


Possible evidence for a large-scale enhancement in the Lyman- α forest power spectrum at redshift $z \geq 4$

Margherita Molaro¹ , Vid Iršič² , James S. Bolton¹ , Maggie Lieu¹ , Laura C. Keating³ ,
Ewald Puchwein⁴ , Martin G. Haehnelt²  & Matteo Viel^{5,6,7,8} 

¹*School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*

²*Kavli Institute for Cosmology and Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK*

³*Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, UK*

⁴*Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany*

⁵*SISSA - International School for Advanced Studies, Via Bonomea 265, I-34136 Trieste, Italy*

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

⁷*INAF - Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, I-34131 Trieste, Italy*

⁸*INFN - National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Inhomogeneous reionization enhances the 1D Ly α forest power spectrum on large scales at redshifts $z \geq 4$. This is due to coherent fluctuations in the ionized hydrogen fraction that arise from large-scale variations in the post-reionization gas temperature, which fade as the gas cools. It is therefore possible to use these relic fluctuations to constrain inhomogeneous reionization with the power spectrum at wavenumbers $\log_{10}(k/\text{km}^{-1} \text{s}) \lesssim -1.5$. We use the Sherwood-Relics suite of hybrid radiation hydrodynamical simulations to perform a first analysis of new Ly α forest power spectrum measurements at $4.0 \leq z \leq 4.6$. These data extend to wavenumbers $\log_{10}(k/\text{km}^{-1} \text{s}) \approx -3$, with a relative uncertainty of 10–20 per cent in each wavenumber bin. Our analysis returns a 2.7σ preference for an enhancement in the Ly α forest power spectrum at large scales, in excess of that expected for a spatially uniform ultraviolet background. This large-scale enhancement could be a signature of inhomogeneous reionization, although the statistical precision of these data is not yet sufficient for obtaining a robust detection of the relic post-reionization fluctuations. We show that future power spectrum measurements with relative uncertainties of $\lesssim 2.5$ per cent should provide unambiguous evidence for an enhancement in the power spectrum on large scales.

Key words: methods: numerical – intergalactic medium – quasars: absorption lines – large scale structure of Universe.

1 INTRODUCTION

Intergalactic neutral hydrogen along the line of sight toward high-redshift quasars leaves an observable spectral signature in the form of Ly α absorption lines. These features are cumulatively referred to as the Ly α forest (see e.g. Rauch 1998; McQuinn 2016), and their observable properties are linked to the physical conditions in the distribution of intergalactic matter on scales of ~ 0.5 –50 comoving Mpc. The factors that determine the observable properties of the Ly α forest absorbers – and hence the physical properties of the intergalactic medium (IGM) – may be grouped into two broad categories: those of cosmological origin, such as the nature of dark matter and the shape of the matter power spectrum (Croft et al. 2002; Seljak et al. 2006; Iršič et al. 2017a; Garzilli et al. 2019; Rogers & Peiris 2021; Villasenor et al. 2022), and those of astrophysical origin, such as feedback processes (Theuns et al. 2002b; Viel et al. 2013b; Gurvich et al. 2017; Chabanier et al. 2019) and the ionization and thermal state of the IGM (Schaye et al. 2000; Bolton et al. 2005; Boera et al. 2014; Hiss et al. 2018; Gaikwad et al. 2021). Canonically, the ionization state of the IGM is determined by the UV photons emitted by stars

and active galactic nuclei. As a result, intergalactic Ly α absorption is also a key probe of the final stages of the reionization era at $z > 5$ (Fan et al. 2006; Becker et al. 2015; Eilers et al. 2018; Bosman et al. 2022)

A widely used statistic for characterising the Ly α forest is the 1D power spectrum of the transmitted flux (McDonald et al. 2000; Palanque-Desabrouille et al. 2013; Iršič et al. 2017b; Walther et al. 2018; Boera et al. 2019; Karaçaylı et al. 2022). When studying the IGM approaching the reionization era, the 1D power spectrum is useful in several different ways. First, the power spectrum amplitude is sensitive to the average Ly α forest transmission, and hence also the IGM neutral hydrogen fraction (Mishra & Gnedin 2022). Second, the shape of the power spectrum on small scales (i.e. for wavenumbers $\log_{10}(k/\text{km}^{-1} \text{s}) \gtrsim -1.5$) depends on the IGM thermal history, through the shape of the Doppler broadening kernel and role that gas pressure plays in setting the physical extent of Ly α absorbers (Nasir et al. 2016). Finally, on large scales ($\log_{10}(k/\text{km}^{-1} \text{s}) \lesssim -1.5$), the power spectrum is sensitive to spatial fluctuations in the thermal and ionization state of the IGM (Cen et al. 2009). These fluctuations – associated with the inhomogeneous heating of the IGM during reionization – will linger for some time after reionization has completed due to the long cooling timescale in the low density IGM (Theuns

* E-mail: margherita.molaro@nottingham.ac.uk

et al. 2002a; Hui & Haïman 2003). The consequence is that observable relics of the reionization era will be imprinted in the Ly α forest power spectrum at redshift $z > 4$.

In this context, Molaro et al. (2022) (hereafter M22) used Sherwood-Relics – a set of hybrid radiation hydrodynamical simulations of the IGM during reionization (Puchwein et al. 2022) – to study the effect of patchy reionization on the 1D Ly α forest power spectrum. In agreement with a number of other studies (Cen et al. 2009; Keating et al. 2018; D’Aloisio et al. 2018; Oñorbe et al. 2019; Wu et al. 2019; Montero-Camacho & Mao 2020), M22 found that remnant patches of hot, highly ionised, low density hydrogen left over following reionization produce large-scale variations in the Ly α forest transmission at $4.2 < z < 5$. These enhance the 1D power spectrum of the transmitted flux by 10–50 per cent on large scales, $\log_{10}(k/\text{km}^{-1} \text{ s}) \lesssim -1.5$, with this effect being largest close to the end point of reionization. However, M22 also demonstrated that these fluctuations will have a limited effect on the recovery of thermal parameters from *existing* measurements of the Ly α forest power spectrum at $z > 4$ (e.g. Boera et al. 2019). One reason for this is that the measurements of Boera et al. (2019) do not include the larger scales at $\log_{10}(k/\text{km}^{-1} \text{ s}) \lesssim -2.2$ where the additional power is expected to be most significant (see e.g. Fig. 6 in M22). Surveys such as the Dark Energy Spectroscopic Instrument (DESI) survey (Vargas-Magana et al. 2019) and the William Herschel Telescope Enhanced Area Velocity Explorer-Quasi-stellar Object (WEAVE-QSO) survey (Pieri et al. 2016) will, however, extend to such large scales with lower resolution spectra, and will measure the 1D Ly α forest power spectrum to a precision of a few per cent. Inhomogeneous reionization effects at $z > 4$ will therefore be an important astrophysical “nuisance” parameter that must be included within the forward modelling frameworks used to constrain the underlying matter power spectrum from these observations. Conversely, the presence of additional large scale power in the Ly α forest data can also provide valuable information on the end stages of inhomogeneous reionization, as this may be consistent with relic fluctuations in the ionization and thermal state of the IGM (e.g. D’Aloisio et al. 2018).

In this work we will build on M22 by investigating the possibility of detecting inhomogeneous reionization using the 1D Ly α forest power spectrum on large scales. For this purpose we will make use of the recent 1D power spectrum measurements presented by Karaçaylı et al. (2022) using data from the Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ) (O’Meara et al. 2017), the Spectral Quasar Absorption Database (SQUAD) (Murphy et al. 2019) and the XQ-100 survey (López et al. 2016). Importantly, these data extend to larger scales, $\log_{10}(k/\text{km}^{-1} \text{ s}) \approx -3$, compared to Boera et al. (2019), although the precision is still rather modest, with relative uncertainties in the power spectrum of 10–20 per cent. As we shall demonstrate, however, this is nevertheless sufficient for providing evidence for enhanced large scale power in the data. A related earlier result was presented by D’Aloisio et al. (2018), who performed an analysis of the 1D power spectrum at slightly higher redshift, $z = 5.2$ – 5.6 , using the sample of 21 quasars presented by McGreer et al. (2015). By comparing these data with a semi-numerical model for the patchy ionization of the IGM, D’Aloisio et al. (2018) found tentative evidence ($\sim 2\sigma$) for an enhancement in large scale power at $\log_{10}(k/\text{km}^{-1} \text{ s}) \approx -3$.

We furthermore introduce two important updates to the Bayesian inference framework used in M22. First, we improve the accuracy and speed of our 1D Ly α forest power spectrum emulator, by replacing the linear interpolation approach we had previously used to build our model grid (e.g. Viel & Haehnelt 2006) with a neural network. Second, we introduce a more flexible, model independent approach

for including the patchy ionization correction quantified by M22 within our Bayesian parameter estimation framework.

