Forecasting constraints on the high-z IGM thermal state from the Lyman-α forest flux autocorrelation function

Wolfson, M.; Hennawi, J.F.; Davies, F.B.; Lukić, Z.; Oñorbe, J.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **540,** 1412–1431 (2025) https://doi.org/10.1093/mnras/staf753
Advance Access publication 2025 May 8

Forecasting constraints on the high-z IGM thermal state from the Lyman-α forest flux autocorrelation function

Molly Wolfson [1]★ Joseph F. Hennawi [1,2] Frederick B. Davies [3] Zarija Lukić[4] and Jose Oñorbe [5]

[1]*Department of Physics, University of California, Santa Barbara, CA 93106, USA*
[2]*Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*
[3]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[4]*Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA*
[5]*Facultad de Física, Universidad de Sevilla, Multidisciplinary Unit for Energy Science, Av. Reina Mercedes s/n, E-41012 Sevilla, Spain*

Accepted 2025 April 24. Received 2025 March 27; in original form 2024 June 22

ABSTRACT

The autocorrelation function of the Lyman-α (Ly α) forest flux from high-z quasars probes the small-scale structure of the intergalactic medium (IGM). The thermal state of the IGM, determined by the physics of reionization, sets the small-scale power observed in the Ly α forest. To explore the sensitivity of the autocorrelation function to the IGM's thermal state, we compute the autocorrelation function from a cosmological hydrodynamical simulation with an instantaneous reionization model and 135 post-processed thermal states. Using mock data sets of 20 quasars, we forecast constraints on T_0 and γ, which characterize the post-processed IGM thermal state, at $5.4 \leq z \leq 6$. While this model simplifies the IGM's thermal state, it serves as a key first step in assessing future observational prospects. We also perform an inference test on mocks and re-weight out posterior distributions to guarantee that they exhibit statistically correct behaviour. At $z = 5.4$, we find that an idealized data set constrains T_0 to 59 per cent and γ to 16 per cent at the 1σ equivalent confidence level. To explore more realistic, non-instantaneous reionization scenarios, we analyse four models combining temperature and ultraviolet background (UVB) fluctuations at $z = 5.8$. We find that mock data generated from a model with both temperature and UVB fluctuations can rule out a model with only temperature fluctuations at the $> 1\sigma$ level 73.9 per cent of the time.

Key words: methods: statistical – intergalactic medium – quasars: absorption lines – dark ages, reionization, first stars.

1 INTRODUCTION

Understanding the epoch of reionization, the time period where the first luminous sources emitted photons that re-ionized the intergalactic medium (IGM), remains a major open problem for studies of the early Universe.The midpoint of reionization has been constrained as $z_{\rm re} = 7.7 \pm 0.7$ from the cosmic microwave background (Planck Collaboration VI 2020). Initial measurements of transmission in the Lyman-α (Lyα) forest (Gunn & Peterson 1965; Lynds 1971) of high redshift quasars suggested that reionization was complete by $z \sim 6$ (Fan et al. 2006; McGreer, Mesinger & Fan 2011; McGreer, Mesinger & D'Odorico 2015). Additional methods used to constrain reionization include observations of Lyα emission from high redshift galaxies (see e.g. Jung et al. 2020; Morales et al. 2021) and large Lyα absorption troughs (see e.g. Becker et al. 2018; Kashino et al. 2020). Measurements of the Lyα forest optical depths scatter on levels that suggest reionization is not actually complete until $z < 6$ (Fan et al. 2006; Becker et al. 2015; Bosman et al. 2018; Eilers, Davies & Hennawi 2018; Yang et al. 2020; Bosman et al. 2022).

An alternative, indirect method to constrain reionization is by looking at the thermal history of the IGM at $z > 5$ (Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2020). During reionization, ionization fronts propagate through the IGM and impulsively heat the reionized gas in the IGM to $\sim 10^4$ K (McQuinn 2012; Davies, Furlanetto & McQuinn 2016; D'Aloisio et al. 2019). The details of the driving sources, the timing, and duration of reionization will determine the precise amount of heat injected. After reionization, the IGM expands and cools through the adiabatic expansion of the Universe and inverse Compton scattering off CMB photons. The combination of these physical processes will allow the IGM gas to relax into a state described by a tight power-law relation between the temperature and density:

$$T = T_0 \Delta^{\gamma - 1}, \tag{1}$$

where $\Delta = \rho/\bar{\rho}$ is the overdensity, $\bar{\rho}$ is the mean density of the Universe, T_0 is the temperature at mean density, and γ is the slope of the relationship (Hui & Gnedin 1997; Puchwein et al. 2015; McQuinn & Upton Sanderbeck 2016). The low-density IGM has long cooling times, so the thermal memory of reionization will persist for hundreds of Myr. This means that thermal state of the IGM at the end and after reionization, $z \sim 5 - 6$, can provide key insights into reionization (Miralda-Escudé & Rees 1994; Hui & Gnedin 1997; Haehnelt & Steinmetz 1998; Theuns et al. 2002; Hui & Haiman 2003; Lidz & Malloy 2014; Oñorbe et al. 2017; Oñorbe, Hennawi & Lukić 2017).

★ E-mail: mawolfson@ucsb.edu

The Lyα optical depth, $\tau_{\rm Ly\alpha}$ is related to the temperature via

$$\tau_{\rm Ly\alpha} = n_{\rm HI}\sigma_{\rm Ly\alpha} \propto T^{-0.7}/\,\Gamma_{\rm UVB}, \quad (2)$$

see Rauch (1998). Several statistics have been used to measure the thermal state of the IGM from the Lyα forest, including the flux probability density (Becker, Rauch & Sargent 2007; Bolton et al. 2008; Viel, Bolton & Haehnelt 2009; Calura et al. 2012; Lee et al. 2015), the curvature (Becker et al. 2011; Boera et al. 2014; Gaikwad et al. 2020), the Doppler parameter distribution (Schaye et al. 1999; Bryan & Machacek 2000; Ricotti, Gnedin & Shull 2000; Schaye et al. 2000; McDonald et al. 2001; Bolton et al. 2010, 2012; Rudie, Steidel & Pettini 2012; Bolton et al. 2014; Rorai et al. 2018; Gaikwad et al. 2020), the joint distribution of the Doppler parameters with the Hydrogen Column Density (Hiss et al. 2018), and wavelets (Lidz et al. 2010; Garzilli et al. 2012; Gaikwad et al. 2020). One of the most commonly used statistics for measuring the structure of the Lyα forest is the 1D flux power spectrum, $P_F(k)$ (Theuns, Schaye & Haehnelt 2000; Zaldarriaga, Hui & Tegmark 2001; Walther et al. 2017; Yèche et al. 2017; Boera et al. 2019; Gaikwad et al. 2020; Wolfson et al. 2021).

The thermal state of the IGM significantly influences the Lyα forest, primarily through two mechanisms: Doppler broadening, which is driven by thermal motions, and Jeans (pressure) smoothing, which affects the distribution of the underlying baryons. To understand Jeans smoothing, it's crucial to consider the role of pressure forces. Pressure forces, influenced by the thermal state, erase gravitational fluctuations at a rate determined by the local sound speed. At low densities, like those of the IGM, this sound-crossing time is approximately the Hubble time. Thus, the Jeans (pressure) smoothing scale serves as a record of the thermal history of the IGM over extensive time-scales (Gnedin & Hui 1998; Kulkarni et al. 2015; Nasir, Bolton & Becker 2016; Oñorbe et al. 2017; Rorai et al. 2017). Both Doppler broadening and Jeans smoothing reduce the small-scale structure of the Lyα forest. These reductions in small-scale structure of the Lyα forest lead to a cut-off in $P_F(k)$ at high-k.

An alternative to the power spectrum is the Lyα forest flux autocorrelation function, its Fourier transform. In this work, we explore its ability to constrain the thermal state of the IGM at $z > 5$. The autocorrelation function of the Lyα forest offers two key advantages over the power spectrum in statistical analysis. First, uncorrelated noise, which is expected for astronomical spectrographs, averages to zero at non-zero lags in the autocorrelation function, eliminating the need for noise correction. In contrast, uncorrelated noise contributes a constant positive value at all scales of the power spectrum, requiring the subtraction of an estimates noise value that introduces additional uncertainty. Secondly, quasar spectra often contain masked regions (e.g. to remove metal lines), which impose a complex window function on power spectrum measurements that must be corrected (see e.g. Walther et al. 2019), further increasing uncertainty. The autocorrelation function, however, remains unaffected, as masking only reduces the number of pixel pairs available at a given velocity lag.

Many previous studies have measured the Lyα forest flux autocorrelation function at lower redshifts for a wide range of applications (McDonald et al. 2000; Rollinde et al. 2003; Becker, Sargent & Rauch 2004; D'Odorico et al. 2006). In addition, the first measurement of the Lyα forest flux autocorrelation function at $z > 5$ was presented in Wolfson et al. (2024) for moderate resolution quasar spectra.

In this work, we investigate constraints on the IGM thermal state under the assumption of a tight power-law relationship between temperature and density, as given by equation (1). This is a simplified model that does not fully capture the expected temperature fluctuations of the IGM at these redshifts (D'Aloisio, McQuinn & Trac 2015; Davies, Becker & Furlanetto 2018b). Additionally, we do not model pressure smoothing, only instantaneous temperature changes. Despite these simplifications, this modelling choice is sufficient to assess whether the autocorrelation function is sensitive enough to the thermal state to justify pursuing more realistic models. Specifically, we will quantify the constraints on T_0 and γ that can be achieved from Lyα forest flux autocorrelation function measurements using mock observational data sets.

More realistically, reionization can lead to significant fluctuations in the temperature of the IGM (D'Aloisio et al. 2015; Davies et al. 2018b). At the same time, fluctuations in the ultraviolet background (UVB) arise during reionization because the ionizing photons produced will be absorbed by the remaining neutral hydrogen at short distances from their initial sources (Davies & Furlanetto 2016; Gnedin, Becker & Fan 2017; D'Aloisio et al. 2018). These distances are characterized by the mean free path of ionizing photons, $\lambda_{\rm mfp}$ (Mesinger & Furlanetto 2009). Various previous studies have investigated the effect of large scale variations in the UVB on the autocorrelation function and power spectrum of the Lyα forest (Zuo 1992a, b; Croft 2004; Meiksin & White 2004; McDonald et al. 2005; Gontcho A Gontcho, Miralda-Escudé & Busca 2014; Pontzen 2014; Pontzen et al. 2014; D'Aloisio et al. 2018; Meiksin & McQuinn 2019; Oñorbe et al. 2019). In particular, Wolfson et al. (2023) showed that the positive fluctuations in the UVB that accompany small $\lambda_{\rm mfp}$ values boost the flux of the Lyα forest on small scales, which can be detected in the autocorrelation function.

We will use additional hydrodynamical simulations that model fluctuations in both the temperature and the UVB to determine the effect on the Lyα forest flux autocorrelation function. Beyond examining the qualitative differences between these models, we will quantify the likelihood ratio for mock data sets, providing a systematic way to compare and constrain a discrete set of models.

The structure of this paper is as follows. We discuss our simulation and grid of T_0 and γ in Section 2. The autocorrelation function and our other statistical methods to constrain these parameters are described in Section 3 with our results being discussed in Section 3.4. We discuss our second set of simulations for models of the IGM with temperature and UVB fluctuations in Section 4 and use the autocorrelation function to quantitatively distinguish between these models in Section 4.3. Finally, we summarize in Section 5.

2 SIMULATION DATA

2.1 Simulation box

In this work we use a simulation box of size $L_{\rm box} = 100$ comoving Mpc (cMpc) h^{-1} run with Nyx code (Almgren et al. 2013). Nyx is a hydrodynamical simulation code that was designed for simulating the Lyα forest with updated physical rates from Lukić et al. (2015). The simulation has 4096^3 dark matter particles and 4096^3 baryon grid cells. It is reionized by a Haardt & Madau (2012) uniform UVB that is switched on at $z \sim 15$. We have two snapshots of this simulation at $z = 5.5$ and $z = 6.0$. In this work, we consider seven redshifts: 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, and 6.0. For redshifts without a corresponding simulation snapshot, we use the density fluctuations and velocities from the nearest available Nyx snapshot. However, we compute the mean density and proper size of the simulation at the desired redshift. To validate this approach, we tested an alternative at $z = 5.7$. While our standard method uses the $z = 5.5$ snapshot for $z = 5.7$, we instead applied the density fluctuations and velocities

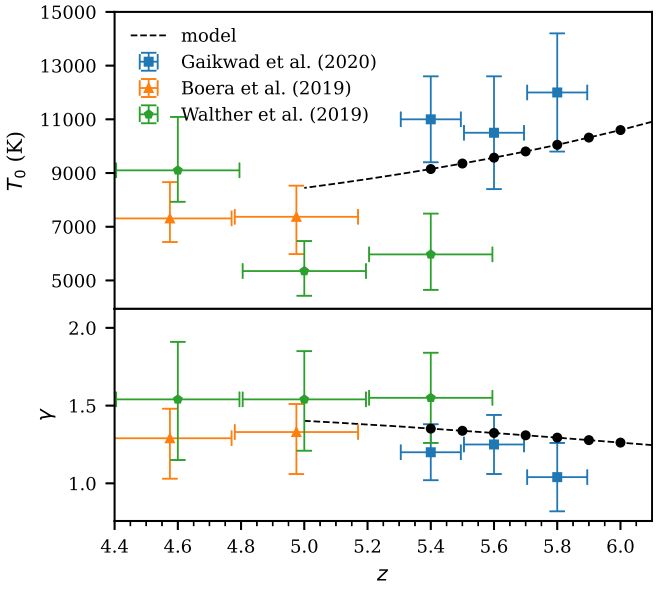

Figure 1. The blue squares, orange triangles, and green pentagons show previous measurements of T_0 and γ at high z from Gaikwad et al. (2020), Boera et al. (2019), and Walther et al. (2019), respectively. The dashed line shows the results for a thermal evolution model calculated with methods similar to Upton Sanderbeck et al. (2016) and Davies et al. (2018b) This model has $z_{\rm reion} = 7.7$, $\Delta T = 20\,000$ K, and $\alpha_{\rm UVB} = 1.5$. We use this model as our true redshift evolution for T_0 and γ in this work. The chosen models are shown as black circles.

from the $z = 6.0$ snapshot and found no significant differences in our final results.

We generate grids of thermal models by post-processing the temperature along the sightlines. For each value of T_0 and γ, we set the temperature of each cell following equation (1) for all densities with no cutoff. Our method does not take into account the full evolution of the thermal state of the IGM, only the instantaneous temperature. This simple model is sufficient to achieve the aim of this paper, which is to see if the autocorrelation function is sensitive to the thermal state. To make our grid we use 15 values of T_0 and 9 values of γ resulting in 135 different combinations of these parameters at each z. The values of T_0 and γ in our grid of thermal models were chosen based on the current models and available data, as shown in Fig. 1. We generate a model for the evolution of the thermal state of the IGM by a method similar to Upton Sanderbeck, D'Aloisio & McQuinn (2016) with $z_{\rm reion} = 7.7$, $\Delta T = 20,000$ K, and $\alpha_{\rm UVB} = 1.5$. For more information on the calculation of the temperature field see Davies et al. (2018b). We select central T_0 and γ values at each redshift from this model, which are shown as black points in Fig. 1 and listed in Table 1. At all z, we use the errors on the measurements reported in Gaikwad et al. (2020) at $z = 5.8$ ($\Delta T_0 = 2200$ K and $\Delta\gamma = 0.22$) and modelled from $T_0 - 4\Delta T_0$ to $T_0 + 4\Delta T_0$ and $\gamma - 4\Delta\gamma$ to $\gamma + 4\Delta\gamma$ in linear bins.

Our simulations do not predict the overall average of the UVB, $\langle\Gamma_{\rm UVB}\rangle$, because this value originates from complicated galaxy physics that are not included in the simulations. Additionally, our method of post-processing different thermal states would affect the resulting $\langle\Gamma_{\rm UVB}\rangle$. Instead, we model a range of possible $\langle\Gamma_{\rm UVB}\rangle$ values through the mean transmitted flux, $\langle F\rangle$, exploiting the relation $\tau_{\rm Ly\alpha} \propto 1/\Gamma_{\rm UVB}$ as seen in equation (2). This is achieved by rescaling the optical depths along each skewer, τ, so that the average transmitted flux across all skewers satisfies $\langle e^{-\tau}\rangle = \langle F\rangle$. These $\langle F\rangle$ model values

Table 1. This table lists the central values of the redshift-dependent thermal state models used in this work. The last column states the central value of $\langle F\rangle$ modelled in this work, which are the measurements from Bosman et al. (2022).

z	T_0 (K)	γ	$\langle F\rangle$
5.4	9149	1.352	0.0801
5.5	9354	1.338	0.0591
5.6	9572	1.324	0.0447
5.7	9804	1.309	0.0256
5.8	10 050	1.294	0.0172
5.9	10 320	1.278	0.0114
6.0	10 600	1.262	0.0089

are centred on the values presented in Bosman et al. (2022) for each redshift bin. We chose a range of models going from $\langle F\rangle - 4\Delta\langle F\rangle$ to $\langle F\rangle + 4\Delta\langle F\rangle$ where the $\Delta\langle F\rangle$ is the largest of the two redshift-dependent values reported in Bosman et al. (2022). These choices of $\langle F\rangle$ are listed in the last column of Table 1.

