The Astrophysical Journal, 965:134 (17pp), 2024 April 20 https://doi.org/10.3847/1538-4357/ad1d5d
© 2024. The Author(s). Published by the American Astronomical Society.
OPEN ACCESS

Constraints on the Evolution of the Ionizing Background and Ionizing Photon Mean Free Path at the End of Reionization

Frederick B. Davies[1], Sarah E. I. Bosman[1,2], Prakash Gaikwad[1], Fahad Nasir[1], Joseph F. Hennawi[3,4], George D. Becker[5], Martin G. Haehnelt[6], Valentina D'Odorico[7,8,9], Manuela Bischetti[10,11], Anna-Christina Eilers[12,20], Laura C. Keating[13], Girish Kulkarni[14], Samuel Lai[15], Chiara Mazzucchelli[16], Yuxiang Qin[17,18], Sindhu Satyavolu[14], Feige Wang[19], Jinyi Yang[19,21], and Yongda Zhu[5]

[1] Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[2] Institute for Theoretical Physics, Heidelberg University, Philosophenweg 12, D-69120, Heidelberg, Germany
[3] Department of Physics, University of California, Santa Barbara, CA 93106, USA
[4] Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
[5] Department of Physics and Astronomy, University of California, Riverside, CA, 92521, USA
[6] Kavli Institute for Cosmology and Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK
[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] IFPU-Institute for Fundamental Physics of the Universe, via Beirut 2, I-34151 Trieste, Italy
[10] INAF-Osservatorio Astronomico di Trieste, Via G.B. Tiepolo, 11, I-34143 Trieste, Italy
[11] Dipartimento di Fisica, Sezione di Astronomia, Universitá di Trieste, via Tiepolo 11, I-34143 Trieste, Italy
[12] MIT Kavli Institute for Astrophysics and Space Research, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[13] Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK
[14] Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India
[15] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[16] Instituto de Estudios Astrofísicos, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Avenida Ejercito Libertador 441, Santiago, Chile
[17] School of Physics, University of Melbourne, Parkville, VIC 3010, Australia
[18] ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
[19] Steward Observatory, University of Arizona, 933 N Cherry Avenue, Tucson, AZ 85721, USA
Received 2023 August 21; revised 2024 January 1; accepted 2024 January 9; published 2024 April 15

Abstract

The variations in Lyα forest opacity observed at $z > 5.3$ between lines of sight to different background quasars are too strong to be caused by fluctuations in the density field alone. The leading hypothesis for the cause of this excess variance is a late, ongoing reionization process at redshifts below six. Another model proposes strong ionizing background fluctuations coupled to a short, spatially varying mean free path of ionizing photons, without explicitly invoking incomplete reionization. With recent observations suggesting a short mean free path at $z \sim 6$, and a dramatic improvement in $z > 5$ Lyα forest data quality, we revisit this latter possibility. Here, we apply the likelihood-free inference technique of approximate Bayesian computation (ABC) to jointly constrain the hydrogen photoionization rate $\Gamma_{\rm HI}$ and the mean free path of ionizing photons $\lambda_{\rm mfp}$ from the effective optical depth distributions at $z = 5.0$–6.1 from XQR-30. We find that the observations are well-described by fluctuating mean free path models with average mean free paths that are consistent with the steep trend implied by independent measurements at $z \sim 5$–6, with a concomitant rapid evolution of the photoionization rate.

Unified Astronomy Thesaurus concepts: Intergalactic medium (813); Reionization (1383)

1. Introduction

The epoch of reionization reflects the cumulative photon output of the first generations of stars and galaxies in the Universe. Determining the precise timing of reionization thus provides crucial clues for understanding structure formation in the first billion years of cosmic time. Observations of the cosmic microwave background suggest a characteristic epoch of $z \sim 7$ (Planck Collaboration et al. 2020), and this rough midpoint is supported by constraints derived from the disappearance of Lyα emission from galaxies at $z > 6$ (e.g., Mason et al. 2018; Hoag et al. 2019; Mason et al. 2019; Weinberger et al. 2019; Jung et al. 2020; although see Wold et al. 2022) and studies of the Lyα damping wing in the highest-redshift quasars known at $z \gtrsim 7$ (e.g., Davies et al. 2018b; Yang et al. 2020a; Wang et al. 2020; Greig et al. 2022).

The endpoint of reionization was originally estimated by the disappearance of widespread transmission in the Lyα forest (Becker et al. 2001; Fan et al. 2006), thought to be due to the onset of Gunn–Peterson (GP) absorption (Gunn & Peterson 1965) of neutral hydrogen in the intergalactic medium (IGM). That interpretation is complicated, however, by the concomitant decrease in the intensity of the ionizing background (Bolton & Haehnelt 2007; Wyithe & Bolton 2011; Davies et al. 2018d; D'Aloisio et al. 2018), which should lead to a mostly opaque Lyα forest with strong sightline-to-sightline variations even without significantly neutral gas (e.g., Lidz et al. 2006). With the discovery of the giant GP trough at $z \sim 5.5$–5.8 toward the quasar ULAS J0148 + 0600 by Becker et al. (2015), however, existing models of the post-reionization IGM were no longer consistent with the distribution of Lyα forest opacity at

[20] Pappalardo Fellow.
[21] Strittmatter Fellow.

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

$z \gtrsim 5.6$. This inconsistency was confirmed by subsequent compilations of Lyα forest opacity measurements (Bosman et al. 2018; Eilers et al. 2018) and has recently been carefully quantified to persist to even later times $z \sim 5.3$ by Bosman et al. (2022).

Several models were put forward to explain the origin of these excess fluctuations, and in particular the existence of the giant GP trough from Becker et al. (2015). In ionization equilibrium, the Lyα forest opacity is directly connected to the residual neutral hydrogen fraction, which depends on the density of the gas, the strength of the ionizing background, and (through the recombination rate) the gas temperature. D'Aloisio et al. (2015) suggested that relic temperature fluctuations from a late-ending, extended, and hot reionization process could lead to a dramatic contrast in IGM temperatures at $z \lesssim 6$, leading to strong variations in Lyα forest opacity due to the temperature dependence of the hydrogen recombination rate. Such models predict that opaque troughs correspond to large-scale galaxy overdensities (Davies et al. 2018a), however, which is disfavored by galaxy surveys toward the largest GP troughs at $z \sim 5.7$ (Becker et al. 2018; Kashino et al. 2020; Christenson et al. 2021). Chardin et al. (2015, 2017) showed that the strong fluctuations in the ionizing background would be a natural consequence of a bright and rare source population, i.e., luminous quasars (see also Meiksin 2020), but the initial suggestions of an overabundant faint quasar population at $z > 5$ (Giallongo et al. 2015) are now generally disfavored (e.g., D'Aloisio et al. 2017; Matsuoka et al. 2018; Parsa et al. 2018). While interest in such quasar-dominated models has begun to resurface in light of the discovery of a substantial population of faint AGNs at $z > 5$ by JWST surveys (e.g., Harikane et al. 2023; Kocevski et al. 2023; Labbe et al. 2023; Maiolino et al. 2023; Matthee et al. 2024), due to their typically highly reddened nature, it is still unknown whether these objects contribute substantially to the ionizing photon budget. Davies & Furlanetto (2016) explored the possibility that galaxy-sourced ionizing background fluctuations could be amplified by a coupling to the mean free path of ionizing photons (see also D'Aloisio et al. 2018) following analytic arguments from McQuinn et al. (2011), but in order to match the broad distribution of Lyα forest opacity, the *average* mean free path would have to be a factor of a few shorter than the extrapolation from lower-redshift measurements (Worseck et al. 2014). All of these models presupposed that reionization was more-or-less complete at $z \lesssim 6$, and required adjusting their corresponding parameters to uncomfortable ends of the parameter space to be even qualitatively consistent with Lyα forest observations.

The most successful model thus far was proposed by Kulkarni et al. (2019; see also Nasir & D'Aloisio 2020), wherein reionization is *not* finished by $z \sim 6$, but instead ends at $z \lesssim 5.5$. In this model, the large variations in the Lyα forest opacity at $z < 6$ are explained by residual "true" (i.e., mostly neutral) GP troughs and shadowing of the ionizing background by large-scale patches of neutral IGM. In particular, late reionization models most readily reproduce observations of rare, extremely large GP troughs (Keating et al. 2020) and so-called "dark gaps" (Zhu et al. 2021, 2022) even down to redshifts $z \lesssim 5.5$. Such models, however, require careful tuning of the ionizing emissivity evolution to self-consistently reproduce both the late Lyα forest fluctuations and other reionization-epoch observables (Kulkarni et al. 2019; Keating et al. 2020), and the efficient moment-based radiative transfer method used in these works may suffer from numerical suppression of large-scale ionizing background fluctuations inside the ionized regions (Wu et al. 2021). Nevertheless, this explanation has proven to be the most natural one so far, and has begun to deliver powerful constraints on the properties of the reionizing sources (Choudhury et al. 2021; Qin et al. 2021).

An important boundary condition to reionization is the strength of the ionizing background, which permeates the Universe after the process is complete, and radiative transfer simulations generally predict a rapid rise in the amplitude of the ionizing background as the ionized bubbles overlap, and the last neutral islands disappear (e.g., Gnedin 2000). Constraints on the hydrogen photoionization rate $\Gamma_{\rm HI}$ derived from Lyα forest observations suggest a rise by a factor of at least a few from $z \sim 6$ to $z \sim 5$ (Bolton & Haehnelt 2007; Calverley et al. 2011; Wyithe & Bolton 2011; Davies et al. 2018d), followed by a much shallower evolution from $z \sim 5$ to $z \sim 2$ (Becker & Bolton 2013), which is potentially suggestive of this picture, although we note that the flattening could instead reflect changes in the nature of the gas responsible for the ionizing opacity (Muñoz et al. 2016). Precision measurements of $\Gamma_{\rm HI}$ at $z \sim 5$–6 are complicated by the low mean transmission of the Lyα forest (Bosman et al. 2018; Eilers et al. 2018) and a degeneracy with the relatively unknown thermal state of the IGM (e.g., Becker & Bolton 2013), but recent improvements in data quality (Bosman et al. 2022) and new constraints on the IGM thermal state at $z > 5$ (e.g., Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2021) suggest that we can now do much better.

Furthermore, the ionizing background at the very end of reionization provides a census of the ionizing photons being produced throughout the Universe, which drove the process to completion, and introduces a strong boundary condition to models of the reionizing sources (e.g., Bouwens et al. 2015a). Connecting the emissivity of ionizing photons $\epsilon_{\rm ion}$ to the photoionization rate $\Gamma_{\rm HI}$, however, requires an additional ingredient: the mean free path of ionizing photons, $\lambda_{\rm mfp}$. While the mean free path can be estimated from the distribution of neutral hydrogen systems at redshifts $z \lesssim 4$ (e.g., Rudie et al. 2013; Prochaska et al. 2014), at higher redshifts, the identification of individual absorption lines becomes more challenging. To overcome this difficulty, and to bypass potential uncertainties related to line blending and absorber clustering (Prochaska et al. 2014), the mean free path can instead be measured *directly* by stacking quasar spectra at their rest-frame Lyman limit (Prochaska et al. 2009). Until recently, such measurements were limited to $z \lesssim 5$ (Prochaska et al. 2009; O'Leary & McQuinn 2012; Fumagalli et al. 2013; Worseck et al. 2014), and so constraints on the ionizing emissivity were limited to a time well after the end of reionization (Becker & Bolton 2013) or required an extrapolation of the mean free path evolution to earlier times (D'Aloisio et al. 2018).

Recently, Becker et al. (2021) measured the mean free path of ionizing photons at $z \sim 6$, improving the spectral stacking method by additionally modeling the impact of the quasar ionizing radiation on the line-of-sight Lyman-series and Lyman limit opacity. They measured a mean free path of $\lambda_{\rm mfp} = 0.75^{+0.65}_{-0.45}$ proper Mpc, much shorter than the extrapolated value implied by the tight power-law evolution measured by Worseck et al. (2014) at $z \lesssim 5$. This mean free path is short enough to

strongly affect the progression of reionization and place higher demands on the ionizing output from galaxies (Cain et al. 2021; Davies et al. 2021), and implicitly requires the structure of the gas to trace that of (cold) dark matter down to very small scales (Emberson et al. 2013; Park et al. 2016; D'Aloisio et al. 2020). Notably, an interpolation between the $z = 5.1$, and $z = 6$ measurements of Becker et al. (2021) now places the mean free path at the level required by the Davies & Furlanetto (2016) model to explain the Lyα forest fluctuations.

Here, we revisit Lyα forest fluctuations in the context of the Davies & Furlanetto (2016) model to constrain parameters of the $z > 5$ IGM, taking advantage of recent improvements in Lyα forest data quality and new constraints on the IGM thermal state (Gaikwad et al. 2021). We apply a likelihood-free inference methodology based on Davies et al. (2018d) to the extended XQR-30 (D'Odorico et al. 2023) Lyα forest data set (Bosman et al. 2022) to constrain the average photoionization rate $\Gamma_{\rm HI}$ and the average mean free path of ionizing photons $\lambda_{\rm mfp}$ from $z = 5.0$ to 6.1. We also assess the degree to which the data are consistent with a draw from the model. Finally, we discuss what our constraints imply for the evolution of the ionizing emissivity and the timing of the reionization epoch.

In this work, we assume a ΛCDM cosmology with $h = 0.685$, $\Omega_m = 0.31$, $\Omega_b = 0.049$, consistent with Planck (Planck Collaboration et al. 2020). The distance units are comoving unless otherwise specified.

2. Simulation Method

We simulate post-reionization Lyα forest fluctuations in a manner very similar to Davies et al. (2018d). The general philosophy is to combine skewers through a cosmological hydrodynamical simulation, which can resolve scales small enough for a converged description of the Lyα forest, with skewers from a separate seminumerical ionizing background simulation, which has a volume sufficiently large to sample the full distribution of large-scale background fluctuations. In the process, we unfortunately decouple the density of the gas responsible for Lyα absorption from the intensity of the ionizing radiation field. We will revisit this assumption in Section 5 and attempt to quantify the resulting biases.

2.1. Nyx Hydrodynamical Simulation

For the small-scale density, temperature, and velocity structure of the gas, we use a hydrodynamical simulation run with the `Nyx` code (Almgren et al. 2013). The simulation we employ is $100/h$ Mpc on a side and has 4096^3 dark matter particles and baryon grid cells, providing a box size and resolution adequate for converged Lyα forest statistics (Lukić et al. 2015; Oñorbe et al. 2019). Our fiducial modeling used three redshift snapshots from $z = 5$, $z = 5.5$, and $z = 6$; for redshifts between these values, we use the overdensity field from the snapshot closest in redshift, and scale the physical densities by $(1 + z)^3$. We extract 100,000 skewers of gas overdensity, temperature, and peculiar velocity from each snapshot starting from random locations in the simulation box and traversing in one of six random directions along grid axes ($\pm x$, y, z). The overdensities are converted into physical densities $\propto (1 + z)^3$ according to the evolving redshift across the range of interest. Lyα forest spectra are computed by assuming an ionization equilibrium to calculate the corresponding neutral hydrogen density and using the Voigt profile approximation of Tepper-García (2006) to deposit Lyα opacity onto a 2 km s^{-1} resolution grid.

