and Discussion

3.1. Models and Mock Spectra

Here we compare our results to predictions from models based on hydrodynamical simulations. We use the following models, which were also used in Paper I:

1. the homogeneous-UVB model from the Sherwood Simulation Suite (Bolton et al. 2017), which uses a homogeneous Haardt & Madau (2012) UV background;
2. late reionization models wherein reionization ends at $z \lesssim 5.3$, including K20-low- τ_{CMB} , K20-low- $\tau_{\text{CMB-hot}}$, and K20-high- τ_{CMB} models from Keating et al. (2020a) and ND20-late-longmfp and ND20-late-shortmfp models from Nasir & D’Aloisio (2020); and
3. an early reionization model wherein the IGM is mostly ionized by $z = 6$ but large-scale fluctuations in the UVB, which are amplified by a short mean free path of ionizing photons ($\lambda_{\text{mfp}}^{912} = 10 h^{-1} \text{ Mpc}$ at $z = 5.6$), persist down to lower redshifts (ND20-early-shortmfp; Nasir & D’Aloisio 2020).

These models were chosen, in part, because they reproduce at least some other observations of the $\text{Ly}\alpha$ forest. The homogeneous-UVB model agrees well with observations at $z < 5$ including the IGM temperature and flux power spectra (Bolton et al. 2017), although it fails to predict the $\text{Ly}\alpha$ opacity distribution at $z > 5$ (e.g., Bosman et al. 2022). The late reionization and fluctuating UVB models are broadly consistent with observations of IGM temperature, mean $\text{Ly}\alpha$ transmission, and fluctuations in $\text{Ly}\alpha$ opacity over the redshift range we

²³ In an extreme case where this foreground absorber links two shorter dark gaps, although very unlikely, one of these two shorter dark gaps would still have a size of $L \sim 25 h^{-1} \text{ Mpc}$.

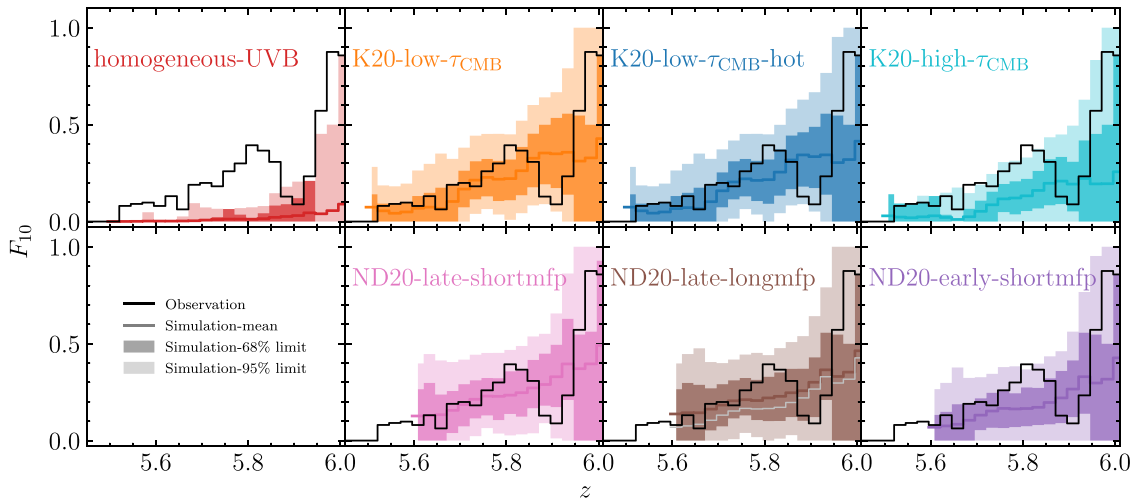


Figure 4. Fraction of QSO spectra showing long ($L \geq 10 h^{-1}$ Mpc) Ly β dark gaps as a function of redshift predicted by different models. The colored lines, dark-shaded regions, and light-shaded regions show the mean, 68% limit, and 95% limit of F_{10} predicted by models based on 10,000 realizations. The black lines plot F_{10} from the observations. For comparison, we also overplot the mean F_{10} predicted by the ND20-early-shortmfp in the panel of the ND20-late-longmfp model with a gray line.

are interested in (Keating et al. 2020a; Nasir & D’Aloisio 2020). Moreover, these models are able to produce long Ly α troughs at $z < 6$ (e.g., Paper I). We note, nevertheless, that none of the models we use can self-consistently predict the mean free path evolution over $5 < z < 6$ as measured in Becker et al. (2021).

In the homogeneous-UVB model, the IGM is instantaneously reionized at $z = 15$. At $z < 6$, therefore, the IGM in this model is fully ionized and the gas is hydrodynamically relaxed. A homogeneous-UVB model that produced a later reionization (e.g., Puchwein et al. 2019; Villaseñor et al. 2021) would mainly alter the temperature and pressure smoothing at $z < 6$. These are small-scale effects, however, that should only minimally impact our measurements. We would generally expect that any homogeneous-UVB model that reionizes by $z = 6$ would produce similar dark gap statistics to the Haardt & Madau (2012) UVB once the ionization rates at $z < 6$ are rescaled to reproduce the observed mean flux.

The K20-low- τ_{CMB} -hot model shares a similar reionization history with the K20-low- τ_{CMB} model, but it has a different thermal history with a volume-weighted mean temperature at the mean density at $z = 6$ of $T_0 = 10,000$ K, compared to that of the latter being 7000 K. The K20-high- τ_{CMB} model has an earlier midpoint of reionization at $z = 8.4$, which is at the upper end of the value suggested by CMB measurements (Planck Collaboration et al. 2020). As for the late reionization models from Nasir & D’Aloisio (2020), the main difference is that the ND20-late-shortmfp model includes stronger post-reionization UVB fluctuations than the ND20-late-longmfp model, while they both have neutral islands surviving at $z < 6$. The mean free paths of ionizing photons at $z = 5.6$ in these two models are $\lambda_{\text{mfp}}^{912} = 10$ and $30 h^{-1}$ Mpc, respectively.

The box sizes we use in this work are $L = 160, 160$, and $200 h^{-1}$ Mpc for simulations in Bolton et al. (2017), Keating et al. (2020a), and Nasir & D’Aloisio (2020), respectively. We note that the K20 models are from radiative transfer simulations run in post-processing and that the ND20 models are seminumeric models. For more details on the models, see Paper I.

We rescale the optical depths in the simulations as needed in order to match the observed mean flux in the Ly α forest (see,

e.g., Section 2.2 in Bolton et al. 2017, and references therein). We scale to the measurements of Bosman et al. (2018), which are consistent with the mean Ly α fluxes obtained from our sample. The same rescaling factor is then applied to both the Ly α and corresponding Ly β optical depths. We note that this rescaling mainly applies to the homogeneous-UVB model, for which scaling by factors of ~ 0.4 – 0.6 is required over $5 < z < 6$. We are therefore explicitly testing only a homogeneous-UVB model that also matches the observed mean Ly α flux. The Keating et al. (2020a) models already produce a mean Ly α transmission consistent with the measurements of Bosman et al. (2018). The mock spectra from this simulation are continuous in redshift, with a smoothly evolving mean flux. Nasir & D’Aloisio (2020) also calibrated their simulations to the observed τ_{eff} from Bosman et al. (2018) but provide one-dimensional skewers at discrete redshifts. For these simulations we therefore only need to rescale the optical depths such that the mock spectra described below have a mean flux that evolves smoothly with redshift.

We derive dark gap predictions from forward-modeled mock spectra that are matched to the observed sample in QSO redshift, resolution, and S/N. Keating et al. (2020a) provide mock spectra of the Ly β forest including the foreground Ly α absorption. For the homogeneous-UVB model and models from Nasir & D’Aloisio (2020) we follow the methods described in Paper I to build the mock Ly β forest and foreground Ly α forest. In all cases we rebin the mock spectra and apply the noise arrays according to each observed spectrum.

3.2. Model Comparisons

3.2.1. Comparisons of F_{10}

We compute the predicted F_{10} of each model based on 10,000 randomly selected sets of mock spectra of the same size as the observed sample. Their mean, 68% limit, and 95% limit are plotted in Figure 4, along with the observations. Similar to F_{30} ²⁴ for the Ly α forest in Paper I, F_{10} shows jagged features

²⁴ F_{30} is defined as the fraction of QSO spectra exhibiting gaps longer than $30 h^{-1}$ Mpc as a function of redshift.