Our simulations are limited to a 100 cMpc h^{-1} box, which does not capture density fluctuations on scales larger than this volume. This introduces cosmic variance that may affect measurements of large-scale Lyman-alpha forest clustering. However, since our analysis focuses on small-scale structure and our UVB and thermal state modelling do not include fluctuations, we expect this effect to be subdominant.

We do not model Lyα forest contaminants, such as damped Lyα systems or metal lines, in this analysis. In observational studies, these contaminants are identified and masked to prevent biasing the results (see e.g. Wolfson et al. 2024). For the autocorrelation function, the focus of this study, such masking reduces the number of pixels in specific velocity bins, effectively lowering the statistical precision of the measurement. Since the impact of masking is well understood, we leave the quantification of missed contaminants to future observational studies.

We draw 1000 skewers from the simulation box. One example skewer at $z = 5.4$ for different T_0 and γ models is shown in Fig. 2. The top panel shows the density of this skewer for all models in black. There are then two pairs of panels each depicting the temperature (top) and flux (bottom) along this skewer.

The second and third panels vary T_0 with constant $\gamma = 1.352$ and $\langle F\rangle = 0.0801$. The coldest model, $T_0 = 2863$ K (blue), has some of the sharpest features. This is seen at $v \sim -300\,{\rm km\,s^{-1}}$ where the low T_0 (blue) model has a secondary sharp peak in the flux. In comparison the hottest model, $T_0 = 15\,435$ K (green), has one wider transmission spike. In addition, increasing T_0 decreases $\tau_{\rm Ly\alpha}$ as described in equation (2), which in turn increases the transmitted flux. For this reason we get the greatest transmission from the $T_0 = 15435$ K (green) model, seen in the transmission spike at $v = 50\,{\rm km\,s^{-1}}$. With fixed $\langle F\rangle$ this leads to greater variation in the flux for higher T_0 models.

The fourth and fifth panels vary γ with constant $T_0 = 9149$ K and $\langle F\rangle = 0.0801$. When $\gamma > 1$ (orange and green) the temperature is directly proportional to the density fluctuations while $\gamma < 1$ (blue) causes the temperature to be inversely proportional to the density fluctuations. When temperature is inversely proportional to density, lower densities have higher temperatures. Low densities and higher temperatures will locally increase the flux so the $\gamma < 1$ (blue) model will lead to transmission spikes with the greatest flux, as seen at $v \sim -100\,{\rm km\,s^{-1}}$.

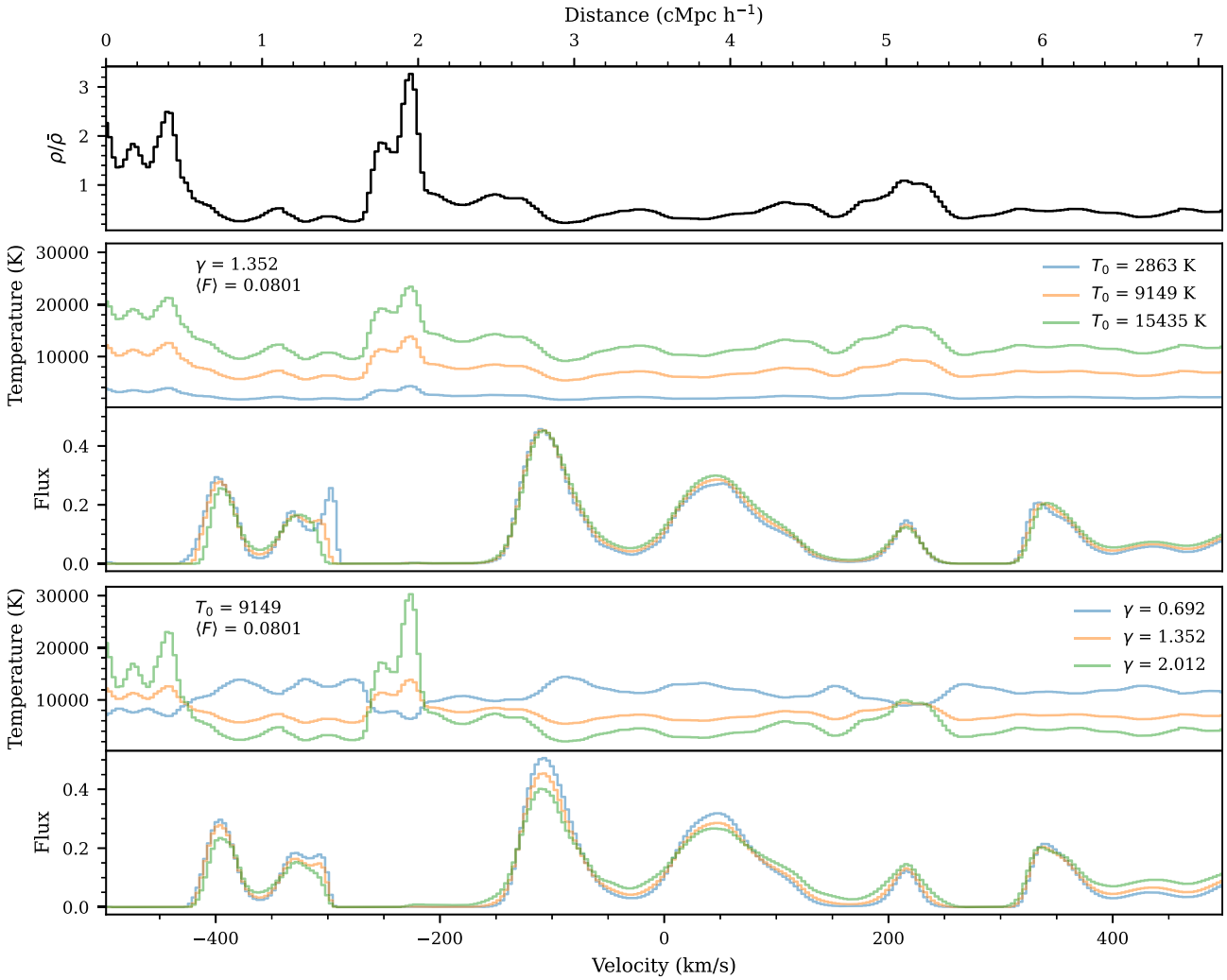

Figure 2. The top panel shows the density along a section of one skewer in black for $z = 5.4$. There are then two pairs of panels each depicting the temperature (top) and flux (bottom) along this skewer. The first pair varies T_0 with constant $\gamma = 1.352$ and $\langle F \rangle = 0.0801$. Shifting T_0 causes a corresponding shift in the temperature values along the skewer. Hotter temperatures (orange and green) smooths the flux, as seen clearly in the loss of a second transmission spike at $z \sim -300\,\mathrm{km\,s^{-1}}$. The second pair varies γ with constant $T_0 = 9148\,\mathrm{K}$ and $\langle F \rangle = 0.0801$. When $\gamma > 1$ (orange and green) the temperature is directly proportional to the density fluctuations while $\gamma < 1$ (blue) causes the temperature to be inversely proportional to the density fluctuations. When temperature is inversely proportional to density, lower densities have higher temperatures. Low densities and higher temperatures will locally increase the flux so the $\gamma < 1$ (blue) model will lead to transmission spikes with the greatest flux, as seen at $v \sim -100\,\mathrm{km\,s^{-1}}$.

2.2 Forward modelling

In order to mimic realistic high-resolution observational data from echelle spectrographs (e.g. from Keck/HIRES, VLT/UVES, and Magellan/MIKE) we forward model our ideal simulation skewers to have imperfect resolution and flux levels. We consider a resolution of $R = 30\,000$ and a signal-to-noise ratio per $10\,\mathrm{km\,s^{-1}}$ pixel (SNR_{10}) of $\mathrm{SNR}_{10} = 30$ at all redshifts.

We model this resolution by smoothing the flux by a Gaussian filter with $\mathrm{FWHM} = 10\,\mathrm{km\,s^{-1}}$. After smoothing we re-sampled the new flux such that the new pixel size was $\Delta v = 2.5\,\mathrm{km\,s^{-1}}$. With this pixel scale, $\mathrm{SNR}_{10} = 30$ corresponds to a signal-to-noise ratio of the pixel size ($\mathrm{SNR}_{\Delta v}$) of 15. For simplicity, we add flux-independent noise in the following way. We generate a single realization of random noise consisting of 1000 skewers, each of length matching the skewer length, with values drawn from a Gaussian distribution with $\sigma_N = 1/\mathrm{SNR}_{\Delta v}$. This noise realization is added to every model at every redshift. By using the same noise realization across all models, we prevent stochastic variations in the noise from introducing additional differences between models. As a result, the noise modelling does not unduly affect the parameter inference.

As discussed in Section 2.1 simulation skewers are 100 cMpc h^{-1} long, much longer than the $\Delta z = 0.1$ redshift bins we have chosen to analyse. Therefore, we split these skewers into two regions of length $\Delta z = 0.1$ and treating these two regions as independent, resulting in a total of 2000 skewers. Note that $\Delta z = 0.1$ corresponds to 33 cMpc h^{-1} at $z = 5.4$ and 29 cMpc h^{-1} at $z = 6.0$.

The initial and forward-modelled flux for one $z = 5.4$ skewer is shown in Fig. 3. This skewer has $T_0 = 9149\,\mathrm{K}$, $\gamma = 1.352$, and $\langle F \rangle = 0.0801$ (our central parameters values at this redshift). The forward modelled skewer, as is always true, uses $R = 30\,000$ and $\mathrm{SNR}_{10} = 30$. The initial flux is plotted as the red dashed line while the forward modelled flux is plotted as the black histogram.

We assume a fiducial data set size of 20 quasar spectra that probe a redshift interval of $\Delta z = 0.1$ per quasar for a total path length of

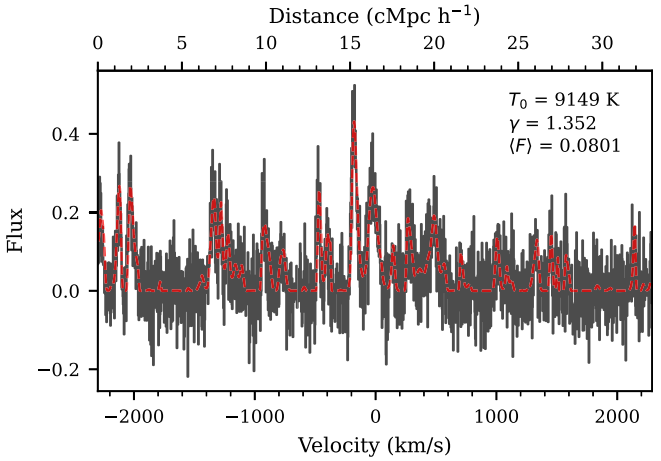

Figure 3. A forward-modelled skewer at $z = 5.4$ with $T_0 = 9149$ K, $\gamma = 1.352$, and $\langle F \rangle = 0.0801$ (our central parameters values at this redshift). This skewer, as is true for all skewers, is forward modelled with $R = 30\,000$ and $\mathrm{SNR}_{10} = 30$. The initial flux from the simulations is a red dashed line while the forward modelled flux is a black histogram.

$\Delta z = 2.0$ at all redshifts. This is a reasonable number of high-z, high resolution quasar observations to consider for a future measurement.

3 METHODS

3.1 Autocorrelation

The autocorrelation function of the flux, $\xi_F(\Delta v)$, is defined as

$$\xi_F(\Delta v) = \langle F(v)F(v + \Delta v)\rangle, \tag{3}$$

where $F(v)$ is the flux of the Lyα forest and the average is performed over all pairs of pixels with the same velocity lag, Δv. Conventionally, the flux contrast field, $\delta_F = (F - \langle F\rangle)/\langle F\rangle$, is used when measuring the power spectrum of the Lyα forest. However, we use the flux directly since $\langle F\rangle$ is small and highly uncertain at high z, where our forecast is focused. This avoids dividing by a small, independently measured quantity, which could otherwise cause δ_F to diverge. The autocorrelation function of the flux contrast can be written as

$$\xi_{\delta_f}(\Delta v) = \frac{\xi_F(\Delta v) - \langle F\rangle^2}{\langle F\rangle^2}. \tag{4}$$

ξ_{δ_f} can be computed via the Fourier transform of the dimensionless power spectrum of the Lyα forest flux contrast, $\Delta^2_{\delta_F}(k) = kP_{\delta_f}(k)/\pi$. In 1D this can be written as:

$$\xi_{\delta_f}(\Delta v) = \int_0^\infty \Delta^2_{\delta_f}(k)\cos(k\Delta v)\mathrm{d}\ln k. \tag{5}$$

The dimensionless power, $\Delta^2_{\delta_f}(k)$, is a smoothly rising function that has a sharp cutoff set by the thermal state of the IGM. Higher temperature values lead to sharper cutoffs as the power at small scales in the Lyα forest is removed. Equation (5) can be particularly useful when building intuition for the trends seen in the autocorrelation function with changing T_0 and γ, which we will discuss later in this section.

We compute the autocorrelation function with the following consideration for the velocity bins. We set the left edge of the smallest bin to be the resolution length, 10 km s^{-1}, and continue with linear bin sizes with a width of the resolution length, 10 km s^{-1}, up to 300 km s^{-1}. Then we switch to logarithmic bin widths where $\log(\Delta v) = 0.029$ out to a maximal distance of 2700 km s^{-1}. This results in 59 velocity bins considered where the first 28 have linear spacing. The centre of our smallest bin is 15 km s^{-1} and the centre of our largest bin is 2295 km s^{-1}. This largest bin corresponds to ~ 16.5 cMpc h^{-1} at $z = 5.4$. We chose to use linear bins on the smallest scales because this is where the thermal state has the greatest effect on the Lyα forest flux. At larger scales we switch to logarithmic binning as this is only sensitive to $\langle F\rangle$ and not the thermal parameters. The main aim of this work is to constrain the thermal parameters so having fine binning at large scales is not as important. To check this we compared our results at $z = 5.4$ to those when using linear bins at all scales and found no significant change to the constraints on the parameters. However, using linear bins at all scales results in 268 total bins, which significantly slowed down our computations. Therefore we used the linear-logarithmic bins at all z throughout the rest of this work.

The model autocorrelation function is computed as the average over all 2000 forward-modelled skewers. Each mock data set is then generated by averaging the autocorrelation function over 20 randomly selected skewers, representing 20 quasar sightlines. Finite resolution affects the autocorrelation function at the smallest velocity lags, but this effect is consistently included in both the models and mock data.

We show the correlation functions calculated for different thermal state parameters in Fig. 4 at $z = 5.4$. The solid lines show the mean values while the shaded regions represent the errors estimated from the diagonal of the covariance matrices. We discuss the computation of these covariance matrices later in this section.

The top panel shows models that vary T_0 with constant γ and $\langle F\rangle$. Varying T_0 results in small changes for the smallest velocity lags, where the second bin centred on 25 km s^{-1} has the largest per cent change in the models. The middle panel has models that vary γ with constant T_0 and $\langle F\rangle$ where the effect of changing γ is strongest on small scales. The bottom panel has models that vary $\langle F\rangle$ with constant T_0 and γ. $\langle F\rangle$ sets the amplitude of the autocorrelation function at all velocity lags. Here the differences between models are linear, with larger $\langle F\rangle$ producing higher autocorrelation values. This scaling follows $\propto \langle F\rangle^2$, as expected from the definition of the autocorrelation function.

For the thermal models, larger T_0 and smaller γ lead to larger correlation function values on small scales. Though these models do not seem to show large differences by eye, we will investigate what statistically rigorous measurements could look like in Section 3.4.

To build intuition for how the autocorrelation function depends on the thermal parameters, we refer to equation (5). In Appendix A, we show the integrand of this equation for $\Delta v = 15$ km s^{-1}. As discussed earlier, $\Delta^2_{\delta_f}(k)$ has a sharp thermal cutoff, which may naively suggest hotter thermal states would result in lower autocorrelation values at small scales. However Fig. 4 shows the opposite trend where larger values of T_0 correspond to higher autocorrelation values at small scales. This behaviour can be explained by two factors. First, hotter thermal states introduce greater flux variation, as noted earlier. Secondly, $\Delta^2_{\delta_f}(k)$, the integrand from equation (5), has higher values at small k for larger T_0 values, particularly when viewed on a linear scale, as shown in Fig. A1.