One fundamental limitation of our Nyx simulation is that, by virtue of its optically thin photoionization and photoheating rates from Haardt & Madau (2012), reionization occurs very early ($z \sim 15$), and with minimal heat injection ($\Delta T \sim 10{,}000$ K). This is in conflict with observational constraints on the timing of reionization (e.g., Davies et al. 2018b) and theoretical models of the expected heating by ionization fronts sweeping through the IGM (Miralda-Escudé & Rees 1994; Abel & Haehnelt 1999; D'Aloisio et al. 2019). As a result, at $z = 5$–6, the IGM is colder, and the temperature–density relation steeper, than our theoretical understanding would otherwise suggest, with further evidence coming from recent constraints on the IGM thermal state at $z \sim 5.5$ (Gaikwad et al. 2020). The combination of these two effects will lead to overestimated recombination rates, and thus overestimated values of $\Gamma_{\rm HI}$ when matched to the observed Lyα forest opacity (e.g., Bolton et al. 2005; Bolton & Haehnelt 2007; Becker & Bolton 2013; D'Aloisio et al. 2018). To overcome this limitation of our hydrodynamical simulations, we adjust the IGM temperatures of our fiducial skewers to match a more realistic average post-reionization thermal state. We model the temperature evolution of the IGM following the analytic method from Davies et al. (2018a; see also Upton Sanderbeck et al. 2016) assuming an instantaneous reionization at $z_{\rm re}$ with a heat injection of $\Delta T = 20{,}000$ K (e.g., D'Aloisio et al. 2019).

In Figure 1, we show the resulting mean temperature T_0 and slope of the temperature–density relation γ for a range of $z_{\rm re}$ compared to recent measurements of the IGM thermal state (Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2020). From this comparison, we adopt $z_{\rm re} = 7.2 \pm 0.5$ as our fiducial range of thermal models. However, we will also show the range of inferred $\Gamma_{\rm HI}$ values corresponding to more extreme thermal models: the cold, native Nyx temperatures corresponding to $z_{\rm re} \sim 15$ and a very hot model with $z_{\rm re} = 6.2$. By manually adjusting the simulated temperatures in post-processing, we neglect the effect of Jeans smoothing on the Lyα transmission, but this is likely a very small effect (see, e.g., Becker & Bolton 2013; Kulkarni et al. 2015).

2.2. Ionizing Background Fluctuations

The volume of the Nyx simulation described above is not entirely sufficient to capture the large-scale ionizing background fluctuations characteristic of the post-reionization IGM, where box sizes of at least a few hundred Mpc are required (e.g., Davies & Furlanetto 2016; D'Aloisio et al. 2018). For our fiducial model, we decided to create a much larger, but independent, cosmological volume within which to simulate the ionizing background—the consequences and biases of this approach will be discussed later.

We computed large-scale simulations of ionizing background fluctuations using the method of Davies & Furlanetto (2016), built off of the seminumerical framework of the `21cmFAST` code from Mesinger et al. (2011). Cosmological initial conditions were generated on a 4096^3 grid in a volume 512 comoving Mpc on a side, then evolved to $z = 5.5$ onto a coarser 512^3 grid via the Zel'dovich approximation (Zel'dovich 1970). Halos were instantiated from the 4096^3 grid via the excursion set halo-filtering approach of Mesinger & Furlanetto (2007) down to a minimum halo mass $M_{h,\rm min} = 2 \times 10^9\ M_\odot$, and shifted to their evolved positions

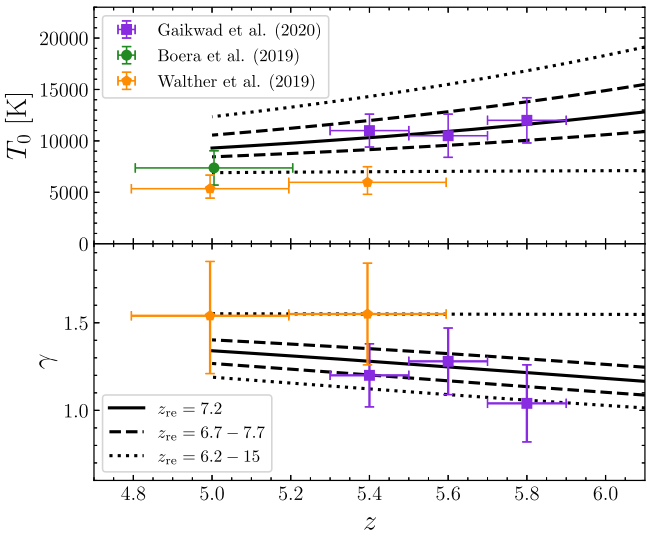

Figure 1. Evolution of the temperature at mean density in the IGM (top) and the slope of the temperature–density relation (bottom). The curves show the analytic models described in the text for our fiducial range of reionization heating models, while the data points show measurements from Boera et al. (2019), Walther et al. (2019), and Gaikwad et al. (2020).

at $z = 5.5$ via the Zel'dovich displacement field evaluated on the 512^3 grid at their initial positions from the linear theory, a procedure, which produces realistic halo clustering down to ∼Mpc scales (Mesinger & Furlanetto 2007). Ultraviolet (UV) luminosities of galaxies corresponding to each halo were assigned by abundance matching to the Bouwens et al. (2015b) UV luminosity functions, resulting in UV magnitudes ranging from $-12.6 \gtrsim M_{\rm UV} \gtrsim -23.3$. For simplicity, we assume that the ionizing luminosity of each galaxy is proportional to its UV luminosity, and leave the constant of proportionality as a free parameter.

We then compute the ionizing background by a brute force radiative transfer algorithm applied to a coarser, 128^3 grid, corresponding to cells 4 Mpc on a side. The ionizing luminosities of the galaxies are deposited onto the coarse grid, and then, the ionizing radiation intensity in cell i is computed via

$$J_{\nu,i} = J_{\nu,i}^{\rm local} + \sum_{j \neq i} \frac{L_{\nu,j}}{(4\pi)^2 d_{ij}^2} e^{-\tau_{ij}(\nu)}, \quad (1)$$

where $J_{\nu,i}^{\rm local}$ is the intensity due to local sources, d_{ij} is the distance between cells i and j, and τ_{ij} is the ionizing photon optical depth between i and j. The sum over all other cells j is performed assuming locally periodic boundary conditions, i.e., the cell i *sees* the other cells j within a volume of the same size as the full box but centered on cell i. The local source intensity is computed assuming a uniform ionizing emissivity $\epsilon_{\nu,i}$ inside the cell,

$$J_{\nu,i}^{\rm local} = \frac{\epsilon_{\nu,i}\lambda_{\nu,i}}{4\pi}(1 - e^{-0.72 l/\lambda_{\nu,i}}), \quad (2)$$

where $\lambda_{\nu,i}$ is the mean free path of cell i at frequency ν, l is the side length of the cell, and the constant factor of 0.72 was found to have good resolution convergence properties in Davies & Furlanetto (2016). The optical depth between cells is computed by integrating the opacity $\kappa = d\tau/dr$ on the 128^3 grid,

$$\tau_{ij}(\nu) = \int_i^j \kappa(\nu, \Gamma, \Delta) dr, \quad (3)$$

where κ is assumed to vary with the local photoionization rate Γ and overdensity relative to the cosmic mean $\Delta = \rho/\bar{\rho}$, according to

$$\kappa = \kappa_0 \times \Delta \left(\frac{\nu}{\nu_{\rm HI}}\right)^{-0.9} \left(\frac{\Gamma}{\Gamma_0}\right)^{-2/3}, \quad (4)$$

where κ_0 is the opacity of the IGM at the ionizing edge of hydrogen $\nu_{\rm HI}$ and at a reference photoionization rate Γ_0, and we adopt the same power-law dependencies on Δ and ν as Davies & Furlanetto (2016). Finally, the photoionization rate is calculated by integrating over J_ν,

$$\Gamma_{{\rm HI},i} = 4\pi \int_{\nu_{\rm HI}}^{\infty} \frac{J_{\nu,i}}{h\nu} \sigma_{\rm HI}(\nu) d\nu \propto \frac{J_{\bar{\nu},i}}{h\bar{\nu}} \sigma_{\rm HI}(\bar{\nu}) \Delta\nu, \quad (5)$$

where $\sigma_{\rm HI}$ is the hydrogen photoionization cross section (from Verner et al. 1996). In practice, for computational efficiency, we perform the calculation with a single ionizing photon frequency $h\bar{\nu} \approx 17.9$ eV, which was found to produce very similar $\Gamma_{\rm HI}$ fluctuations to a more comprehensive multi-frequency calculation (Davies & Furlanetto 2016).

We computed fluctuating $\Gamma_{\rm HI}$ fields for 14 different average mean free path values $\lambda_{\rm mfp} = \kappa_0^{-1} = 5$, 6, 8, 10, 15, 20, 25, 30, 40, 50, 60, 80, 100, 150 (comoving) Mpc at a fixed redshift $z = 5.5$. This relatively coarse sampling was necessitated by the computational inefficiency of the Davies & Furlanetto (2016) method; see Gaikwad et al. (2023) for a recent implementation of a more efficient algorithm. We first computed the $\Gamma_{\rm HI}$ field assuming a uniform $\kappa = \kappa_0(\bar{\nu}/\nu_{\rm HI})^{-0.9}$, and then determine a normalization factor (corresponding to a combination of an ionizing emissivity normalization and a correction for the monochromatic approximation) to initialize the mean $\Gamma_{\rm HI}$ to the value corresponding to the prescribed ionizing opacity ,i.e., Γ_0 in Equation (4). In the following, we will use only the *relative* fluctuations in Γ, i.e., the relative fluctuating field $\tilde{\Gamma}_{\rm HI} \equiv \Gamma_{\rm HI}/\bar{\Gamma}_{\rm HI}$, so the exact value of Γ_0 is unimportant. Note that, due to the dependence of κ on $\Gamma_{\rm HI}$, the calculation must be iterated many times to achieve self-consistency. We iterate the calculation until the average change in $\Gamma_{\rm HI}$ falls below 0.1%, requiring $\gtrsim$20 iterations for $\lambda_{\rm mfp} \lesssim 10$ Mpc and ∼5 iterations for $\lambda_{\rm mfp} \gtrsim 60$ Mpc.

In Figure 2, we show a slice through four of the fluctuating ionizing background fields with $\lambda_{\rm mfp} = 10$–80 Mpc. The spatial structure of the fluctuations is very similar, driven by the large-scale clustering of the ionizing sources, but the fluctuations are stronger at shorter average mean free path. We take advantage of this coherence and linearly interpolate in $\log\tilde{\Gamma}$ between adjacent models to produce skewers at an arbitrary average mean free path from 5 to 150 Mpc.

As discussed in the following section, our inference procedure requires many millions of simulated Lyα forest spectra. To reduce the computational requirements for producing the mock spectra, we adopt an approach analogous to the fluctuating Gunn–Peterson approximation (Weinberg et al. 1997) to both adjust the mean $\Gamma_{\rm HI}$ and introduce the fluctuating

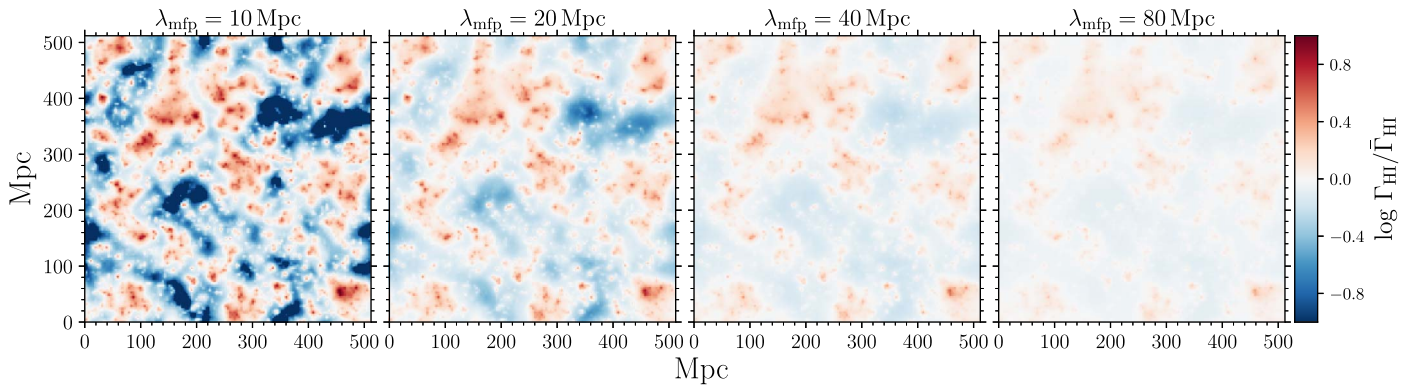

Figure 2. Slices through a subset of our fluctuating $\Gamma_{\rm HI}$ fields, 4 Mpc thick and 512 Mpc on a side, showing a representative range of $\lambda_{\rm mfp}$ values of 10, 20, 40, and 80 Mpc from left to right.

$\Gamma_{\rm HI}$ field. We first compute an initial set of 100,000 Lyα forest spectra from the Nyx simulation skewers with a uniform $\Gamma_{\rm HI} = \Gamma_{\rm init}$, where we set $\Gamma_{\rm init}$ to the midpoint of the redshift-dependent uniform prior on $\Gamma_{\rm HI}$ used during the inference procedure (see Section 4). We then introduce $\Gamma_{\rm HI}$ fluctuations by rescaling the pixel optical depths along the skewer,

$$\tau_\alpha(z_{{\rm Ly}\alpha,i}) \propto \left(\frac{\Gamma_{\rm new} \times \tilde{\Gamma}_{\rm HI}(z_i)}{\Gamma_{\rm init}}\right)^{-1}, \quad (6)$$

where $\Gamma_{\rm new}$ is the new average value of $\Gamma_{\rm HI}$, and $\tilde{\Gamma}_{\rm HI}$ is evaluated at z_i without taking into account redshift-space distortions. We find that this procedure is $\sim$50 times faster than a more accurate reevaluation of the spectra from the neutral hydrogen distribution, without substantially affecting our main results.

While the ionizing background is inherently fluctuating in our simulations, references to "$\Gamma_{\rm HI}$" in the rest of the text, particularly in the context of our constraints from observations, will represent the volume-averaged value of $\Gamma_{\rm HI}$.