This paper is structured as follows. In Section 2 we briefly re-describe the simulations used in M22 and introduce the new machine-learning generated power spectrum emulator that will be adopted in this work. In Section 3 we consider the effect of patchy reionization on the Ly α power spectrum, and introduce a new parameterisation that allows the power spectrum template presented by M22 to be implemented within our Markov Chain Monte Carlo analysis in a more general, model independent way. In Section 4, we apply our framework to the data presented by Karaçaylı et al. (2022) and discuss the possible evidence for an enhancement in power on large scales. We furthermore consider the possibility of applying our analysis to future, higher precision data before concluding in Section 5. We assume a flat Λ CDM cosmology throughout this work, with $\Omega_{\Lambda} = 0.692$, $\Omega_{\text{m}} = 0.308$, $\Omega_{\text{b}} = 0.0482$, $\sigma_8 = 0.829$, $n_s = 0.961$, $h = 0.678$ (Planck Collaboration et al. 2014), and a primordial helium fraction by mass of $Y_{\text{p}} = 0.24$ (Hsyu et al. 2020).

2 EMULATING THE Ly α FOREST 1D POWER SPECTRUM USING NEURAL NETWORKS

2.1 Numerical simulations

In this work – with the exception of the neural network described in Section 2.2 – we use the same Bayesian inference and simulation set-up adopted in M22. This is based on a Monte Carlo Markov Chain (MCMC) sampler combined with a Metropolis Hastings algorithm, as introduced by Hrišić et al. (2017a). The backbone of our Ly α forest power spectrum emulator consists of 12 hydrodynamical simulations that form part of the Sherwood-Relics simulation suite (Puchwein et al. 2022). Each model follows a $20h^{-1}$ cMpc cosmological volume with 2×1024^3 particles using variations of the Puchwein et al. (2019) spatially homogeneous UV background synthesis model (see Table 1 in M22). For completeness we briefly recapitulate some key information about the simulations here, but refer the reader to M22 for further details.

The IGM thermal history in each of the 12 simulations is characterised by three parameters which, as in M22, will be used to describe the IGM thermal state: the cumulative energy per unit mass deposited into gas at the mean background density, u_0 , the temperature at mean density, T_0 , and γ , the index in the power-law relation describing the temperature-density relation, where $T = T_0[\rho/\langle\rho\rangle]^{\gamma-1}$ (Hui & Gnedin 1997; McQuinn 2016), where $\rho/\langle\rho\rangle$ is the local gas overdensity. The 12 simulations, each characterised by a different u_0 value, are then post-processed to finely sample the $T_0 - \gamma$ parameter-space: gas particles in the simulations are translated and rotated in the temperature-density plane to obtain different combinations of these parameters (see Boera et al. (2019), Gaikwad et al. (2020), and M22 for more details).

The resulting Ly α optical depths are also rescaled in post processing to give different values for the effective optical depth, τ_{eff} , where $\tau_{\text{eff}} = -\ln\langle F \rangle$ and $\langle F \rangle$ is the mean Ly α forest transmission. In each redshift bin, the parameter grid uniformly spans T_0 between 5 000 K and 14 000 K in steps of 1 000 K, γ between 0.9 to 1.8 in steps of 0.1, and τ_{eff} from 0.3 to 1.7 times the observed effective optical depth $\tau_{\text{eff,obs}}$ (Viel et al. 2013a) in steps of $0.1\tau_{\text{eff,obs}}$ for each redshift bin considered. This leads to $10 \times 10 \times 15 \times 12 = 18\,000$ post-processed sets of Ly α forest spectra in the parameter grid for each redshift considered. The 1D Ly α forest power spectrum is then calculated from each set of spectra assuming a constant $\log_{10}(k/\text{km}^{-1} \text{ s})$ step-size

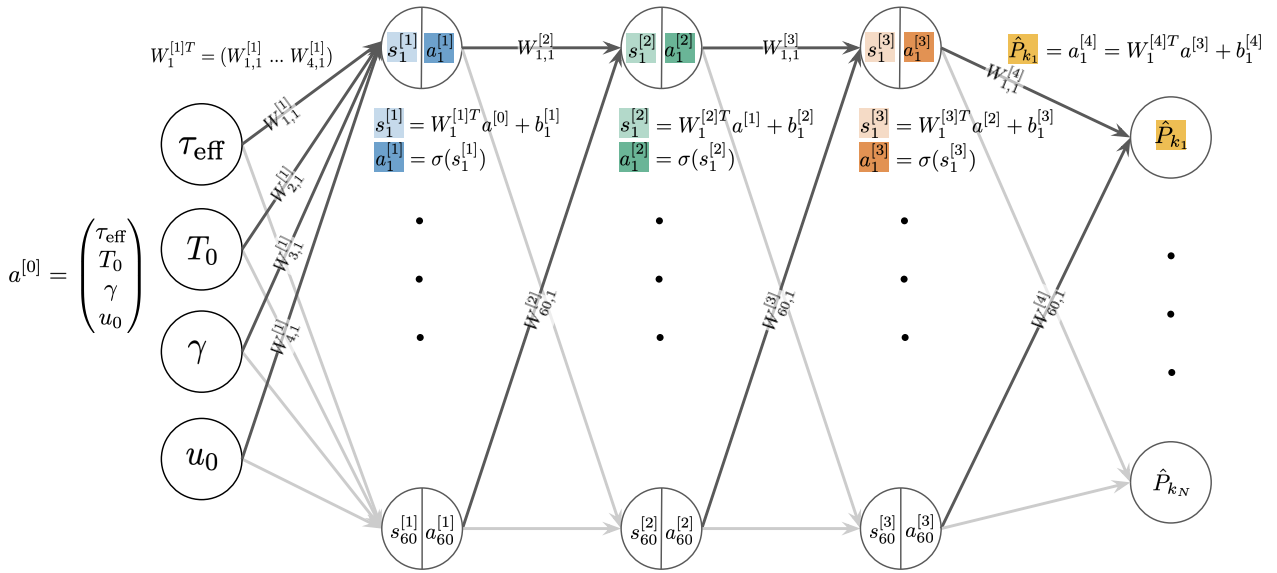


Figure 1. Schematic illustrating the neural network used to predict a $\text{Ly}\alpha$ forest power spectrum (\hat{P}_k) sampled with N bins in wavenumber, k , given four input astrophysical parameters (τ_{eff} , T_0 , γ , u_0). In addition to the input layer with four nodes for the input parameters, the network has three hidden layers (blue, green and orange) with 60 nodes each, and an output layer (yellow) with N nodes, one for each k -bin. \mathbf{W} and \mathbf{b} are the weight and bias matrices respectively, σ is the ReLU activation function (see text for details), \mathbf{a} is the value of the node, and s is an intermediate value used in the calculation of \mathbf{a} . The upper index in the square bracket refers to the layer (0 to 4, where 0 is the input layer and 4 the output layer), while the lower indices refer to the node. In the case of the weights, which “connect” two layers, the first lower index refers to the node in the first layer, and the second to the second layer. Equations are provided only for the top node of each layer.

of 0.1, following Boera et al. (2019) and M22 (although we will adopt a different binning strategy when analysing observational data in Section 4). We refer to this grid of simulated power spectra as $\mathbf{G}_{\text{homog}}$, because it relies on simulations performed with a spatially homogeneous ultraviolet (UV) background.

Finally, as discussed in M22, for each redshift considered the 18 000 realisations of the power spectrum in $\mathbf{G}_{\text{homog}}$ can be modified using a template that includes the effect of inhomogeneous reionization on the large scale power at $\log_{10}(k/\text{km}^{-1} \text{ s}) \leq -1.5$. This template uses the data compiled in Table 2 and eq. (3) of M22 (see also Section 3.2 later). We refer to the resulting grid of modified power spectra as $\mathbf{G}_{\text{patchy}}$, because these now also include the expected (model-dependent) effect of patchy reionization on large scales. Furthermore, for all mock realisations of the power spectrum the off-diagonal elements in the covariance matrix were obtained from the $80h^{-1}$ cMpc, homogeneous-UVB Sherwood-Relics simulation with 2048^3 gas particles described in Puchwein et al. (2022). A larger box size was chosen for constructing the covariance in this work to limit any artificial correlations at $\log_{10}(k/\text{km}^{-1} \text{ s}) < -2.3$.

2.2 Neural network interpolator

How to interpolate between different simulations to obtain a theory prediction at any point in parameter space is an issue common to all MCMC-based Bayesian inference studies of the $\text{Ly}\alpha$ forest. Methods currently in use are either based on linear interpolation techniques (Viel et al. 2013a; Iršič et al. 2017a; Yèche et al. 2017; Palanque-Delabrouille et al. 2020), as were adopted in M22, or on alternative techniques such as Latin hypercube sampling and Gaussian process interpolation (e.g. Bird et al. 2019; Pedersen et al. 2021; Rogers & Peiris 2021; Fernandez et al. 2022).

In this work, we propose an approach based on machine learning techniques. We train a supervised neural network to generate

the 1D power spectrum from the grid of available simulations with very high levels of accuracy, making it a reliable emulator to use in parameter recovery from forthcoming high-precision data. While a neural network trained on our set of simulations may not be immediately generalisable to simulations used in other studies, similar neural networks could very easily be trained using independent parameter grids, and transfer-learning could be invoked to avoid having to retrain the networks completely from scratch (Pan & Yang 2009).