We compute the covariance matrices for the models by averaging over randomly drawn mock data sets, each consisting of 20 skewers (quasar sightlines):

$$\Sigma(T_0, \gamma, \langle F\rangle) = \frac{1}{N_{\mathrm{mocks}}}\sum_{i=1}^{N_{\mathrm{mocks}}}(\boldsymbol{\xi}_i - \boldsymbol{\xi}_{\mathrm{model}})(\boldsymbol{\xi}_i - \boldsymbol{\xi}_{\mathrm{model}})^{\mathrm{T}}, \tag{6}$$

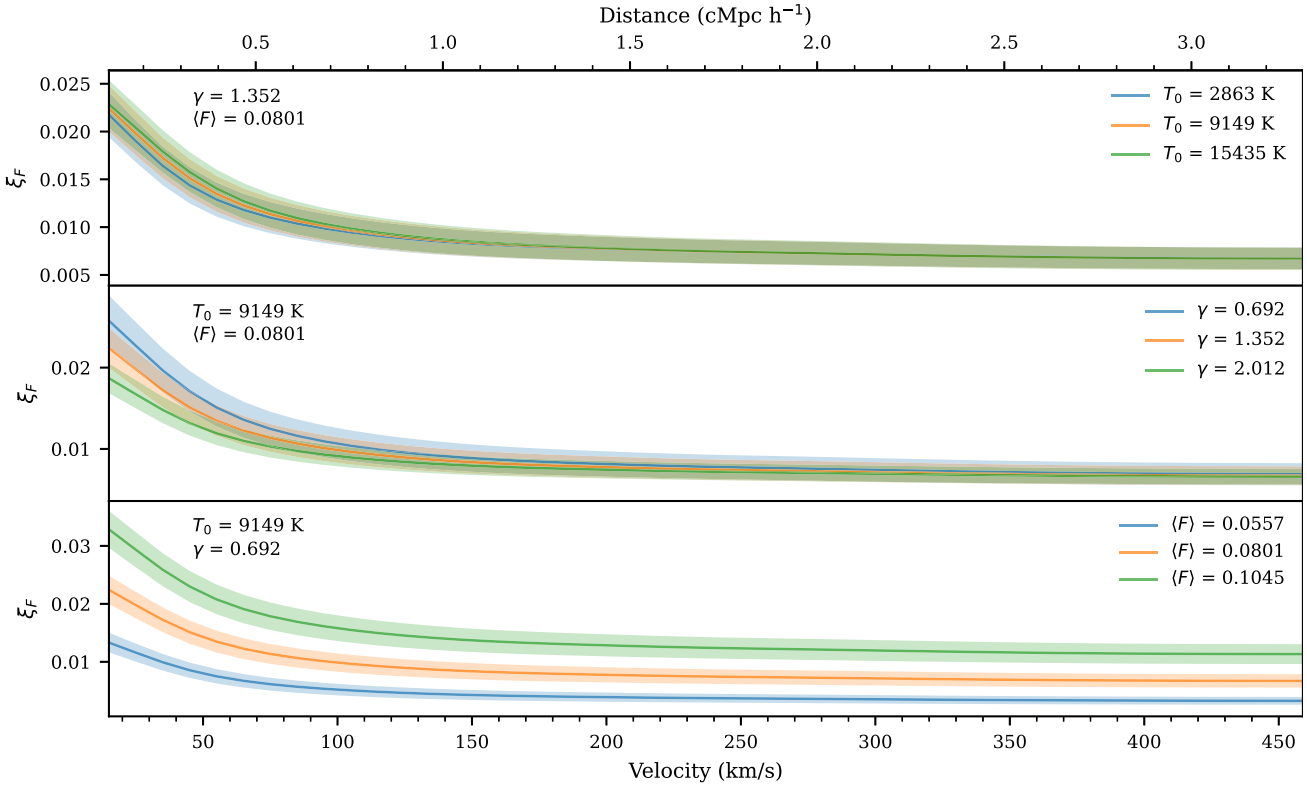

Figure 4. This figure demonstrates the effects of varying the parameters on the autocorrelation function from the simulations at $z = 5.4$. Each of the three panels varies one parameter from T_0, γ, and $\langle F \rangle$ while keeping the others constant. The constant parameter values are written in the top left of each panel. The solid lines show the model values and the shaded regions show errors estimated by the diagonals of the covariance matrices. T_0 (top panel) and γ (middle panel) affect the autocorrelation function on small scales. $\langle F \rangle$ (bottom panel) affects the autocorrelation function on all scales.

where $\boldsymbol{\xi}_i = \boldsymbol{\xi}_i(T_0, \gamma, \langle F \rangle)$ is the i-th mock autocorrelation function, $\boldsymbol{\xi}_{\rm model} = \boldsymbol{\xi}_{\rm model}(T_0, \gamma, \langle F \rangle)$ is the model value of the autocorrelation function, and $N_{\rm mocks}$ is the number of forward-modelled mock data sets used. Since the covariance is computed as an average over multiple mock data sets (each consisting of 20 skewers), it naturally incorporates the effects of sample variance. We use $N_{\rm mocks} = 500\,000$ for all models and redshifts in this work, see Appendix B for a discussion on the convergence of the covariance matrix. Note that $\boldsymbol{\xi}_i(T_0, \gamma, \langle F \rangle)$ and $\Sigma(T_0, \gamma, \langle F \rangle)$ are computed at each point on the grid of T_0, γ, and $\langle F \rangle$, resulting in 1215 separate computations.

To visualize the covariance matrix, we define the correlation matrix, C, which expresses the covariances between jth and kth bins in units of the the diagonal elements of the covariance matrix. Specifically, the jth, kth element of the correlation matrix is given by:

$$C_{jk} = \frac{\Sigma_{jk}}{\sqrt{\Sigma_{jj}\Sigma_{kk}}}. \tag{7}$$

One example correlation matrix is shown in Fig. 5 for $z = 5.4$ with $T_0 = 9149\,\mathrm{K}$, $\gamma = 1.352$, $\langle F \rangle = 0.0801$. We find that all bins of the autocorrelation function are highly correlated, a result of each pixel in the Lyα forest contributing to multiple (in fact, almost all) bins in the autocorrelation function.

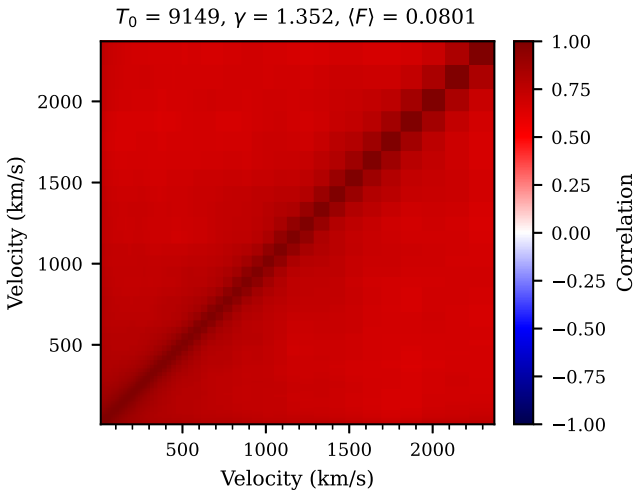

Figure 5. This figure shows the correlation matrix calculated with equation (7) for the model at $z = 5.4$ with $T_0 = 9149\,\mathrm{K}$, $\gamma = 1.352$, $\langle F \rangle = 0.0801$. The colour bar is fixed to span from -1 to 1, which is all possible values of the correlation matrix. This illustrates that all bins in the autocorrelation function are highly correlated with each other.

3.2 Parameter estimation

To quantitatively constrain the parameters we modelled $\boldsymbol{\theta} = (T_0, \gamma, \langle F \rangle)$, we use Bayesian inference with a multivariate Gaussian likelihood and a flat prior over the parameters. This likelihood, $\mathcal{L} = \mathcal{L}(\boldsymbol{\xi}|\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\xi}|T_0, \gamma, \langle F \rangle)$, has the form:

$$\mathcal{L} = \frac{1}{\sqrt{\det(\Sigma)(2\pi)^n}} \exp\left(-\frac{1}{2}(\boldsymbol{\xi} - \boldsymbol{\xi}_{\rm model})^{\rm T}\Sigma^{-1}(\boldsymbol{\xi} - \boldsymbol{\xi}_{\rm model})\right), \tag{8}$$

where $\boldsymbol{\xi}$ is the autocorrelation function from our mock data, $\boldsymbol{\xi}_{\rm model} = \boldsymbol{\xi}_{\rm model}(T_0, \gamma, \langle F \rangle)$ is the model value of the autocorrelation function, $\Sigma = \Sigma(T_0, \gamma, \langle F \rangle)$ is the model-dependent covariance matrix estimated by equation (6), and $n = 59$ is the number of points in the autocorrelation function. We discuss the assumption of using a multivariate Gaussian likelihood in Appendix C. This discussion shows that our mock data does not exactly follow a Guassian distribution. This discrepancy may affect our parameter inference, we investigate the consequences of this assumption in a later section.

Our models are defined by three parameters $\boldsymbol{\theta} = (T_0, \gamma, \langle F \rangle)$. Following Bayes' theorem, we have

$$P(T_0, \gamma, \langle F \rangle | \boldsymbol{\xi}) = \frac{\mathcal{L}(\boldsymbol{\xi} | T_0, \gamma, \langle F \rangle) P(T_0, \gamma, \langle F \rangle)}{P(\boldsymbol{\xi})}, \tag{9}$$

where $P(T_0, \gamma, \langle F \rangle | \boldsymbol{\xi})$ is the posterior distribution for $\boldsymbol{\theta} = T_0, \gamma, \langle F \rangle$ given observed autocorrelation function $\boldsymbol{\xi}$. $\mathcal{L}$ is the likelihood defined in equation (8), $P(T_0, \gamma, \langle F \rangle)$ is the prior on our parameters, and $P(\boldsymbol{\xi})$ is the probability of the data which can be interpreted as a normalization constant since it is independent of $\boldsymbol{\theta}$.

We compute these posteriors using Markov Chain Monte Carlo (MCMC) with the `emcee` (Foreman-Mackey et al. 2013) package. We assumed uniform priors spanning the range of each parameter we considered. Because we considered different central values at each z, the priors are also z dependent. At $z = 5.4$ these are $T_0 \in$ Uniform(349, 17949), $\gamma \in$ Uniform(0.472, 2.232), and $\langle F \rangle \in$ Uniform(0.0557, 0.1045). For $\mathcal{L}$, we linearly interpolate the model values and covariance matrix elements onto a finer 3D grid of T_0, γ, and $\langle F \rangle$ then use the nearest model during the MCMC. This fine grid has 29 values of T_0, 33 values of γ, and 41 values of $\langle F \rangle$ which corresponds to adding 1, 3, and 4 points between the existing grid points, respectively. Our MCMC was run with 16 walkers taking 3500 steps each and skipping the first 500 steps of each walker as a burn-in, resulting in 48 000 samples.

Fig. 6 shows the result of our inference procedure for one mock data set at $z = 5.4$. The top panel shows the mock data set along with various lines relating to the inference procedure as follows. The green dotted line and accompanying text show the model value for the simulation from which the mock data was drawn. The mock data set is plotted as black points with error bars derived from the diagonal elements of the covariance matrix of the model nearest to the inferred model. The inferred model, represented by the red line and accompanying text, is based on the median of each parameter's samples, determined by the 50th percentile of the MCMC chains. The error on the inferred model written in the text is given by the 16th and 84th percentiles of the MCMC chains. The blue lines represent 100 random draws from the MCMC chain, illustrating the variety of models in the posterior distribution. The bottom left panel shows a corner plot of the posteriors for T_0, γ, and $\langle F \rangle$.

3.3 Inference test and re-weighting

To assess the reliability of our inference method, we perform a statistical validation test (inference test). This test evaluates whether our posterior distributions are statistically robust and whether the assumptions underlying our likelihood function are valid. The procedure is described in detail in Hennawi et al. (2025), though see also Wolfson et al. (2023). Below, we provide a brief summary of the method and its motivation.

Statistically, the true parameter values, $\boldsymbol{\theta}_{\rm true}$, should fall within the nth credibility contour n per cent of the time when the experiment is repeated (see e.g. Prangle et al. 2014; Ziegel & Gneiting 2014; Morrison & Simon 2018; Sellentin & Starck 2019). Mathematically, this condition can be expressed as follows: the probability, α, obtained by integrating the posterior probability density from equation (9) over a volume, V, of parameter space, $\boldsymbol{\theta}$,

$$\alpha = \int_{V_\alpha} P(\boldsymbol{\theta} | \boldsymbol{\xi}) {\rm d}\boldsymbol{\theta}, \tag{10}$$

should correspond to the coverage probability, $C(\alpha)$, which is the fraction of cases in which the true parameters, $\boldsymbol{\theta}_{\rm true}$, lie within V_α, the region enclosed by the α-th credibility contour. Thus, by testing whether $C(\alpha) = \alpha$ for multiple values of α, we can assess the statistical validity of our posteriors.

Overall, our inference test is done as follows:

(i) Draw $N = 300$ parameter vectors $\boldsymbol{\theta}_{\rm true} = (T_0, \gamma, \langle F \rangle)_{\rm true}$, from their uniform priors, which were described in 3.2.

(ii) Generate $N = 300$ mock autocorrelation data set corresponding to these $(T_0, \gamma, \langle F \rangle)_{\rm true}$, as described in 3.1.

(iii) Perform parameter inference on each mock data set as described in 3.2, resulting in 48 000 samples each in the posterior distributions.

(iv) Consider a set of M credibility contour levels $\alpha \in [0, 1]$. For each value α and each mock, test whether the true values, $\boldsymbol{\theta}_{\rm true} = (T_0, \gamma, \langle F \rangle)_{\rm true}$, reside within the volume V_α enclosed by α-th contour. For each α, the coverage probability $C(\alpha)$ is the fraction of the N mock data sets for which the true values lie within the volume V_α defined by equation (10).

The coverage probability, $C(\alpha)$, is determined by the fraction of trials in which $\boldsymbol{\theta}_{\rm true}$ falls within V_α. Since this corresponds to counting successes in N independent trials, $C(\alpha)$ follows a Binomial distribution, $B(N, C(\alpha))$. We estimate the uncertainty on $C(\alpha)$ using the 16th and 84th percentiles of this distribution.

We plot the results of our inference procedure at $z = 5.4$ from 300 posteriors in the left panel of Fig. 7. The grey shaded regions around our resulting line show the 1σ errors on $C(\alpha)$ from $B(N, C(\alpha))$. We expect $C(\alpha) = \alpha$, shown as the red dashed line. To interpret this plot, consider a point such as $\alpha \approx 0.6$, which represents the 60th percentile contour, calculated from the 60th percentile of probabilities from the MCMC chain draws for each mock data set. Here, the true parameters fall within the 60th percentile contour only $C(\alpha) \sim 52$ per cent of the time, implying that our posteriors are too narrow. Consequently, we are underestimating the errors and need to widen them so that the true parameters will fall within the 60th percentile contour more often. We run this inference test at all z considered in this work and find the deviation from the $C(\alpha) = \alpha$ line is larger at higher redshifts. For more details, see Appendix D for the inference test at $z = 6$. We also run the inference test for mock data generated from a multivariate Gaussian distribution in Appendix E, where the results align with the $C(\alpha) = \alpha$ line. This suggests that the discrepancy when using forward-modelled data arises because the data distribution is not perfectly Gaussian.

Recent work has focused on correcting posteriors that fail the coverage probability test (see e.g. Prangle et al. 2014; Grünwald & van Ommen 2017; Sellentin & Starck 2019). Specifically, Sellentin & Starck (2019) proposed relabelling the credibility contours, α, using the calculated $C(\alpha)$. In this work, we adopt the method of Hennawi et al. (2025), which involves calculating a set of weights for the MCMC draws that mathematically broaden the posteriors in a rigorous way. This approach is described in detail in Wolfson et al. (2023) and Hennawi et al. (2025), and we refer the reader to those papers for specifics on computing the weights. Here, we will only discuss the effect of adding these weights to the posteriors.