2.2.1. Self-consistent Background Fluctuations

For comparison and bias assessment purposes, we also produced a similar suite of simulations of ionizing background fluctuations using the halo catalog of the 100 Mpc$/h$ Nyx simulation with the same set of average mean free path values. We abundance-matched dark matter halos down to the same minimum mass, $M_{\rm min} = 2 \times 10^9\ M_\odot$, but we note that, due to the coarse force resolution of the uniform grid particle-mesh scheme used to evolve the dark matter particles, the number of halos below $M_h \sim 10^{10}\ M_\odot$ is significantly underestimated (e.g., Lukić et al. 2007; Almgren et al. 2013). As a result, the UV magnitude range is somewhat restricted compared to our larger volume, $-13.9 \gtrsim M_{\rm UV} \gtrsim -22.7$.

For these complementary simulations, we computed the background fluctuations on a 64^3 grid, corresponding to a spatial resolution of $\sim$2 Mpc. These simulations tend to have weaker large-scale fluctuations than the fiducial ones for relatively short mean free paths, but for long mean free paths, they exhibit stronger large-scale fluctuations. The former is due to the suppression of large-scale modes in the density field due to the smaller volume; the latter is due to a limitation of the algorithm used to compute the radiation field. Specifically, the size of the *local volume* seen by any cell sets an upper limit to the effective mean free path of roughly half the box size, or in this case $\sim$73 Mpc.

We will refer to this suite of simulations as the *self-consistent* model, as it provides ionizing background fluctuations sourced by the same underlying density field as the Lyα forest. As mentioned above, the fiducial model relies on an independent volume for ionizing background fluctuations, which inherently decouples them from the density field. We will compare the constraints obtained in the fiducial model to the self-consistent model, as well as a decorrelated version of the self-consistent model, where the ionizing background fluctuations are (as in the fiducial model) drawn from a random region in the 100 Mpc$/h$ volume.

3. Likelihood-free Inference with Approximate Bayesian Computation

Armed with the simulation framework described above, we aim to constrain two parameters: $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$. The distribution of Lyα forest opacity at $z > 5$ is not well-described by Gaussian statistics (e.g., Bosman et al. 2018) that would typically be adopted in Bayesian parameter inference, and at $z \gtrsim 5.6$, one must also sensibly account for nondetections that represent the most constraining GP troughs. We adopt a likelihood-free inference technique known as ABC (Rubin 1984; Tavaré et al. 1997; Pritchard et al. 1999) to overcome these challenges with a principled approach.[22] Our method is very similar to that described in Davies et al. (2018d), but with some important differences that we describe below.

The approximation fundamental to the *approximate* nature of ABC is in the definition of the likelihood of the data vector $\boldsymbol{d}$ given the model parameters θ, $p(\boldsymbol{d}|\theta)$. In ABC, the likelihood is approximated by (e.g., Marin et al. 2012)

$$p(\boldsymbol{d}|\theta) \approx p(\rho(\boldsymbol{d}, \boldsymbol{x}(\theta)) < \epsilon|\theta) \approx p(\rho(\boldsymbol{s_d}, \boldsymbol{s_x}(\theta))|\theta) \quad (7)$$

where ρ is a (nearly arbitrary) distance metric between the data and a mock data draw $\boldsymbol{x}$, and ϵ is a distance threshold below which the data and mock data are deemed "similar enough." As ϵ approaches zero, the approximate posterior probability density function (PDF) will converge toward the true posterior PDF (e.g., Blum 2010). Typically the raw data are first transformed into a lower dimensional summary statistic $\boldsymbol{s}(\boldsymbol{d}) \equiv \boldsymbol{s_d}$, and the same procedure is applied to make mock observations of the summary statistic $\boldsymbol{s}(\boldsymbol{x}) \equiv \boldsymbol{s_x}$. The ABC

[22] We note that more sophisticated and efficient likelihood-free inference methods aided by machine learning have recently been developed (e.g., Alsing et al. 2018; Cole et al. 2022), but we adopt ABC for conceptual simplicity.

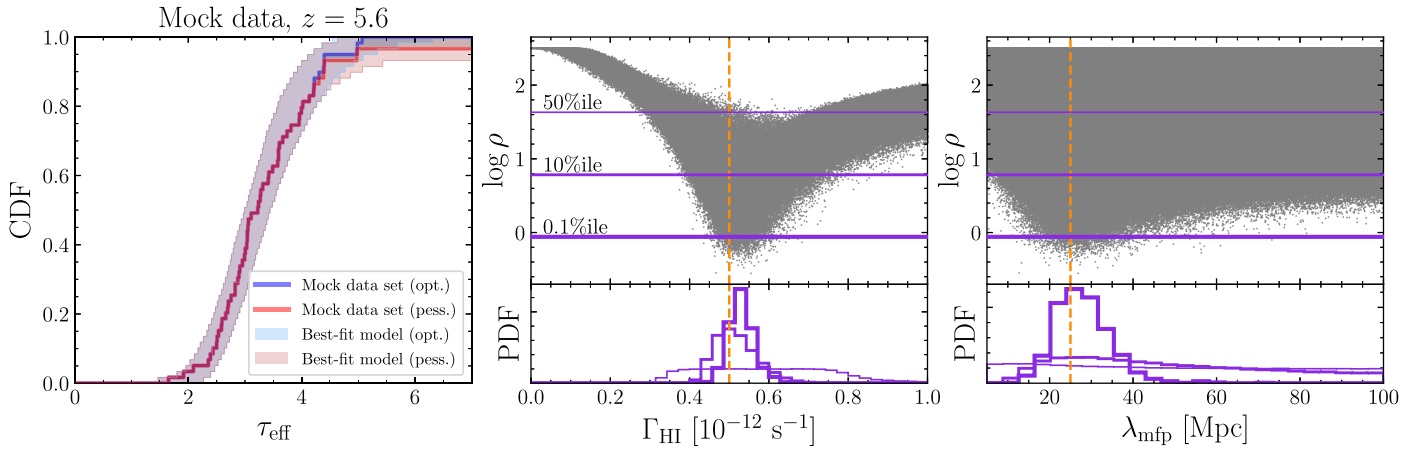

Figure 3. Demonstration of ABC on a mock data set. The blue and red curves in the left panel correspond to the cumulative distribution functions of Lyα forest effective optical depth from a mock XQR-30 data set, with nondetections treated as optimistic ($F = 2 \times \sigma_F$) or pessimistic ($F = 0$), respectively, following Bosman et al. (2018, 2022). The shaded regions correspond to the central 95% of the scatter of additional mock data sets with $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ set to their mean posterior estimates. The gray points in the upper halves of the middle and right panels show the distance metric ρ (Equation (8)) computed from 1,000,000 mock data sets. The horizontal lines show three different ρ thresholds below which lie 50%, 10%, and 0.1% of the mock data samples from top to bottom. The corresponding posterior PDFs on $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ are shown in the lower panels, with the true values indicated by the vertical dashed lines.

procedure involves computing many such mock data sets, with parameter values θ drawn from the prior $p(\theta)$, and selecting some number of samples with distances below some threshold ϵ—those samples then (approximately) represent samples from the posterior PDF $p(\theta|\boldsymbol{d})$. As discussed below, the threshold ϵ is typically chosen such that one retains a given (small) fraction of the parameter samples.

There is substantial freedom in choosing both the distance metric ρ and the summary statistic $\boldsymbol{s}$. We choose to summarize the Lyα forest data by their $\Delta z = 0.1$ binned effective optical depths, $\boldsymbol{s} = -\ln\langle \exp -\tau_\alpha \rangle = \tau_{\rm eff}$, and compute the Euclidean (L_2 norm) distance between the *rank-order* set of observed and mock $\tau_{\rm eff}$ (see also Davies et al. 2018d). Specifically, we have

$$\rho(\boldsymbol{s_d}, \boldsymbol{s_x}) = \sqrt{\sum_i^{N_{\rm q}} (\tau_{{\rm eff},i}^{\rm obs} - \tau_{{\rm eff},i}^{\rm mock})^2}\,, \tag{8}$$

where $\tau_{{\rm eff},i}^{\rm obs} i$ and $\tau_{{\rm eff},i}^{\rm mock}$ are the ith highest $\tau_{\rm eff}$ values in the set of $N_{\rm q}$ observed and mock Lyα forest spectra, respectively. While the transformation to effective optical depth provides sensitivity to strongly absorbed regions of the Lyα forest, observational noise can lead to fully opaque GP troughs with a mean transmission below zero, leading to undefined $\tau_{\rm eff}$. In both the observed data and mock data, we set $\tau_{\rm eff} = -\ln(2 \times \sigma_F)$ if the mean transmission is below twice the statistical error σ_F; a common transformation in the $z > 5$ Lyα forest literature (Becker et al. 2015; Bosman et al. 2018, 2022; Eilers et al. 2018).

In Figure 3, we demonstrate an example of the procedure described above applied to a mock Lyα forest data set at $z = 5.6$. We assume true values of $\Gamma_{\rm HI} = 5 \times 10^{-13}$ s^{-1}, and $\lambda_{\rm mfp} = 25$ Mpc, and a data set consisting of 50 quasar sightlines with a statistical error $\sigma_F = 0.01$, on the $\Delta z = 0.1$ binned transmitted flux, and a continuum error $\sigma_C = 0.1$. The cumulative distribution function (CDF) of this "observed" data set is shown by the thick curves in the left panel of Figure 3, where the left (blue) and right (red) curves display undetected transmission with $F < 2\sigma_F$ as $\tau_{\rm eff} = -\ln(2\sigma_F)$, or $\tau_{\rm eff} = \infty$, respectively, similar to Bosman et al. (2018, 2022). We draw 1,000,000 values of $\theta = \{\Gamma_{\rm HI}, \lambda_{\rm mfp}\}$ assuming a uniform prior from $\Gamma_{\rm HI} = 0$–10^{-12} s^{-1}, and $\lambda_{\rm mfp} = 5$–150 Mpc, and for each θ, we compute a mock data set including the same noise properties as the observed one. The resulting distances $\rho(\boldsymbol{s_d}, \boldsymbol{s_x})$ versus $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ are shown in the (upper) middle and right panels of Figure 3, respectively. We then choose a series of distance thresholds ϵ such that 50%/10%/0.1% of the mock data sets have $\rho(\boldsymbol{s_d}, \boldsymbol{s_x}) < \epsilon$, shown by the horizontal lines, and retain the corresponding θ values as samples from the posterior PDFs in the lower panels. The blue and red shaded regions' curves in the left panel of Figure 3 show the central 95% of CDFs computed from mock data sets adopting the lowest ϵ posterior mean $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$, demonstrating good consistency with the *observed* one.

4. Constraints on $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ from the Extended XQR-30 Data Set

We apply the ABC inference methodology described above to the extended Lyα forest data set from Bosman et al. (2022). The data consist of 51 optical quasar spectra observed with the X-Shooter spectrograph (Vernet et al. 2011), primarily sourced from the Extended XQR-30 sample (D'Odorico et al. 2023), as well as 16 spectra taken with the Echellette Spectrograph and Imager (ESI; Sheinis et al. 2002) spectrograph, i.e., 67 quasar spectra in total. All quasars are located at $z > 5.7$, and their spectra possess signal-to-noise ratios (S/Ns) larger than 10 per pixel. Intervening damped Lyα absorbers are masked based on the detection of associated metal absorbers in the metal catalog of Davies et al. (2023). In a few quasars, portions of the spectra with suspected broad absorption line contamination are also masked. We refer interested readers to Bosman et al. (2022) for more details of the sample.

In order to measure the effective optical depths, the intrinsic quasar continua are reconstructed based on their unabsorbed emission features at $\lambda_{\rm rest} > 1215.16$ Å. We use the principal component analysis (PCA) approach of Davies et al. (2018c) with the improvements described in Bosman et al. (2021, 2022). The PCA approach reproduces the quasar continua at $\lambda_{\rm rest} < 1180$ Å with a wavelength-dependent ~8% accuracy, and subpercent precision. Finally, the effective optical depths

Table 1
Derived Constraints on the Hydrogen Photoionization Rate $\Gamma_{\rm HI}$ in Units of 10^{-12} s^{-1} and Mean Free Path $\lambda_{\rm mfp}$ in Units of Comoving Mpc

z	$N_{\rm q}$	$\Gamma_{\rm HI}$ (10^{-12} s^{-1})	Stat. Err.	Tot. Err	S-c Corr.	$\lambda_{\rm mfp}$ (Mpc)	Stat. Err	Tot. Err	S-c Corr.
5.0	37	0.599	$^{+0.035}_{-0.031}$	$^{+0.109}_{-0.085}$	1.034	⋯	>27.7	>23.3	0.828
5.1	48	0.582	$^{+0.028}_{-0.024}$	$^{+0.115}_{-0.081}$	1.043	⋯	>33.7	>26.1	0.800
5.2	55	0.547	$^{+0.022}_{-0.023}$	$^{+0.104}_{-0.080}$	1.035	⋯	>42.6	>33.9	0.857
5.3	58	0.503	$^{+0.024}_{-0.021}$	$^{+0.113}_{-0.074}$	1.066	79.1	$^{+43.9}_{-29.7}$	$^{+43.9}_{-52.3}$	0.546
5.4	64	0.447	$^{+0.025}_{-0.024}$	$^{+0.135}_{-0.069}$	1.135	29.0	$^{+6.9}_{-5.2}$	$^{+6.9}_{-17.5}$	0.633
5.5	64	0.381	$^{+0.019}_{-0.017}$	$^{+0.104}_{-0.060}$	1.103	44.3	$^{+25.0}_{-11.3}$	$^{+25.0}_{-31.7}$	0.380
5.6	59	0.343	$^{+0.024}_{-0.021}$	$^{+0.110}_{-0.055}$	1.137	23.5	$^{+6.9}_{-4.7}$	$^{+6.9}_{-13.9}$	0.686
5.7	51	0.255	$^{+0.019}_{-0.016}$	$^{+0.084}_{-0.042}$	1.139	24.0	$^{+7.8}_{-6.2}$	$^{+7.8}_{-15.1}$	0.654
5.8	45	0.208	$^{+0.021}_{-0.019}$	$^{+0.076}_{-0.036}$	1.161	15.9	$^{+6.5}_{-4.8}$	$^{+6.5}_{-10.1}$	0.652
5.9	28	0.185	$^{+0.021}_{-0.020}$	$^{+0.063}_{-0.035}$	1.125	23.5	$^{+19.0}_{-10.6}$	$^{+19.0}_{-17.8}$	0.474
6.0	19	0.167	$^{+0.032}_{-0.029}$	$^{+0.079}_{-0.037}$	1.195	⋯	<22.6	<22.6	⋯
6.1	10	0.147	$^{+0.055}_{-0.039}$	$^{+0.097}_{-0.044}$	1.229	⋯	<21.6	<22.6	⋯

Note. The nominal values represent the median of the posterior PDF, with the statistical error (stat. err.) representing the 16th–84th percentiles of the posterior PDF. The total error (tot. err.) includes contributions from systematic uncertainties (see text for details), and should be considered highly covariant between redshift bins. The self-consistency correction (s-c corr.) is the ratio of the posterior medians of the self-consistent model and the fiducial model; the gray points in Figures 6 and 8 can be recovered by multiplying the fiducial constraints by the value in this column. Upper and lower limits on $\lambda_{\rm mfp}$ roughly correspond to 2σ limits; see Section 4 for details.

are measured in bins of $dz = 0.1$, and only bins that are >50% unmasked are retained. The number of retained quasar sightlines $N_{\rm q}$ at each redshift is shown in Table 1. The masking, observational uncertainties, and wavelength-dependent continuum uncertainties are forward-modeled in all simulated mocks. Bosman et al. (2022) showed that mock data generated in this manner are statistically indistinguishable from the real data in the post-reionization regime ($5.0 < z < 5.3$), solidifying our confidence in the noise model.