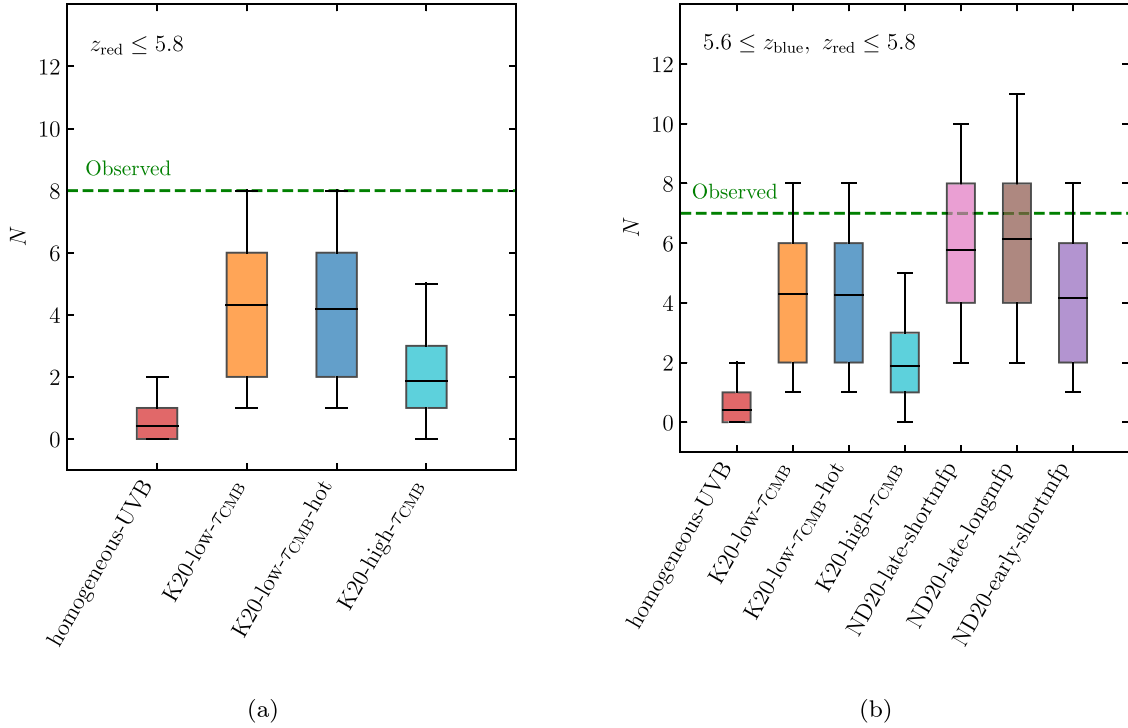


Figure 5. (a) Number of long ($L \geq 10 h^{-1}$ Mpc) Ly β dark gaps that lie entirely below $z = 5.8$ in the mock sample from each model. (b) Number of long dark gaps entirely over $5.6 \leq z \leq 5.8$. In both panels the center lines, boxes, and error bars show the mean, 68% limit, and 95% limit, respectively. The observed numbers of long dark gaps in each redshift range are indicated by dashed green lines.

due to the combined effects of step changes in the number of sight lines with redshift and the quantization of F_{10} for a finite sample size. We note that 68th and 95th percentile limits can share their upper and/or lower bounds at some redshifts, for the same reason. These features, on the other hand, show the constraining ability of the current sample size and data quality. The drop at $z \sim 5.9$ seen in the observed F_{10} is also broadly included in the 95% limits for most of the models.

The homogeneous-UVB model is not supported by the data at the $\geq 95\%$ level. This is consistent with the conclusion based on the Ly α forest in Paper I that a fully ionized IGM with a homogeneous-UVB scenario is disfavored by the data at $z < 6$ down to $z \simeq 5.3$. In contrast, the late reionization models are still consistent with the data, except for the K20-high- τ_{CMB} model, which covers the observed F_{10} just within its 95% upper limit. This supports the conclusion of Paper I that this extended reionization model is less favored by the data owing to its insufficient neutral hydrogen and/or UVB fluctuations at $z < 6$.

Our results further show that dark gaps in Ly β are more sensitive probes of neutral regions than gaps in Ly α . For dark gaps in the Ly α forest, we see little difference between the ND20-early-shortmfp model and the ND20-late models (Paper I). In the Ly β forest, however, the former predicts a smaller F_{10} than the latter by ~ 0.05 at most redshifts. This difference is not enough for us to distinguish them based on the current sample, although the Ly β gaps put some pressure on the early reionization model. Nasir & D’Aloisio (2020) note that these models are different in their Ly β dark gap length distributions, while they cannot be distinguished in the Ly α forest. Nevertheless, we compute the differential dark gap length distribution for individual ΔL bins, $L\Delta P(L)/\Delta L$, and find that their differences are minor compared to the scatter of the data. Looking ahead to the era of Extremely Large

Telescopes, we forecast that ~ 100 lines of sight with the Ly β forest covering $z \sim 5.8$ would be needed to distinguish the two models at $\sim 95\%$ confidence based on F_{10} . An S/N of 50 per 10 km s^{-1} for the spectra would be adequate according to our tests using mock spectra.

3.2.2. Total Number of Long Dark Gaps at $z \leq 5.8$

To further illustrate the differences between models, we use our mock data to calculate the total number of long dark gaps predicted to lie entirely at $z < 5.8$. Figure 5 compares the model results to the observations. Given that the ND20 models only have outputs down to $z = 5.6$, we exclude these models when counting the total number of long dark gaps below $z = 5.8$. We nevertheless include the ND20 models for dark gaps that fall entirely over $5.6 \leq z \leq 5.8$ for reference.

The results are consistent with those from F_{10} . As shown in Figure 5(a), the 95% upper limit of the predicted number of long dark gaps by the homogeneous-UVB model is 2. This is a factor of 4 smaller than the observed value, which is 8. The K20-high- τ_{CMB} model is also disfavored by the data at $>95\%$ confidence given its deficit of long dark gaps. In contrast, the rapid late reionization models, i.e., K20-low- τ_{CMB} models, agree with the observations within their 95% limits.

Over $5.6 \leq z \leq 5.8$ the observed number of long dark gaps decreases by 1 while the simulation predictions show little change. In this case, rapid late reionization models from Keating et al. (2020a) are still consistent with the data. The observations also support both fluctuating UVB and late reionization models from Nasir & D’Aloisio (2020). We note that the difference between the predicted mean numbers and the observed value is smallest for the ND20-late models, wherein $\langle x_{\text{HI}} \rangle$ is still higher than $\sim 5\%$ by $z = 5.6$. On the

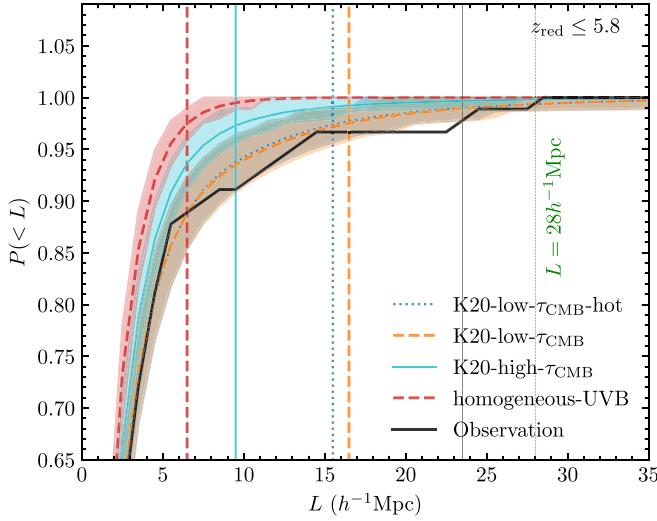


Figure 6. Cumulative distribution of dark gaps that are entirely below $z = 5.8$. Vertical colored lines indicate the 95% limit for each model. The color shaded regions plot the 68% limit for each model. From left to right are the homogeneous-UVB model, the K20-high- τ_{CMB} model, and the K20-low- τ_{CMB} -hot model (almost completely overlapped with the K20-low- τ_{CMB} model).

other hand, the K20-high- τ_{CMB} model is disfavored by the data also in this redshift range. This would suggest that very extended reionization scenarios in which insufficient neutral hydrogen and/or UVB fluctuations remain at $z < 6$ may be disfavored.