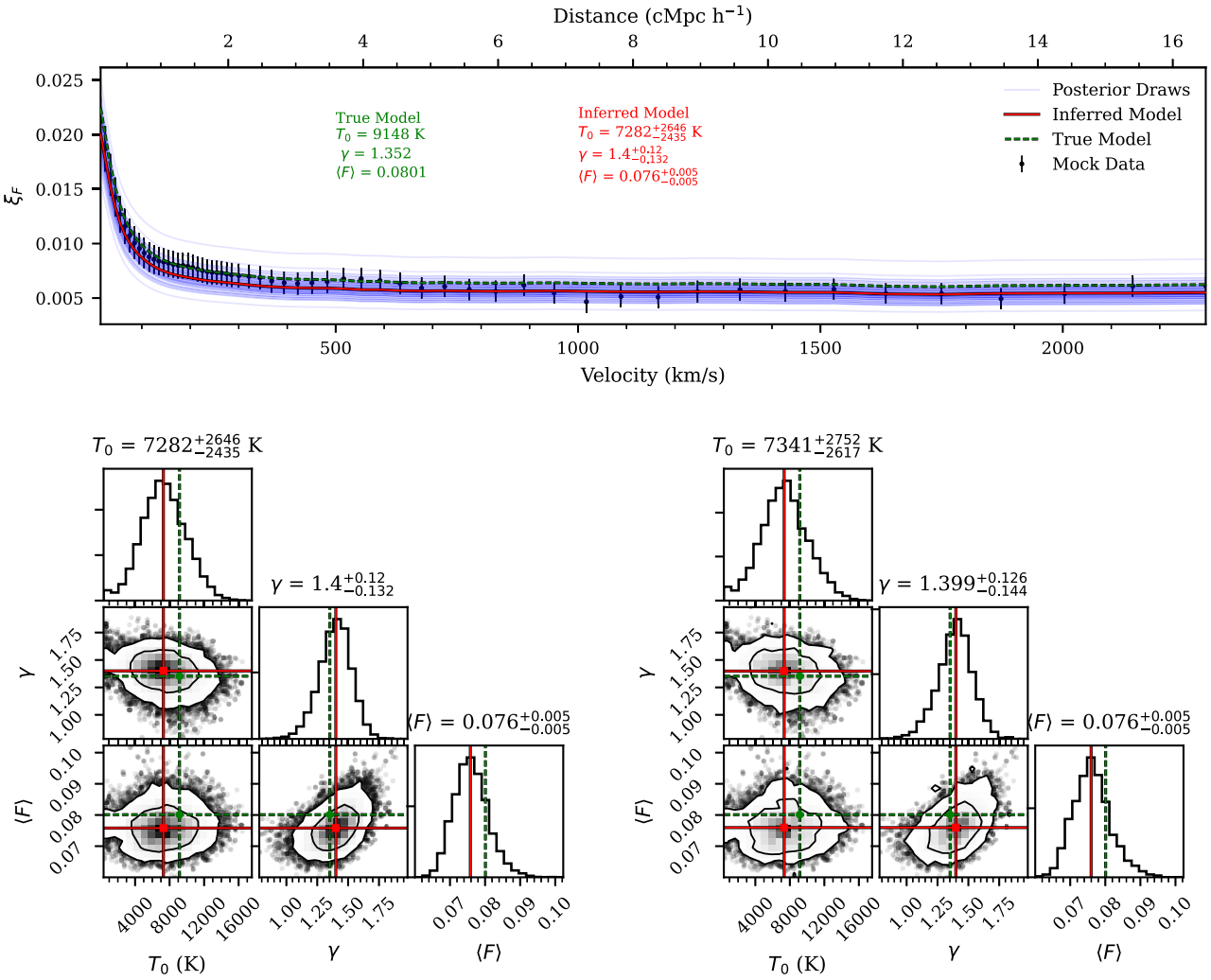

Figure 6. This figure illustrates the results of our inference procedure applied to one mock data set at $z = 5.4$. The top panel shows the data and models that resulted from our inference procedure, the bottom left has the corner plot resulting from the fit, and the bottom right has the same corner plot which has been re-weighted to pass our inference test. In the top panel, the black points are the mock data with error bars from the diagonals of the covariance matrix of the inferred model. The inferred model was calculated by the median (50th percentile) of the MCMC chains of each parameter independently. The inferred model is shown as a red line while the accompanying red text reports errors calculated from the 16th and 84th percentiles of each parameter. In comparison, the true model, which was used to generate the data, is shown as a green dotted line. The parameters for this model is written in the accompanying green text. To demonstrate the width of the posterior, multiple faint blue lines are shown which are the models corresponding to the parameters from 100 random draws of the MCMC chain. The bottom left panel shows a corner plot of the values of T_0, γ, and $\langle F \rangle$ that immediately result from our inference procedure. The bottom right panel shows the corner plot of the values of T_0, γ, and $\langle F \rangle$ from our inference procedure that has been re-weighted with the weights calculated from our inference test as described in Section 3.3. For this mock data set, the true model parameters fall within the 68th percentile contours.

We show the re-weighted posteriors for T_0, γ, and $\langle F \rangle$ in the bottom right panel of Fig. 6. The weights give greater importance to values of T_0, γ, and $\langle F \rangle$ that lie outside the 68 per cent contour, effectively broadening the posteriors and increasing the errors on the fit. For the mock data set in Fig. 6, the marginalized re-weighted posterior for T_0 gives $T_0 = 7341^{+2752}_{-2617}$ K, resulting in a ~ 6 per cent increase in the error when compared to the previous measurement of 7282^{+2646}_{-2435} K. Similarly, the re-weighted posterior for γ gives $\gamma = 1.399^{+0.126}-0.144$, resulting in a ~ 6 per cent increase in the error when compared to whereas the previous measurement of $1.400^{+0.120}-0.132$. The error on $\langle F \rangle$ does not change. Examining the 2D distributions in this corner plot, such as the $(\gamma, \langle F \rangle)$ distribution in the middle panel of the bottom row, we observe small regions outside the main 95 per cent contour that are up-weighted. This arises from giving one particular draw a higher weight, illustrating how the weights introduce an additional source of noise to the posterior distribution. We also note that since the changes are around 6–7 per cent, this difference is difficult to discern visually in the corner plots.

The need for re-weighting, or some method to correct our posteriors to pass an inference test, comes from our incorrect (though frequently used) assumption of a multivariate Gaussian likelihood. The values of the autocorrelation function at these high z do not adequately follow a multivariate Gaussian distribution, which should serve as a cautionary note for other studies of the Lyα forest at these z. Using a more appropriate form of the likelihood, such as a skewed distribution, or likelihood-free inference [e.g. approximate Bayesian computation as used in Davies et al. (2018a) or other

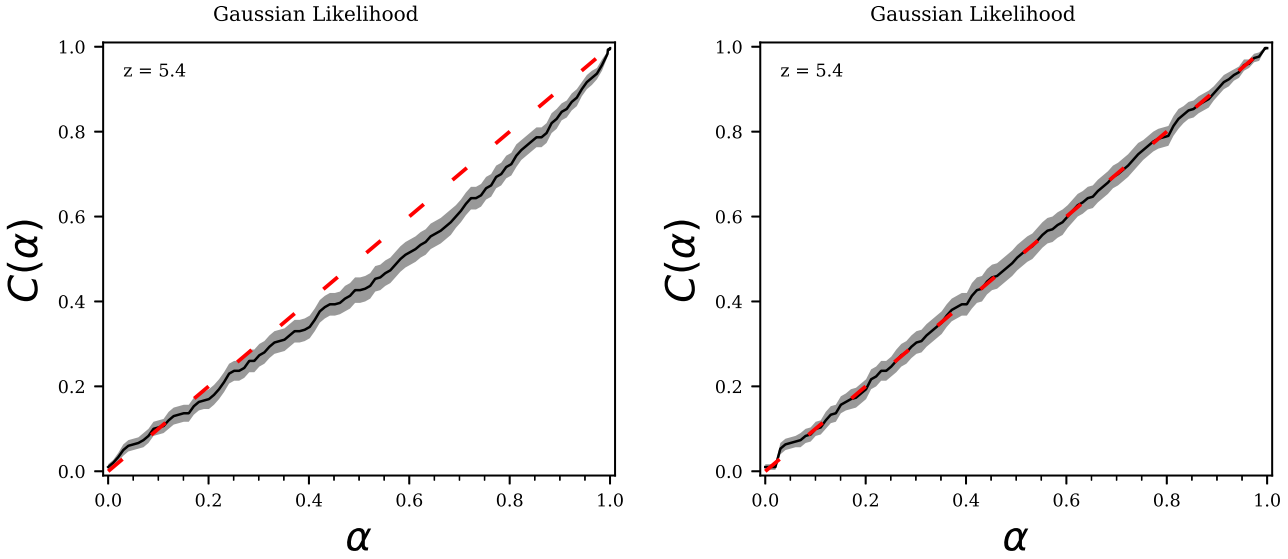

Figure 7. The left panel shows the coverage resulting from the inference test at $z = 5.4$ from 300 mock data sets with parameters drawn from our priors on T_0, γ, and $\langle F \rangle$. This shows that, for example, the true parameters fall above the 60th percentile in the MCMC chain ~ 50 per cent of the time. The line falls below the $C(\alpha) = \alpha$ line, meaning that the posteriors are overconfident (too narrow). The right panel of this figure shows the coverage resulting from the inference test with the use of one set of weights to re-weight the posteriors, which passes.

machine learning methods] may yield more optimal posteriors that better capture the information content of the autocorrelation function. Exploring these approaches has been left for future work.

3.4 Thermal state measurements

Now that we have established the reliability of our posterior distributions, we can discuss constraints on the thermal state. We investigate the impact of sample variance by analysing the distribution of measurements across 100 mock data sets, all generated with the same true parameters, $\boldsymbol{\theta}_{\rm true} = (T_0, \gamma, \langle F \rangle)_{\rm true}$. For each z, we adopt the values of T_0, γ, and $\langle F \rangle$ from Table 1. As stated earlier, each mock data set is created by randomly selecting and averaging the autocorrelation function over 20 skewers. For each mock data set, we perform MCMC sampling as described in Section 3.2 and then apply the re-weighting procedure from Section 3.3. Using these weighted posteriors, we compute the marginalized distributions for T_0 and γ.

In addition to analysing the 100 mock data sets, we consider an idealized data set where the autocorrelation function is averaged over all 2000 skewers instead of 20. The measurement errors for this data set are still derived from the covariance matrices assuming a 20 quasar sample. Notably, this 2000 skewer average is also used for the models shown in Fig. 4 and in the likelihood defined by equation (8). Averaging over all 2000 skewers removes random fluctuations from the choice of 20 skewers, providing an estimate with optimal precision. This idealized mock data set is then fit using MCMC and re-weighted following the same procedure as the 20 quasar mock data sets.

At $z = 5.4$, Fig. 8 shows all 100 marginalized re-weighted posteriors as thin blue curves for T_0 (top panel) and γ (bottom panel). The thick blue histogram shows the posterior from the idealized data set described above. The corresponding idealized measurement is written in blue text in Fig. 8. The measurements of the idealized data sets at every redshift are reported in Table 2.

The re-weighted posteriors appear noisy, similar to what is observed in the bottom right panel of Fig. 6. This noise is a

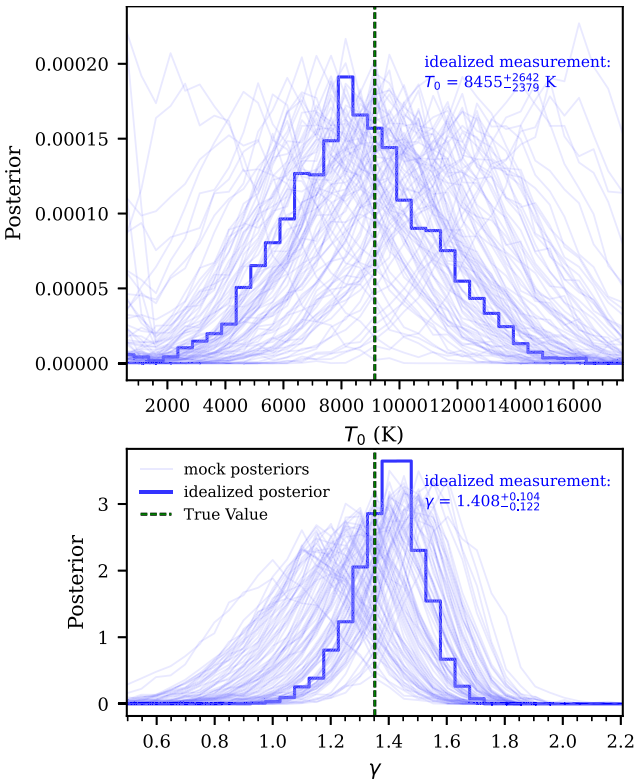

Figure 8. 100 marginalized re-weighted posteriors of T_0 and γ at $z = 5.4$ from mock data sets with true $T_0 = 9149$, $\gamma = 1.352$, and $\langle F \rangle = 0.0801$ (faint blue lines). The top panel shows the marginalized posteriors for T_0 and the bottom panel shows the marginalized posteriors for γ. Both panels also show the re-weighted posterior from the idealized data set (thick blue histograms). The measurement resulting from fitting the idealized data set are written in blue text. The idealized autocorrelation function was taken as the average over 2000 skewers with error from the covariance matrices computed from 20 quasar (skewer) data sets. This demonstrates the different possible behaviours the posterior can have for different mock data sets with the same T_0, γ, and $\langle F \rangle$ values.

Table 2. The results of fitting to the idealized mock data sets described in Section 3.4. The value of the autocorrelation functions for these mock data sets was taken as the average over 2000 skewers and the error was set by the covariance matrices computed for a 20 quasar (skewer) data set. The first (third) column contains the modelled value of T_0 (γ) at each z. The second (fourth) column contains the measurements for T_0 (γ) calculated by the 16th, 50th, and 84th percentiles. In general the trend of the errors is to increase with increasing redshift.

z	Model T_0	Measured T_0	Model γ	Measured γ
5.4	9149	8455^{+2642}_{-2379}	1.352	$1.408^{+0.104}_{-0.122}$
5.5	9354	8643^{+3152}_{-3054}	1.338	$1.422^{+0.116}_{-0.141}$
5.6	9572	8480^{+3720}_{-3642}	1.324	$1.433^{+0.121}_{-0.151}$
5.7	9804	8222^{+5188}_{-4176}	1.309	$1.460^{+0.139}_{-0.166}$
5.8	10 050	8346^{+4926}_{-4576}	1.294	$1.485^{+0.157}_{-0.204}$
5.9	10 320	7892^{+6111}_{-4655}	1.278	$1.513^{+0.170}_{-0.223}$
6.0	10 600	9574^{+6219}_{-5133}	1.262	$1.511^{+0.196}_{-0.256}$

direct consequence of our re-weighting procedure and is expected to improve with future developments in likelihood-free inference methods. For T_0, the posterior obtained from the model fit has a width comparable to the typical widths of the mock data set posteriors. For γ, the model-fit posterior is slightly narrower than most individual mock data set posteriors. Additionally, the posteriors for γ show an asymmetry: those peaking at lower γ values tend to be broader than those peaking at higher γ values. The model-derived posteriors (thick blue lines) for both T_0 and γ contain the true parameter values within their 1σ confidence intervals. Across the 100 mock data sets, the model value of T_0 falls within the 1σ confidence region 75 times, while the model value of γ falls within it 69 times. Statistically, we expect the true values to fall within these intervals 68 times, which is within 1.5σ of the expectation for the binomial distribution $B(100, .75)$ and well within 1σ for $B(100, .69)$. This agreement again demonstrates the reliability of our inference method.

Table 2 reports the parameter constraints obtained when using the model values of the autocorrelation function as our data at all z. This represents an idealized scenario that eliminates the impact of statistical fluctuations from individual mock data sets. The first (third) column contains the modelled value of T_0 (γ) at each z used in this measurement. The second (fourth) column contains the corresponding constraints on T_0 (γ) given by the 16th, 50th, and 84th percentiles of the marginalized re-weighted posterior distributions. In general, the uncertainties increase with redshift. At $z = 5.4$, the model constrains T_0 to 59 per cent and γ to 16 per cent, providing the strongest constraints. By $z = 6.0$, these constraints weaken to 119 per cent for T_0 and 30 per cent for γ.

To visualize the differences between measurements at different redshifts, we plot the results for two random mock data sets in Fig. 9. The first and third panels show the marginalized posteriors for T_0, while the second and fourth panels show the marginalized posteriors for γ. Each violin illustrates the marginalized re-weighted posterior for one randomly selected mock data set at the corresponding redshift. The light blue shaded region indicates the 2.5th and 97.5th percentiles (2σ) of the MCMC draws, while the darker blue shaded region indicates the 16th and 84th percentiles (1σ) of the MCMC draws. The dot dashed line shows the simulated model value evolution, as shown in Fig. 1 and reported in Table 1.

Looking at the posteriors for a given redshift (one column in the figure), the only difference between the posteriors is the random mock data set drawn. This results in varying precision, as seen in Fig. 8 for $z = 5.4$, due to sample variance. Across the different z values, both the random mock data set and the true values of T_0 and γ differ. As mentioned, the individual posteriors are noisy due to the re-weighting procedure described in Section 3.3. The behaviour here echos the results found with the idealized measurements, where the precision of the constraints on T_0 and γ decreases with increasing redshift. For the highest redshift bins ($z > 5.7$), the posteriors for the mock data sets tend to show large values at the boundary of our prior more frequently.