For each redshift bin—the same $dz = 0.1$ bins defined in Bosman et al. (2022)—we compute 10,000,000 mock data sets, drawing from uniform priors in $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ in uniform prior windows[23] where the posterior PDF has substantial support. We then define a relative threshold $\epsilon_{\rm th}$ by the fraction of retained mocks with distances below the actual ABC threshold, e.g., with $\epsilon_{\rm th} = 0.1$ we would retain the parameter values whose mock data sets resulted in a distance smaller than the 10th percentile of the entire distribution. For our fiducial posterior PDFs, we retain the 1000 samples with the smallest distances in our set of 10,000,000, i.e., $\epsilon_{\rm th} = 10^{-4}$.

We show the resulting 2D posterior PDFs of $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ from $z = 5.0$–6.1 in Figure 4. The posteriors show only modest correlations between the two parameters, and both decrease steadily to higher redshift. While $\Gamma_{\rm HI}$ is well-constrained at all redshifts, along the $\lambda_{\rm mfp}$ dimension, the posterior clearly runs into the edge of the prior (set by our grid of models) at $z \lesssim 5.3$ and $z \gtrsim 5.9$. To understand this apparent lack of constraining power, note that our constraints on $\lambda_{\rm mfp}$ are driven by a difference between the observed Lyα forest fluctuations and the fluctuations one would expect given a uniform ionizing background. At the low-redshift end, the weak constraints are due to the fact that the observed Lyα forest fluctuations are fully consistent with a uniform ionizing background (Bosman et al. 2022), while, at the high-redshift end, the strength of fluctuations is comparable to our shortest $\lambda_{\rm mfp}$ model.

[23] In practice, we ran small numbers of mocks with very wide priors, and then progressively truncated them to optimize the sampling of the posterior core.

We note that the width of the posterior PDFs in the $\lambda_{\rm mfp}$ dimension are particularly large at $z = 5.5$, and $z = 5.9$ compared to their neighboring redshifts. In general, these variations come about due to the nonlinear connection between $\lambda_{\rm mfp}$ and the width of the $\tau_{\rm eff}$ distribution; the strength of radiation field fluctuations varies only modestly for $\lambda_{\rm mfp} \gtrsim 40$ Mpc (e.g., Figure 2), so any upward tail of the posterior will inevitably be elongated. In addition, at $z = 5.5$ where this apparent difference is most pronounced, it was already noted in Bosman et al. (2022) that the disagreement between the $\tau_{\rm eff}$ CDF and simulations assuming a uniform ionizing background was smaller than at any other redshift $z > 5.3$. We can thus understand the relatively weak constraint, as this disagreement between the observations and uniform background simulations is the source of the constraining power of our analysis.

In Figure 5, we show the posterior medians (black circles) and central 68% credible intervals (black error bars) of $\Gamma_{\rm HI}$, marginalized over $\lambda_{\rm mfp}$. We have also adjusted the posterior constraints by a few percent to correct for the coarse snapshot sampling employed in the simulated spectra (see Appendix); the *raw* uncorrected posterior medians are shown as open squares. Note that these statistical error bars apply only to the fiducial IGM thermal model with $z_{\rm re} = 7.2$. The dark gray shaded regions in Figure 5 show the range of posterior medians obtained for different IGM thermal models with $z_{\rm re} = 6.7$, and $z_{\rm re} = 7.7$, while the light shaded region shows the range corresponding to extreme thermal models with $z_{\rm re} = 6.2$, and $z_{\rm re} \sim 15$ (see Section 2). The uncertainty in $\Gamma_{\rm HI}$ resulting from the IGM thermal state is much larger than the statistical uncertainty except for $z \gtrsim 5.9$, where the transmitted flux is much lower and sampled across relatively few sightlines. Tighter statistical constraints could likely be achieved at these redshifts from more informative summary statistics (e.g., the flux PDF on smaller scales and/or Lyβ transmission; see Davies et al. 2018d), but we leave a more detailed look at the $z \gtrsim 6$ IGM in XQR-30 to future work.

The crosses in Figure 5 show the posterior medians in the self-consistent model (Section 2.2.1), where the ionizing

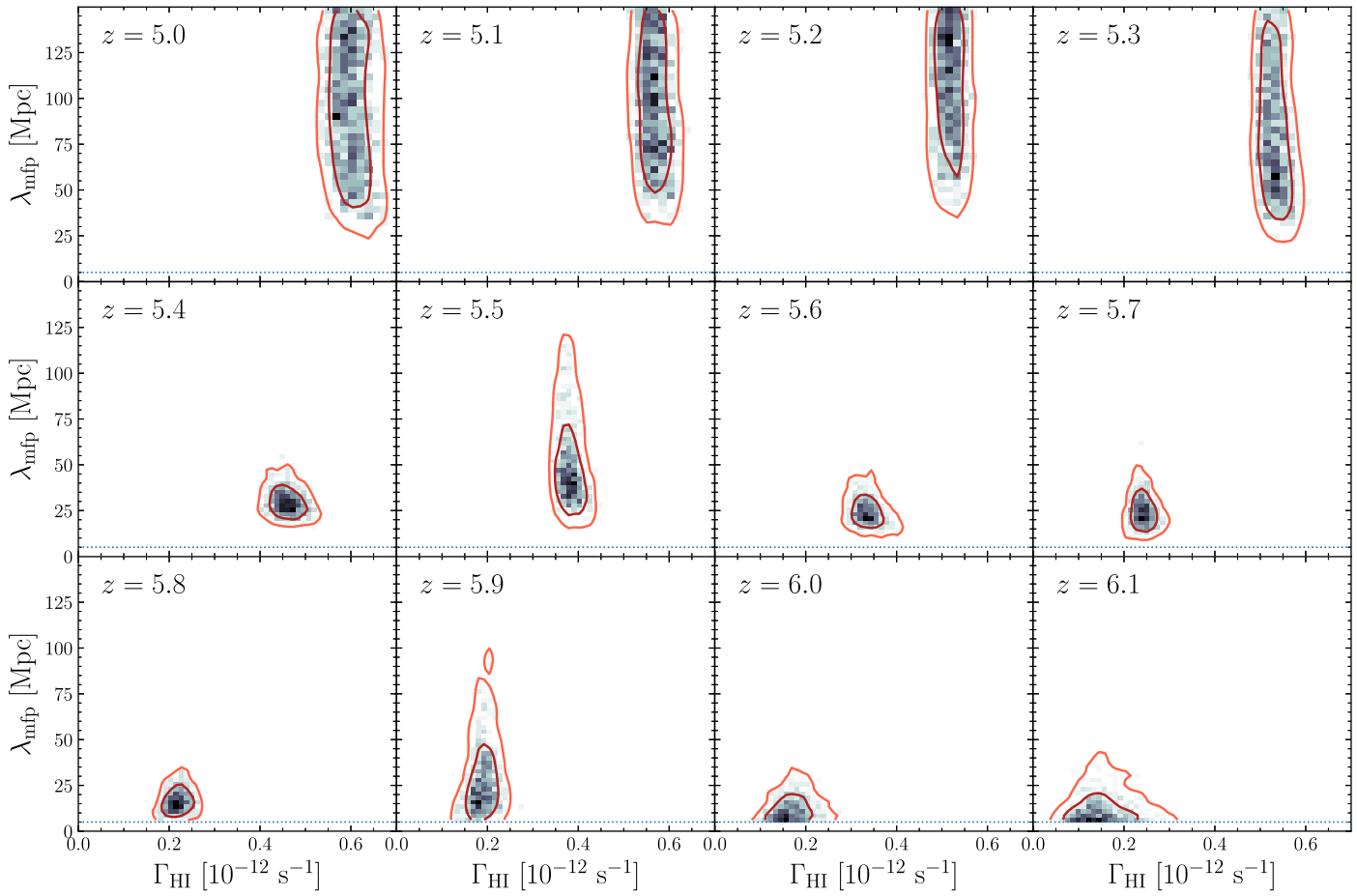

Figure 4. Joint posterior PDF of $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ at $z = 5.0$–6.1. The inner and outer contours contain 68% and 95% of the distribution, respectively. The blue dashed line shows the lower edge of the $\lambda_{\rm mfp}$ prior.

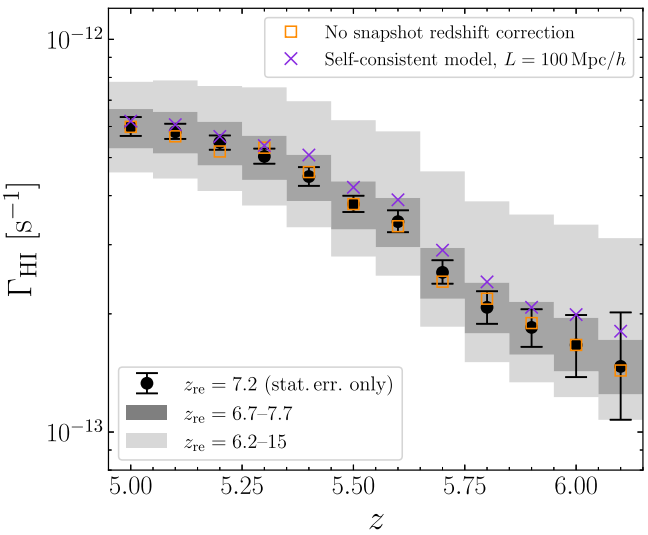

Figure 5. Posterior medians (black circles) and central 68% credible intervals (black thin error bars) on $\Gamma_{\rm HI}$ from the XQR-30 data set assuming $z_{\rm re} = 7.2$. The dark gray shaded region shows the deviation of the posterior medians for $z_{\rm re} = 6.7$, and $z_{\rm re} = 7.7$, while the light gray shaded region shows the range from more extreme thermal models with $z_{\rm re} = 6.2$ and $z_{\rm re} \sim 15$. The open orange points show the constraints without the correction for the coarse redshift snapshot sampling (see Appendix). The purple crosses show the posterior medians from the self-consistent model in the $L = 100$ Mpc$/h$ hydrodynamical simulation volume; see Section 2.2.1.

background fluctuations are drawn from the same physical location in the Nyx hydrodynamical box as the density field when computing the Lyα forest absorption. We see a gradual positive offset of the $\Gamma_{\rm HI}$ values in the self-consistent model, which increases from a few percent at $z = 5$ to $\sim$25% at $z \geqslant 6$. This offset comes about due to correlation between ionizing background fluctuations, and the density field–regions with high $\Gamma_{\rm HI}$ tend to have higher density, and the regions with low $\Gamma_{\rm HI}$ tend to have lower density, shifting the mean Lyα forest transmission to lower values. Thus, the mean $\Gamma_{\rm HI}$ must be higher to reproduce the mean Lyα forest opacity. If we randomize the ionizing background fluctuations with respect to the density field in the 100 Mpc$/h$ box, we find that the resulting $\Gamma_{\rm HI}$ constraints are indistinguishable from the fiducial model.

In the left panel of Figure 6, we compare our constraints on $\Gamma_{\rm HI}$ (solid points) to literature values (open points) after combining the statistical uncertainties with the thermal state uncertainties in quadrature, along with an additional 0.03 dex of systematic uncertainty to approximate the uncertainty due to Jeans smoothing (e.g., Becker & Bolton 2013) and an additional upward systematic uncertainty given by the bias between the fiducial and self-consistent model constraints. We also show gray points, which represent the posterior medians shifted upward by the same amount. Our constraints are consistent with previous measurements by Calverley et al. (2011),

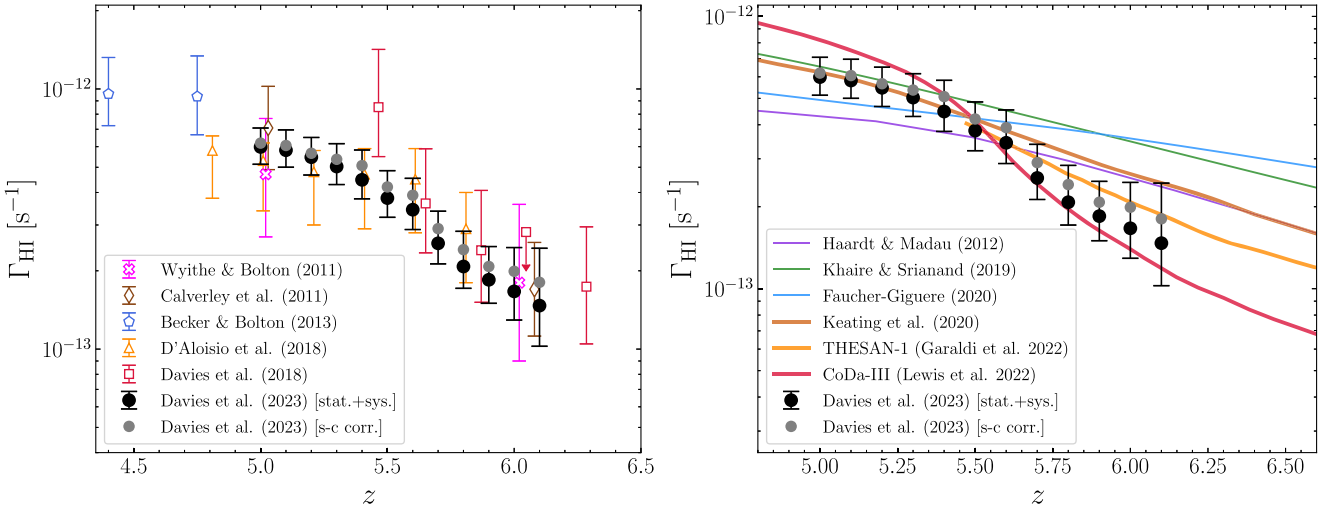

Figure 6. Left: Comparison of our $\Gamma_{\rm HI}$ constraints (black points with error bars, including statistical and systematic uncertainty) to previous measurements from the literature, Calverley et al. (2011; brown diamonds; quasar proximity zone profiles), Wyithe & Bolton (2011), Becker & Bolton (2013), D'Aloisio et al. (2018; pink crosses, blue pentagons, and orange triangles; mean Lyα forest transmission), and Davies et al. (2018d; red squares; Lyα and Lyβ transmission spikes). No corrections have been made for differences in cosmology or assumptions of the IGM thermal state between these works. The gray points show our constraints with an approximate correction for a bias due to the lack of self-consistency. Right: comparison to theoretical models of the ionizing background (curves), computed from 1D cosmological radiative transfer calculations by Haardt & Madau (2012; purple), Khaire & Srianand (2019; green), and Faucher-Giguère (2020; blue) and 3D radiation-hydrodynamic cosmological simulations by Keating et al. (2020; brown), Garaldi et al. (2022; orange; THESAN), and Lewis et al. (2022; red; CoDa-III).