The machine learning model that we adopt here is an artificial neural network – sometimes known as a feed-forward neural network – which consists of a series of mathematical operations applied on the input parameters (see Fig. 1). Such operations are dependent on the choice of an “activation function” which we denote by σ , and the particular sequence in which the operations are applied. This sequence is best described by combinations of nodes organised in layers, linked so that every node in a given layer becomes the input to every node in the following layer. This gives rise to the “neural network” which lends its name to this method.

The output of node m in layer l , denoted by $a_m^{[l]}$, is given by

$$a_m^{[l]} = \sigma(a^{[l-1]} W_m^{[l]T} + b_m^{[l]}), \quad (1)$$

where the weight, \mathbf{W} , and bias, \mathbf{b} , are matrices whose values are iteratively constrained during the training process, and where T denotes the transpose of the matrix. A neural network without an activation function is essentially a linear regression model, with the activation function adding non-linearity. The number of nodes in each layer and the total number of layers considered are free parameters that can be modified to improve the network’s performance (Hornik et al. 1989).

In this work, the neural network will be used to obtain predictions of the $\text{Ly}\alpha$ forest power spectrum at parameter values in between those already included in our simulation grid. In our case, therefore, the input parameters \mathbf{a}^0 are the four astrophysical parameters that our Bayesian inference MCMC method seeks to constrain, that is $\mathbf{a}^0 =$

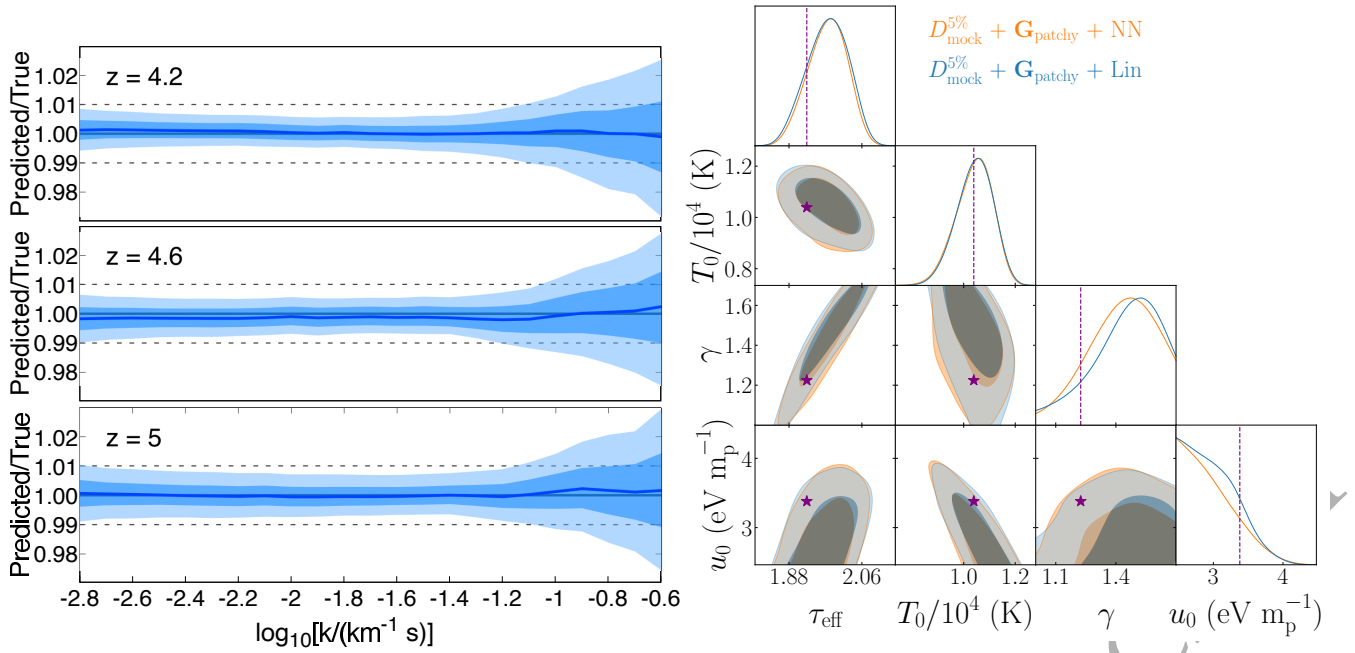


Figure 2. *Left panel:* The distribution of the ratio between the neural-network-predicted and true power spectra in the 10-fold cross validation test. The shaded regions show the 68 and 95 per cent confidence intervals, while the solid blue line shows the median. The solid grey line indicates the perfect recovery case and the dashed lines show a ± 1 per cent range. We have verified that the error distribution is well approximated by a Gaussian in all k bins, and in all redshift bins. *Right panel:* The one and two dimensional posterior distributions for τ_{eff} , T_0 , γ and u_0 at $z = 5$ from our analysis of a mock Ly α forest power spectrum drawn from the RT-late simulation. We assume 5 per cent relative uncertainties on the mock data for $-2.8 \leq \log_{10}[(k/(\text{km}^{-1} \text{s}))] \leq -0.7$. Results are displayed for the linear interpolator (blue contours) or the neural-network interpolator (orange contours) using the simulation grid $\mathbf{G}_{\text{patchy}}$. The purple stars show the true parameters used in the mock data. The bin at $z = 5$ has the poorest match between the neural network and linear interpolator, which nonetheless is very good. All parameters are recovered to within 1-2 σ .

$\{\tau_{\text{eff}}, T_0, \gamma, u_0\}$, where τ_{eff} was converted into $\langle F \rangle = e^{-\tau_{\text{eff}}}$ prior to training. The output of the network will be the values of the power spectrum at a discrete number, N , of bins in wavenumber, k . As parameter recovery is performed independently using power spectrum data in separate redshift bins, we always consider independent neural networks trained, validated, and tested on 1D power spectra at the redshift under consideration. We furthermore independently train networks on both $\mathbf{G}_{\text{homog}}$ and $\mathbf{G}_{\text{patchy}}$.

In each redshift bin, the data for the training, validation, and testing of the neural network are the 18 000 realisations of the power spectrum obtained from the simulations described in Section 2.1. In each case, a randomly selected 90 per cent of these power spectra are used to train the neural network, while 10 per cent are saved for validation. The loss function L used for training is the mean squared error function, such that for power spectrum i in the training sample,

$$L_t = \frac{1}{N} \sum_{i=1}^N (\hat{P}_{k_i,t} - P_{k_i,t})^2, \quad (2)$$

where $\hat{P}_{k_i,t}$ is the neural network prediction for the power spectrum, and $P_{k_i,t}$ is the true value of the power spectrum in k -bin i for training sample t . After experimenting with different activation functions we settled on the Rectified Linear Unit (ReLU) (Fukushima 1975), and chose a neural structure of [60,60,60]. Training was performed with batching (with batch size 12) to computationally optimise the process (LeCun et al. 2012).

The precision of our neural network was tested using k -fold cross validation (Stone 1974) as follows: the complete set of simulated realisations available in each redshift bin (18 000 power spectra) was first shuffled, and then divided into 10 sets of equal size. The initial

randomization ensures that the thermal parameters of the power spectra making up each set are randomly chosen. One of the sets was then held back and the neural network was trained on the remaining 9 sets. The neural network was then tested on the held-back set, resulting in a distribution of neural-network-predicted/true values for the 1 800 power spectra in the set not used in training. This process was then iteratively repeated for all the ten sets, resulting in ten independent distributions of neural-network-predicted/true values.

In Fig. 2 (left panel), we show the *average* distribution of the ten neural-network-predicted/true distributions independently obtained for neural networks trained on $\mathbf{G}_{\text{patchy}}$. The contour plots show (in blue) the average 68 and 95 per cent confidence intervals, while the blue solid line shows the average of the median of each distribution. The recovery error in the $\log_{10}[(k/(\text{km}^{-1} \text{s}))] < -1.2$ range is around 0.5 per cent at the 68 per cent confidence level, increasing up to ~ 1 per cent at largest wavenumbers considered here ($\log_{10}[(k/(\text{km}^{-1} \text{s}))] = -0.6$), representing a very high level of precision. Furthermore, we find the uncertainty in the recovery to be well approximated by a Gaussian distribution in all k -bins considered.