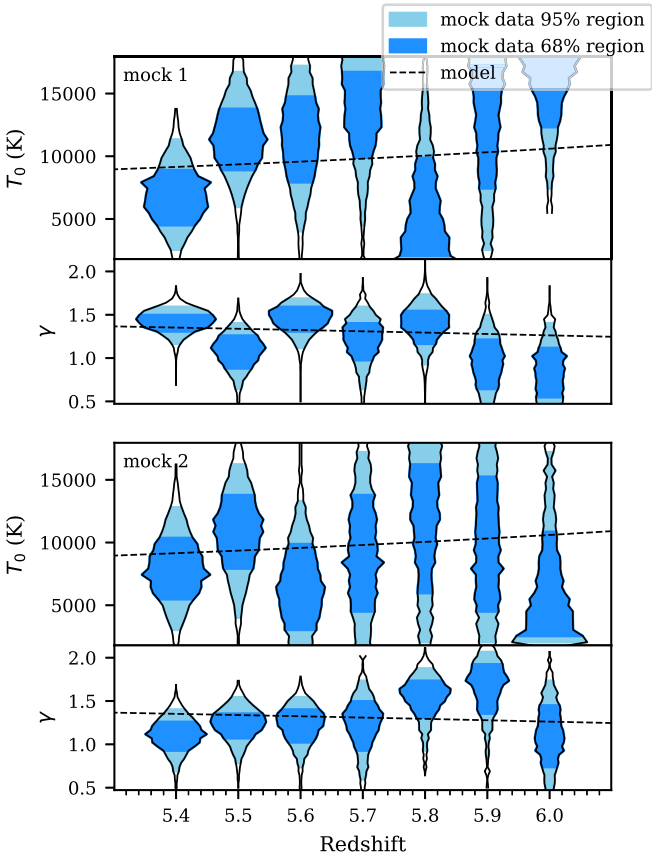

Figure 9. The marginalized posteriors for two random mock data sets at each z for T_0 and γ. The first and third panels show the marginalized posteriors for T_0 while the second and fourth panels show the same for γ. For each posterior, the light blue shaded region demarcates the 2.5th and 97.5th percentile of the weighted MCMC draws while the darker blue shaded region demarcates the 17th and 83rd percentile of the weighted MCMC draws. There are 14 total random mock data sets used to make this figure. For a given T_0 and γ posterior pair (in the first and second or third and fourth panels) the mock data set is the same. The behaviour of each posterior is partially determined by the specific mock data set considered, due to sample variance. The size of the data set is consistent across z but the true parameter values of the mock data varies as shown by the black dot dashed line. This black dot dashed line was also shown in Fig. 1 and the values at each z are reported in Table 1.

4 INHOMOGENEOUS REIONIZATION

So far in this work, we have post-processed simulations to have different thermal states following a tight temperature–density relationship. This is a simple yet necessary first step in exploring the sensitivity of the Lyα forest flux autocorrelation function to the thermal state of the IGM at high-redshifts. However, as previously discussed, recent measurements of the Lyα optical depth at $z > 5.5$ have shown scatter

that cannot be explained by density fluctuations alone (Fan et al. 2006; Becker et al. 2015; Bosman et al. 2018; Eilers et al. 2018; Bosman et al. 2022). It is possible that these fluctuations come from fluctuations in the temperature field (D'Aloisio et al. 2015; Davies et al. 2018b) or fluctuations in the UVB (Davies & Furlanetto 2016; Gnedin et al. 2017; D'Aloisio et al. 2018). Fluctuations in either of these fields can arise if reionization is extended or patchy.

On top of the measurements of fluctuations in the Lyα forest optical depth at $z > 5.5$, recent measurements of the mean free path of ionizing photons at $z > 5$ suggest a UVB that cannot be well described by uniform fields (Becker et al. 2021; Bosman 2021; Gaikwad et al. 2023; Zhu et al. 2023).

In order to explore the effect of temperature and UVB fluctuations on the Lyα forest flux autocorrelation function, we consider a set of four simulation models. These simulations have two different reionization models (one of which causes temperature fluctuations) and two UVB models (one of which has UVB fluctuations). These simulations and their results will be described in detail in the following sections.

4.1 Simulation box

For these models we use an additional Nyx simulation box with a size of $L_{\rm box} = 40$ cMpc h^{-1} and 2048^3 resolution elements at $z = 5.8$. We comment on the effect of this box size on the resulting autocorrelation function at the end of Section 4.2. A slice through the density field of this simulation is shown in the top left panel of Fig. 10.

We consider two reionization models: an instantaneous model and an extended, inhomogeneous model [the 'flash' and inhomogeneous methods described in Oñorbe et al. (2019), respectively]. The instantaneous model of reionization assigns all resolution elements the same redshift of reionization, $z_{\rm reion,\,HI}$. For this work we use $z_{\rm reion,\,HI} = 7.75$. A brief summary of the inhomogeneous model of reionization is as follows, each resolution element is assigned its own redshift of reionization such that reionization has a given midpoint, $z^{\rm median}_{\rm reion,\,HI}$, and duration, $\Delta z_{\rm reion,\,HI}$. For this work we use $z^{\rm median}_{\rm reion,\,HI} = 7.75$ and $\Delta z_{\rm reion,\,HI} = 4.82$. It is possible for cells to be ionized before the redshift of reionization through other processes such as collisional reionization. In both models, at the redshift of reionization for a given resolution element heat, ΔT, is injected. In both of our reionization models $\Delta T = 2 \times 10^4$ K. These two models result in two different temperature fields. We say that the instantaneous reionization model has 'no temperature fluctuations' and the inhomogeneous reionization model has 'temperature fluctuations'.

The bottom row of Fig. 10 shows slices through the resulting temperature field from these two simulations: one with no temperature fluctuations on the left and one with temperature fluctuations on the right. From this figure we see that model with temperature fluctuations has a larger scatter in the temperature with the greater abundance of colder (darker blue) regions. These cold regions correspond to the regions of higher density in the top left panel. This follows from the model of reionization where the denser regions reionize (and are heated) first and thus have more time to cool to a lower temperature by $z = 5.8$.

In addition to a constant UVB model, we have a model with UVB fluctuations. This UVB model was generated by the same method presented in Oñorbe et al. (2019) with $\lambda_{\rm mfp} = 15$ cMpc. The method follows the approach of Davies & Furlanetto (2016) where we consider modulations in the ionization state of optically thick absorbers assuming that $\lambda_{\rm mfp} \propto \Gamma^{2/3}_{\rm UVB}/\Delta$ where Δ is the local matter density. For the fluctuating UVB, $\Gamma_{\rm UVB}$ was calculated on a uniform grid of 64^3 at $z = 6$ and then linearly interpolated the $\log \Gamma_{\rm UVB}$ field to match the hydrodynamical simulation with 2048^3. The top right panel of Fig. 10 shows a slice through the UVB model with fluctuations. The largest UVB values are in the same location as the high density areas shown in the top left panel. These are the densest regions of the simulation which contain the majority of the sources of ionizing photons. We do not show the model without UVB fluctuations as this is a constant field.

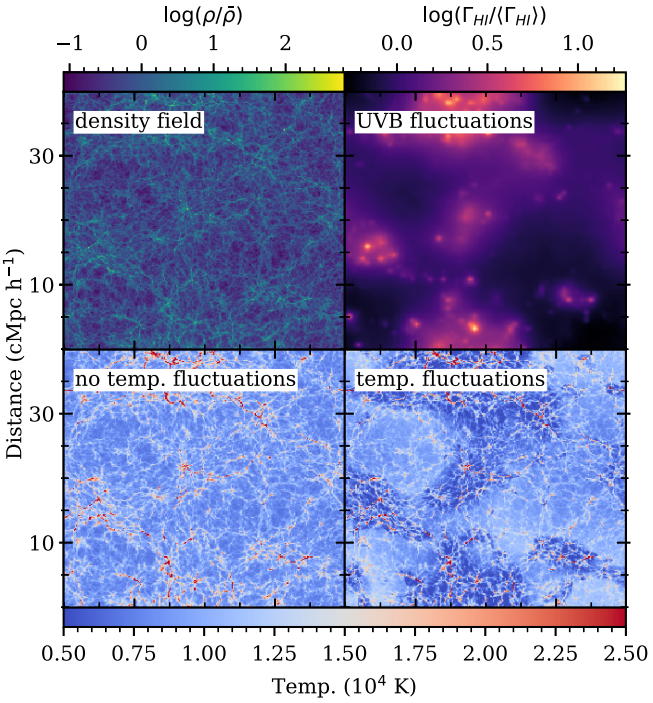

Figure 10. This figure shows slices of density field (top left), the temperature field (bottom row), and UVB (top right) for the Nyx simulation described in Section 4. The bottom left panel shows the temperature field without fluctuations. The bottom right panel right shows the temperature field with fluctuations. The model with temperature fluctuations has a greater scatter in the temperature field, as can be seen by the greater abundance of colder (darker blue) regions. These cold regions correspond to the regions of higher density in the top left panel. The top right panel shows a slice through the UVB field of the simulation with $\lambda_{\rm mfp} = 15$ cMpc, which gives a fluctuating UVB. The largest UVB values are in the same location as the high density areas shown in the top left panel. These are the densest regions of the simulation which contain the majority of the sources of ionizing photons.

Thus our four models of reionization are (1) no temperature fluctuations and no UVB fluctuations; (2) no temperature fluctuations with UVB fluctuations; (3) temperature fluctuations with no UVB fluctuations, and (4) both temperature and UVB fluctuations. All four models are normalized to $\langle F \rangle = .0172$, which is the measured value at $z = 5.8$ from Bosman et al. (2022). We do not consider multiple values of $\langle F \rangle$ for these models since they represent four discrete models and we will not try to constrain any parameters.

We now consider the effect of these four simulation models on the transmitted flux. Fig. 11 shows one skewer from each of the four different reionization models at $z = 5.8$. The top panel shows the resulting Lyα forest flux. The second panel shows the density field along the skewer. The third panel shows the temperature along the skewer. The bottom panel shows the UVB background values. Each panel has four lines representing models with no temperature and no UVB fluctuations (solid blue), no temperature fluctuations with UVB

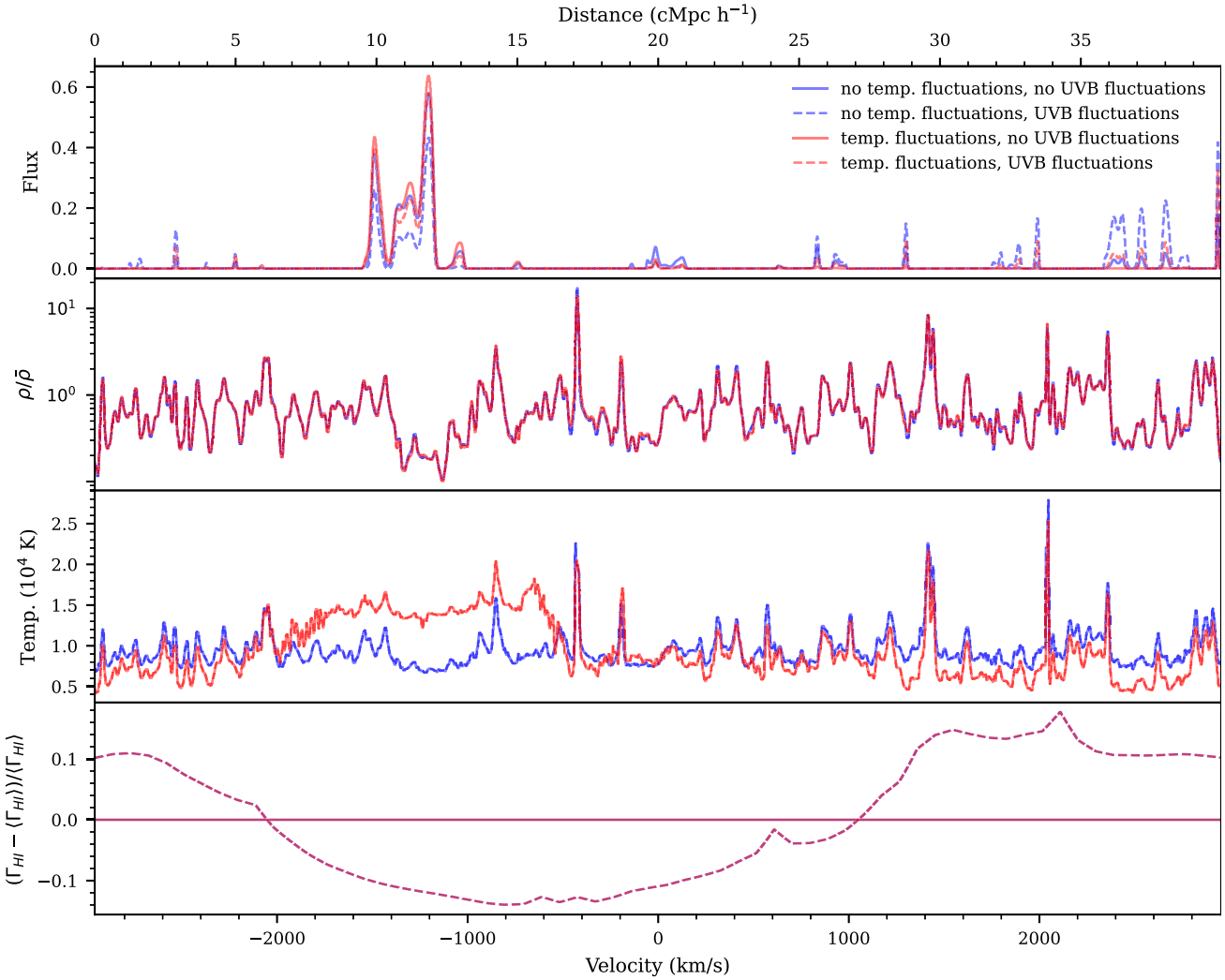

Figure 11. This figure shows one skewer from the four various reionization models at $z = 5.8$. The top panel shows the resulting Lyα forest flux. The second panel shows the density field along the skewer. The third panel shows the temperature along the skewer. The bottom panel shows the UVB background values. Each panel has four lines representing models with no temperature and no UVB fluctuations (solid blue), no temperature fluctuations with UVB fluctuations (dashed blue), temperature fluctuations with no UVB fluctuations (solid red), and both temperature and UVB fluctuations (dashed red). Comparing the solid lines to each other isolates the effect of temperature fluctuations only. When comparing these two models, we see that a positive scatter in the temperature of the IGM leads to increased flux over $-1600\,\mathrm{km\,s^{-1}} < v < -1000\,\mathrm{km\,s^{-1}}$. Comparing the dashed lines to the solid lines of the same colour isolates the effect of UVB fluctuations. For example consider $v > 1000\,\mathrm{km\,s^{-1}}$ where the models with UVB fluctuations (dashed) in the bottom panel are constantly greater than the models without UVB fluctuations (solid). In the top panel, these positive fluctuations in the UVB boost the flux in these dashed lines over the solid lines of the same colour.

fluctuations (dashed blue), temperature fluctuations with no UVB fluctuations (solid red), and both temperature and UVB fluctuations (dashed red). Comparing the solid lines to each other isolates the effect of temperature fluctuations only. When comparing these two models, we see that a positive scatter in the temperature of the IGM leads to increased flux over $-1600\,\mathrm{km\,s^{-1}} < v < -1000\,\mathrm{km\,s^{-1}}$. Comparing the dashed lines to the solid lines of the same colour isolates the effect of UVB fluctuations. For example consider $v > 1000\,\mathrm{km\,s^{-1}}$ where the models with UVB fluctuations (dashed) in the bottom panel are constantly greater than the models without UVB fluctuations (solid). In the top panel, these positive fluctuations in the UVB boost the flux in these dashed lines over the solid lines of the same colour.

In general, UVB fluctuations are anticorrelated with temperature fluctuations. This can be explained by the fact that dense regions in the simulations are cooler because they reionized earlier, but they have higher UVB values due to the increased number of ionizing photon sources, as discussed earlier. An example of this anticorrelation is seen at $-1600\,\mathrm{km\,s^{-1}} < v < -1000\,\mathrm{km\,s^{-1}}$, where there is a positive temperature fluctuation and a negative UVB fluctuation. Overall, this anticorrelation causes the effects of these two fluctuating fields to cancel each other out. As a result, the flux from the model with both temperature and UVB fluctuations (dashed red line) closely resembles the flux from the model with no temperature fluctuations and no UVB fluctuations (solid blue line). This behaviour is consistently observed in the flux panel of Fig. 11.