Wyithe & Bolton (2011), and D'Aloisio et al. (2018), as well as constraints from the pilot study of Davies et al. (2018d), although we note that we have not carefully accounted for the different IGM thermal states (or ranges of thermal states) assumed by those works. In the right panel of Figure 6, we compare to predictions of $\Gamma_{\rm HI}(z)$ from empirically motivated 1D cosmological radiative transfer models by Haardt & Madau (2012), Khaire & Srianand (2019), and Faucher-Giguère (2020), and the 3D coupled radiation-hydrodynamics simulations from Keating et al. (2020), THESAN (Garaldi et al. 2022), and CoDa-III (Lewis et al. 2022). While the 1D radiative transfer models are all in rough agreement with the average $\Gamma_{\rm HI}$ from $z\sim5$–5.5, they fail to reproduce the steep downturn toward $z\sim6$. In contrast, the $\Gamma_{\rm HI}(z)$ in the 3D simulations rises more rapidly, and is more consistent with the trend of our constraints. In particular, the evolution from the CoDa-III simulation closely reproduces the rapid rise from $z\sim6$ to $z\sim5.4$, although it overshoots to a more highly ionized IGM at $z\sim5$.

In Figure 7, we show the derived constraints on $\lambda_{\rm mfp}$. At $z=5.0$–5.2, and $z=6.0$–6.1, the posteriors show no clear peak interior to the boundaries of the prior. We define 95% lower and upper limits, respectively, by the $\lambda_{\rm mfp}$ at which the posterior first falls a factor e^{-2} from its peak value. At $z=5.3$, and $z=5.9$, the posterior PDFs are clearly peaked, but the posterior PDF does not quite fall below a factor of e^{-2} from its peak value at the upper and lower edges of the prior boundary, respectively. We show these strongly prior-influenced constraints as open points with error bars.

The open squares in Figure 7 show the posterior medians for the (random) ionizing background fluctuations drawn from the 100 Mpc$/h$ box. At redshifts where the mean free path is constrained to be quite small, $\lambda_{\rm mfp}\lesssim30$ Mpc, this model prefers a shorter mean free path. This is due to the general decrease in the amplitude of background fluctuations in the smaller ionizing background simulation at fixed $\lambda_{\rm mfp}$. Quantitatively, we find that the width of the central 68% distribution of fluctuations (in $\log\tilde{\Gamma}$) is roughly a factor of 2 narrower in the 100 Mpc$/h$ box than the 512 Mpc box for $\lambda_{\rm mfp}=20$ Mpc. For

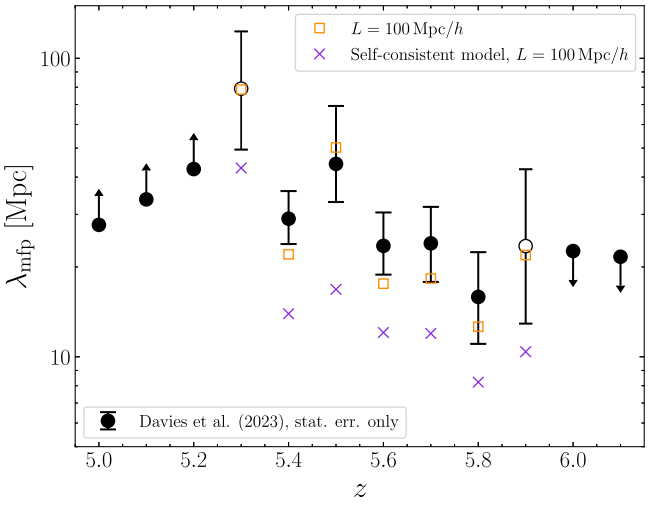

Figure 7. Posterior medians (black circles) and central 68% credible intervals (black thin error bars) on $\lambda_{\rm mfp}$ from the XQR-30 data set, with open circles corresponding to marginal constraints and arrows corresponding to 2σ limits (see text for details). Open orange points show the posterior medians from the 100 Mpc$/h$ simulation with uncorrelated density and ionizing background, while the purple crosses show the posterior medians from the fully self-consistent simulation.

larger mean free paths, the situation can be reversed—for $\lambda_{\rm mfp}\gtrsim50$ Mpc, the larger-scale background fluctuations in the smaller volume exceed those in the larger one due to the truncation of the periodic boundary conditions in the Davies & Furlanetto (2016) method.

The crosses in Figure 7 show the posterior medians for the self-consistent model. In this case, we see a stark factor of $\sim$2 decrease in the preferred $\lambda_{\rm mfp}$ at all redshifts. This is due to the ionizing background fluctuations now having to overcome the large-scale fluctuations in the density field in order to increase the scatter in the Lyα forest opacity. That is, the regions with a stronger ionizing background will preferentially lie in dense environments that will have a higher baseline Lyα opacity, while the regions with weak ionizing background will

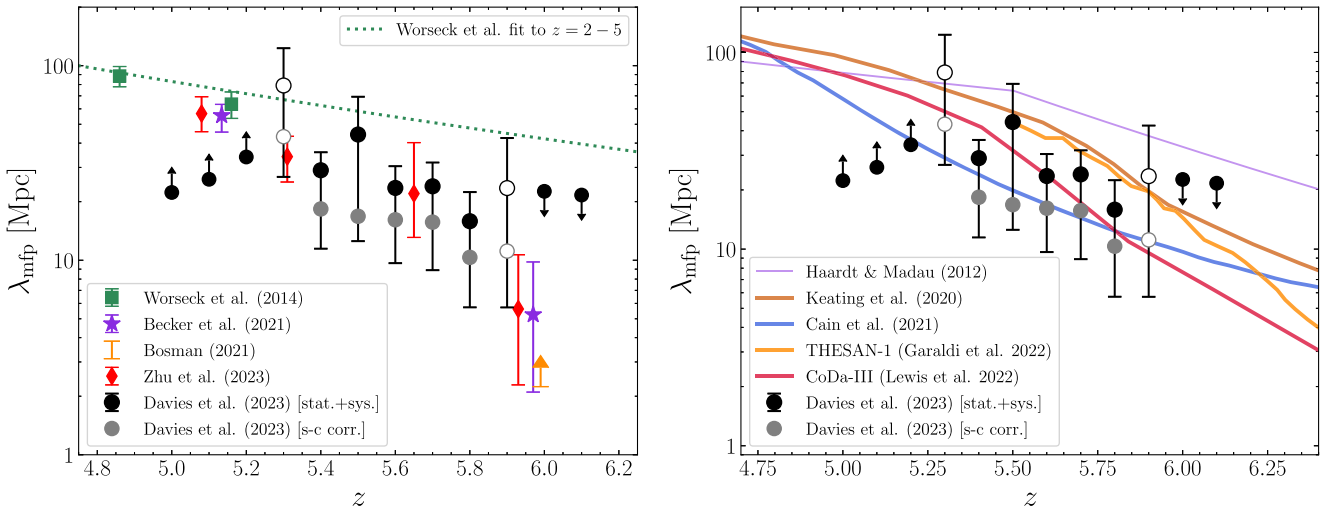

Figure 8. Left: Our constraints on $\lambda_{\rm mfp}$ compared to observations are shown by the black points, while the gray points show the constraints shifted to approximately correct for a bias due to the lack of self-consistency between the density field and ionizing background. Measurements from Worseck et al. (2014; green squares), Becker et al. (2021; purple pentagons), and Zhu et al. (2023; red diamonds), and the lower limit from Bosman (2021; orange arrow) are shown for comparison. The dotted line shows the extrapolated fit to the full redshift range of measurements from Worseck et al. (2014). Right: comparison to the mean free path in theoretical models, either prescribed by extrapolation of empirical constraints on the column density distribution of hydrogen absorbers (Haardt & Madau 2012, purple) or directly measured in the simulations by Keating et al. (2020; brown), Cain et al. (2021; blue), Garaldi et al. (2022; orange; THESAN), and Lewis et al. (2022; red; CoDa-III).

correspond to voids. This then increases the requirements on the fluctuations in the radiation field relative to the uncorrelated case. However, the difference in $\lambda_{\rm mfp}$ is likely exaggerated by the relatively small volume of the Nyx simulation, 100 Mpc$/h$, compared to the fiducial 512 Mpc model, which lacks fluctuations in the source field on the largest coherence scales of the radiation field ($\gtrsim$100 Mpc; see Figure 2), thus maintaining a stronger correlation between the background and the gas density on the smaller $dz = 0.1 \sim 50$ Mpc scale of the $\tau_{\rm eff}$ measurements. As we cannot assess the full strength of this effect, and how a much larger (and thus computationally very expensive) self-consistent model might mitigate this offset, we conservatively opt to extend the lower envelope of the uncertainty on $\lambda_{\rm mfp}$ by the magnitude of the measured offset between the self-consistent and fiducial method constraints from the 100 Mpc$/h$ box. We also show the posterior medians shifted by this offset as gray points. At $z = 5.5$, where the fiducial analysis using the 100 Mpc$/h$ box results in a larger value for $\lambda_{\rm mfp}$ due to the additional effect of the truncated boundary conditions mentioned above, we instead adopt the difference between the self-consistent model and the 512 Mpc model.

In the left panel of Figure 8, we compare our constraints to the measurements from Worseck et al. (2014), Becker et al. (2021), and Zhu et al. (2023), as well as the lower limit from Bosman (2021). The single power-law evolution with redshift that Worseck et al. (2014) find to be a good fit at $z = 2$–5 is disfavored at $z \gtrsim 5.4$, with our constraints indicating a more rapid decline to higher redshift. Intriguingly, our constraints are in good agreement with the $z = 6$–5 trend found by Becker et al. (2021), Zhu et al. (2023), effectively bridging the gaps between the more direct quasar-stacking measurements. In the right panel of Figure 8, we compare our constraints to theoretical predictions for the evolution of $\lambda_{\rm mfp}$. The empirical absorber model from Haardt & Madau (2012), also adopted in a similar form by later ionizing background calculations (e.g., Puchwein et al. 2019; Faucher-Giguère 2020), lies well above our constraints at $z > 5.3$, suggesting that the distribution of H I absorbers must evolve more rapidly at high redshift. Meanwhile, the subgrid opacity model of the Cain et al. (2021) simulations as well as the radiation-hydrodynamic simulations from Keating et al. (2020), THESAN (Garaldi et al. 2022), and CoDa-III (Lewis et al. 2022) are in much better agreement, with perhaps a modest overprediction of $\lambda_{\rm mfp}$ in the Keating et al. (2020) simulations and THESAN relative to our constraints.

Our constraints on $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$, including the statistical-only and total uncertainties, are summarized in Table 1.

4.1. Consistency between Data and Model

As in most other Bayesian parameter inference methods, our ability to place constraints on model parameters does not require that the model actually provides a good description of the data. Here, we investigate the degree to which the data are consistent with being a draw from our model.

In Figure 9, we show the CDFs of the observed Lyα forest data from Bosman et al. (2022) as blue and red solid lines, corresponding to the "optimistic" and "pessimistic" definitions from Bosman et al. (2018), similar to Figure 3. In the former case, Lyα mean fluxes below twice the (statistical) noise σ_F are set to $\tau_{\rm eff} = -\ln(2 \times \sigma_F)$, while, in the latter case, they are assumed to have $\tau_{\rm eff} > 10$. The value of the red curves at the right-hand edge of the panels in Figure 9 thus corresponds to the fraction of Lyα forest sightlines with detected (i.e., $>2\sigma_F$) flux. The shaded regions correspond to the central 95% of the distribution of mock data sets drawn from the posterior mean values of $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$, where blue and red similarly correspond to the optimistic and pessimistic CDF definitions. We can see that, in the majority of cases, the red and blue curves fit neatly within the model distributions, with the only exceptions being a handful of the most highly opaque sightlines at $z = 5.6$, and $z = 5.8$. In the highest-redshift bins, $z = 6.0$, and $z = 6.1$, the blue curves lie close to the upper end of the model distributions—however, this is the expected behavior in the regime where most observations result in nondetections, as the

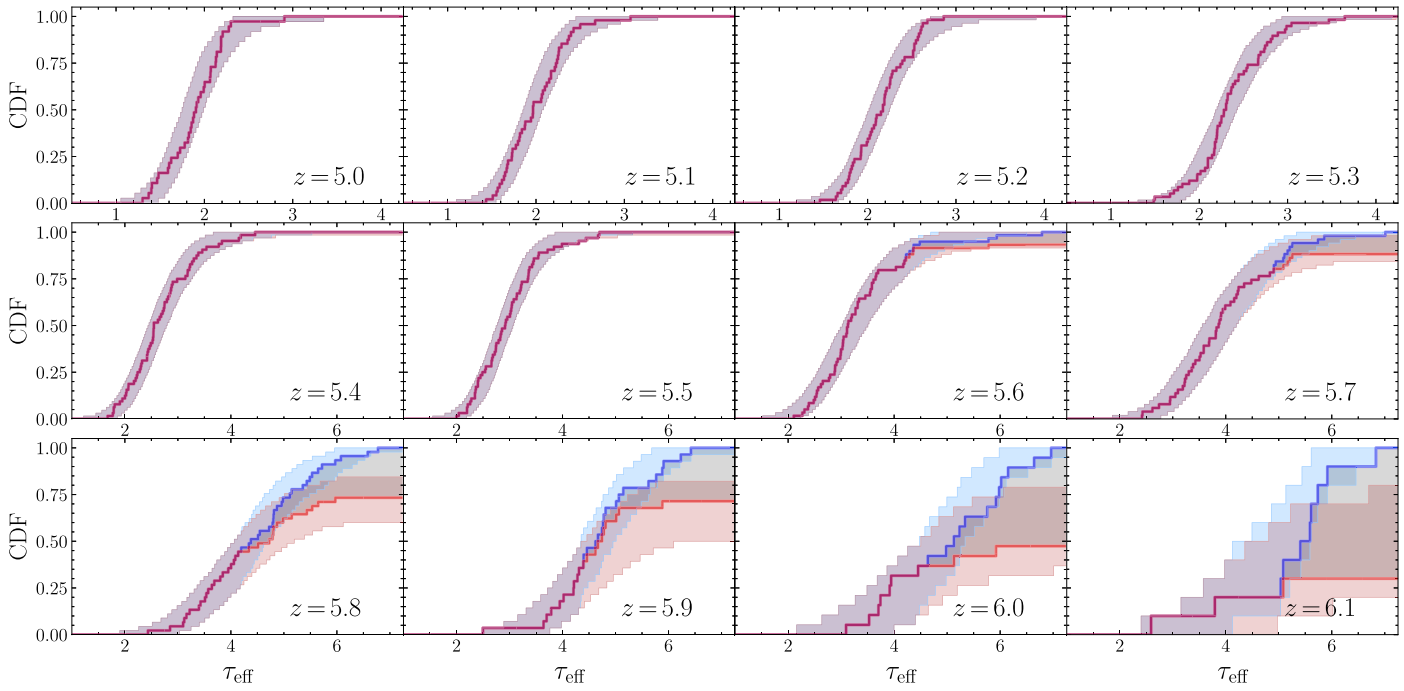

Figure 9. Comparison between the Lyα forest $\tau_{\rm eff}$ CDFs from Bosman et al. (2022; solid curves) and the central 95% of the distribution of mock CDFs from our simulations when adopting the posterior mean values of $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ (shaded regions). The red color corresponds to CDFs where Lyα fluxes detected at less than 2σ significance are assumed to be zero (i.e., $\tau_{\rm eff} = \infty$), while the blue color corresponds to those fluxes being set to twice the statistical noise (i.e., $\tau_{\rm eff} = -\ln(2 \times \sigma_N)$).