In the right panel of Fig. 2, we compare the one and two dimensional posterior distributions for τ_{eff} , T_0 , γ and u_0 at $z = 5$ from our analysis of a mock Ly α forest power spectrum drawn from the RT-late simulation described in M22. We assume 5 per cent relative uncertainties on the mock data for $-2.8 \leq \log_{10}[(k/(\text{km}^{-1} \text{s}))] \leq -0.7$. The results are obtained using either a linear interpolator (blue contours) or the neural-network interpolator introduced in this work (orange contours) for our $\mathbf{G}_{\text{patchy}}$ grid of models. The purple stars show the true parameters of the mock data, showing that we recover the input parameters within 1-2 σ . We see that the choice of interpolator makes little difference to the parameter recovery. However, along with the

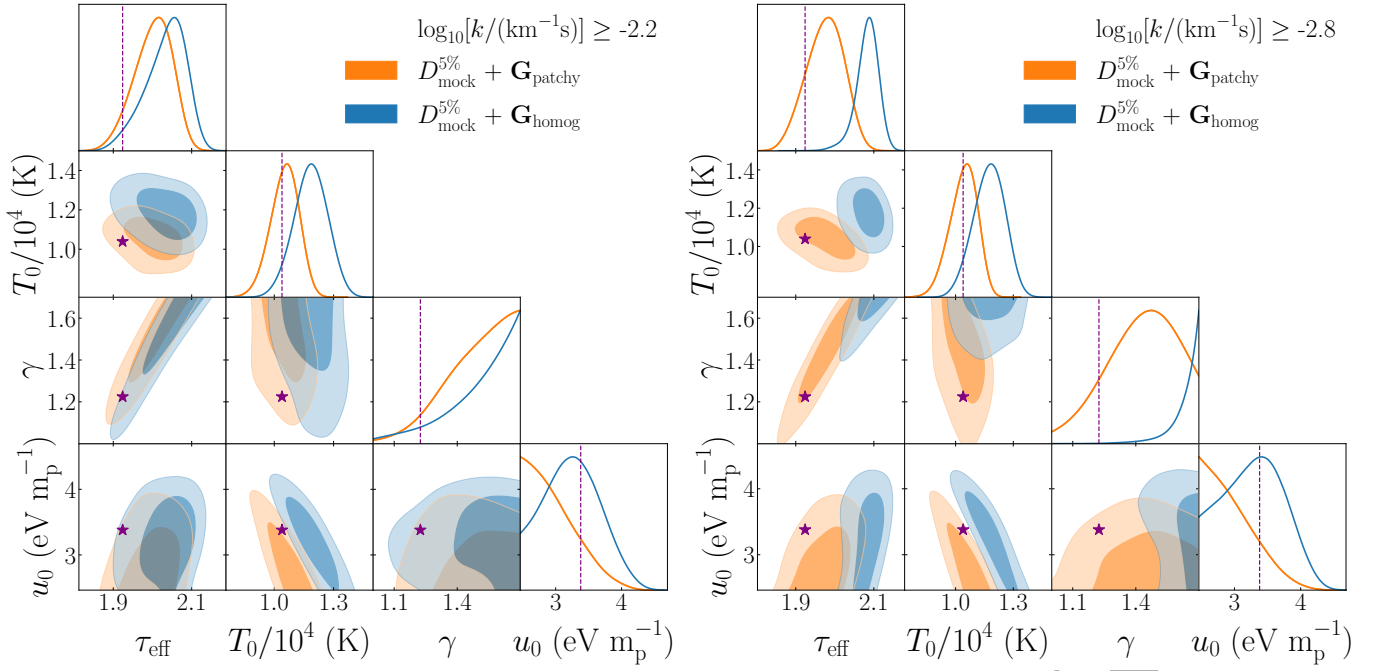


Figure 3. The one and two dimensional posterior distributions for τ_{eff} , T_0 , γ , and u_0 from an analysis of mock 1D power spectrum data drawn from the M22 RT-late simulation at $z = 5$. The mocks assume a 5 per cent relative error on the power spectrum. The purple stars and vertical dashed lines correspond to the true parameter values in the simulation. The results are obtained using either our original grid of homogeneous UVB models ($\mathbf{G}_{\text{homog}}$, blue contours) or that same grid including the M22 template for the effect of inhomogeneous reionization on the power spectrum ($\mathbf{G}_{\text{patchy}}$, orange contours). The left panel is for mock data spanning wavenumbers $-2.2 \leq \log_{10}[(k/(\text{km}^{-1} \text{s}))] \leq -0.7$, whereas the right panel is for mock data extending to smaller wavenumbers/larger scales, $-2.8 \leq \log_{10}[(k/(\text{km}^{-1} \text{s}))] \leq -0.7$. Note how the parameter recovery is more sensitive to the effect of inhomogeneous reionization when including data on larger scales.

significantly improved control of interpolation uncertainties, the neural network interpolator reduces the computational requirements of the MCMC sampler. Using the same computational hardware, we find the neural network requires ~ 1.5 per cent of the CPU time required by the linear interpolator for an equal number of steps. We will leverage the speed of this method in Section 4.4, where we perform an analysis of multiple realisations to assess the variance in our parameter recovery.

3 THE EFFECT OF INHOMOGENEOUS REIONIZATION ON THERMAL PARAMETER RECOVERY

3.1 Revisiting M22 using the 1D power spectrum on large scales

We now revisit the analysis of the 1D power spectrum presented by M22. This earlier work concluded that the effect of inhomogeneous reionization on the shape of the power spectrum at $4.2 < z < 5$ should not strongly bias *existing* constraints on IGM thermal parameters obtained from an analysis of high resolution Ly α forest data (see also Wu et al. 2019). This conclusion was based on assuming a relative uncertainty of 10 per cent on power spectrum measurements (i.e. similar or smaller than the level of precision currently achieved for the high resolution data presented by Boera et al. (2019)). On the other hand, M22 noted that, for data with relative uncertainties at the 5 per cent level, patchy reionization effects that change the shape of the power spectrum at small scales, $\log_{10}(k/\text{km}^{-1} \text{s}) > -1.5$ will introduce a modest ($\sim 1\sigma$) shift in the recovery of IGM thermal parameters. M22 found this was primarily driven by divergent peculiar velocity gradients and variations in the thermal broadening kernel that alter the shape of the power spectrum on small scales.

However, the power spectrum analysis presented by M22 only considered wavenumbers $\log_{10}[k/(\text{km}^{-1} \text{s})] \geq -2.2$, matching the measurement range presented by Boera et al. (2019). As discussed earlier, this excludes the larger scales where the enhancement in the power spectrum due to relic fluctuations in the neutral hydrogen fraction from reionization are expected to be largest. In this context, surveys such as DESI (Vargas-Magana et al. 2019) and WEAVE-QSO (Pieri et al. 2016) will not only achieve higher levels of precision, but will also probe larger physical scales in the redshift range $2 \leq z \leq 4.5$, where enhancements to the 1D power spectrum arising from patchy reionization are expected to be most prominent. Here, we therefore quantify how patchy reionization will impact on the recovery of IGM parameters by applying the M22 analysis to smaller wavenumbers/larger scales where the patchy correction is largest.

In Fig. 3, we use our parameter estimation framework to consider parameter recovery from mock data drawn from the RT-late simulation used in M22 (see also Puchwein et al. 2022). This simulation follows inhomogeneous reionization ending at redshift $z_{\text{R}} = 5.3$, where we choose to define the end point of reionization as the redshift when the volume-averaged HI fraction first falls below $x_{\text{HI}} = 10^{-3}$ in the model. We assume relative uncertainties of 5 per cent on the power spectrum and consider two different wavenumber ranges: $-2.2 \leq \log_{10}(k/\text{km}^{-1} \text{s}) \leq -0.7$ in the left panel (i.e. the range considered by M22 and Boera et al. (2019)), and a range that extends to larger scales, $-2.8 \leq \log_{10}(k/\text{km}^{-1} \text{s}) \leq -0.7$, in the right panel.

The panels in Fig. 3 show the one and two dimensional posterior distributions recovered for τ_{eff} , T_0 , γ and u_0 in a single redshift bin at $z = 5$. The results are obtained using our neural network power spectrum emulator, for either the grid of homogeneous UV background models used in M22 ($\mathbf{G}_{\text{homog}}$, blue contours), or that

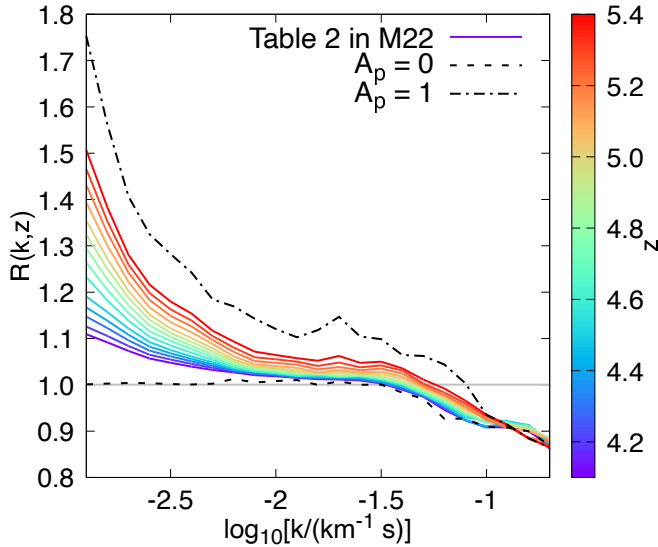


Figure 4. The coloured curves show the inhomogeneous reionization correction to the power spectrum, $R(k, z)$, presented in Table 2 of M22 (note that in this work we drop the ‘mid’ label used by M22). The redshift at which this correction is drawn from the reionization simulation used in M22 is indicated in the colour bar. For comparison, the dashed and dot-dashed curves show $R(k, z)$ for $A_p = 0$ and $A_p = 1$, respectively. This follows the updated parameterisation we use in this work, given by Eq. (4). The range corresponding to the tabulated M22 template is $A_p = [0.255, 0.765]$.