3.2.3. Detection Rate of an $L \geq 28 h^{-1} \text{ Mpc}$ Dark Gap

Perhaps the most conspicuous feature in the observations is the $L = 28 h^{-1} \text{ Mpc}$ dark gap toward PSO J025–11 that extends down to $z \approx 5.5$. The appearance of this gap may indicate that significant neutral hydrogen islands and/or UVB fluctuations persist down to $z \approx 5.5$ and provide further leverage for discriminating between models. As the outputs of the ND20 models have no redshift coverage for this dark gap, we only compare the K20 models and the homogeneous-UVB model for this section.

For each model, we use 10,000 bootstrapping realizations to calculate the CDF of dark gap length, $P(<L)$. Figure 6 compares the observed and predicted $P(<L)$ for dark gaps that are entirely below $z = 5.8$. As indicated by the vertical lines, the observed dark gap with $L = 28 h^{-1} \text{ Mpc}$ is well beyond the 95% limits of all the models shown here. These results suggest that the $L = 28 h^{-1} \text{ Mpc}$ gap we observed in the Ly β forest toward PSO J025–11 is extremely rare in these models. We perform Mann–Whitey U tests (Mann & Whitney 1947) for the hypotheses that the distributions of L in the data and predicted by models are equal, for each model respectively. The hypothesis is rejected with p -values < 0.05 for the homogeneous-UVB model.

We further calculate the detection rate of at least one $L = 28 h^{-1} \text{ Mpc}$ gap entirely below $z = 5.61$ in the mock samples from each model with the required redshift coverage. We note that in the data there are 10 QSO spectra where the Ly β forest (excluding the proximity zone) covers the full central redshift range of the $L = 28 h^{-1} \text{ Mpc}$ dark gap. We find zero detections of such long dark gaps in the homogeneous-UVB model out of 10,000 trials. The K20-high- τ_{CMB} model yields a detection rate of 4%. Both the K20-low- τ_{CMB} and

K20-low- τ_{CMB} -hot models give a higher detection rate of 10%. These results suggest that in the case of a late reionization, models with a volume-weighted average neutral hydrogen fraction $\langle x_{\text{HI}} \rangle \gtrsim 5\%$ at $z = 5.6$ are more consistent with the observations. In addition, the relatively rare presence of $L \geq 28 h^{-1} \text{ Mpc}$ gaps in the models could also be related to the size of the simulation volume ($160 h^{-1} \text{ Mpc}$ for the K20 simulations). Simulations run in larger volumes may be needed to compute more accurate statistics on the incidence of these rare, long troughs in late reionization models.

3.3. Neutral Hydrogen Fraction

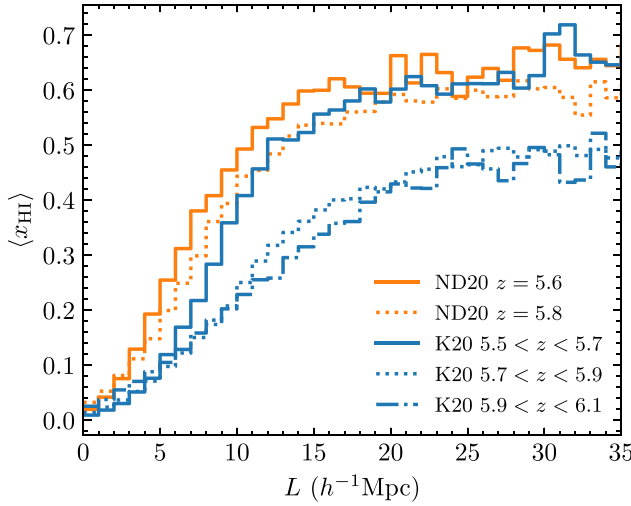
We can further use dark gaps to infer an upper limit on $\langle x_{\text{HI}} \rangle$. One can set a strict upper limit on the neutral fraction by assuming that all dark gaps correspond to neutral gas (e.g., McGreer et al. 2011, 2015). At the end of reionization, however, a combination of density and UVB fluctuations will tend to produce dark gaps even once the gas is ionized. We therefore wish to use insights from reionization models to derive a more physically motivated but still conservative upper limit on $\langle x_{\text{HI}} \rangle$ from the covering fraction of dark gaps. As described below, we use the fact that dark gaps in the late reionization models tend to show a correlation between the volume-averaged neutral fraction within a gap and the gap length. By applying this relationship to the observed gap length distribution, we can set constraints on $\langle x_{\text{HI}} \rangle$.

Our goal is to set physically reasonable constraints on $\langle x_{\text{HI}} \rangle$ while minimizing the model dependency. We therefore wish to identify the maximum average neutral fraction for a given gap length that is allowed by the models. We first explore the distribution of neutral fractions for a given dark gap length, $f(x_{\text{HI}}|L)$. We focus on two models wherein neutral regions contribute significantly to forming dark gaps, the ND20-late-longmfp model and the K20-low- τ_{CMB} model. Using the mock data, we calculate x_{HI} for each dark gap by averaging the neutral fraction pixel-wise. Figure 7(a) plots the mean neutral fraction of dark gaps as a function of length, $\langle x_{\text{HI}} \rangle_L$, at different redshifts. It is related to $f(x_{\text{HI}}|L)$ by

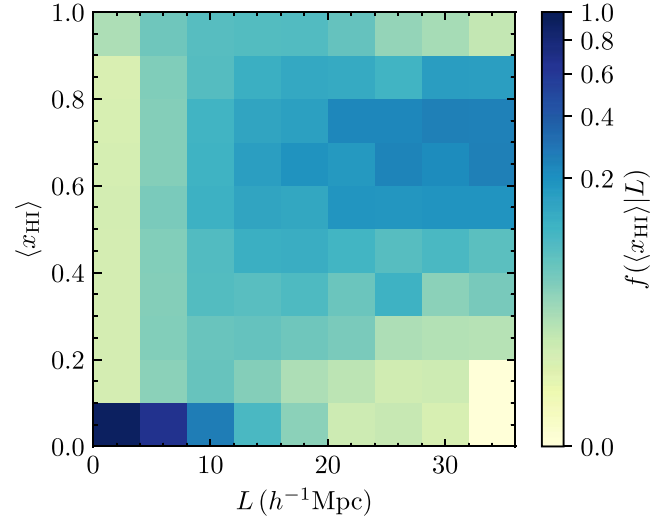
$$\langle x_{\text{HI}} \rangle_L = \int_0^1 x_{\text{HI}} f(x_{\text{HI}}|L) dx_{\text{HI}}. \quad (1)$$

As shown in the figure, dark gaps of a given length tend to be more neutral as redshift decreases. This is largely because the opacity of the ionized IGM tends to decrease with decreasing redshift, making it more difficult to produce long gaps through density and/or UVB fluctuations alone. In order to set conservative upper limits of $\langle x_{\text{HI}} \rangle$, we adopt the $\langle x_{\text{HI}} \rangle_L$ relationship from ND20-late-longmfp at $z = 5.6$. The normalized $f(x_{\text{HI}}|L)$ for each dark gap length interval is plotted in Figure 7(b). This is similar to but slightly higher than the relationship from K20-low- τ_{CMB} at the same redshift. We also note that the redshift evolution in $\langle x_{\text{HI}} \rangle_L$ in these models is relatively modest, up to a factor of ~ 2 in the K20-low- τ_{CMB} model between $z \sim 6$ and 5.6.

In order to translate the observed gap length distribution into an $\langle x_{\text{HI}} \rangle$ constraint, we calculate \mathcal{F}_L , the fraction of QSO spectra showing dark gaps with length L as a function of redshift. At a certain redshift, the total mean neutral hydrogen



(a)



(b)

Figure 7. (a) The mean volume-weighted neutral fraction ($\langle x_{\text{HI}} \rangle$) over a $\text{Ly}\beta$ dark gap for a given dark gap length in the mock data. In this figure, “ND20” and “K20” refer to the ND20-late-longmfp model and the K20-low- τ_{CMB} model, respectively. (b) Distribution of $\langle x_{\text{HI}} \rangle$ for a given $\text{Ly}\beta$ dark gap length, $f(\langle x_{\text{HI}} | L)$, in the ND20-late-longmfp model at $z = 5.6$; $f(\langle x_{\text{HI}} | L)$ is normalized for each L interval.