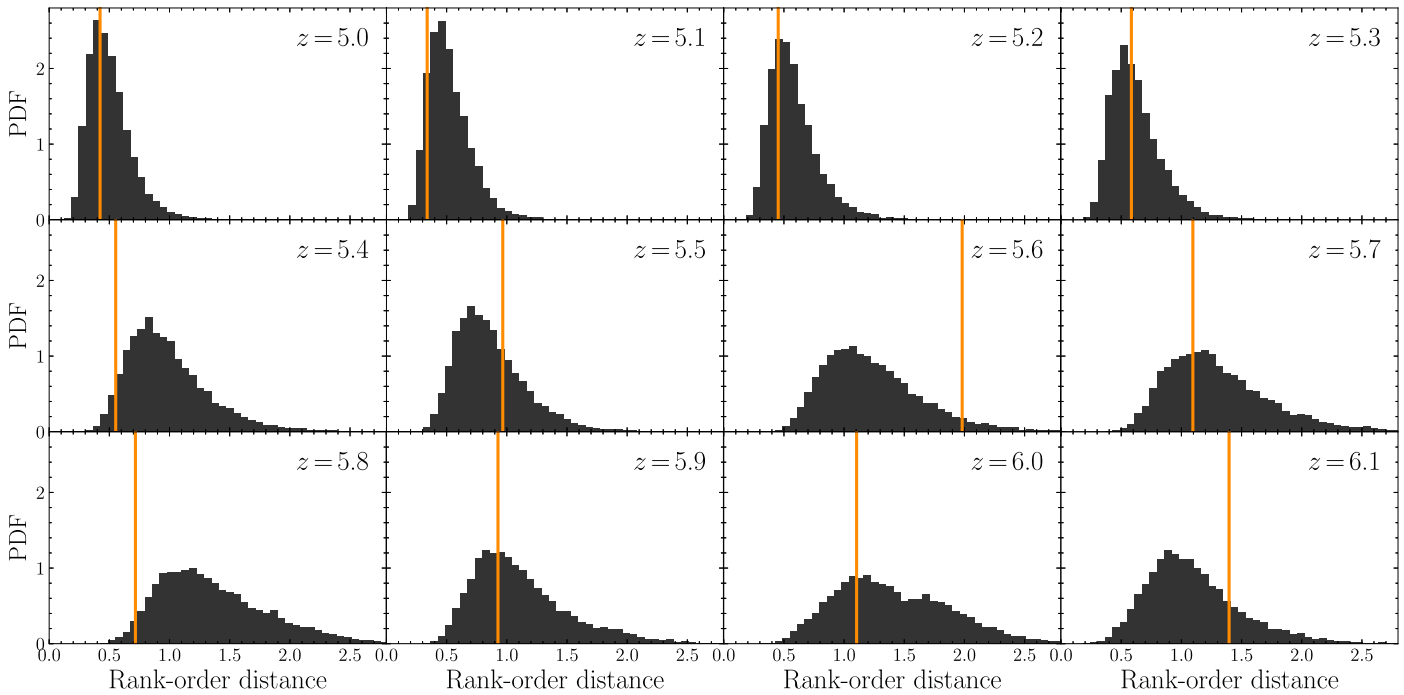

Figure 10. Distribution of distances (Equation (8)) between the average set of rank-order $\tau_{\rm eff}$ and individual sets from 10,000 mock observations, both adopting the posterior mean values of $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ at each redshift. The vertical orange lines show the distance between the average simulated rank-order $\tau_{\rm eff}$ and the observed Lyα forest data.

optimistic CDF has an upper bound in $\tau_{\rm eff}$ given by the distribution of noise values of the quasar spectra, i.e., the CDF of $-\ln(2 \times \sigma_F)$.

To further quantify the goodness-of-fit, we compare the ABC distance between the average simulated CDF at the posterior mean values of $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ and the observed data to the distribution of distances to 10,000 mock data sets generated from the same model. We show the distribution of the resulting consistency metric in Figure 10. In general, the distance to the data is well within the range of distances to mock data sets.

5. Caveats and Limitations of Our Analysis

As mentioned previously, our fiducial model for post-reionization ionizing background fluctuations suffers from important limitations. In this section, we reiterate and

summarize these limitations, and discuss their potential consequences for the interpretation of our results.

5.1. Uncorrelated Density and Ionization Rate Fields

In the fiducial model, by virtue of using separate volumes for the ionizing background calculation and the hydrodynamical simulation, the density and ionizing radiation fields of the mock spectra used for inference are decoupled. The two should generally be correlated on large scales (e.g., Mesinger & Furlanetto 2009). Because the Lyα forest opacity roughly scales as $\tau_{\rm Ly\alpha} \propto \Delta^2\Gamma_{\rm HI}^{-1}$ (e.g., Weinberg et al. 1997), the variance of $\tau_{\rm Ly\alpha}$ should behave as $\sigma_\tau^2 \sim \tau^2[\sigma_{\Delta^2}^2/\Delta^2 + \sigma_\Gamma^2/\Gamma - 2\sigma_{\Delta^2\Gamma}/\Delta^2\Gamma]$, i.e., the presence of correlated fluctuations in the density and ionization rate ($\sigma_{\Delta^2\Gamma}$) should suppress the strength of Lyα forest fluctuations at fixed $\lambda_{\rm mfp}$. Because the fluctuations become stronger with decreasing $\lambda_{\rm mfp}$, we expect our constraints to be biased high.

We tested this hypothesis by constructing a self-consistent model of ionizing background fluctuations within the smaller $100\,{\rm Mpc}/h$ volume of our hydrodynamical simulation (Section 2.2.1). As expected, in the regime of strong Lyα forest fluctuations, we recover much shorter mean free paths. A fraction of this difference is due to the suppression of the amplitude of fluctuations by the relatively small volume of the Nyx simulation, $100\,{\rm Mpc}/h$ on a side, compared to the fiducial ionizing background model, 512 Mpc on a side. Without a much larger self-consistent model, we are unable to entirely disentangle these two effects, so we instead opt to conservatively allow for a systematic error that encompasses the constraints derived from the self-consistent model, as seen by the large lower error bars in Figure 8.

5.2. Uncertainties in the IGM Thermal State

As previously mentioned in Section 2, our hydrodynamical simulation was run with an optically thin UV background, which reionized and heated the volume at very early times, $z_{\rm re} \sim 15$. We have chosen to adjust the temperatures in post-processing to better reflect current constraints on the timing and heat injection of reionization. In particular, we assume a fixed heat injection of $\Delta T = 20{,}000$ K, which is a representative temperature of the gas after the passage of the ionization front (Miralda-Escudé & Rees 1994; D'Aloisio et al. 2019), and tuned the range of $z_{\rm re}$ to be consistent with the IGM thermal state measured from the distribution of Lyα transmission spike widths by Gaikwad et al. (2020; see also Bolton et al. 2012). In principle, one could treat ΔT and $z_{\rm re}$ as additional parameters, introduce priors, and add the Gaikwad et al. (2020) measurements to the computation of the likelihood. Due to the computational demands of our likelihood-free inference method, and because, in this work, we are not attempting to constrain ΔT or $z_{\rm re}$, we opted to instead impose a plausible range rather than include them directly in the parameter inference.

We note, however, that the IGM thermal state at $z > 5$ is still substantially uncertain. The measurements using the 1D Lyα forest flux power spectrum by Walther et al. (2019), Boera et al. (2019) suggest somewhat lower temperatures with a steeper temperature–density relation at $z \sim 5$. If the IGM was reionized earlier, or if the heat injection was much lower, the upper envelope of the light gray region in Figure 5 shows that the $\Gamma_{\rm HI}$ required to reproduce the observed Lyα forest transmission could be substantially higher, with less evolution required from $z \sim 6$ to $z \sim 5$.

While our hydrodynamical simulation includes the effect of Jeans smoothing on the gas via its fiducial thermal history, by post hoc altering the IGM thermal state of the simulation, we neglect the differences in this smoothing that the different thermal histories would have otherwise imprinted (e.g., Gnedin & Hui 1998; Peeples et al. 2010a, 2010b; Kulkarni et al. 2015; Nasir et al. 2016; Oñorbe et al. 2017). The effect of Jeans smoothing on measurements of $\Gamma_{\rm HI}$ was thoroughly investigated over a wide range of thermal histories by Becker & Bolton (2013), who found that at $z \lesssim 5$ the effect was minor, contributing $\lesssim$0.03 dex to the error budget. We adopt a somewhat higher 0.03 dex systematic uncertainty to account for the trend of larger error at higher redshift, and add this additional error (in quadrature) to our fiducial constraints.

5.3. Lack of Post-reionization Temperature Fluctuations

The thermal state of our hydrodynamical simulation assumes a homogeneous heat injection, but the reionization process is patchy with different reionization timing in different locations (Furlanetto et al. 2004). This should lead to a highly inhomogeneous IGM thermal state immediately after reionization is complete, which persists to much later times (e.g., Trac et al. 2008; Keating et al. 2018; D'Aloisio et al. 2019; Oñorbe et al. 2019; Wu et al. 2019). The thermal state in our simulation is thus too uniform, and lacks any (anti)correlation with large-scale density. Due to the expected anticorrelation between large-scale density and post-reionization temperature fluctuations (D'Aloisio et al. 2015), our Lyα forest fluctuations are likely overestimated at fixed $\lambda_{\rm mfp}$ (Davies et al. 2018a; D'Aloisio et al. 2018). This would act to bias our $\lambda_{\rm mfp}$ measurements high; with a more realistic simulation, we would require stronger ionizing background fluctuations to counteract the effect of thermal state fluctuations and reproduce the large variations in the Lyα forest, and thus estimate a lower $\lambda_{\rm mfp}$.

5.4. Parameter Choices in the Fluctuating Ionizing Background Simulations

Our model for ionizing background fluctuations has several fixed parameters that we have not explored in detail. For example, we assume that $\lambda_{\rm mfp} \propto \Delta^{-1}\Gamma_{\rm HI}^{2/3}$, but both power-law indices are uncertain. Chardin et al. (2017) found a shallower density dependence $\lambda_{\rm mfp} \propto \Delta^{-0.4}$ when post-processing the Sherwood simulations (Bolton et al. 2017) with the self-shielding prescription from Rahmati et al. (2013). Incorporating this weaker density dependence would lead to stronger ionizing background fluctuations at fixed $\lambda_{\rm mfp}$. Our choice of the $\Gamma_{\rm HI}$ dependence is motivated by McQuinn et al. (2011), but both higher (up to $\sim$1) and lower values (down to $\sim 1/3$) are plausible (see the discussion in Becker et al. 2021), which would increase or decrease the $\lambda_{\rm mfp}$ required to match the observed Lyα forest fluctuations, respectively.

In addition, our model for the ionizing sources assumes that only halos more massive than $2 \times 10^9\ M_\odot$ produce ionizing photons, and that these halos can be assigned a UV luminosity via abundance matching, and further that the ionizing luminosity is proportional to the UV luminosity. All of these assumptions impact the effective bias of the ionizing emissivity field in the fluctuating ionizing background simulations. Other works have adopted different combinations of these

assumptions—for example, Kulkarni et al. (2019) prescribe ionizing luminosities proportional to halo mass for $M_h > 10^9\ M_\odot$, which results in more ionizing photons coming from lower mass halos and thus a lower bias of the emissivity field relative to our approach. The possible parameter space of source models is quite large, requiring far more efficient approaches to statistical inference than ours to constrain these additional parameters (e.g., Qin et al. 2021).

6. Discussion

With the caveats above in mind, the constraints presented in this work show a substantial improvement in the statistical precision of $\Gamma_{\rm HI}(z)$ at $z > 5$, and provide the first quantitative estimates of $\lambda_{\rm mfp}$ from the excess fluctuations in the Lyα forest alone. In this section, we discuss the implications of these measurements for our understanding of the $z > 5$ IGM.

6.1. Is Late Reionization Required by z ∼ 5–6 Lyα Forest Fluctuations?

Since the successful reproduction of the large-scale $z \sim 5$–6 Lyα forest variations by Kulkarni et al. (2019), it has commonly been understood that reionization is incomplete at least down to $z \sim 5.5$ (and more recently, $z \sim 5.3$, see, Zhu et al. 2021; Bosman et al. 2022; Zhu et al. 2022). However, here, we have demonstrated good agreement between the data and a model, which does not require incomplete reionization, which at first glance goes against this consensus. Crucially, due to the lack of full self-consistency in our model described in Section 2, we cannot make a strong claim that reionization must be complete at $z \lesssim 6$. With that said, previous claims of a complete incompatibility between the Lyα forest opacity distribution at $z \sim 5.5$–6 and a fully ionized IGM may not be entirely conclusive.

The IGM models employed by works in the literature that have suggested that incomplete reionization is required at $z \sim 5.5$–6 are not without their own limitations. The moment-based radiative transfer method used by Kulkarni et al. (2019), Keating et al. (2020) has been suggested to exhibit suppressed fluctuations in the radiation field at the end of reionization (Wu et al. 2021; see also Gaikwad et al. 2023); thus, they may require more large-scale neutral islands from incomplete reionization in order to achieve strong Lyα forest fluctuations. In addition, the spatial resolution of their radiative transfer models ($\sim$100 kpc) is too coarse to resolve self-shielding in Lyman limit systems (e.g., McQuinn et al. 2011), potentially resulting in an artificially elongated mean free path in ionized regions and further reducing the strength of fluctuations. The seminumerical method of Qin et al. (2021) similarly requires a substantial neutral fraction ($x_{\rm HI} \sim 0.15$) to reproduce the Lyα forest transmission statistics at $z \sim 5.8$ from Bosman et al. (2018), but they employed an approximate spatial filtering approach to compute ionizing background fluctuations. In contrast, the seminumerical method of Choudhury et al. (2021) employed a somewhat more sophisticated (but similar) treatment of the radiation field, and required roughly half as much neutral gas to explain the same Lyα forest data set. Similar to our findings, Zhu et al. (2021, 2022) found that the early reionization/short mean free path model from Nasir & D'Aloisio (2020) is consistent with the distribution of dark gaps in the Lyα and Lyβ forests.