From here, we forward model the skewers in the same way as discussed in Section 2.2 with $R = 30\,000$ and $\mathrm{SNR}_{10} = 30$. The only difference is that we leave the skewers with the full 40 cMpc h^{-1} length and then use only 15 (where before we used 20) skewers when calculating mock data sets. The mock data sets here and in the previous section contain the same path length corresponding to 20 observed

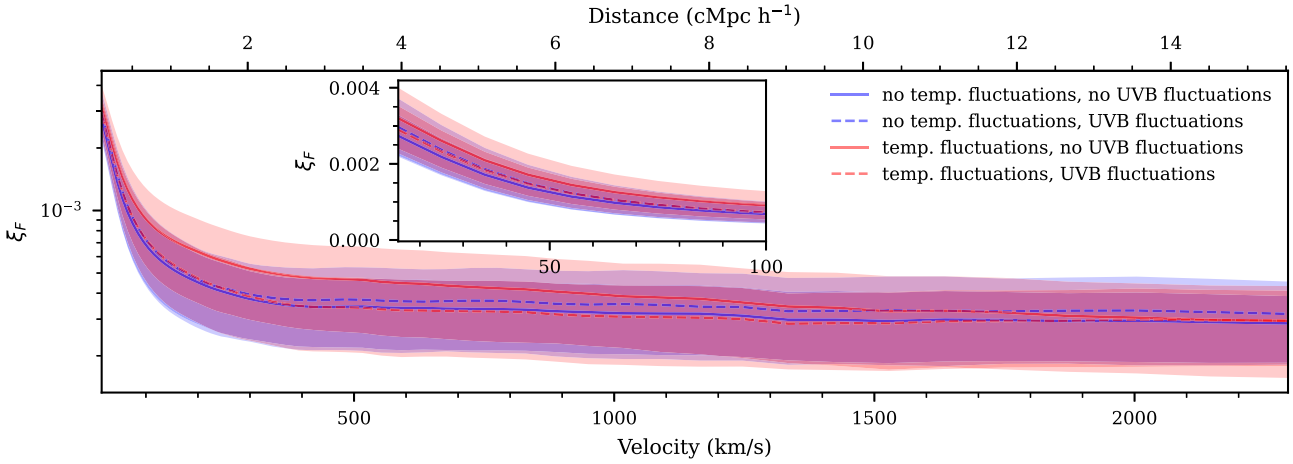

Figure 12. This figure shows the correlation function for the four reionization models at $z = 5.8$ with a logarithmic y-axis. The lines show the model values of the correlation function while the shaded region shows the errors estimated from the diagonal of the covariance matrices. The colours and line styles here match those in Fig. 11 with the model with no temperature fluctuations and no UVB fluctuations (solid blue), no temperature fluctuations with UVB fluctuations (dashed blue), temperature fluctuations with no UVB fluctuations (solid red), and both temperature and UVB fluctuations (dashed red). Comparing the red to the blue lines with the same style isolates the effect of temperature fluctuations while comparing the dashed to the solid line with the same colour isolates the effect of UVB fluctuations. Note that the shaded regions are about the same size for all four models. The inset shows the first 100 km s^{-1} of the autocorrelation functions with a linear y-axis to see the differences at small scales.

quasars with $\Delta z = 0.1$. We do not show an example of the forward modelled skewer here as they are very similar to that shown in Fig. 3.

4.2 Autocorrelation

The autocorrelation functions are computed via equation (3) and the covariance matrices are computed via equation (6).

Fig. 12 shows the correlation function for the four reionization models at $z = 5.8$ with a logarithmic y-axis. The inset shows the first 100 km s^{-1} of the autocorrelation functions with a linear y-axis to highlight the differences at small scales. The lines show the model value while the shaded regions are the error estimated from the diagonals of the covariance matrices. The colours and line styles here match those in Fig. 11 with the model with no temperature fluctuations and no UVB fluctuations (solid blue), no temperature fluctuations with UVB fluctuations (dashed blue), temperature fluctuations with no UVB fluctuations (solid red), and both temperature and UVB fluctuations (dashed red). Comparing the red to the blue lines with the same style isolates the effect of temperature fluctuations while comparing the dashed to the solid line with the same colour isolates the effect of UVB fluctuations. Note that the shaded regions are about the same size for all four models.

First compare the model with no temperature fluctuations and no UVB fluctuations (solid blue) and the model with temperature fluctuations with no UVB fluctuations (solid red), which isolates the effect of temperature fluctuations. The model values for these models show that adding temperature fluctuations boosts the value of the autocorrelation function for $\Delta v < 1800$ km s^{-1}. This follows from the additional variation added by the temperature fluctuations.

Now consider the model with no temperature fluctuations and no UVB fluctuations (solid blue) and the model with no temperature fluctuations with UVB fluctuations (dashed blue), which adds UVB fluctuations to a model without temperature fluctuations. Comparing these line in the inset shows that adding UVB fluctuations increases the value of the autocorrelation function on small scales. This result falls in line with that found in Wolfson et al. (2023) which says that a shorter $\lambda_{\rm mfp}$ value leads to greater boosts on small scales of the autocorrelation function. At larger scales there is a slight boost in the model with no temperature fluctuations with UVB fluctuations (dashed blue) seen with the logarithmic scale.

Finally consider the model with temperature fluctuations with no UVB fluctuations (solid red) and the model with both temperature and UVB fluctuations (dashed red), which compares adding UVB fluctuations to a model with temperature fluctuations. In this case adding UVB fluctuations decreases the value of the autocorrelation function for $\Delta v < 1800$ km s^{-1}. This is the opposite effect as adding UVB fluctuations to a model without temperature fluctuations (seen in comparing the blue lines) and the results from Wolfson et al. (2023). However, there is an anticorrelation between the UVB and temperature fluctuations resulting from the correlations with the density field. For a fluctuating UVB, the UVB is highest where the density is greatest, since this is where ionizing photon sources are located. For a fluctuating temperature model, the temperature is lowest where the density is greatest, which decreases the transmitted flux. This causes more constant flux levels and decreases the autocorrelation function values at these small scales, as seen in these lines. Ultimately, the correlations with density cause the model with both temperature and UVB fluctuations (dashed red) to be most similar to the model with no temperature fluctuations and no UVB fluctuations (solid blue). Note that on small scales there is still a boost in the model with both temperature and UVB fluctuations (dashed red) over the model without both fluctuations (solid blue), which comes from increased variation in the flux.

Note that the UVB models used in this section were computed in a relatively small box (40 cMpc h^{-1}). This box size was shown to suppress UVB fluctuations on all scales in Wolfson et al. (2023). As a result, suppressing UVB fluctuations leads to a lower autocorrelation signal at small scales. It is reasonable to assume that temperature fluctuations would also be suppressed in this box size, following an analogous argument to the UVB fluctuation suppression, though investigating this further is left for future work. Additionally, understanding the interplay between temperature and UVB fluctuations in a larger box is another aspect left to future research. This work will be crucial for assessing the realistic impact of these fluctuations on

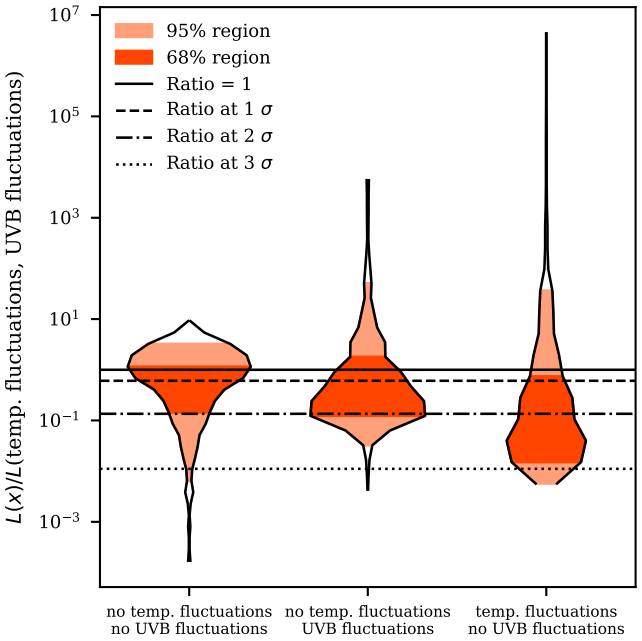

Figure 13. This figure shows the distribution of likelihood ratios from 1000 mock data sets where the mock data originates from the model with both temperature and UVB fluctuations. The violin plots show the full distribution where the light orange shaded region demarcates the 2.5th and 97.5th percentiles (2σ) of the ratio values while the darker orange shaded region demarcates the 16th and 84th percentiles (1σ) of the ratio values. The solid black line shows where the ratio is equal to 1, which is where both models are just as likely given the mock data. The dashed, dot-dashed, and dotted back lines show the value where the alternative models are ruled out at the 1, 2, and 3σ levels respectively.

observations of the Lyα forest at high-z and for understanding the potential of these observations to constrain reionization.

4.3 Ruling-out reionization scenarios

For these four reionization models, there is no grid of parameters that can be constrained via MCMC. Instead, we will investigate how confidently other models can be ruled out given mock data from a single model. We will rule out models via the likelihood ratio, $\mathcal{R}$, which is defined as

$$\mathcal{R} = \frac{\mathcal{L}(\text{model})}{\mathcal{L}(\text{reference model})} \tag{11}$$

Again for this we assume the likelihood, $\mathcal{L}$, is the multivariate Gaussian likelihood from equation (8).

Here we assume that the mock data comes from the model with both temperature and UVB fluctuations (red dashed lines in the Figs 11 and 12). Therefore, we will be looking at the value of the likelihood for the mock data sets using the other three reionization models divided by the likelihood for what we know is the true mock data model (with both temperature and UVB fluctuations). To investigate the distribution of potential likelihood ratio values, we use 1000 mock data sets.

The distribution of the 1000 likelihood ratio values for each of the alternative reionization models are shown in Fig. 13. The violin plots show the full distribution where the light orange shaded region demarcates the 2.5th and 97.5th percentiles (2σ) of the ratio values while the darker orange shaded region demarcates the 16th and 84th percentiles (1σ) of the ratio values. The solid black line shows where the ratio is equal to 1, which is where both models are just as likely given the mock data. The dashed, dot-dashed, and dotted back lines show the value where the alternative models are ruled out at the 1, 2, and 3 σ levels, respectively.

Overall, it is most difficult to rule out the model with no temperature fluctuations and no UVB fluctuations (solid blue lines in previous plots), as is seen in the left most violin in Fig. 13. This distribution has 44.6 per cent of the mock data sets that favour the incorrect, alternative reionization scenario than the true model with both temperature and UVB fluctuations. Then only 40.4 per cent, 17.4 per cent, and 3.4 per cent of mock data sets can be ruled out at the 1, 2, and 3 σ levels, respectively. This follows from the autocorrelation values for these models seen in Fig. 12 and the discussion there about how the temperature fluctuations and UVB fluctuations are anticorrelated and thus produce an autocorrelation function most similar to the model which lacks both of these fluctuations.

The next most difficult model to rule out is the model with no temperature fluctuations but with UVB fluctuations (dashed blue lines in the previous plot) as seen in the central violin in Fig. 13. This distribution has 26.5 per cent of the mock data sets that favour the incorrect, alternative reionization scenario than the true model with both temperature and UVB fluctuations. Then 60.6 per cent, 23.8 per cent, and only 0.3 per cent of mock data sets can be ruled out at the 1, 2, and 3 σ levels, respectively. Between this and the left plot there are fewer mock data sets here that can be ruled out at least at the 3σ level but over half of them can be ruled out at 1σ.

The easiest model to rule out is the model with temperature fluctuations but with no UVB fluctuations (solid red lines in the previous plots) as seen in the right most violin in Fig. 13. This distribution has only 21.8 per cent of the mock data sets that favour the incorrect, alternative reionization scenario than the true model with both temperature and UVB fluctuations. Then 73.9 per cent, 54.0 per cent, and 7.9 per cent of mock data sets can be ruled out at the 1, 2, and 3 σ levels, respectively, which is the greatest percentages out of the three alternative models. This also follows from the differences between these models in Fig. 12. The model with temperature fluctuations but no UVB fluctuations has the greatest values of the autocorrelation function at most scales, making it the easiest to distinguish.

This is the distribution of the likelihood ratio for 1000 mock data sets. The actual given observational data set would ultimately determine if it is possible to rule out each model. It is possible for the incorrect models to be favoured over the true model from which the mock data was drawn, though this was always true for less than half of the mock data sets.

5 CONCLUSIONS

In this work we investigated the precision of possible constraints on the thermal state of the IGM from the autocorrelation function of Lyα forest flux in high resolution quasar observations. This came in two forms: constraining T_0 and γ when the IGM thermal state follows a tight power law of the form of equation (1) and investigating the likelihood ratio for models with temperature fluctuations from different reionization scenarios.

We discussed the constraints on T_0 and γ in Section 3. Overall, we found that the autocorrelation function is sensitive to both T_0 and γ across multiple redshift bins for realistic mock data sets of 20 quasars with a resolution of $R = 30\,000$. We computed the marginalized re-weighted posteriors for 100 mock data sets at $5.4 \leq z \leq 6.0$. These re-weighted posteriors showed a variety of behaviours, depending on the specific mock data set chosen and the true values of T_0 and γ

for the mock data set. We also considered an idealized data set with the model value of the autocorrelation function and a covariance matrix derived from 20 quasar sightlines. The errors on both T_0 and γ increase with redshift, which may come from the lower values of $\langle F \rangle$ at higher z. At $z = 5.4$, the idealized data constrained T_0 to 59 per cent and γ to 16 per cent. At higher redshifts, these constraints weaken. Specifically at $z = 6.0$, the idealized data constrained T_0 to 119 per cent and γ to 30 per cent.

Note that our procedure uses a multivariate Gaussian likelihood, MCMC, and a set of weights for the MCMC chains that ensures our posteriors pass an inference test. This guarantees that our reported errors are statistically correct and can be trusted. The initial failure of our procedure to pass this test arose from the incorrect assumption that the autocorrelation function follows a multivariate Gaussian distribution, as discussed in Appendix C. This highlights the caution needed when using a multivariate Gaussian likelihood with statistics of the Lyα forest at high z, including both the autocorrelation function and the power spectrum, as similar issues with non-Gaussian data may arise. Improved likelihoods or likelihood-free inference methods may offer a more optimal inference procedure (see e.g. Davies et al. 2018a; Alsing et al. 2019). We have left the exploration of these methods to future work.

We discussed the likelihood ratios for four different reionization models in Section 4, assuming a Gaussian distribution of the data. Looking at mock data from a model which has temperature and UVB fluctuations, we found that it is easiest to rule out a model with temperature fluctuations and no UVB fluctuations and it is most difficult to rule out a model with no temperature or UVB fluctuations. The actual ability to distinguish between models depends on the specific mock data set considered. In the most difficult case, we found that 40.4 per cent of mock data sets could rule out a model without temperature or UVB fluctuations at $> 1\sigma$ level. In the easiest case, we found that 73.9 per cent of mock data sets could rule out a model with only temperature fluctuations at $> 1\sigma$ level.

As discussed in Section 4.2, the temperature and UVB fluctuations were generated in a relatively small simulation box, which may suppress their impact on the autocorrelation function. Wolfson et al. (2023) demonstrated this suppression of the autocorrelation function on small scales for UVB fluctuations in a small box size. Consequently, distinguishing between models with and without UVB fluctuations would likely be easier if they were generated in a larger simulation box. The analogous effect on temperature fluctuations from a larger box has not yet been studied. Therefore, future work on fluctuating temperature and UVB models will be essential for obtaining the best possible constraints on reionization.

Both the thermal state and the UVB fluctuations affect the Lyα forest flux autocorrelation function. Modelling both of these physical effects by varying multiple parameters in a larger box will allow the autocorrelation function to constrain the two simultaneously. This will allow us to put quantitative constraints on the thermal state of the IGM, the $\lambda_{\rm mfp}$ that describes the UVB, and ultimately reionization. We leave this exploration to future work.

This work assumed 20 high-resolution quasar observations in our forecasting. There are currently over 100 known quasars above a redshift of 6, a subset of which already have high resolution spectroscopic observations. Thus the 20 quasars used in this work is reasonable for a near-future observational constraint. In addition, the number of known quasars with high resolution observations is expected to continue to grow in the coming years which would only improve the prospects of this constraint.

Here we used the autocorrelation function of the Lyα forest flux. We chose to look at this clustering statistic for a couple statistical properties: namely that uncorrelated noise averages out and spectral masking is easy to implement. In comparison to the power spectra, the autocorrelation function has a covariance matrix with large off-diagonal values which makes it more difficult to intuitively look at the resulting fits to data (e.g. Fig. 6). In addition to the intuition, these large off-diagonal values also make it more difficult to trouble-shoot the inference procedure using a Gaussian likelihood. Using a statistic with a more diagonal covariance matrix, like the power spectra, is easier to implement when fitting data.

Constraining the thermal state of the IGM at $z > 5$ is a crucial method for probing the end of reionization. However, measuring the thermal state at these redshifts is challenging, and only a few methods have been employed thus far (Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2020). This work demonstrates that the autocorrelation function of the Lyα forest flux provides a new, competitive approach to constrain the thermal state across multiple redshift bins at $z \geq 5.4$. Specifically, it does so with a simplified thermal state model characterized by a tight power-law relation with parameters T_0 and γ. This represents an essential first step toward using this method to constrain the IGM's thermal state at high redshifts.