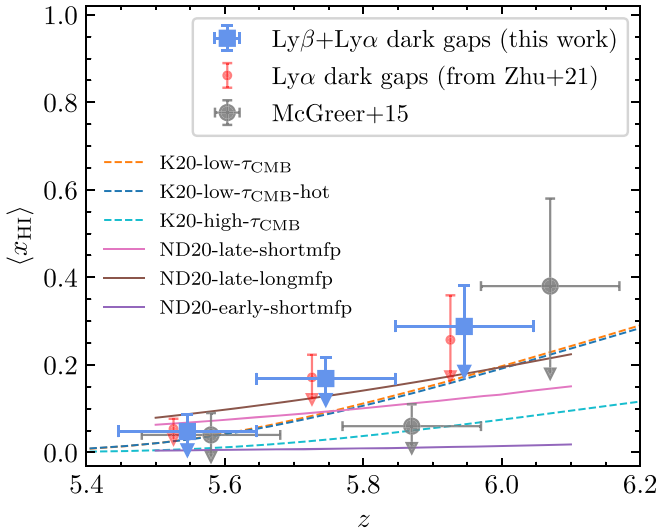


Figure 8. Inference on the neutral hydrogen fraction ($\langle x_{\text{HI}} \rangle$) from $\text{Ly}\beta$ dark gaps, which are also dark in the $\text{Ly}\alpha$ forest by definition. We show constraints based on $\text{Ly}\alpha$ dark gaps from Paper I with red error bars, which are shifted by -0.02 in redshift for display. Gray markers plot the constraints based on fraction of pixels that are dark in both the $\text{Ly}\alpha$ and $\text{Ly}\beta$ forests from McGreer et al. (2015). The vertical error bars show the 68% (1σ) limits. The horizontal error bars indicate the $\Delta z = 0.2$ redshift bins. For reference, colored lines plot the redshift evolution of $\langle x_{\text{HI}} \rangle$ for the reionization models used in this work.

fraction is then given by

$$\langle x_{\text{HI}} \rangle = \sum_{L=1}^{\infty} \mathcal{F}_L \langle x_{\text{HI}} \rangle_L. \quad (2)$$

Here we use a sum for L instead of an integral because we measure dark gap lengths in increments of $1 h^{-1} \text{Mpc}$. To estimate the uncertainty in $\langle x_{\text{HI}} \rangle$, we randomly select the observed sight lines with replacement and calculate the corresponding \mathcal{F}_L . We use bootstrapping to randomly sample the neutral hydrogen fraction from $f(\langle x_{\text{HI}} | L)$ given by models and multiply by the observed \mathcal{F}_L of this sample, and then we

sum up for all dark gap lengths. The final uncertainty in $\langle x_{\text{HI}} \rangle$ is calculated by repeating this process 10,000 times.

The results are shown in Figure 8. We calculate $\langle x_{\text{HI}} \rangle$ in Equation (2) over $\Delta z = 0.2$ bins. The inferred upper limits on $\langle x_{\text{HI}} \rangle$ are $0.05^{+0.04}_{-0.04}$, $0.17^{+0.05}_{-0.05}$, and $0.29^{+0.09}_{-0.10}$ at $z \simeq 5.55$, 5.75 , and 5.95 , respectively. We also calculate $\langle x_{\text{HI}} \rangle$ following the same method based on $\text{Ly}\alpha$ dark gaps presented in Paper I, as shown with red symbols in Figure 8. The $\text{Ly}\alpha$ dark gaps yield $\langle x_{\text{HI}} \rangle \leq 0.05$, 0.17 , and 0.26 at $z \simeq 5.55$, 5.75 , and 5.95 , respectively. The measurements based on $\text{Ly}\alpha$ and $\text{Ly}\beta$ dark gaps are highly consistent with each other. Compared to the measurements based on the fraction of dark pixels by McGreer et al. (2015), our results potentially allow a higher neutral fraction over $5.6 \lesssim z \lesssim 6.0$ and a later reionization. The difference in $\langle x_{\text{HI}} \rangle$ might be due to cosmic variance and/or the different definitions of dark gaps and dark pixels used in these works. The $\langle x_{\text{HI}} \rangle$ measurement at $z \sim 5.9$ in McGreer et al. (2015), moreover, may be biased by transmission peaks in the QSO proximity zone given that their wavelength range for both the $\text{Ly}\alpha$ and $\text{Ly}\beta$ forests ends at $z_{\text{QSO}} - 0.1$, which is less than 6.5 pMpc from the QSO at $z \sim 6$ (see proximity zone size measurements in, e.g., Eilers et al. 2017, 2020).

4. Conclusion

In this work, we explore the IGM near the end of reionization using dark gaps in the $\text{Ly}\beta$ forest over $5.5 \lesssim z \lesssim 6.0$. We show that about 10%, 40%, and 80% of QSO spectra exhibit long ($L \geq 10 h^{-1} \text{Mpc}$) dark gaps in their $\text{Ly}\beta$ forest at $z \simeq 5.6$, 5.8 , and 6.0 , respectively. Among these gaps, we detect a very long ($L = 28 h^{-1} \text{Mpc}$) and dark ($\tau_{\text{eff}} \gtrsim 6$) $\text{Ly}\beta$ gap extending down to $z \sim 5.5$ toward the $z_{\text{em}} = 5.85$ QSO PSO J025–11.

A comparison between the observed $\text{Ly}\beta$ dark gap statistics for the whole sample of 42 lines of sight and predictions from multiple reionization models (Bolton et al. 2017; Keating et al. 2020a; Nasir & D’Aloisio 2020) confirms that evidence of reionization in the form of neutral islands and/or a fluctuating

UV background persists down to at least $z \sim 5.5$. This supports the conclusions in Paper I and Bosman et al. (2022). In Paper I we noted a possible tension between $\text{Ly}\alpha$ gap statistics and a model wherein reionization ends by $z < 6$ but has a relatively early midpoint of $z = 8.4$ (Keating et al. 2020a). With $\text{Ly}\beta$ this tension becomes more significant ($>95\%$ level) based on the count of long dark gaps at $z \leq 5.8$, suggesting that very extended reionization scenarios with insufficient remaining neutral hydrogen and/or UVB fluctuations at $z < 6$ may be disfavored. In contrast, rapid late reionization models with $\langle x_{\text{H I}} \rangle \gtrsim 5\%$ at $z = 5.6$ (Keating et al. 2020a; Nasir & D’Aloisio 2020) are consistent with the observations. A model wherein reionization ends early but retains large-scale fluctuations in the ionizing UV background (Nasir & D’Aloisio 2020) is also permitted by the dark gap data. We note, however, that recent IGM temperature measurements from Gaikwad et al. (2020) disfavor this model.

A caveat is that we are testing only specific reionization models, including only one with a fluctuating UVB in which reionization ends at $z > 6$. By comparison, Gnedin et al. (2017) showed that their full radiative transfer simulations, which reionized near $z \sim 7$, were able to reproduce the $\text{Ly}\alpha$ dark gap distribution measured from ESI spectra of a set of 12 $z \sim 6$ QSOs. Because $\text{Ly}\beta$ dark gaps are correlated with $\text{Ly}\alpha$ opacities (Appendix A), it is possible that some early reionization scenarios with UVB fluctuations can reproduce our $\text{Ly}\beta$ dark gap distributions while also matching the observed evolution of the mean $\text{Ly}\alpha$ transmission.

Finally, we use the observed $\text{Ly}\beta$ gaps to place constraints on the neutral hydrogen fraction based on the association between neutral islands and dark gaps seen in reionization simulations. Our results are broadly consistent with, but more permissive than, the constraints from McGreer et al. (2011, 2015) that are based on the dark pixel fraction. Notably, we find an upper limit at $z \simeq 5.75$ of $\langle x_{\text{H I}} \rangle \leq 0.17$. This constraint is consistent with scenarios wherein reionization extends significantly below $z = 6$.

We thank Elisa Boera and Fahad Nasir for their help with simulated data and for useful discussions. We also thank the anonymous referee for their constructive comments.