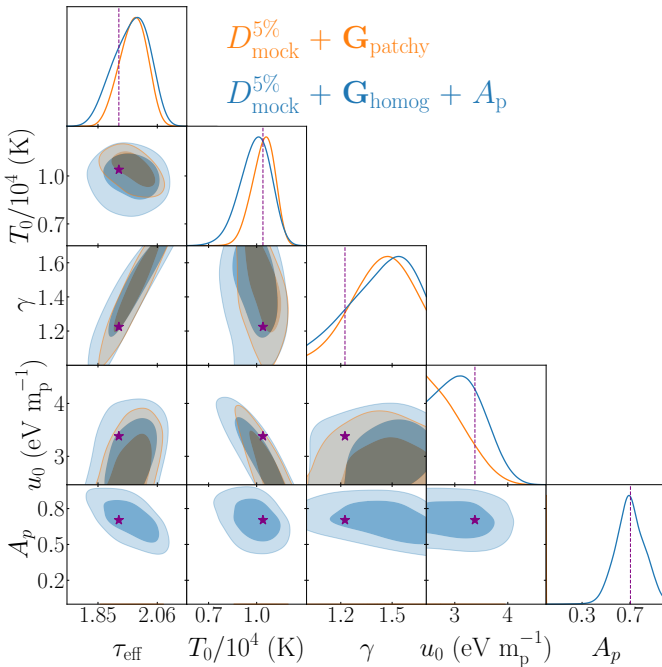


Figure 5. As for the right panel of Fig. 3, but now showing the results for $\mathbf{G}_{\text{homog}} + A_p$ instead of $\mathbf{G}_{\text{homog}}$, where the former now uses the parameterisation for the patchy reionization template introduced in Eq. (4) (see text in Section 3.2 for further details). These two approaches lead to similar parameter constraints with overlapping contours. This indicates our updated $\mathbf{G}_{\text{homog}} + A_p$ approach remains accurate, while simultaneously affording a greater degree flexibility and model independence to our analysis.

same grid including a template that captures the effect of inhomogeneous reionization on the power spectrum, based on eq. (3) in M22 ($\mathbf{G}_{\text{patchy}}$, orange contours). The purple stars and vertical dashed lines correspond to the true parameter values used in the simulation.

First, as already noted in M22 (see their fig. 10), for wavenumbers $-2.2 \leq \log_{10}(k/\text{km}^{-1} \text{ s}) \leq -0.7$ patchy reionization introduces a modest $\sim 1\sigma$ shift between the true and recovered parameters (i.e. the blue contours and purple stars in the left panel of Fig. 3) when assuming a uniform UV background ($\mathbf{G}_{\text{homog}}$). However, when including larger scales in the right panel, the systematic bias in the parameter recovery is much larger and is at the level of $\sim 6\sigma$ for the total χ^2 . The largest shift is for τ_{eff} ; this is unsurprising, given that τ_{eff} sets the amplitude of the power spectrum on the scales where the patchy reionization effects are largest. There is also a large shift in the temperature-density relation, however this is entirely due to the very strong degeneracy between τ_{eff} and γ . We also observe that the simulation parameters are generally well recovered (within $\sim 1-2\sigma$) when applying our patchy reionization template to the grid of homogeneous UV background models ($\mathbf{G}_{\text{patchy}}$), even when extending our analysis to larger scales. This is expected and serves as a useful consistency test, given that these mock data were drawn from one of the simulations used in M22 to obtain the patchy reionization template.

Parameter recovery is therefore much more sensitive to the effect of inhomogeneous reionization when including data on larger scales. As has been noted elsewhere (e.g. Cen et al. 2009; Keating et al. 2018; D’Aloisio et al. 2018; Oñorbe et al. 2019; Montero-Camacho & Mao 2020), this raises the interesting possibility that – given a significant detection of enhanced large scale power – the end stages of reionization may be constrained with precision measurements of the 1D Ly α forest power spectrum on scales $\log_{10}(k/\text{km}^{-1} \text{ s}) \simeq -3$ at $z > 4$.

3.2 A generalised approach to modelling the enhanced large scale power from patchy reionization

Before turning to consider observational data, we additionally modify the framework first introduced by M22 to provide a more flexible, model independent approach for quantifying the amount of excess large scale power in data. The motivation for this is two-fold. First, it avoids using a parameterisation that is directly tied to the redshift evolution of the reionization models used in M22. Second, it allows for a general parameterisation of the excess of large-scale power that can capture a larger variety of models. In the same manner as cosmological parameter data compression (e.g. McDonald et al. 2000; Pedersen et al. 2022), the interpretation of the measured excess large scale power can then be mapped back onto physical models.

We implement this updated approach by continuing to use the template provided in Table 2 of M22, where the shape of the correction to the power spectrum for patchy reionization – obtained from radiative transfer simulations – is given by the quantity $R(k, z)$. Here

$$R(k, z) = \frac{P_{\text{RT}}(k, z)}{P_{\text{homog}}(k, z)}, \quad (3)$$

where $P_{\text{RT}}(k, z)$ ($P_{\text{homog}}(k, z)$) represent the 1D flux power spectrum from simulations that include (ignore) the effect of inhomogeneous reionization on the Ly α forest, but otherwise have the same *volume-averaged* reionization history (see Puchwein et al. 2022, for further details). However, instead of assuming a model dependent redshift evolution for this template that is tied to the adopted reionization history (i.e. eq. (3) in M22), we now allow $R(k, z)$ to vary *freely* across each redshift bin within our analysis. We implement

this by introducing the free parameter A_p with a flat prior in the range $[0, 1]$. This parameter controls the amplitude of the imprint of patchy reionization on the 1D flux power spectrum. This is achieved by relating it to the redshift, z , at which the correction is linearly interpolated from the data in Table 2 of M22 by

$$z = 3.45 + 2.55A_p, \quad \text{for } 0 \leq A_p \leq 1. \quad (4)$$

This linear mapping has been obtained using the patchy reionization template $R(k, z)$ at $4.1 \leq z \leq 5.4$ in table 2 of M22, and linearly extrapolating to a lower and upper redshift limit of $z = 3.45$ and $z = 6$, respectively. The lower redshift is chosen such that the extrapolated template correction is negligible on large scales, which we find occurs at $z = 3.45$ (i.e. for $A_p = 0$). The extrapolation similarly gives $A_p = 1$ by $z = 6$. In practice, however, the exact parameterisation makes little difference to our results, as our main aim is to establish if $A_p \neq 0$ is preferred in observational data. This is illustrated further in Fig. 4, where the coloured curves show the data presented in Table 2 of M22, and the dashed (dot-dashed) curves correspond to $A_p = 0$ ($A_p = 1$) within our new parameterisation. We will refer to this approach as $\mathbf{G}_{\text{homog}+A_p}$ within our analysis framework.

In Fig. 5 we perform a brief consistency test of this approach, by once again examining the one and two dimensional posterior distributions obtained from the mock data used in Fig. 3. Recall the mock data are drawn from the RT-late simulation in M22 at $z = 5$, and assume a 5 per cent relative error on the power spectrum measurement. The orange contours are identical to those shown in Fig. 3 ($\mathbf{G}_{\text{patchy}}$), while the blue contours show the results for the generalised $\mathbf{G}_{\text{homog}+A_p}$ approach. The purple stars and vertical lines, as usual, represent the true parameters in the mock data. Note that $\mathbf{G}_{\text{patchy}}$ corresponds to a fixed value of A_p (as obtained by comparing Eq. (4) to eq. (3) of M22), such that $R(k, z)$ is assumed to exactly match the shape expected for a model with reionization ending at $z_R = 5.3$, as is appropriate for the model from which the mock data was produced. We observe that the two sets of contours overlap and A_p is recovered within 1σ , indicating this approach maintains the accuracy of our parameter recovery, while also affording a greater degree of generality.

4 A HINT OF ENHANCED LARGE SCALE POWER IN THE FLUX POWER SPECTRUM

4.1 Observational data

We now apply our updated analysis framework to the observational measurements of the 1D Ly α forest power spectrum at $4 \leq z \leq 4.6$ presented recently by Karaçaylı et al. (2022). The Karaçaylı et al. (2022) flux power spectrum has been obtained using a collection of high resolution and high signal-to-noise data from VLT/XSHOOTER (XQ-100, López et al. (2016)), VLT/UVES (SQUAD, Murphy et al. (2019)) and Keck/HIRES (KODIAQ, O’Meara et al. (2017)). The flux power spectrum measurements are made in a total of 15 redshift bins, in the range of $z = 1.8 - 4.6$ with bin size $\Delta z = 0.2$. This work focuses on high redshifts only ($z > 4$), because the relic fluctuations of reionization dissipate at lower redshifts (see M22 and Fig. 4). In each of the redshift bins, the flux power is measured in 21 k -bins, from $k = 0.0011 \text{ km}^{-1} \text{ s}$ to $k = 0.39569 \text{ km}^{-1} \text{ s}$ using non-uniform spacing (the bins are equidistant in k on large scales, and equidistant in $\log_{10} k$ on small scales, with the divide at $k \sim 0.01 \text{ km}^{-1} \text{ s}$). Following the discussion of Karaçaylı et al. (2022) on the effects of noise and resolution on the smallest scales, and because we are interested in isolating the effect of inhomogeneous reionization on large scales, we further restrict our analysis by limiting the wavenumber range to

$k < 0.1 \text{ km}^{-1} \text{ s}$. This leaves 15 k -bins in each of the 4 redshift bins, for a total of 60 data points.