ACKNOWLEDGEMENTS

We acknowledge helpful conversations with the ENIGMA group at UC Santa Barbara and Leiden University. JFH acknowledges support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 885301) and from the National Science Foundation under Grant No. 1816006. JO acknowledges support from grants CNS2022-135878 and PID2022-138855NB-32 funded by MICIU/AEI/10.13039/501100011033 and EU NextGenerationEU/PRTR and project PPIT2024-31833 funded by UE Ministerio de Hacienda y Función Pública–Fondos Europeos–Junta de Andalucía–Consejería de Universidad, Investigación e Innovación.

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231.

DATA AVAILABILITY

The simulation data analysed in this article will be shared on reasonable request to the corresponding author.

REFERENCES

Almgren A. S., Bell J. B., Lijewski M. J., Lukić Z., Van Andel E., 2013, ApJ, 765, 39
Alsing J., Charnock T., Feeney S., Wandelt B., 2019, MNRAS, 488, 4440
Becker G. D., Sargent W. L. W., Rauch M., 2004, ApJ, 613, 61
Becker G. D., Rauch M., Sargent W. L. W., 2007, ApJ, 662, 72
Becker G. D., Bolton J. S., Haehnelt M. G., Sargent W. L. W., 2011, MNRAS, 410, 1096
Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, MNRAS, 447, 3402
Becker G. D., Davies F. B., Furlanetto S. R., Malkan M. A., Boera E., Douglass C., 2018, ApJ, 863, 92
Becker G. D., D'Aloisio A., Christenson H. M., Zhu Y., Worseck G., Bolton J. S., 2021, MNRAS, 508, 1853
Boera E., Murphy M. T., Becker G. D., Bolton J. S., 2014, MNRAS, 441, 1916
Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, ApJ, 872, 101

Bolton J. S., Viel M., Kim T.-S., Haehnelt M. G., Carswell R. F., 2008, MNRAS, 386, 1131
Bolton J. S., Becker G. D., Wyithe J. S. B., Haehnelt M. G., Sargent W. L. W., 2010, MNRAS, 406, 612
Bolton J. S., Becker G. D., Raskutti S., Wyithe J. S. B., Haehnelt M. G., Sargent W. L. W., 2012, MNRAS, 419, 2880
Bolton J. S., Becker G. D., Haehnelt M. G., Viel M., 2014, MNRAS, 438, 2499
Bosman S. E. I., 2021, preprint (arXiv:2108.12446)
Bosman S. E. I., Fan X., Jiang L., Reed S., Matsuoka Y., Becker G., Haehnelt M., 2018, MNRAS, 479, 1055
Bosman S. E. I. et al., 2022, MNRAS, 514, 55
Bryan G. L., Machacek M. E., 2000, ApJ, 534, 57
Calura F., Tescari E., D'Odorico V., Viel M., Cristiani S., Kim T. S., Bolton J. S., 2012, MNRAS, 422, 3019
Croft R. A. C., 2004, ApJ, 610, 642
D'Aloisio A., McQuinn M., Trac H., 2015, ApJ, 813, L38
D'Aloisio A., McQuinn M., Davies F. B., Furlanetto S. R., 2018, MNRAS, 473, 560
D'Aloisio A., McQuinn M., Maupin O., Davies F. B., Trac H., Fuller S., Sanderbeck P. R. U., 2019, ApJ, 874, 154
D'Odorico V. et al., 2006, MNRAS, 372, 1333
Davies F. B., Furlanetto S. R., 2016, MNRAS, 460, 1328
Davies F. B., Furlanetto S. R., McQuinn M., 2016, MNRAS, 457, 3006
Davies F. B., Hennawi J. F., Eilers A.-C., Lukić Z., 2018a, ApJ, 855, 106
Davies F. B., Becker G. D., Furlanetto S. R., 2018b, ApJ, 860, 155
Eilers A.-C., Davies F. B., Hennawi J. F., 2018, ApJ, 864, 53
Fan X. et al., 2006, AJ, 132, 117
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gaikwad P. et al., 2020, MNRAS, 494, 5091
Gaikwad P. et al., 2023, MNRAS, 525, 4093
Garzilli A., Bolton J. S., Kim T. S., Leach S., Viel M., 2012, MNRAS, 424, 1723
Gnedin N. Y., Hui L., 1998, MNRAS, 296, 44
Gnedin N. Y., Becker G. D., Fan X., 2017, ApJ, 841, 26
Gontcho A Gontcho S., Miralda-Escudé J., Busca N. G., 2014, MNRAS, 442, 187
Grünwald P., van Ommen T., 2017, Bayesian Anal., 12, 1069
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Haardt F., Madau P., 2012, ApJ, 746, 125
Haehnelt M. G., Steinmetz M., 1998, MNRAS, 298, L21
Hennawi J. F., Kist T., Davies F. B., Tamanas J., 2025, MNRAS, 539, 2621
Hiss H., Walther M., Hennawi J. F., Oñorbe J., O'Meara J. M., Rorai A., Lukić Z., 2018, ApJ, 865, 42
Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27
Hui L., Haiman Z., 2003, ApJ, 596, 9
Jung I. et al., 2020, ApJ, 904, 144
Kashino D., Lilly S. J., Shibuya T., Ouchi M., Kashikawa N., 2020, ApJ, 888, 6
Kulkarni G., Hennawi J. F., Oñorbe J., Rorai A., Springel V., 2015, ApJ, 812, 30
Lee K.-G. et al., 2015, ApJ, 799, 196
Lidz A., Malloy M., 2014, ApJ, 788, 175
Lidz A., Faucher-Giguère C.-A., Dall'Aglio A., McQuinn M., Fechner C., Zaldarriaga M., Hernquist L., Dutta S., 2010, ApJ, 718, 199
Lukić Z., Stark C. W., Nugent P., White M., Meiksin A. A., Almgren A., 2015, MNRAS, 446, 3697
Lynds R., 1971, ApJ, 164, L73
McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, ApJ, 543, 1
McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., 2001, ApJ, 562, 52
McDonald P., Seljak U., Cen R., Bode P., Ostriker J. P., 2005, MNRAS, 360, 1471
McGreer I. D., Mesinger A., Fan X., 2011, MNRAS, 415, 3237
McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499
McQuinn M., 2012, MNRAS, 426, 1349
McQuinn M., Upton Sanderbeck P. R., 2016, MNRAS, 456, 47
Meiksin A., McQuinn M., 2019, MNRAS, 482, 4777
Meiksin A., White M., 2004, MNRAS, 350, 1107
Mesinger A., Furlanetto S., 2009, MNRAS, 400, 1461
Miralda-Escudé J., Rees M. J., 1994, MNRAS, 266, 343
Morales A. M., Mason C. A., Bruton S., Gronke M., Haardt F., Scarlata C., 2021, ApJ, 919, 120
Morrison J., Simon N., 2018, J. Comput. Graph. Stat., 27, 648
Nasir F., Bolton J. S., Becker G. D., 2016, MNRAS, 463, 2335
Oñorbe J., Hennawi J. F., Lukić Z., 2017, ApJ, 837, 106
Oñorbe J., Davies F. B., Lukić Z., Hennawi J. F., Sorini D., 2019, MNRAS, 486, 4075
Oñorbe J., Hennawi J. F., Lukić Z., Walther M., 2017, ApJ, 847, 63
Planck Collaboration VI, 2020, A&A, 641, A6
Pontzen A., 2014, Phys. Rev. D, 89, 083010
Pontzen A., Bird S., Peiris H., Verde L., 2014, ApJ, 792, L34
Prangle D., Blum M. G. B., Popovic G., Sisson S. A., 2014, Aust. New Zealand J. Stat., 56, 309
Puchwein E., Bolton J. S., Haehnelt M. G., Madau P., Becker G. D., Haardt F., 2015, MNRAS, 450, 4081
Rauch M., 1998, ARA&A, 36, 267
Ricotti M., Gnedin N. Y., Shull J. M., 2000, ApJ, 534, 41
Rollinde E., Petitjean P., Pichon C., Colombi S., Aracil B., D'Odorico V., Haehnelt M. G., 2003, MNRAS, 341, 1279
Rorai A. et al., 2017, Science, 356, 418
Rorai A., Carswell R. F., Haehnelt M. G., Becker G. D., Bolton J. S., Murphy M. T., 2018, MNRAS, 474, 2871
Rudie G. C., Steidel C. C., Pettini M., 2012, ApJ, 757, L30
Schaye J., Theuns T., Leonard A., Efstathiou G., 1999, MNRAS, 310, 57
Schaye J., Theuns T., Rauch M., Efstathiou G., Sargent W. L. W., 2000, MNRAS, 318, 817
Sellentin E., Starck J.-L., 2019, J. Cosmol. Astropart. Phys., 2019, 021
Theuns T., Schaye J., Haehnelt M. G., 2000, MNRAS, 315, 600
Theuns T., Schaye J., Zaroubi S., Kim T.-S., Tzanavaris P., Carswell B., 2002, ApJ, 567, L103
Upton Sanderbeck P. R., D'Aloisio A., McQuinn M. J., 2016, MNRAS, 460, 1885
Viel M., Bolton J. S., Haehnelt M. G., 2009, MNRAS, 399, L39
Walther M., Hennawi J. F., Hiss H., Oñorbe J., Lee K.-G., Rorai A., O'Meara J., 2017, ApJ, 852, 22
Walther M., Oñorbe J., Hennawi J. F., Lukić Z., 2019, ApJ, 872, 13
Wolfson M., Hennawi J. F., Davies F. B., Oñorbe J., Hiss H., Lukić Z., 2021, MNRAS, 508, 5493
Wolfson M. et al., 2024, MNRAS, 531, 3069
Wolfson M., Hennawi J. F., Davies F. B., Oñorbe J., 2023, MNRAS, 521, 4056
Yang J. et al., 2020, ApJ, 904, 26
Yèche C., Palanque-Delabrouille N., Baur J., du Mas des Bourboux H., 2017, J. Cosmol. Astropart. Phys., 2017, 047
Zaldarriaga M., Hui L., Tegmark M., 2001, ApJ, 557, 519
Zhu Y. et al., 2023, ApJ, 955, 115
Ziegel J. F., Gneiting T., 2014, Electron. J. Stat., 8, 2619
Zuo L., 1992a, MNRAS, 258, 45
Zuo L., 1992b, MNRAS, 258, 36

APPENDIX A: POWER SPECTRA MODELS

As explained in Section 3.1, the dimension-less power, $\Delta^2_{\delta_f}(k)$, can be written as the Fourier transform of the autocorrelation function of the flux contrast, $\xi_{\delta_f}(\Delta v)$. $\xi_{\delta_f}(\Delta v)$ is explicitly written in terms of $\Delta^2_{\delta_f}(k)$ in equation (5), which says $\xi_{\delta_f}(\Delta v)$ is the integral of $\Delta^2_{\delta_f}(k)\cos(k\Delta v)$ in logarithmic k bins. We refer to $\Delta^2_{\delta_f}(k)\cos(k\Delta v)$ as the integrand for the rest of this discussion. To build intuition for the autocorrelation function at small scales we show the integrand for $\Delta v = r = 15\,\mathrm{km\,s^{-1}}$ in Fig. A1.

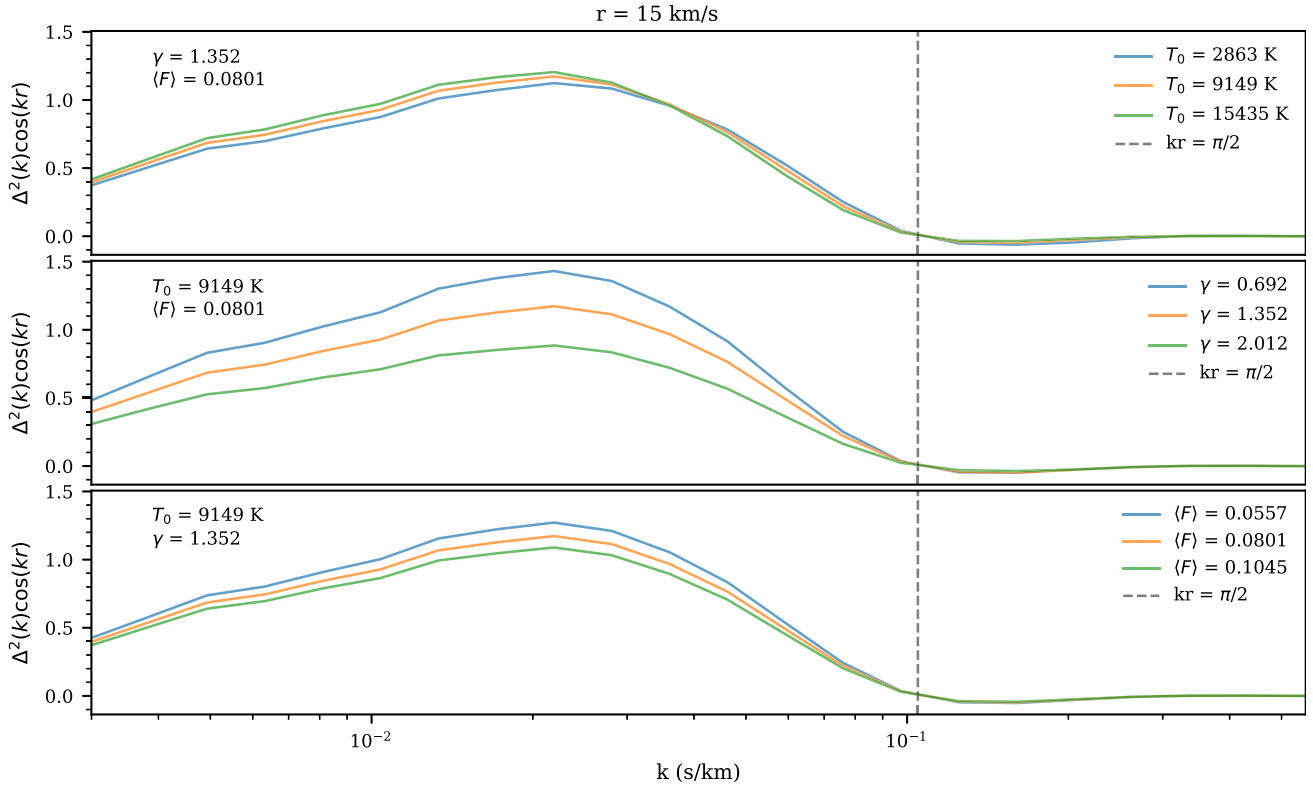

Figure A1. This figure shows the mean value of $\Delta^2 \cos(kr)$ where $r = 15\,\mathrm{km\,s^{-1}}$ for different sets of parameters. Each panels varies one parameter while keeping the others constant with T_0, γ, and $\langle F \rangle$ varying in the top, middle, and bottom panels, respectively. For these panels the solid lines show the model values calculated by averaging Δ^2 from all forward modelled skewers available. This figure is meant to explain the behaviour of the autocorrelation seen in Fig. 4 at $\Delta v = r = 15\,\mathrm{km\,s^{-1}}$ due to the relation in equation (5).

This figure mimics the set up of Fig. 4 for the autocorrelation function where each panels varies one parameter while keeping the others constant. For these panels the solid lines show the model values calculated by averaging Δ^2 from all forward modelled skewers available. The vertical grey dashed line shows where $\cos(kr) = 0$.

In the top panel T_0 varies while γ and $\langle F \rangle$ are constant. At small k the larger values of T_0 have larger values of the integrand while at small k there is thermal cutoff and smaller values of T_0 now have larger values of the integrand. When integrating over these logarithmic bins the greater T_0 values end up with more area and thus the autocorrelation functions are also greater.

APPENDIX B: CONVERGENCE OF THE COVARIANCE MATRICES

We calculate the covariance matrices for our models with mock draws, as defined in equation (6). Using mock draws is inherently noisy and it should converge as $1/\sqrt{N}$ where N is the number of draws used. As stated in the text, we used 500 000 mock draws. To check that this number is sufficient to minimize the error in our calculation, we looked at the behaviour of elements of one covariance matrix in Fig. B1. This covariance matrix is for the model with $z = 5.4$, $T_0 = 9149$ K, $\gamma = 1.352$, and $\langle F \rangle = 0.0801$, which is the central model at this redshift. The correlation matrix for this model is also shown in Fig. 5.

The values in the plot have been normalized to 1 at 10^6 draws. The four elements have been chosen such that there is one diagonal value and three off-diagonal values in different regions of the matrix. At all values of the number of mock draws considered, the covariance elements fall within 2 per cent of their final value. By around $\sim$ 100 000 draws, the values fall within 0.5 per cent of the final value. For this reason, using 500 000 mock draws is sufficient to generate the covariance matrices used in this study. In Fig. B1, 500 000 mock draws is represented by the vertical dashed grey line.