Y.Z., G.D.B., and H.M.C. were supported by the National Science Foundation through grants AST-1615814 and AST-1751404. H.M.C. was also supported by the National Science Foundation Graduate Research Fellowship Program under grant No. DGE-1326120. S.E.I.B. and F. Walter acknowledge funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No. 740246 “Cosmic Gas”). L.C.K. was supported by the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 885990. M.B. acknowledges support from PRIN MIUR project “Black Hole Winds and the Baryon Life Cycle of Galaxies: The Stone-guest at the Galaxy Evolution Supper,” contract No. 2017PH3WAT. F.B. 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). H.C. thanks the support by NASA through the NASA FINESST grant NNH19ZDA005K. A.-C.E. acknowledges support by NASA

through the NASA Hubble Fellowship grant No. 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. X.F. and J.Y. acknowledge support from the NSF grants AST 115-15115 and AST 19-08284. M.G.H. acknowledges support from the UKRI STFC (grant Nos. ST/N000927/1 and ST/S000623/1). G.K.’s research is partly supported by the Max Planck Society via a partner group grant. A.P. acknowledges support from the ERC Advanced Grant INTERSTELLAR H2020/740120. Parts of this work were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project No. CE170100013. F. Wang thanks the support provided by NASA through the NASA Hubble Fellowship grant No. 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.

Based on observations collected at the European Southern Observatory under ESO programmes 060.A-9024(A), 084.A-0360(A), 084.A-0390(A), 084.A-0550(A), 085.A-0299(A), 086.A-0162(A), 086.A-0574(A), 087.A-0607(A), 088.A-0897(A), 091.C-0934(B), 096.A-0095(A), 096.A-0418(A), 097.B-1070(A), 098.A-0111(A), 098.B-0537(A), 0100.A-0243(A), 0100.A-0625(A), 0101.B-0272(A), 0102.A-0154(A), 0102.A-0478(A), 1103.A-0817(A), and 1103.A-0817(B).

Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. Finally, this research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExSci), under contract with the National Aeronautics and Space Administration.

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

Facilities: Keck:II (ESI), VLT:Kueyen (X-Shooter).

Software: Astrocook (Cupani et al. 2020), Astropy (Astropy Collaboration et al. 2013), Matplotlib (Hunter 2007), NumPy (van der Walt et al. 2011), SpectRes (Carnall 2017).

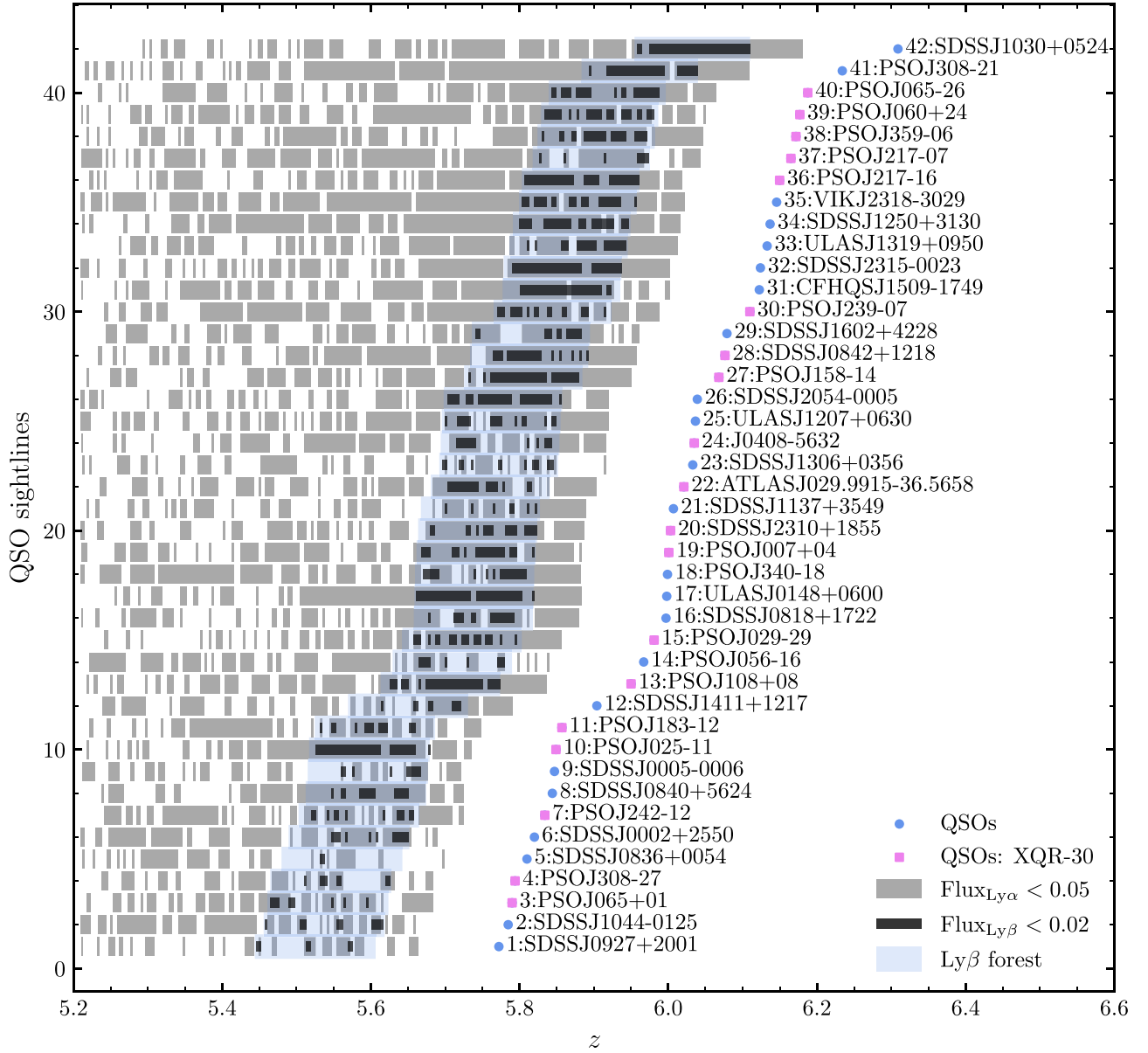


Figure A1. Overview of Ly β -opaque regions and Ly α dark gaps from our sample of 42 QSO lines of sight. Black bars show Ly β -opaque regions, where normalized flux in the Ly β forest $\text{Flux}_{\text{Ly}\beta} < 0.02$ per $1 h^{-1} \text{Mpc}$ bin. Gray bars show Ly α dark gaps as defined in Paper I. Light-blue shaded regions highlight the redshift ranges of the Ly β forest. The overlap between the gray bars and black bars yields Ly β dark gaps as defined in this work.

Appendix A

Relationship between Ly β Dark Gaps and Ly α Dark Gaps

To illustrate the effects of requiring Ly β gaps to also be dark in the Ly α forest, here we explore the relationship between Ly β dark gaps and Ly α dark gaps. In Figure A1 we overplot Ly β -opaque regions ($\text{Flux}_{\text{Ly}\beta} < 0.02$ per $1 h^{-1} \text{Mpc}$ bin) on Ly α dark gaps as defined in Paper I. Although Ly β -opaque regions overlap strongly with Ly α dark gaps, there do exist regions that are dark only in the Ly β forest, e.g., the long Ly β -opaque region toward CFHQS J1509–1749 that bridges two Ly α dark gaps, as shown in the figure. These cases are due to

foreground Ly α absorption in the Ly β forest. Requiring Ly β gaps to also be dark in the Ly α forest partially avoids this kind of foreground contamination.

We further plot the length of Ly α dark gaps versus the length of corresponding Ly β -opaque regions in Figure A2. Most of the long ($\geq 10 h^{-1} \text{Mpc}$) Ly β dark gaps appear in $L \geq 30 h^{-1} \text{Mpc}$ Ly α dark gaps. Only 1 out of 23 long Ly β -opaque regions contains transmission in Ly α and is split into two Ly β dark gaps.