By virtue of combining several data sets, these measurements boast one of the highest precision measurements to date of the flux power spectrum using high resolution Lyman- α forest data. A typical uncertainty on the flux power spectrum at $z = 4.2$ (4.6) is at the level of $\sim 10 - 15$ per cent (20 - 30 per cent), ranging from intermediate to large scales. This is similar to the less sparsely sampled measurements of Boera et al. (2019), who use bin sizes $\Delta z = 0.4$ and report uncertainty levels on the flux power spectrum of $\sim 10 - 20$ per cent (8 - 10 per cent) at $z = 4.2$ (4.6). Furthermore, the Karaçaylı et al. (2022) power spectrum analysis is performed jointly on the quasar spectra from all three samples using the optimal quadratic estimator (Karaçaylı et al. 2020). An advantage over the more common fast Fourier transform estimators found in the literature (e.g. Viel et al. 2013a; Iršič et al. 2017b; Walther et al. 2019) lies in better control of the masked regions and therefore better recovery of the large-scale modes. This is critical for our analysis, as the effect of inhomogeneous reionization on the power spectrum is greatest on the largest scales probed by the data. Note also that although the flux power spectrum measurements from large spectroscopic surveys such as the extended Baryon Oscillation Spectroscopic Survey (eBOSS) reach a higher statistical precision (e.g. Chabanier et al. 2019), the typical spectral resolution prevents these surveys from measuring small-scale power. At large scales, the IGM thermal parameters are highly degenerate with the additional enhancement of power from reionization fluctuations, and adding small-scale information can significantly help in breaking this degeneracy (e.g. Viel et al. 2013a; Iršič et al. 2017a; Yèche et al. 2017). Furthermore, different systematic uncertainties can impact on the power spectrum measurements from low and high resolution data (see Section 4.3), such that a joint analysis with eBOSS data would be more complicated. For these reasons, we limit our investigation to the high-resolution data of Karaçaylı et al. (2022) and leave a joint analysis with low-resolution data to future work.

4.2 Data analysis

In order to compare our models to the flux power spectrum measurements of Karaçaylı et al. (2022), we use the same strategy and grid of simulations described in Section 2 and in M22. However, we make two minor but important modifications. First, because we are comparing our models to observational data, we must apply a mass resolution correction to all of our simulated models, which have been performed in $20h^{-1} \text{ cMpc}$ boxes with a gas particle mass of $M_{\text{gas}} = 9.97 \times 10^4 h^{-1} M_{\odot}$ (see Table 1 in M22). The resolution correction is obtained using a grid of simulations that have parameters matched to those used in M22, but with a mass resolution that is improved by a factor of 8.¹ This correction is largest at small scales, and is at most 12 per cent at $k = 0.1 \text{ km}^{-1} \text{ s}$ ($z = 5$), but is very small (< 1 per cent) on the large scales where the effect of patchy reionization is largest. Secondly, we now train the neural network on post-processed power spectra that use the same k -binning as the Karaçaylı et al. (2022) data. We have verified that the accuracy and precision of the trained neural network remains at the same level demonstrated in Fig. 2 earlier.

¹ Specifically, we use simulations performed with box sizes of $10h^{-1} \text{ cMpc}$ using 2×1024^3 particles, yielding a gas particle mass of $M_{\text{gas}} = 1.25 \times 10^4 h^{-1} M_{\odot}$ (see Table 1 of Puchwein et al. 2022, for full details).

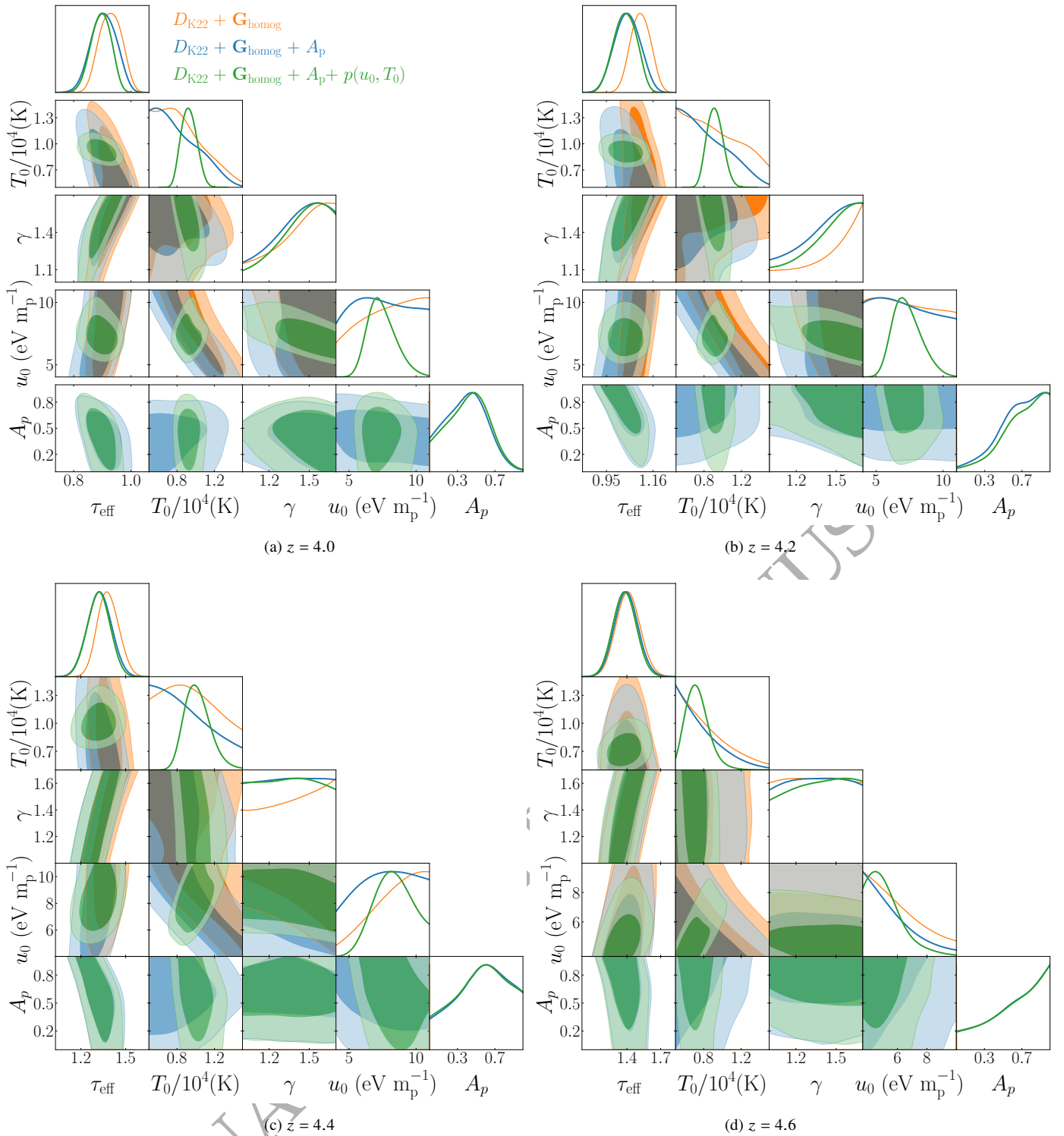


Figure 6. The one and two dimensional posterior distributions for τ_{eff} , T_0 , γ , u_0 and A_p from the analysis of the [Karaçaylı et al. \(2022\)](#) data using $\mathbf{G}_{\text{homog}}$ (orange contours) and $\mathbf{G}_{\text{homog}} + A_p$ (blue contours). The shaded regions show the 68 and 95 per cent confidence intervals for the joint distribution. For the latter, we also show in green the contours obtained when including the $u_0 - T_0$ prior described in Section 4. The four panels are for the four [Karaçaylı et al. \(2022\)](#) redshift bins analysed in this study, namely $z = 4.0, 4.2, 4.4, 4.6$.