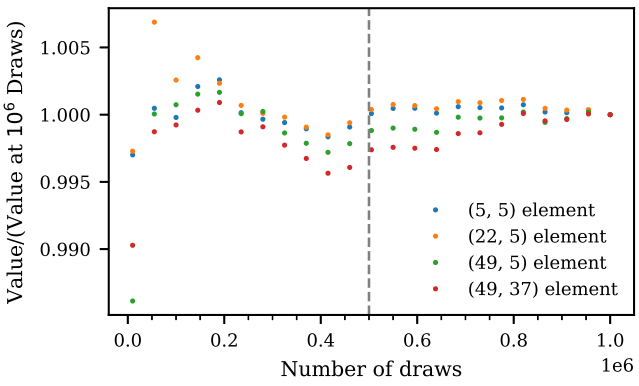

Figure B1. This figure shows the behaviour of four elements of the model covariance matrix ($z = 5.4$, $T_0 = 9149$ K, $\gamma = 1.352$, and $\langle F \rangle = 0.0801$) for different numbers of mock draws. At all values of the number of mocks considered, the covariance elements fall within 2 per cent of their final value. By around $\sim$ 100 000 draws, all of the values fall within 0.5 per cent of the final value. For this reason, using 500 000 mock draws is sufficient to generate the covariance matrices used in this study. 500 000 mock draws is represented by the vertical dashed grey line.

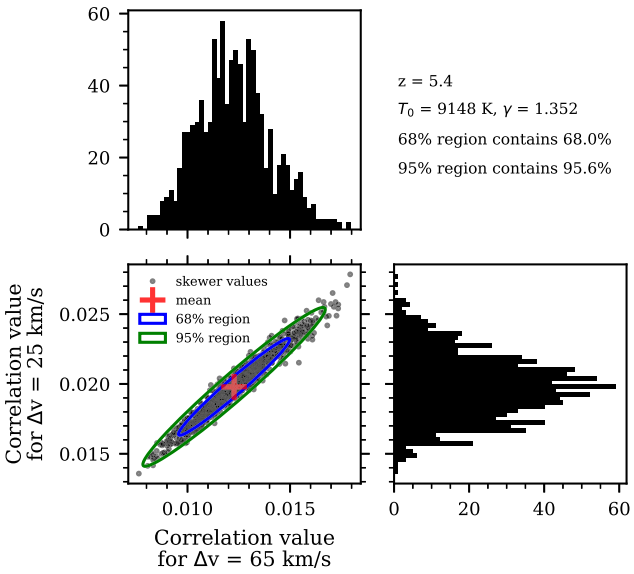

Figure C1. This figure shows the distribution 1000 mock draws from two bins of the autocorrelation function ($\Delta v = 25.0\,\mathrm{km\,s^{-1}}$ and $\Delta v = 65.0\,\mathrm{km\,s^{-1}}$) for one model ($z = 5.4$, $T_0 = 9148$ K, $\gamma = 1.352$, and $\langle F \rangle = 0.0801$). The top panel shows the distribution of only the $\Delta v = 65.0\,\mathrm{km\,s^{-1}}$ bin while the right panel shows the distribution of only the $\Delta v = 25.0\,\mathrm{km\,s^{-1}}$ bin. The blue (green) circle represents the 68 per cent (95 per cent) ellipse calculated from the covariance matrix calculated for this model from equation (6). The red plus shows the calculated mean. Additionally the per cent of mock draws that fall within each of these contours is written in the top right. Both the 1D and 2D distributions seem relatively well described by a Gaussian distribution. In the 2D plot, there are more points outside the 95 per cent contour to the top right than on any other side but it is not extreme.

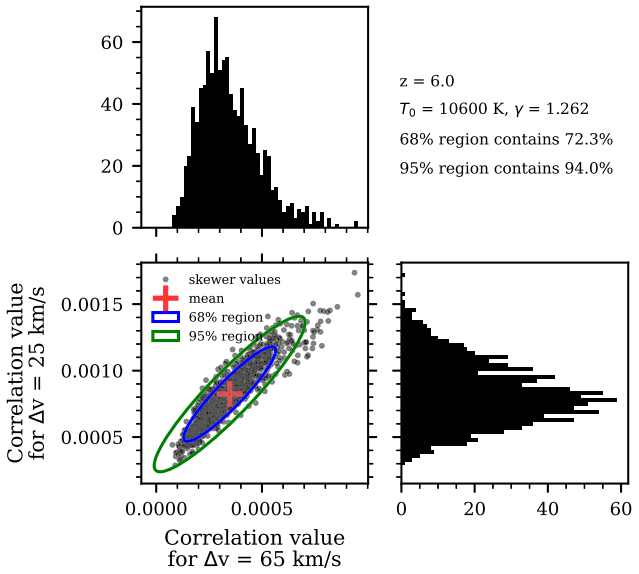

Figure C2. This figure shows the distribution 1000 mock draws from two bins of the autocorrelation function ($\Delta v = 25.0\,\mathrm{km\,s^{-1}}$ and $\Delta v = 65.0\,\mathrm{km\,s^{-1}}$) for one model ($z = 6$, $T_0 = 10\,600$ K, $\gamma = 1.262$, and $\langle F \rangle = 0.0089$). The top panel shows the distribution of only the $\Delta v = 65.0\,\mathrm{km\,s^{-1}}$ bin while the right panel shows the distribution of only the $\Delta v = 25.0\,\mathrm{km\,s^{-1}}$ bin. The blue (green) circle represents the 68 per cent (95 per cent) ellipse calculated from the covariance matrix calculated for this model from equation (6). The red plus shows the calculated mean. Additionally the per cent of mock draws that fall within each of these contours is written in the top right. Both the 1D and 2D distributions are not well described by a Gaussian with 72.3 per cent of the mock draws falling within the 68 per cent contour and 94. per cent of the mock draws falling within the 95 per cent contour.

APPENDIX C: NON-GAUSSIAN DISTRIBUTION OF THE VALUES OF THE AUTOCORRELATION FUNCTION

For our inference, we used the multivariate Gaussian likelihood defined in equation (8). This functional form assumes that the distribution of mock draws of the autocorrelation function is Gaussian distributed about the mean for each bin. In order to visually check this we will look at the distribution of mock draws from two bins of the autocorrelation function for two different models.

Both Figs C1 and C2 show the distribution of 1000 mock data sets from the velocity bins of the autocorrelation function with $\Delta v = 25.0\,\mathrm{km\,s^{-1}}$ and $\Delta v = 65.0\,\mathrm{km\,s^{-1}}$. The bottom left panels show the 2D distribution of the autocorrelation values from these bins. The blue (green) ellipses represents the theoretical 68 per cent (95 per cent) percentile contour calculated from the covariance matrix calculated for each model from equation (6). The red crosses shows the calculated mean. The top panels show the distribution of only the $v = 65.0\,\mathrm{km\,s^{-1}}$ bins while the right panels show the distribution of only the $v = 25.0\,\mathrm{km\,s^{-1}}$ bins.

Fig. C1 shows mock values of two bins of the autocorrelation function for the model at $z = 5.4$ with $T_0 = 9148$ K, $\gamma = 1.352$, and $\langle F \rangle = 0.0801$. Both the 1D and 2D distributions seem relatively well described by Gaussian distributions by eye though they do show some evidence of non-Gaussian tails to larger values. The number of points falling in each contour both fall within 1 per cent of the expected values. In the bottom left panel with the 2D distribution there are more mock values falling outside the 95 per cent contour to the top right (higher values) than in any other direction. For this reason the distribution is not exactly Gaussian but a Gaussian visually appears as an acceptable approximation.

Fig. C2 shows mock values of two bins of the autocorrelation function for the model at $z = 6$ with $T_0 = 10\,600$ K, $\gamma = 1.262$, and $\langle F \rangle = 0.0089$. In both the top and right panels, which show the distribution of values for one bin of the autocorrelation function, the distribution of mock draws is skewed with tails to the right. This is quantified by the per cent of points in the two ellipses from the bottom left panel labelled in the top right with 72.3 per cent of the mock draws falling within the 68 per cent contour and 94.0 per cent of the mock draws falling within the 95 per cent contour. The points outside of the contours are highly skewered towards the top right (higher values). It is only possible for the autocorrelation function to be negative due to noise, which generally averages to very small values approaching zero at the non-zero lags of the autocorrelation function. This can be seen in the black points and histogram do not go below 0, though the 95 per cent ellipse shown in green in the bottom left panel does go negative for $\Delta v = 65\,\mathrm{km\,s^{-1}}$.

Figs C1 and C2 show the changing distribution of the autocorrelation value with z, T_0, γ, and $\langle F \rangle$. There is a greater deviation from a multivariate Gaussian distribution at higher z. It is possible that adding additional sightlines will cause the autocorrelation function to better follow a multivariate Gaussian distribution due to the central limit theorem, though investigating this in detail is beyond the scope of the paper. However, even with more sightlines $\langle F \rangle$ will be low at high-z so we still expect the distribution to be skewed as the values mainly will not fall below 0. The incorrect assumption of the multivariate Gaussian likelihood thus contributes to the failure

of our method to pass an inference test as discussed in Section 3.3 for $z = 5.4$ and Appendix D for $z = 6$. For our final constraints, we calculated weights for our MCMC chains such that the resulting posteriors do pass our inference test, as discussed in Section 3.3. The whole method of assuming a multivariate Gaussian then re-weighting the posteriors in non-optimal and future work using a more correct likelihood or likelihood-free inference will improve our results.

APPENDIX D: INFERENCE TEST AT HIGH REDSHIFT

Here we present the results of the inference test at $z = 6$. This calculation was done following the procedure described in Section 3.3. Fig. D1 shows the results for $z = 6$ and can be compared to the $z = 5.4$ results in Fig. 7. The left panel here shows the initial coverage plot which deviates greatly from the expected $C(\alpha) = \alpha$ line, much more so than the $z = 5.4$. This likely comes from a greater deviation from the assumption of a multivariate Gaussian likelihood as described in Appendix C. The $z = 6$ mock data show highly skewed distributions that are not well described by a Gaussian likelihood. The inference lines at other redshifts are available upon request.

APPENDIX E: GAUSSIAN DATA INFERENCE TEST

As shown in Appendix C, the distribution of mock values of the autocorrelation function is not exactly Gaussian distributed. In order to confirm the failure of our mock data to pass an inference test (as discussed in Section 3.3 and Appendix D) comes from the use of a multivariate Gaussian likelihood, we generate Gaussian-distributed data and run inference tests. For one value of T_0, γ, and $\langle F \rangle$, we randomly generate a mock data set from a multivariate Gaussian with the given mean model and covariance matrix calculated from our mock data as described in Section 3.1. We can then continue with the inference test as described in Section 3.3. The results for this inference test for $z = 5.4$ and $z = 6.0$ are shown in Fig. E1. Here both redshifts inference lines fall along the $C(\alpha) = \alpha$ line that is expected for all α values. This behaviour is also seen at the other redshifts. The fact that perfectly Gaussian data passes an inference test with the same likelihood, priors, and method as was used on mock data confirms that the failure of our mock data to pass an inference test is due to the non-Gaussian distribution of the mock data.

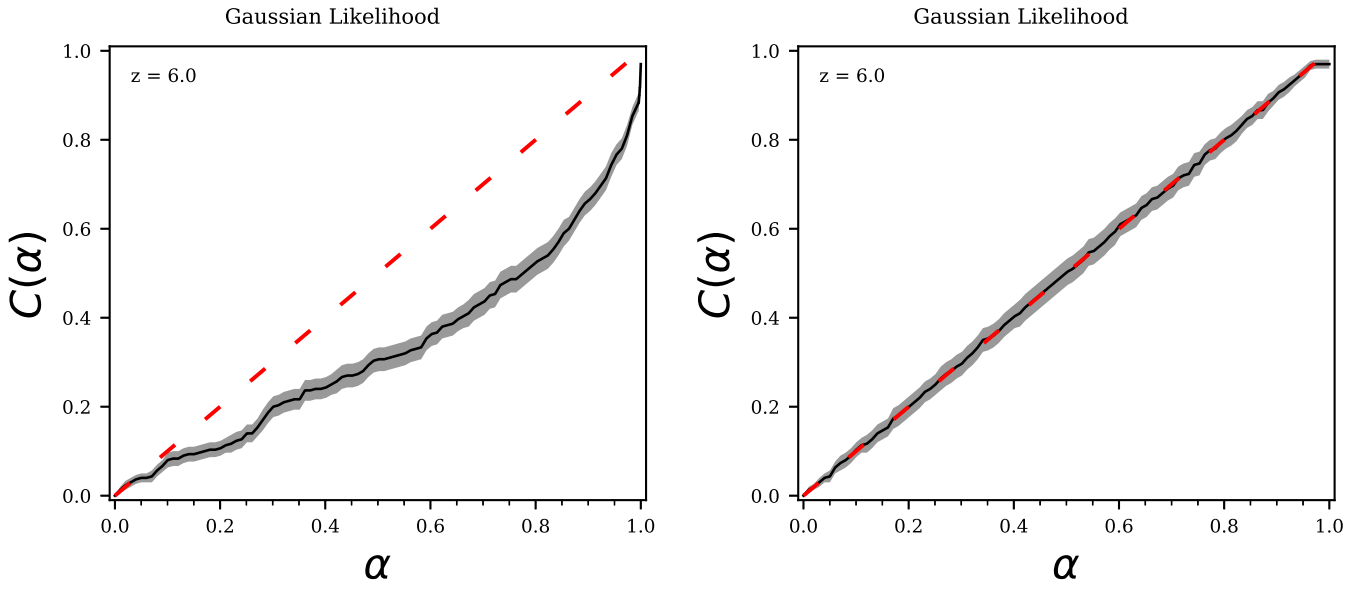

Figure D1. The left panel of this figure shows the coverage resulting from the inference test from 300 models at $z = 6$. drawn from our priors on T_0, γ, and $\langle F \rangle$. Here we see that the true parameters for the models fall above the 60th percentile in the MCMC chain ~ 35 per cent of the time, for example. The right panel of this figure shows the coverage resulting from the inference test with the use of one set of weights to re-weight the posteriors. With these weights we are able to pass the inference test.

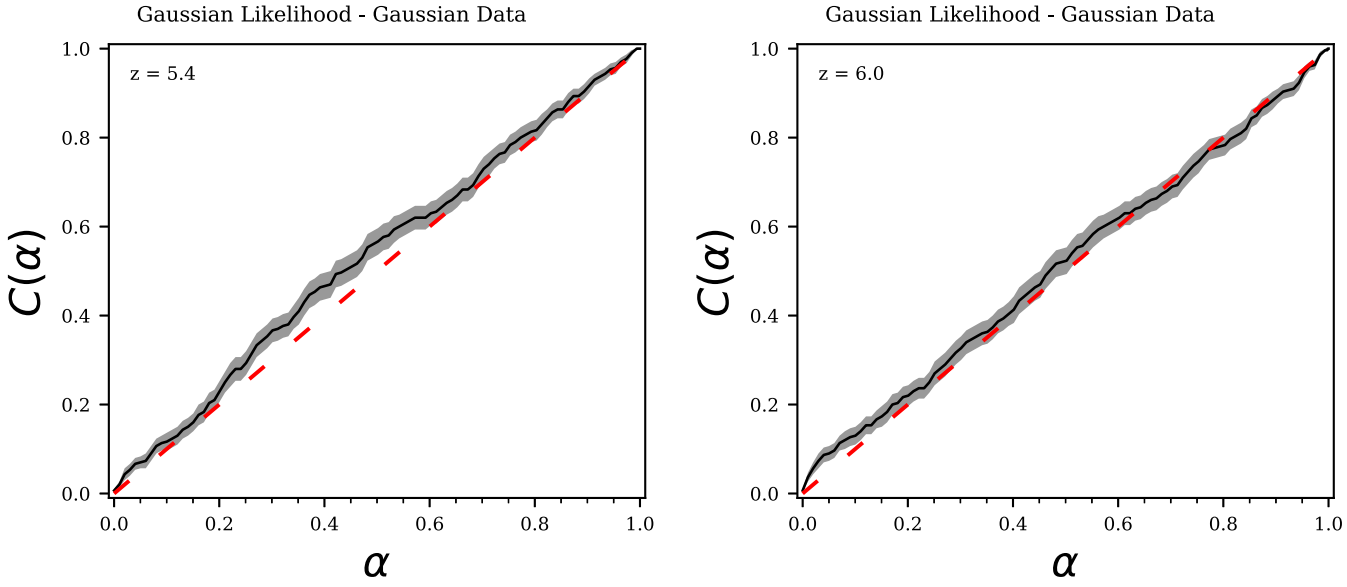

Figure E1. Both panels of this figure show the coverage plot resulting from the inference test from 300 data sets generated by randomly drawing points from the mean model and covariance matrix. The means and covariance matrices used come from $z = 5.4$ in the left panel and $z = 6.0$ in the right panel. The true parameter values for both panels were drawn from our priors on T_0, γ, and $\langle F \rangle$. In both panels, the Gaussian mock data produced inference lines that fall on top of the $C(\alpha) = \alpha$ line within errors, as expected for the statistically correct posteriors.

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