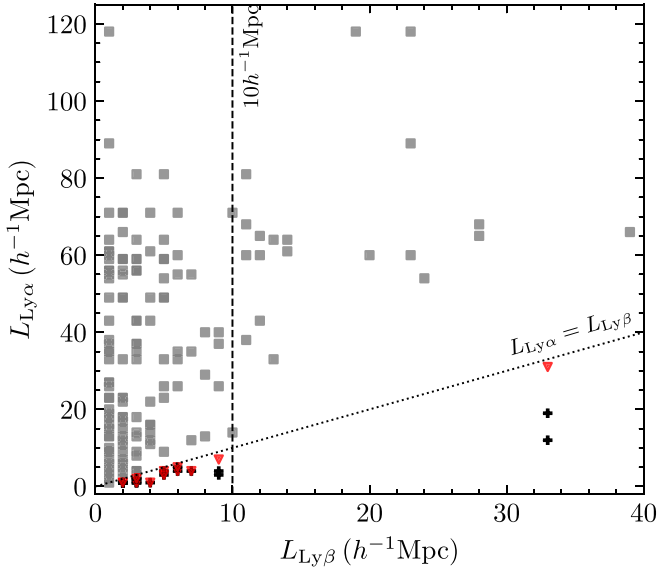


Figure A2. Length of $\text{Ly}\alpha$ dark gaps vs. length of $\text{Ly}\beta$ -opaque regions. For $\text{Ly}\beta$ dark gaps that are entirely within the redshift range of the $\text{Ly}\alpha$ dark gap, we plot the length of the $\text{Ly}\alpha$ gap vs. the length of the $\text{Ly}\beta$ gap with a gray square. Red triangles denote situations where not all $1 h^{-1} \text{Mpc}$ pixels in a $\text{Ly}\beta$ -opaque region have $\text{Flux}_{\text{Ly}\alpha} < 0.05$ ($\text{Ly}\alpha$ dark gaps). The path lengths of $\text{Ly}\alpha$ dark gaps inside these $\text{Ly}\beta$ -opaque regions are marked with black crosses.

Appendix B Metal-enriched Systems

In Figure B1 we display an overview of dark gaps with metal-enriched systems overplotted for the 27 QSO sight lines in our

sample where the identification of metals is relatively complete and consistent. We label metal systems with redshifts in the $\text{Ly}\beta$ forest and in the foreground $\text{Ly}\alpha$ forest separately. These systems are included in a metal absorber catalog that will be presented in R. L. Davies et al. (2022, in preparation). Briefly, the Python application *Astrocook* was used to perform an automated search for Mg II, Fe II, C IV, Si IV, and N V absorbers and DLA-like systems probed by C II and other low-ionization species. Candidate absorbers were identified using a cross-correlation algorithm within *Astrocook* that searches for redshifts where significant absorption is present in all relevant transitions. Custom filtering algorithms and visual inspection were then used to remove false positives and produce the final absorber list.

We then investigate the correlation between long ($L \geq 10 h^{-1} \text{Mpc}$) dark gaps and metal systems. We find that the probability for a metal system in the $\text{Ly}\beta$ forest to lie in a long dark gap is $15\% \pm 9\%$, where the 68% confidence limit comes from bootstrapping these 27 sight lines 10,000 times. This probability is $31\% \pm 9\%$ in the case of a system in the corresponding foreground $\text{Ly}\alpha$ forest. In these calculations we count clustered metal absorbers with a separation of $< 1 h^{-1} \text{Mpc}$ as one system. By comparison, the probability that a randomly chosen point lies in a long dark gap is $22\% \pm 9\%$. Our results suggest that the correlation between long dark gaps and (foreground) metal systems is not highly significant, at least for this subsample. The relatively lower probability of finding metal absorbers within the redshifts of long dark gaps nevertheless potentially favors the association between high IGM $\text{Ly}\alpha$ opacities and galaxy underdensities (see also Becker et al. 2018; Kashino et al. 2020; Christenson et al. 2021).

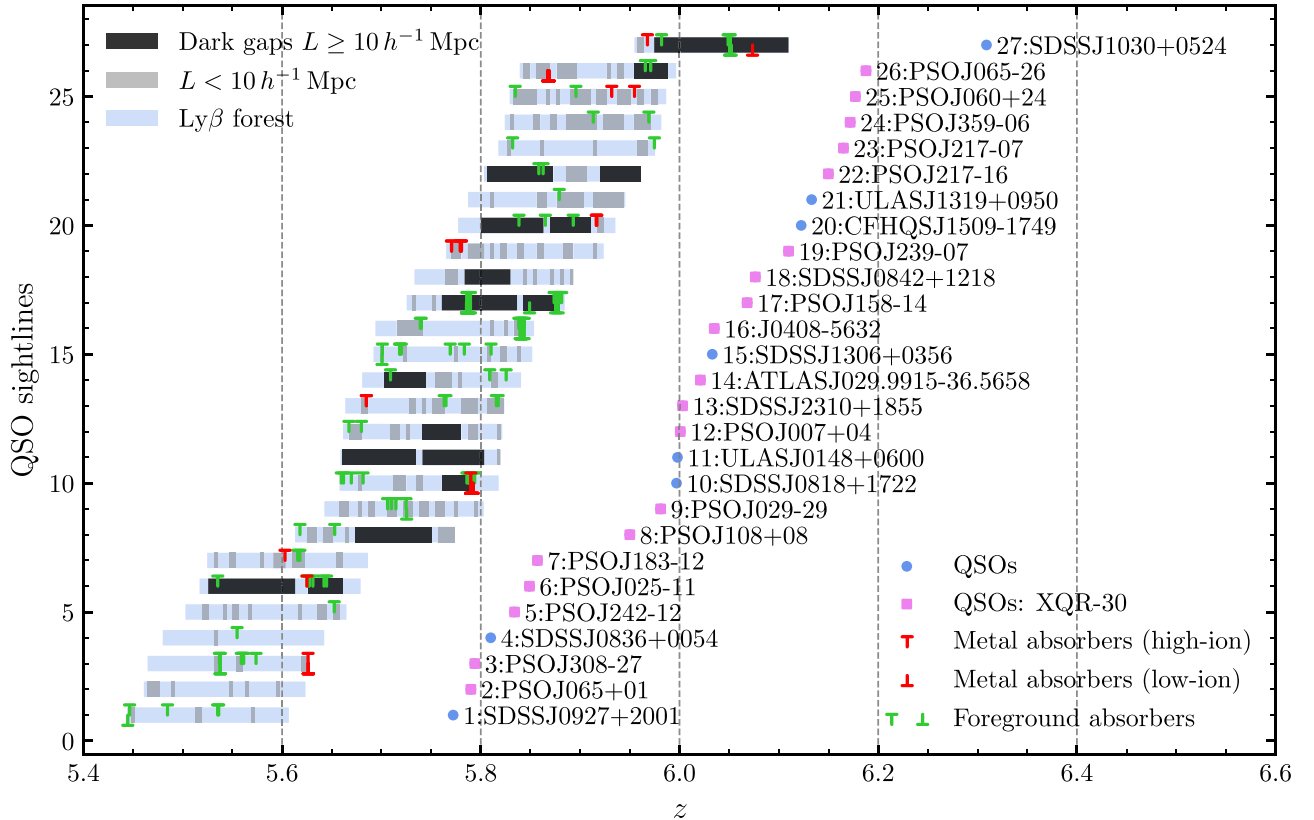


Figure B1. Similar to Figure 2, but with metal-enriched absorbers overplotted for 27 sight lines that have a relatively complete and consistent identification of metal-enriched absorbers (R. L. Davies et al. 2022, in preparation). We label high-ionization and low-ionization systems within the redshift of the $\text{Ly}\beta$ forest with red “T” and “L” symbols, respectively. Foreground metal-enriched absorbers whose $\text{Ly}\alpha$ absorption would fall within the $\text{Ly}\beta$ forest are labeled with green symbols at the corresponding $\text{Ly}\beta$ redshifts.

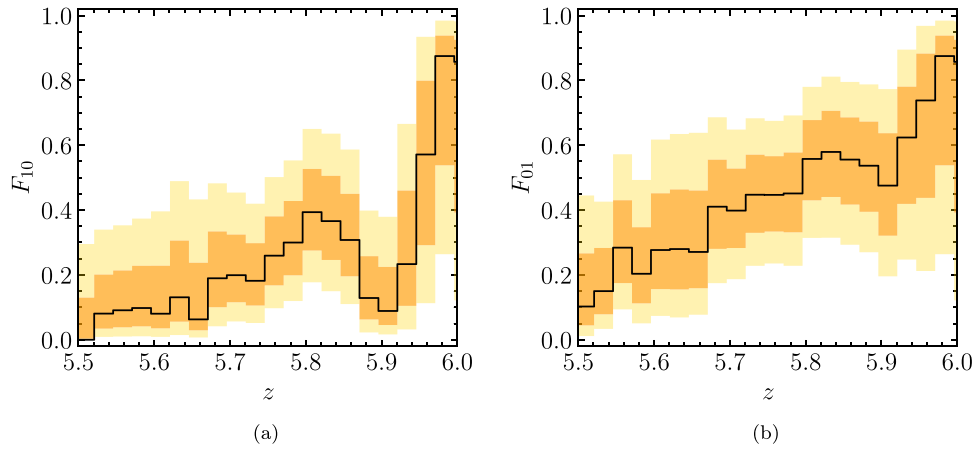


Figure C1. (a) Statistical uncertainty estimation for F_{10} shown in Figure 3(c). Dark- and light-shaded regions mark the 68% and 95% limits, respectively, of F_{10} based on Beta distribution. (b) Fraction of QSO spectra showing dark gaps with $L \geq 1 h^{-1}$ Mpc.