In Fig. 6 we show the one and two dimensional posterior distributions obtained from the analysis of the [Karaçaylı et al. \(2022\)](#) data excluding ($\mathbf{G}_{\text{homog}}$, orange contours) or including ($\mathbf{G}_{\text{homog}} + A_p$, blue contours) our parameterisation for the effect of inhomogeneous reionization on large scales. For the latter, we also consider a third

case where a non-flat $u_0 - T_0$ prior is added to the analysis. We add this because the posterior on $u_0 - T_0$ is wide and some regions of parameter space (e.g. where T_0 is very large and u_0 is very low) are unphysical. Hence, instead of allowing u_0 and T_0 to vary freely, we apply a prior that encompasses a physically plausible region within

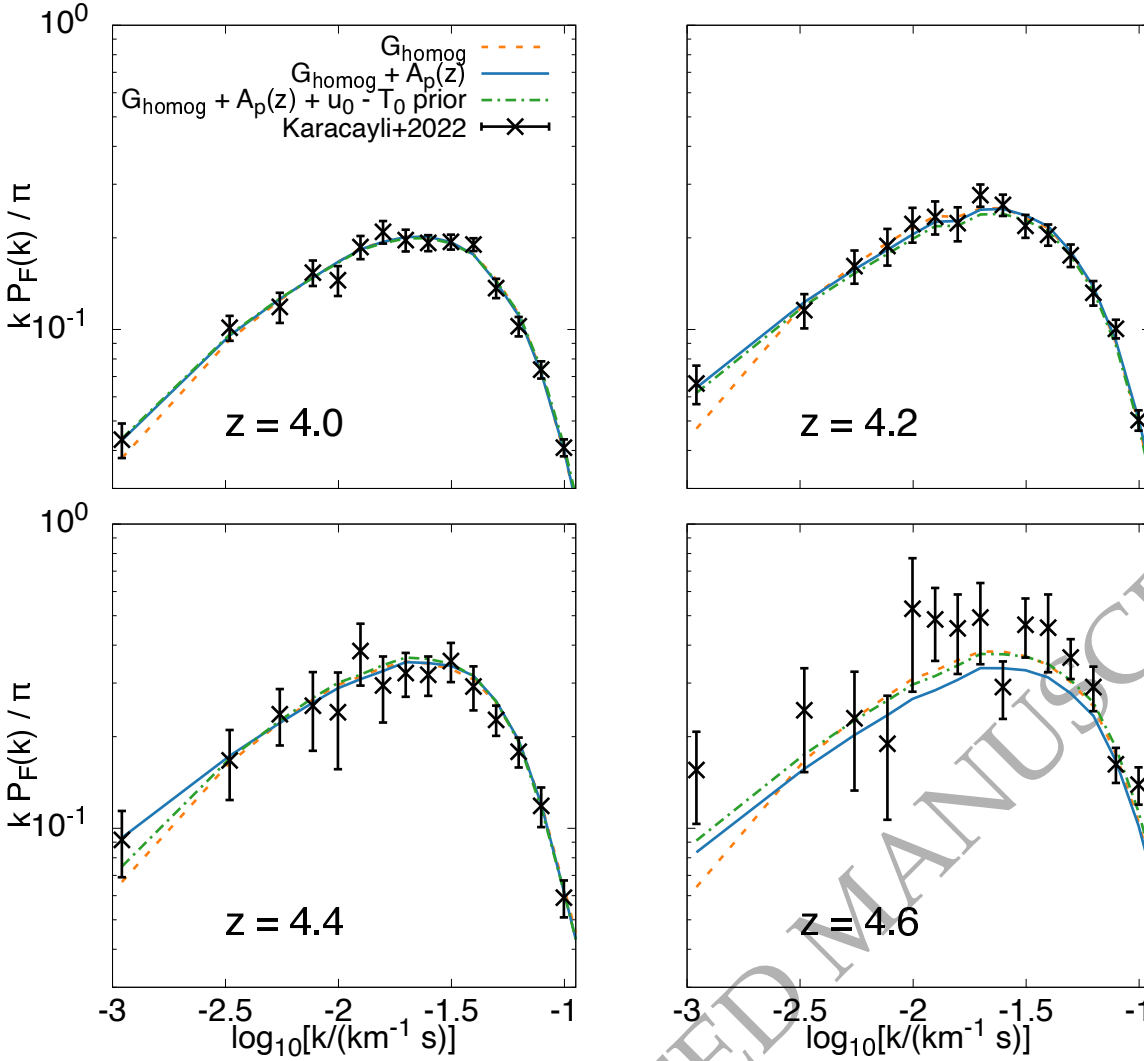


Figure 7. The best-fit $P_F(k)$ models obtained from our MCMC analysis of the Karaçaylı et al. (2022) data for the three cases displayed in Fig. 6: G_{homog} shown by the orange dashed curves, $G_{\text{homog}} + A_p$ shown by the solid blue curves, and $G_{\text{homog}} + A_p + u_0 - T_0$ prior shown by the dotted-dashed green curves. The data points with error bars show the Karaçaylı et al. (2022) power spectrum measurements in the four redshift bins considered here, $z = 4.0, 4.2, 4.4, 4.6$. The best-fit $\chi^2/\text{d.o.f.}$ values for the three cases are 55.4/44, 41.6/40, and 43.3/40 respectively, corresponding to p-values of $p = 0.12, 0.40$ and 0.33 . The data therefore exhibit a preference for an enhancement in the large scale power relative to models that assume a spatially homogeneous UV background.

the $u_0 - T_0$ plane. We base this on the thermal histories used in our hydrodynamical simulations, corresponding to a region bounded by $6 \leq (u_0/eV \text{ m}_p^{-1})(T_0/10^4 \text{ K})^{-1.7} \leq 12$ and $0.5 \leq T_0/10^4 \text{ K} \leq 1.5$. This prior enforces a tighter correlation between $u_0 - T_0$ that is almost perpendicular to the degeneracy axis between u_0 and T_0 found in the data. We note, however, that the $u_0 - T_0$ prior only affects parameter recovery from the power spectrum at the smallest scales, $k \sim 0.1 \text{ km}^{-1} \text{ s}^{-1}$, that are most sensitive to pressure smoothing and the thermal broadening kernel (Nasir et al. 2016); there is little effect on the recovery of A_p as a consequence. This result further suggests that the patchy reionization information contained in the A_p parameter is sensitive to the flux power spectrum on large scales only.

Finally, in Fig. 7, we compare the best-fit power spectrum models obtained from this analysis to the Karaçaylı et al. (2022) data in each redshift bin. Note that the goodness-of-fit varies across individual redshift bins, with the poorest at $z = 4.6$ where – although the error bars are generally larger – there is some tension between the models and the change of the amplitude of the observed power

spectrum from small to large scales. Note also that adjacent redshift bins are correlated, such that performing a simple “chi-by-eye” can be misleading. The joint fit across all four redshift bins is reasonable, with best-fit $\chi^2/\text{d.o.f.}$ values of 55.4/44, 41.6/40, and 43.3/40 for the G_{homog} , $G_{\text{homog}} + A_p$, and $G_{\text{homog}} + A_p + u_0 - T_0$ prior cases respectively, corresponding to p-values of $p = 0.12, 0.40$ and 0.33 . The improved $\chi^2/\text{d.o.f.}$ values for the $G_{\text{homog}} + A_p$ cases with or without the $u_0 - T_0$ prior indicate that introducing the parameter A_p in the MCMC analysis – which introduces a boost to the large scale 1D power spectrum and a small suppression at small scales (see Fig. 4 and M22) – leads to a better fit to the observational data. There is a preference for A_p being required by the data at a significance of 2.7σ (i.e. for an enhancement in the 1D power spectrum at large scales, $\log(k/\text{km}^{-1} \text{ s}) \sim -3$, relative to Ly α forest models that assume a homogeneous UV background at $z > 4$). This is in qualitative agreement with the independent analysis presented by D’Aloisio et al. (2018), although the formal significance we derive is higher. This may hint at the presence of relic fluctuations in the ionization and thermal state of the IGM following the completion of

reionization (e.g. Cen et al. 2009; Keating et al. 2018; Oñorbe et al. 2019; Wu et al. 2019; Montero-Camacho & Mao 2020; Puchwein et al. 2022).

4.3 Alternative explanations for enhanced large scale power

Although the preference for non-zero A_p may be a signature of patchy reionization, other factors may be contributing to the boost in the 1D power spectrum at large physical scales. We briefly discuss some other possibilities here.

First, the enhanced flux power spectrum at large scales could have a non-astrophysical origin. Estimators used for the statistical analysis of fluctuations in the transmitted Ly α flux require modelling of the intrinsic quasar continuum (e.g. Francis et al. 1992; Suzuki et al. 2005; Āurovčřková et al. 2020; Bosman et al. 2021). The modelling of this intrinsic quasar property is a complex procedure, and can leave residual contamination in the transmitted flux fluctuations. This can be due to the diversity in quasar continua when employing statistical methods to reconstruct the continuum, or due to absorption that can remove a large fraction of the intrinsic flux. The second of the two is especially of note for high-redshift quasars, where the larger IGM neutral fraction causes on average stronger absorption. A third possibility is linked to the echelle spectrographs commonly used in high-resolution spectroscopy. Observations made at higher echelle orders produce distortions in the spectrum over each order that are subsequently corrected for during flux calibration. This procedure can leave residual fluctuations imprinted on the absorption features. As has already been pointed out elsewhere (e.g. McDonald et al. 2005; Iršič et al. 2017b; Karaçaylı et al. 2022), any of the above mentioned sources of quasar continuum fluctuations could lead to enhanced power on larger scales, and this enhancement will be similar to the effect of patchy reionization. With this specifically in mind, Karaçaylı et al. (2022) marginalized over the parameters of their quasar continuum model in order to mitigate any scale-dependence from variations in the continuum. The preference for an enhancement in the large scale power we find in this work is therefore obtained *already assuming* a reasonable range of variation in the continuum. The marginalization procedure used by Karaçaylı et al. (2022), however, significantly increases the uncertainty on the power spectrum on large scales, and therefore effectively limits the increase in statistical power gained from combining different data sets.