Finally, we emphasize that our results do not in any way rule out the presence of a significant neutral hydrogen fraction in the $z < 6$ IGM. Rather, they show that imposing the existence of neutral islands is not explicitly required to match the particular summary statistic we are considering, namely the distribution of $\tau_{\rm eff}$ on $dz = 0.1$ scales. In reality, the $\lambda_{\rm mfp}$ we measure is only representative of the *actual* mean free path if the universe is fully ionized as assumed in our simulations. Our $\lambda_{\rm mfp}$ may thus be better interpreted as an effective parameter, reflecting both the existence of neutral islands in the deepest, large-scale voids, but also the strong fluctuations in the ionizing background due to a short, fluctuating mean free path of ionizing photons inside of ionized regions. Indeed, in our shortest mean free path models, large-scale regions with the weakest ionizing background are consistent with the gas remaining mostly neutral. However, compared to Gaikwad et al. (2023) who see a similar effect, we estimate lower volume-averaged neutral fractions of $\sim$0.02%–7% at $z = 5.8$ ($\sim$0.02%–8% at $z = 5.9$), with a large systematic uncertainty between our fiducial and self-consistent models. We will explore the nature of these regions more closely in future work.

6.2. Constraints on the Ionizing Emissivity

While the mean free path we measure is highly uncertain, and may be more of an effective parameter as described above, we can nevertheless cautiously explore an interpretation of our $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ constraints in terms of the ionizing emissivity at the hydrogen-ionizing edge, $\epsilon_{912} = \epsilon_\nu(\lambda = 912\ \text{Å})$. Assuming the local source approximation (Meiksin & White 2003), the relationship between $\Gamma_{\rm HI}$, λ, and ϵ can be expressed as

$$\Gamma_{\rm HI} = 4\pi \int_{\nu_{\rm HI}}^{\infty} \frac{J_\nu}{h\nu} \sigma_{\rm HI}(\nu) d\nu \approx \int_{\nu_{\rm HI}}^{\infty} \frac{\epsilon_\nu \lambda_\nu}{h\nu} \sigma_{\rm HI}(\nu) d\nu. \quad (9)$$

In practice, at $z > 5$, the integral over frequency effectively extends only to $4 \times \nu_{\rm HI}$ due to the onset of strong He II absorption. Approximating the ionizing emissivity as a power law $\epsilon_\nu = \epsilon_{912}(\nu/\nu_{\rm HI})^{-\alpha}$, and recalling that we also treat the mean free path as a power law with frequency (Equation (4)), we numerically solve this expression for ϵ_{912} at each redshift using our measured $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ values. We assume a fiducial emissivity spectral index of $\alpha = 2$, following Becker & Bolton (2013).

Our ionizing emissivity estimates are subject to the statistical and systematic uncertainties in the $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ measurements described above, as well as an additional systematic uncertainty in the spectral index α assumed for the ionizing sources. Following Becker & Bolton (2013), we adopt a range of $\alpha = 1.0$–3.0. For our fiducial uncertainty estimates shown in Figure 11, we propagate the statistical uncertainties assuming $\epsilon_{912} \propto \Gamma_{\rm HI}/\lambda_{\rm mfp}$, and combine them in quadrature with the systematic uncertainties on $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ as well as the systematic uncertainty from the spectral index variations $\alpha = 1.0$–3.0. For simplicity, here, we treat the self-consistent model corrections for $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ by inflating the systematic error terms in the upper and lower directions, respectively.

The resulting ionizing emissivities and their corresponding uncertainties are shown in Figure 11, compared to literature constraints from Becker & Bolton (2013), D'Aloisio et al. (2018), and Becker et al. (2021). Note that ϵ_{912} is related to the number of

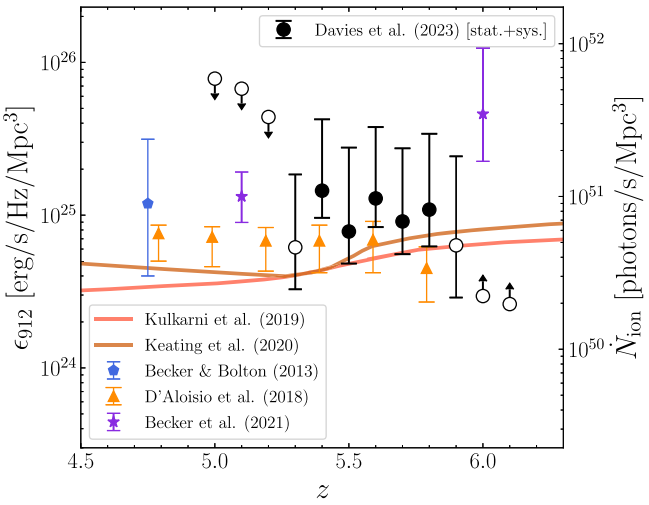

Figure 11. Implied ionizing emissivity ϵ_{912} evolution from our $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ constraints (black circles) compared to literature values from Becker & Bolton (2013; blue pentagon), D'Aloisio et al. (2018; orange triangles), and Becker et al. (2021; purple stars). The right axis shows the corresponding photon emissivity $\dot{N}_{\rm ion}$ assuming an ionizing source spectral index $\alpha = 2$.

ionizing photons emitted per comoving volume $\dot{N}_{\rm ion}$ by $\dot{N}_{\rm ion} = h\alpha\epsilon_{912}$ (e.g., Becker & Bolton 2013), shown on the right axis assuming $\alpha = 2$. We find that the ionizing emissivity is consistent with a roughly constant $\epsilon_{912} \sim 10^{25}$ erg s^{-1} Hz^{-1} Mpc^{-3} from $z \sim 6$–5, consistent with previous work.

6.3. Comparison to XQR-30 Analysis by Gaikwad et al. (2023)

In a recent complementary work, Gaikwad et al. (2023) have also constrained the photoionization rate and ionizing photon mean free path from the same XQR-30 Lyα forest data that we employ here. Here, we compare our results and discuss the differences in analysis methodology.

Gaikwad et al. (2023) use a self-consistent model (in our terminology, see Section 2.2.1) consisting of smoothed particle hydrodynamics simulations with 2048^3 baryon and dark matter particles in a volume 160 Mpc$/h$ on a side. They compute ionizing background fluctuations using an independently developed version of the Davies & Furlanetto (2016) method, which is heavily optimized by using a tree decomposition of the emissivity field, allowing them to evaluate the radiation field at a much higher resolution (512^3, $\sim$0.5 Mpc). The statistical comparison between their dense model grid and the XQR-30 Lyα forest data is performed using the Anderson–Darling test on the $\tau_{\rm eff}$ CDF, where they set a p-value threshold to map out a 1σ contour around the best-fit model.

The largest difference between the two works is primarily a philosophical one—Gaikwad et al. (2023) perform inference on a different definition of the mean free path $\lambda_{\rm mfp}$ than adopted here. Specifically, while we treat the mean free path as a fully subgrid quantity that arises from unresolved gas clumping, Gaikwad et al. (2023) measure $\lambda_{\rm mfp}$ from the density field of their simulation after applying the fluctuating $\Gamma_{\rm HI}$ field. That is, for a large number of skewers, they evaluate the density of neutral hydrogen $n_{\rm HI}$ given the local gas density and $\Gamma_{\rm HI}$, and then calculate the ionizing opacity at the Lyman limit $\tau_{\rm HI}$ by integrating the contribution from each resolution element,

$$\tau_{\rm HI}(x) = \int_0^x n_{\rm HI}\sigma_{\rm HI}(\nu_{\rm HI})dl. \quad (10)$$

They then compute the corresponding mean free path by fitting an exponential profile to the average transmission ($F(x) = \exp[-\tau_{\rm HI}(x)] = \exp[-x/\lambda_{\rm mfp}]$) of the simulated skewers. This procedure imposes a physical prior on the possible combinations of $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ via the gas density distribution of their hydrodynamical simulation. In addition, this allows Gaikwad et al. (2023) to obtain tight constraints on $\lambda_{\rm mfp}$ even without an excess in Lyα forest fluctuations, as it can be obtained from the density field even if the ionizing background is entirely uniform. In contrast, our method infers the mean free path solely from the excess in Lyα forest fluctuations over the uniform case, i.e., from the character of the radiation field fluctuations alone.

In Figure 12, we compare our constraints on $\Gamma_{\rm HI}$ (left) and $\lambda_{\rm mfp}$ (right) to Gaikwad et al. (2023). In general, the two agree very well, suggesting that, despite our rather different statistical and modeling methodologies, the rapid evolution in both $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ is robustly indicated by the XQR-30 Lyα forest data. In detail, our method recovers a smaller uncertainty on $\Gamma_{\rm HI}$, primarily due to the broader range of IGM thermal parameters marginalized over by Gaikwad et al. (2023), while they achieve tighter constraints on $\lambda_{\rm mfp}$ more uniformly across the full redshift range of their study, subject to the difference in $\lambda_{\rm mfp}$ definition described above.

7. Conclusion

In this work, we have constrained the evolution of the photoionization rate $\Gamma_{\rm HI}$ and the mean free path of ionizing photons $\lambda_{\rm mfp}$ in the IGM at $z = 5.0$–6.1 using the extended XQR-30 (D'Odorico et al. 2023) Lyα forest data set from Bosman et al. (2022). We assume that the excess fluctuations in the $z > 5.3$ Lyα forest are due to a strong coupling between the ionizing background and mean free path (Davies & Furlanetto 2016), and constrain parameters using the likelihood-free inference technique of ABC.

We recover a smooth evolutionary trend in $\Gamma_{\rm HI}$, which increases by a factor of $\sim$4 from $z = 6$, to $z = 5$; in agreement with past observations but with a finer redshift sampling and smaller statistical uncertainty. The increase is much stronger than predicted by empirically motivated 1D cosmological radiative transfer models, but is largely consistent with the evolution found by state-of-the-art 3D radiation-hydrodynamic cosmological simulations. We similarly find consistency with recent measurements of the mean free path $\lambda_{\rm mfp}$ from stacked quasar spectra and recent radiation-hydrodynamic models.

We find that the statistical constraining power of the coarsely binned $\tau_{\rm eff}$ measurements is very strong, to the extent that for both $\Gamma_{\rm HI}$ and $\lambda_{\rm mfp}$ we are strongly limited by systematic uncertainties. For the former, we are limited by our knowledge of the IGM thermal state, while for the latter (and the former, to a lesser extent) we are limited by the computational resources required to simulate a converged Lyα forest in a large enough volume to capture the full intensity of ionizing background fluctuations. Additionally, the relic thermal fluctuations left by reionization at earlier times (e.g., D'Aloisio et al. 2015) will also confound efforts to precisely constrain the parameters of the $z = 5$–6 IGM without a complete model of the entire reionization process. Reducing these systematic uncertainties will require extracting information from both the small-scale (e.g., Gaikwad et al. 2020) and large-scale properties of the Lyα forest, and then performing statistical inference on full

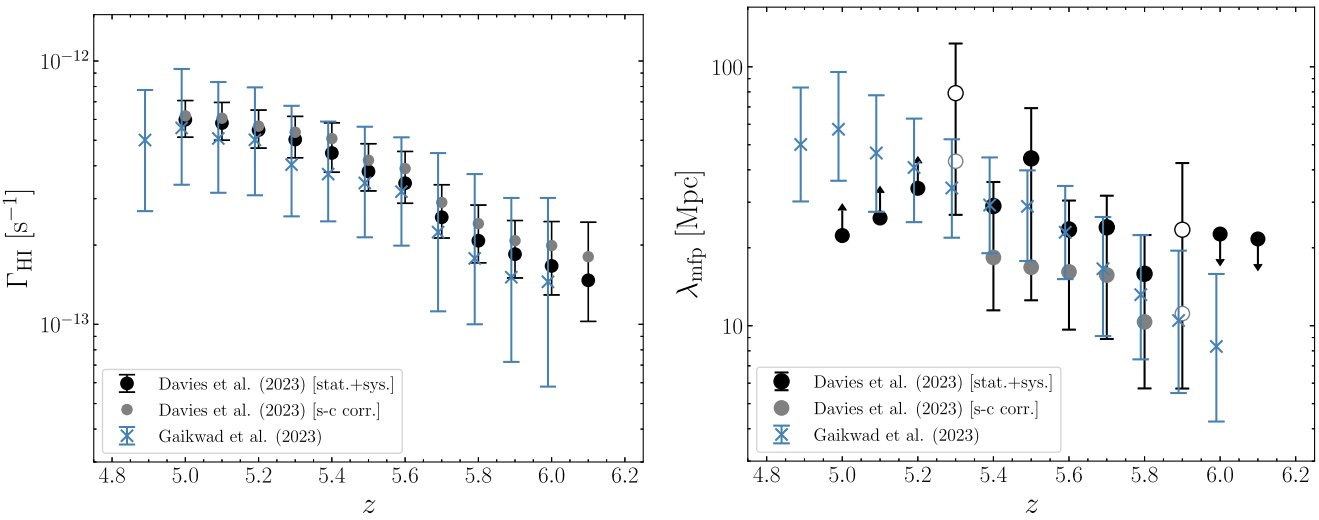

Figure 12. Left: Our constraints on $\Gamma_{\rm HI}$ (black and gray circles) compared to Gaikwad et al. (2023; blue crosses), where we have shifted the latter slightly in redshift for clarity. The mean trend agrees very well, but the wider assumed range in IGM thermal state parameters leads to larger error bars in Gaikwad et al. (2023). Right: Comparison to the mean free path constraints from Gaikwad et al. (2023). Despite the substantial philosophical differences in the definition of $\lambda_{\rm mfp}$ (see text for details), our constraints agree quite well.

reionization lightcones with a detailed model for the sources and sinks (e.g., Qin et al. 2021).

At the low-redshift end of our data set, $z \sim 5$, our ability to constrain $\lambda_{\rm mfp}$ is limited by the insensitivity of the Lyα forest to long mean free paths, as the resulting fluctuations in the radiation field become weak compared to the intrinsic fluctuations already imprinted by the IGM density field. However, at these redshifts, the direct stacking of quasar spectra appears to be a far more sensitive probe with fewer underlying assumptions and systematic uncertainties (Worseck et al. 2014; Becker et al. 2021; Zhu et al. 2023). At the high-redshift end, $z \sim 6$, we are instead limited by the relative sparsity of sightlines and the poor spatial resolution of our ionizing background simulations. The latter can already be rectified by adopting more efficient methods of computing the radiation field (e.g., Gaikwad et al. 2023), while the former will require additional deep spectroscopy of yet higher-redshift quasars than the XQR-30 sample (e.g., Yang et al. 2020b).

Acknowledgments

S.E.I.B. is funded by the Deutsche Forschungsgemeinschaft (DFG) under Emmy Noether grant No. BO 5771/1-1. G.K. is partly supported by the Department of Atomic Energy (Government of India) research project with project identification No. RTI 4002, and by the Max Planck Society through a Max Planck Partner Group. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any author accepted manuscript version arising from this submission.