Appendix C

Uncertainties in the Fraction of QSO Spectra Showing Dark Gaps

The evolution in F_{10} shown in Figure 4 shows a large drop near $z = 5.9$. To estimate the statistical fluctuations in F_{10} , we treat the “hit rate” of long dark gaps at individual redshifts as a binomial experiment defined by the number of hits (number of long dark gaps, n_{dark}) inside a different number of trials (number of QSO sight lines, n_{qso}). At a certain redshift, the posterior probability distribution function for the true “hit rate,” x , can be expressed as a Beta distribution, $f(x; \alpha, \beta) \propto x^{\alpha-1}(1-x)^{\beta-1}$, with $\alpha = n_{\text{dark}} + 0.5$ and $\beta = n_{\text{qso}} - n_{\text{dark}} + 0.5$, assuming a Jeffreys prior. As shown in Figure C1(a), the evolution of F_{10} is consistent with a monotonic increase with z within the 95% confidence intervals. We caution that the analysis here assumes that the “hit rates” at different redshifts are independent from each other.

While the dip could be due to statistical fluctuations, we nevertheless wish to check whether it may relate to possible biases in the data related to $\text{Ly}\beta$ absorption near that redshift. To check for possible systematic effects, we calculate the fraction of QSO spectra showing dark gaps of any length ($L \geq 1 h^{-1}$ Mpc) as a function of redshift, F_{01} . As shown in Figure C1, the drop in F_{10} at $z \sim 5.9$ is not present in F_{01} . Instead, the evolution of $\text{Ly}\beta$ -opaque regions with redshift appears relatively smooth. We thus find no evidence of systematic effects in the data that would suggest lower absorption overall near $z = 5.9$.

Appendix D

Dark Gaps in a Lower-redshift Sample

Here we examine the extent to which strong, clustered absorbers associated with galaxies may be able to produce

long dark gaps in the $\text{Ly}\beta$ forest. These (typically metal-enriched) absorbers may produce discrete absorption in either $\text{Ly}\beta$ over the trough or $\text{Ly}\alpha$ at the corresponding foreground redshifts. They may also connect otherwise short dark gaps to form longer gaps. Of particular interest are very long gaps analogous to the $L = 28 h^{-1}$ Mpc gap toward PSO J025–11. To test whether such gaps may be due to (circum)galactic absorbers rather than the IGM, we search for dark gaps at $z \lesssim 5.5$ in a sample of QSO lines of sight that lie at somewhat lower redshifts than our main sample. Because the IGM becomes increasingly transparent toward lower redshifts, any long dark gaps in this sample might signal a significant contribution from discrete systems associated with galaxies.

Our lower-redshift sample includes 27 ESI and X-Shooter spectra of QSOs over $5.0 < z_{\text{em}} < 5.7$ from the Keck and VLT archives. The selection of targets is based on their redshift and is independent of foreknowledge of dark gaps. QSO spectra in this lower-redshift sample have S/N greater than 20 pixel^{-1} in the $\text{Ly}\beta$ forest. In order to account for the increased mean transmission at low redshifts, we conservatively use a higher flux threshold of 0.08 when searching for dark gaps. The ratio of mean flux in the $\text{Ly}\beta$ forest at $z = 4.8$ and 5.6 is about 3.2 (e.g., Fan et al. 2006; Eilers et al. 2019; Bosman et al. 2021); thus, a flux threshold of 4 times the high-redshift value is used.

Figure D1 shows dark gaps detected in this lower-redshift sample. No dark gaps longer than $10 h^{-1}$ Mpc are detected. The lack of any long gaps in this sample suggests that extended gaps created largely by strong, discrete absorbers are rare, at least over $5 \lesssim z \lesssim 5.5$, which is reasonably close in redshift to our main sample. This increases our confidence that the $L = 28 h^{-1}$ Mpc dark gap toward PSO J025–11 is likely to mainly arise from IGM absorption.

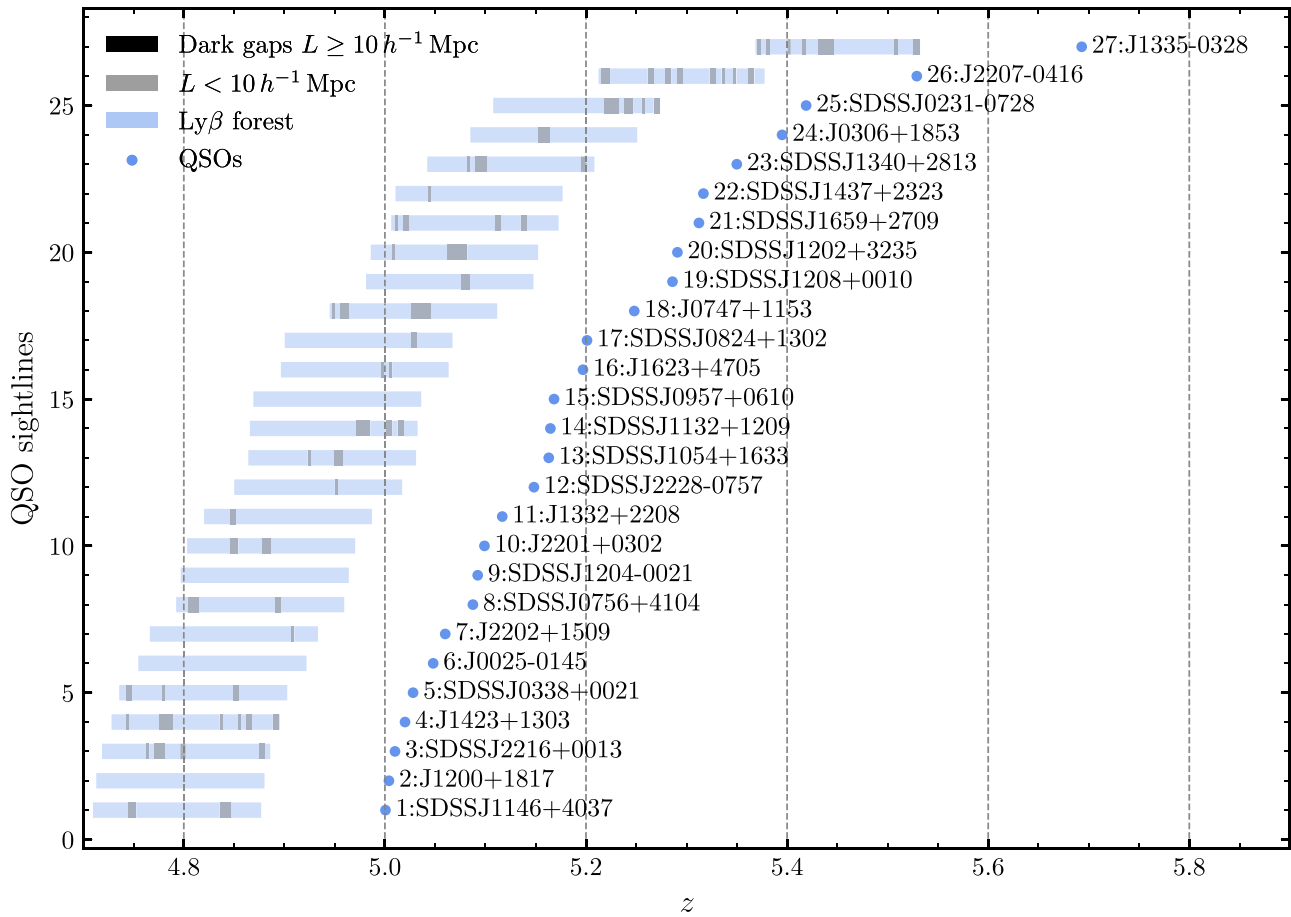


Figure D1. Similar to Figure 2, but showing dark gaps identified in the Ly β forest from a lower-redshift sample.