Appendix
Correcting for Coarsely Spaced Snapshot Redshifts

While the data we consider arise from a range of redshifts from $z = 5.0$, to $z = 6.1$, in steps of $dz = 0.1$, we only have outputs from our hydrodynamical simulation in steps of $dz = 0.5$. There is thus the potential for biased or incorrect inference of model parameters at redshifts between snapshots.

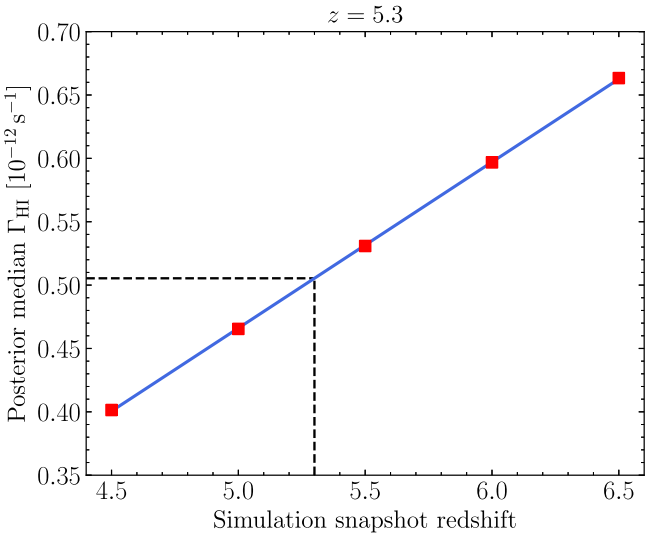

Figure 13. Variation of the posterior median $\Gamma_{\rm HI}$ at $z = 5.3$ as a function of hydrodynamical simulation snapshot redshift (red points). The blue curve shows a linear fit to $\Gamma_{\rm HI}$ as a function of snapshot redshift. In this example, the posterior median $\Gamma_{\rm HI}$ at $z = 5.3$ is adjusted from 0.532×10^{-12} s^{-1} to 0.505×10^{-12} s^{-1}.

We correct our posterior constraints by performing the full inference procedure described in the text but with five fixed simulation snapshots with redshifts $z_{\rm snap} = 4.5$, 5.0, 5.5, 6.0, and 6.5. We found that different snapshots resulted in substantially shifted constraints on $\Gamma_{\rm HI}$, with only a minor shift in $\lambda_{\rm mfp}$. For $\Gamma_{\rm HI}$, we found that the shift of the posterior percentiles with snapshot redshift was extremely well-described by a linear fit in $\Gamma_{\rm HI}(z_{\rm snap})$, as shown for the posterior median $\Gamma_{\rm HI}$ at $z = 5.3$ in Figure 13. For our fiducial constraints presented in Figure 6 and in Table 1, we have shifted the reported posterior percentiles by evaluating the corresponding polynomial fit at the true redshift of the Lyα forest data. These corrections amount to a maximum of $\sim$6%, and as seen by the very tight relation in Figure 13, they contribute a negligible additional source of uncertainty.

ORCID iDs

Frederick B. Davies https://orcid.org/0000-0003-0821-3644
Sarah E. I. Bosman https://orcid.org/0000-0001-8582-7012
Prakash Gaikwad https://orcid.org/0000-0002-2423-7905
Fahad Nasir https://orcid.org/0000-0003-0294-8674
Joseph F. Hennawi https://orcid.org/0000-0002-7054-4332
George D. Becker https://orcid.org/0000-0003-2344-263X
Martin G. Haehnelt https://orcid.org/0000-0001-8443-2393
Valentina D'Odorico https://orcid.org/0000-0003-3693-3091
Manuela Bischetti https://orcid.org/0000-0002-4314-021X
Anna-Christina Eilers https://orcid.org/0000-0003-2895-6218
Laura C. Keating https://orcid.org/0000-0001-5211-1958
Girish Kulkarni https://orcid.org/0000-0001-5829-4716
Samuel Lai https://orcid.org/0000-0001-9372-4611
Chiara Mazzucchelli https://orcid.org/0000-0002-5941-5214
Yuxiang Qin https://orcid.org/0000-0002-4314-1810
Sindhu Satyavolu https://orcid.org/0000-0001-5818-6838
Feige Wang https://orcid.org/0000-0002-7633-431X
Jinyi Yang https://orcid.org/0000-0001-5287-4242
Yongda Zhu https://orcid.org/0000-0003-3307-7525

References

Abel, T., & Haehnelt, M. G. 1999, ApJL, 520, L13
Almgren, A. S., Bell, J. B., Lijewski, M. J., Lukić, Z., & Van Andel, E. 2013, ApJ, 765, 39
Alsing, J., Wandelt, B., & Feeney, S. 2018, MNRAS, 477, 2874
Becker, G. D., & Bolton, J. S. 2013, MNRAS, 436, 1023
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, R. H., Fan, X., White, R. L., et al. 2001, AJ, 122, 2850
Blum, M. G. B. 2010, J. Am. Stat. Assoc., 105, 1178
Boera, E., Becker, G. D., Bolton, J. S., & Nasir, F. 2019, ApJ, 872, 101
Bolton, J. S., Becker, G. D., Raskutti, S., et al. 2012, MNRAS, 419, 2880
Bolton, J. S., & Haehnelt, M. G. 2007, MNRAS, 382, 325
Bolton, J. S., Haehnelt, M. G., Viel, M., & Springel, V. 2005, MNRAS, 357, 1178
Bolton, J. S., Puchwein, E., Sijacki, D., et al. 2017, MNRAS, 464, 897
Bosman, S. E. I. 2021, arXiv:2108.12446
Bosman, S. E. I., Davies, F. B., Becker, G. D., et al. 2022, MNRAS, 514, 55
Bosman, S. E. I., Fan, X., Jiang, L., et al. 2018, MNRAS, 479, 1055
Bosman, S. E. I., Ďurovčíková, D., Davies, F. B., & Eilers, A. C. 2021, MNRAS, 503, 2077
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015a, ApJ, 811, 140
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015b, ApJ, 803, 34
Cain, C., D'Aloisio, A., Gangolli, N., & Becker, G. D. 2021, ApJL, 917, L37
Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S. 2011, MNRAS, 412, 2543
Chardin, J., Haehnelt, M. G., Aubert, D., & Puchwein, E. 2015, MNRAS, 453, 2943
Chardin, J., Puchwein, E., & Haehnelt, M. G. 2017, MNRAS, 465, 3429
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
Cole, A., Miller, B. K., Witte, S. J., et al. 2022, JCAP, 2022, 004
D'Aloisio, A., McQuinn, M., Davies, F. B., & Furlanetto, S. R. 2018, MNRAS, 473, 560
D'Aloisio, A., McQuinn, M., Maupin, O., et al. 2019, ApJ, 874, 154
D'Aloisio, A., McQuinn, M., Trac, H., Cain, C., & Mesinger, A. 2020, ApJ, 898, 149
D'Aloisio, A., McQuinn, M., & Trac, H. 2015, ApJL, 813, L38
D'Aloisio, A., Upton Sanderbeck, P. R., McQuinn, M., Trac, H., & Shapiro, P. R. 2017, MNRAS, 468, 4691
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., & Furlanetto, S. R. 2016, MNRAS, 460, 1328
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
Davies, F. B., Hennawi, J. F., Eilers, A. C., & Lukić, Z. 2018d, ApJ, 855, 106
Davies, R. L., Ryan-Weber, E., D'Odorico, V., et al. 2023, MNRAS, 521, 314
D'Odorico, V., Bañados, E., Becker, G. D., et al. 2023, MNRAS, 523, 1399
Eilers, A. C., Davies, F. B., & Hennawi, J. F. 2018, ApJ, 864, 53
Emberson, J. D., Thomas, R. M., & Alvarez, M. A. 2013, ApJ, 763, 146
Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117
Faucher-Giguère, C. A. 2020, MNRAS, 493, 1614
Fumagalli, M., O'Meara, J. M., Prochaska, J. X., & Worseck, G. 2013, ApJ, 775, 78
Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2004, ApJ, 613, 1
Gaikwad, P., Haehnelt, M. G., Davies, F. B., et al. 2023, MNRAS, 525, 4093
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
Garaldi, E., Kannan, R., Smith, A., et al. 2022, MNRAS, 512, 4909
Giallongo, E., Grazian, A., Fiore, F., et al. 2015, A&A, 578, A83
Gnedin, N. Y. 2000, ApJ, 542, 535
Gnedin, N. Y., & Hui, L. 1998, MNRAS, 296, 44
Greig, B., Mesinger, A., Davies, F. B., et al. 2022, MNRAS, 512, 5390
Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Harikane, Y., Zhang, Y., Nakajima, K., et al. 2023, ApJ, 959, 39
Hoag, A., Bradač, M., Huang, K., et al. 2019, ApJ, 878, 12
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., Puchwein, E., & Haehnelt, M. G. 2018, MNRAS, 477, 5501
Keating, L. C., Weinberger, L. H., Kulkarni, G., et al. 2020, MNRAS, 491, 1736
Khaire, V., & Srianand, R. 2019, MNRAS, 484, 4174
Kocevski, D. D., Onoue, M., Inayoshi, K., et al. 2023, ApJL, 954, L4
Kulkarni, G., Hennawi, J. F., Oñorbe, J., Rorai, A., & Springel, V. 2015, ApJ, 812, 30
Kulkarni, G., Keating, L. C., Haehnelt, M. G., et al. 2019, MNRAS, 485, L24
Labbe, I., Greene, J. E., Bezanson, R., et al. 2023, arXiv:2306.07320
Lewis, J. S. W., Ocvirk, P., Sorce, J. G., et al. 2022, MNRAS, 516, 3389
Lidz, A., Oh, S. P., & Furlanetto, S. R. 2006, ApJL, 639, L47
Lukić, Z., Heitmann, K., Habib, S., Bashinsky, S., & Ricker, P. M. 2007, ApJ, 671, 1160
Lukić, Z., Stark, C. W., Nugent, P., et al. 2015, MNRAS, 446, 3697
Maiolino, R., Scholtz, J., Curtis-Lake, E., et al. 2023, arXiv:2308.01230
Marin, J. M., Pudlo, P., Robert, C. P., & Ryder, R. J. 2012, Statistics and Computing, 22, 1167
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
Matsuoka, Y., Strauss, M. A., Kashikawa, N., et al. 2018, ApJ, 869, 150
Matthee, J., Naidu, R. P., Brammer, G., et al. 2024, ApJ, 963, 129
McQuinn, M., Oh, S. P., & Faucher-Giguère, C. A. 2011, ApJ, 743, 82
Meiksin, A. 2020, MNRAS, 491, 4884
Meiksin, A., & White, M. 2003, MNRAS, 342, 1205
Mesinger, A., & Furlanetto, S. 2007, ApJ, 669, 663
Mesinger, A., & Furlanetto, S. 2009, MNRAS, 400, 1461
Mesinger, A., Furlanetto, S., & Cen, R. 2011, MNRAS, 411, 955
Miralda-Escudé, J., & Rees, M. J. 1994, MNRAS, 266, 343
Muñoz, J. A., Oh, S. P., Davies, F. B., & Furlanetto, S. R. 2016, MNRAS, 455, 1385
Nasir, F., Bolton, J. S., & Becker, G. D. 2016, MNRAS, 463, 2335
Nasir, F., & D'Aloisio, A. 2020, MNRAS, 494, 3080
O'Leary, R. M., & McQuinn, M. 2012, ApJ, 760, 4
Oñorbe, J., Davies, F. B., Lukić, Z., et al. 2019, MNRAS, 486, 4075
Oñorbe, J., Hennawi, J. F., & Lukić, Z. 2017, ApJ, 837, 106
Park, H., Shapiro, P. R., Choi, J. h., et al. 2016, ApJ, 831, 86
Parsa, S., Dunlop, J. S., & McLure, R. J. 2018, MNRAS, 474, 2904
Peeples, M. S., Weinberg, D. H., Davé, R., Fardal, M. A., & Katz, N. 2010a, MNRAS, 404, 1281
Peeples, M. S., Weinberg, D. H., Davé, R., Fardal, M. A., & Katz, N. 2010b, MNRAS, 404, 1295
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., & Feldman, M. W. 1999, Mol. Biol. Evol., 16, 1791
Prochaska, J. X., Madau, P., O'Meara, J. M., & Fumagalli, M. 2014, MNRAS, 438, 476
Prochaska, J. X., Worseck, G., & O'Meara, J. M. 2009, ApJL, 705, L113

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
Rahmati, A., Pawlik, A. H., Raičević, M., & Schaye, J. 2013, MNRAS, 430, 2427
Rubin, D. B. 1984, AnSta, 12, 1151
Rudie, G. C., Steidel, C. C., Shapley, A. E., & Pettini, M. 2013, ApJ, 769, 146
Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, PASP, 114, 851
Tavaré, S., Balding, D. J., Griffiths, R. C., & Donnelly, P. 1997, Genetics, 145, 505
Tepper-García, T. 2006, MNRAS, 369, 2025
Trac, H., Cen, R., & Loeb, A. 2008, ApJL, 689, L81
Upton Sanderbeck, P. R., D'Aloisio, A., & McQuinn, M. J. 2016, MNRAS, 460, 1885
Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, ApJ, 465, 487
Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, A&A, 536, A105
Walther, M., Oñorbe, J., Hennawi, J. F., & Lukić, Z. 2019, ApJ, 872, 13
Wang, F., Davies, F. B., Yang, J., et al. 2020, ApJ, 896, 23
Weinberg, D. H., Hernsquit, L., Katz, N., Croft, R., & Miralda-Escudé, J. 1997, in Proc. of the 13th IAP Astrophysics Coll., Structure and Evolution of the Intergalactic Medium from QSO Absorption Line System, ed. P. Petitjean & S. Charlot (Paris: Editions Frontieres), 133
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
Worseck, G., Prochaska, J. X., O'Meara, J. M., et al. 2014, MNRAS, 445, 1745
Wu, X., McQuinn, M., & Eisenstein, D. 2021, JCAP, 2021, 042
Wu, X., McQuinn, M., Kannan, R., et al. 2019, MNRAS, 490, 3177
Wyithe, J. S. B., & Bolton, J. S. 2011, MNRAS, 412, 1926
Yang, J., Wang, F., Fan, X., et al. 2020a, ApJL, 897, L14
Yang, J., Wang, F., Fan, X., et al. 2020b, ApJ, 904, 26
Zel'dovich, Y. B. 1970, A&A, 5, 84
Zhu, Y., Becker, G. D., Bosman, S. E. I., et al. 2021, ApJ, 923, 223
Zhu, Y., Becker, G. D., Bosman, S. E. I., et al. 2022, ApJ, 932, 76
Zhu, Y., Becker, G. D., Christenson, H. M., et al. 2023, ApJ, 955, 115