ORCID iDs

Yongda Zhu <https://orcid.org/0000-0003-3307-7525>
 George D. Becker <https://orcid.org/0000-0003-2344-263X>
 Sarah E. I. Bosman <https://orcid.org/0000-0001-8582-7012>
 Laura C. Keating <https://orcid.org/0000-0001-5211-1958>
 Valentina D’Odorico <https://orcid.org/0000-0003-3693-3091>
 Rebecca L. Davies <https://orcid.org/0000-0002-3324-4824>
 Holly M. Christenson <https://orcid.org/0000-0002-0421-065X>
 Eduardo Bañados <https://orcid.org/0000-0002-2931-7824>
 Fuyan Bian <https://orcid.org/0000-0002-1620-0897>
 Manuela Bischetti <https://orcid.org/0000-0002-4314-021X>
 Huanqing Chen <https://orcid.org/0000-0002-3211-9642>
 Frederick B. Davies <https://orcid.org/0000-0003-0821-3644>
 Anna-Christina Eilers <https://orcid.org/0000-0003-2895-6218>
 Xiaohui Fan <https://orcid.org/0000-0003-3310-0131>
 Prakash Gaikwad <https://orcid.org/0000-0002-2423-7905>
 Bradley Greig <https://orcid.org/0000-0002-4085-2094>
 Martin G. Haehnelt <https://orcid.org/0000-0001-8443-2393>
 Girish Kulkarni <https://orcid.org/0000-0001-5829-4716>
 Samuel Lai <https://orcid.org/0000-0001-9372-4611>
 Andrea Pallottini <https://orcid.org/0000-0002-7129-5761>
 Yuxiang Qin <https://orcid.org/0000-0002-4314-1810>
 Emma V. Ryan-Weber <https://orcid.org/0000-0002-5360-8103>
 Fabian Walter <https://orcid.org/0000-0003-4793-7880>

Feige Wang <https://orcid.org/0000-0002-7633-431X>
 Jinyi Yang <https://orcid.org/0000-0001-5287-4242>

References

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
 Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, *Natur*, **553**, 473
 Becker, G. D., Bolton, J. S., Madau, P., et al. 2015, *MNRAS*, **447**, 3402
 Becker, G. D., D’Aloisio, A., Christenson, H. M., et al. 2021, *MNRAS*, **508**, 1853
 Becker, G. D., Davies, F. B., Furlanetto, S. R., et al. 2018, *ApJ*, **863**, 92
 Becker, G. D., Pettini, M., Rafelski, M., et al. 2019, *ApJ*, **883**, 163
 Boera, E., Becker, G. D., Bolton, J. S., & Nasir, F. 2019, *ApJ*, **872**, 101
 Bolton, J. S., Puchwein, E., Sijacki, D., et al. 2017, *MNRAS*, **464**, 897
 Bosman, S. E. I., Davies, F. B., Becker, G. D., et al. 2022, *MNRAS*, **514**, 55
 Bosman, S. E. I., Āurovčíková, D., Davies, F. B., & Eilers, A. C. 2021, *MNRAS*, **503**, 2077
 Bosman, S. E. I., Fan, X., Jiang, L., et al. 2018, *MNRAS*, **479**, 1055
 Cain, C., D’Aloisio, A., Gangolli, N., & Becker, G. D. 2021, *ApJL*, **917**, L37
 Carnall, A. C. 2017, arXiv:1705.05165
 Choudhury, T. R., Paranjape, A., & Bosman, S. E. I. 2021, *MNRAS*, **501**, 5782
 Christenson, H. M., Becker, G. D., Furlanetto, S. R., et al. 2021, *ApJ*, **923**, 87
 Cupani, G., D’Odorico, V., Cristiani, S., et al. 2020, *Proc. SPIE*, **11452**, 114521U
 Davies, F. B., Becker, G. D., & Furlanetto, S. R. 2018a, *ApJ*, **860**, 155
 Davies, F. B., Bosman, S. E. I., Furlanetto, S. R., Becker, G. D., & D’Aloisio, A. 2021, *ApJL*, **918**, L35
 Davies, F. B., Hennawi, J. F., Bañados, E., et al. 2018b, *ApJ*, **864**, 142
 Davies, F. B., Hennawi, J. F., Bañados, E., et al. 2018c, *ApJ*, **864**, 143
 de Belsunce, R., Gratton, S., Coulton, W., & Efsthathiou, G. 2021, *MNRAS*, **507**, 1072
 Eilers, A.-C., Davies, F. B., Hennawi, J. F., et al. 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., Hennawi, J. F., Decarli, R., et al. 2020, *ApJ*, **900**, 37
- Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, *AJ*, **132**, 117
- Furlanetto, S. R., Hernquist, L., & Zaldarriaga, M. 2004, *MNRAS*, **354**, 695
- Gaikwad, P., Rauch, M., Haehnelt, M. G., et al. 2020, *MNRAS*, **494**, 5091
- Gaikwad, P., Srianand, R., Haehnelt, M. G., & Choudhury, T. R. 2021, *MNRAS*, **506**, 4389
- Gallerani, S., Ferrara, A., Fan, X., & Choudhury, T. R. 2008, *MNRAS*, **386**, 359
- Gnedin, N. Y., Becker, G. D., & Fan, X. 2017, *ApJ*, **841**, 26
- Greig, B., Mesinger, A., Davies, F. B., et al. 2022, *MNRAS*, **512**, 5390
- Haardt, F., & Madau, P. 2012, *ApJ*, **746**, 125
- Hoag, A., Bradač, M., Huang, K., et al. 2019, *ApJ*, **878**, 12
- Hu, W., Wang, J., Zheng, Z.-Y., et al. 2019, *ApJ*, **886**, 90
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Jung, I., Finkelstein, S. L., Dickinson, M., et al. 2020, *ApJ*, **904**, 144
- Kashino, D., Lilly, S. J., Shibuya, T., Ouchi, M., & Kashikawa, N. 2020, *ApJ*, **888**, 6
- Keating, L. C., Kulkarni, G., Haehnelt, M. G., Chardin, J., & Aubert, D. 2020a, *MNRAS*, **497**, 906
- Keating, L. C., Weinberger, L. H., Kulkarni, G., et al. 2020b, *MNRAS*, **491**, 1736
- Kulkarni, G., Keating, L. C., Haehnelt, M. G., et al. 2019, *MNRAS*, **485**, L24
- Mann, H. B., & Whitney, D. R. 1947, *Ann. Math. Stat.*, **18**, 50
- Mason, C. A., Fontana, A., Treu, T., et al. 2019, *MNRAS*, **485**, 3947
- Mason, C. A., Treu, T., Dijkstra, M., et al. 2018, *ApJ*, **856**, 2
- McGreer, I. D., Mesinger, A., & D’Odorico, V. 2015, *MNRAS*, **447**, 499
- McGreer, I. D., Mesinger, A., & Fan, X. 2011, *MNRAS*, **415**, 3237
- Muñoz, J. B., Qin, Y., Mesinger, A., et al. 2022, *MNRAS*, **511**, 3657
- Nasir, F., & D’Aloisio, A. 2020, *MNRAS*, **494**, 3080
- Paschos, P., & Norman, M. L. 2005, *ApJ*, **631**, 59
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, *A&A*, **571**, A16
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, *A&A*, **641**, A6
- Puchwein, E., Haardt, F., Haehnelt, M. G., & Madau, P. 2019, *MNRAS*, **485**, 47
- Qin, Y., Mesinger, A., Bosman, S. E. I., & Viel, M. 2021, *MNRAS*, **506**, 2390
- Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, *PASP*, **114**, 851
- Songaila, A., & Cowie, L. L. 2002, *AJ*, **123**, 2183
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22
- Vernet, J., Dekker, H., D’Odorico, S., et al. 2011, *A&A*, **536**, A105
- Villasenor, B., Robertson, B., Madau, P., & Schneider, E. 2021, arXiv:2111.00019
- Wang, F., Davies, F. B., Yang, J., et al. 2020, *ApJ*, **896**, 23
- Weinberger, L. H., Haehnelt, M. G., & Kulkarni, G. 2019, *MNRAS*, **485**, 1350
- Wold, I. G. B., Malhotra, S., Rhoads, J., et al. 2022, *ApJ*, **927**, 36
- Yang, J., Wang, F., Fan, X., et al. 2020a, *ApJ*, **904**, 26
- Yang, J., Wang, F., Fan, X., et al. 2020b, *ApJL*, **897**, L14
- Zhu, Y., Becker, G. D., Bosman, S. E. I., et al. 2021, *ApJ*, **923**, 223