There can also be other astrophysical effects that contribute to the shape of the large-scale power spectrum. Outflows driven by active galactic nuclei can *suppress* power on large scales, although this effect only becomes important toward lower redshifts, $z \lesssim 2.5$, where the volume filling factor of the outflows is larger and the Ly α forest is sensitive to higher density gas (Viel et al. 2013b; Chabanier et al. 2020). At higher redshifts, enhancements to the power spectrum on large scales are more likely to arise from high column density systems, such as damped Ly α absorbers and super Lyman limit systems with column densities $N_{\text{HI}} \gtrsim 10^{19} \text{ cm}^{-2}$. These systems have large Lorentzian damping wings that correlate on scales of order of $\sim 1000 \text{ s}^{-1} \text{ km}$ or $\sim 10 h^{-1} \text{ Mpc}$. Measurements of the Ly α forest flux power spectrum therefore typically compile a catalogue of such systems, and the affected regions are subsequently masked (Iršič et al. 2017b; Walther et al. 2019; Boera et al. 2019), but at the cost of a loss of information and hence statistical power (Palanque-DeLabrouille et al. 2020; Karaçaylı et al. 2022). However, due to incompleteness in the catalogues, particularly in low signal-to-noise data, residual effects from damping wings can remain. This effect was studied in detail by Rogers et al. (2018) (and see also McDonald et al. 2005; Font-Ribera & Miralda-Escudé 2012), where it

was found that damped systems introduce a large-scale enhancement in the power spectrum that is similar to that expected from patchy reionization. A more complex analysis would therefore seek to further marginalize over the parameters of a model for contamination by damped absorbers.

The onset of inhomogeneous He II reionisation at redshifts $z \lesssim 4$ (e.g. Worseck et al. 2016) could also potentially impact on the Ly α forest power spectrum, although the effect on the 1D power spectrum at large scales is minimal (McQuinn et al. 2011). The largest effect is instead expected at small scales due to increased line widths associated with He II photo-heating (La Plante et al. 2018; Upton Sanderbeck & Bird 2020). Furthermore the bulk of the additional He II photo-heating should occur at lower redshifts than those we consider here, particularly if He II reionization does not fully complete until $z \sim 3$.

Lastly, in this work we have assumed a fixed Λ CDM cosmology, but the enhancement of large-scale power by reionization also has important consequences for the inference of cosmological parameters from the Ly α forest power spectrum. This will be relevant for parameters that derive most of the constraining power from large scales. For example, parameters that change the amplitude of matter fluctuations, such as σ_8 and the sum of the neutrino masses, Σ_ν , could be biased high if the impact of reionization on large scales is not accounted for. Higher precision measurements of the flux power spectrum at large scales could alleviate this tension, as the change in the shape of power spectrum will be different when varying either σ_8 or the reionization model. However, such reasoning is more complicated for the value of the spectral index, n_s , and the running of the spectral index, α_s . A combination of a lower value for n_s and a lower mean Ly α forest transmission can mimic the shape of the reionization signal when averaged over too narrow a range of wavenumbers. Such an analysis therefore runs the risk of n_s being biased low, or instead trading a low n_s value for a non-zero running of the spectral index. Further study of these effects will be required when inferring cosmological parameters from the Ly α forest power spectrum at $z > 4$.

4.4 Implications for future constraints on inhomogeneous reionization

We now briefly turn to consider how future, high precision measurements of the Ly α forest power spectrum on large scales may be used to detect the signature of inhomogeneous reionization. We achieve this by constructing mock realisations of the power spectrum using the RT-late simulation described in M22. We then vary the size of the relative uncertainty on the mock realisations, and apply our parameter inference framework to obtain A_p . This enables us to assess the precision that the true underlying value of A_p used in the simulation – which we refer to as A_p^{true} – is recovered from the mock data.

In Fig. 8, we obtain a distribution for the recovered best fit A_p parameter after performing our MCMC analysis assuming different relative uncertainties on the power spectrum measurements on 500 mock realisations. We show the median (solid curves) and the 68 per cent confidence intervals (shaded regions) for A_p/A_p^{true} in three different redshift bins for different relative uncertainties on the power spectrum. We start from a 20 per cent relative uncertainty – comparable to that in Karaçaylı et al. (2022) – and extend this to higher precision cases that may be within reach of future surveys such as WEAVE-QSO and DESI. Note the recovered 68 per cent scatter is smaller in the higher redshift bins because the large scale enhancement in the 1D power spectrum due to patchy reionization is more prominent approaching the end-point of reionization.

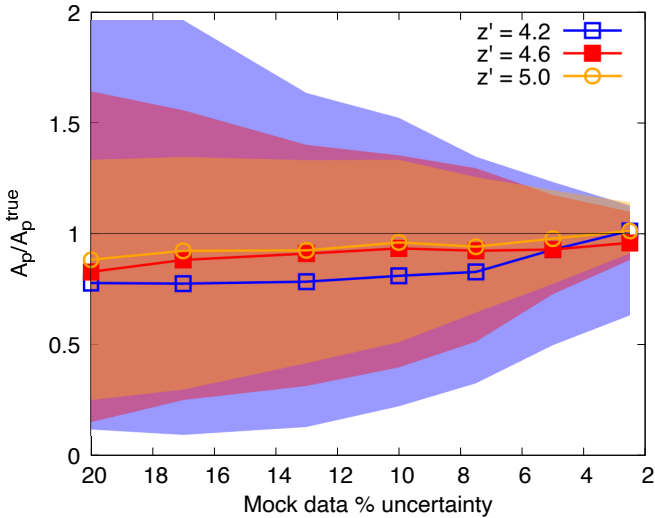


Figure 8. Median (solid curves) and 68 per cent confidence interval (shaded regions) for A_p/A_p^{true} obtained from 500 mock realisations of the 1D Ly α forest power spectrum with varying relative uncertainties. The results are shown separately for the three redshift bins considered. The solid grey line shows the perfect recovery case, where $A_p/A_p^{\text{true}} = 1$.

Fig. 8 highlights how the precision of the A_p^{true} recovery significantly improves as the relative uncertainty on the mock data decreases. For future, high-precision observations of the 1D Ly α power spectrum, this opens the possibility of using the power spectrum on large scales to unambiguously recover the signature of post-reionisation fluctuations in the IGM thermal state. We expect that this will be most useful when combined with a joint analysis of the transmitted flux distribution (e.g. Bosman et al. 2022) to obtain self-consistent constraints on the timing of reionization.

5 CONCLUSIONS

Inhomogeneous reionization impacts on the shape of the 1D Ly α forest power spectrum at redshifts $z \geq 4$. This is because of coherent fluctuations in the ionized hydrogen fraction on large scales (see also Cen et al. 2009; Keating et al. 2018; D’Aloisio et al. 2019; Oñorbe et al. 2019; Wu et al. 2019; Montero-Camacho & Mao 2020), and spatial variations in the thermal broadening kernel and divergent peculiar velocities at small scales (Wu et al. 2019; Oñorbe et al. 2019; Molaro et al. 2022).

In this work, we have used the Sherwood-Relics simulations (see Puchwein et al. 2022, for an overview) to assess the possibility of detecting the relic signature of inhomogeneous reionization at $4.0 \leq z \leq 4.6$ using measurements of the 1D power spectrum at wavenumbers $\log_{10}(k/\text{km}^{-1} \text{s}) \leq -2.2$ (Karaçaylı et al. 2022), corresponding to larger scales than those considered in recent analyses of the intergalactic medium (IGM) thermal state (Boera et al. 2019; Molaro et al. 2022, although see also D’Aloisio et al. (2018) for earlier work using data at higher redshift). We update the analysis framework introduced by Molaro et al. (2022) by introducing a neural-network-based interpolator in our Bayesian inference analysis. This approach improves the control of interpolation uncertainties in our analysis (typically < 1 per cent), and is more computationally efficient than our previous linear interpolation technique. We also use a more general, model independent approach for parameterising

the power spectrum template for inhomogeneous reionization first presented by Molaro et al. (2022). Our main results are as follows:

- As already discussed by Molaro et al. (2022), we find that if considering mock realisations of the Ly α forest power spectrum at wavenumbers $-2.2 \leq \log_{10}(k/\text{km}^{-1} \text{s}) \leq -0.7$ with 5 per cent relative uncertainties, patchy reionization introduces a modest $\sim 1\sigma$ shift between the true and recovered IGM parameters if (incorrectly) assuming a spatially uniform UV background. However, when including larger scales with wavenumbers $\log_{10}(k/\text{km}^{-1} \text{s}) \approx -3$, the bias in the parameter recovery is much larger and is at the level of $\sim 6\sigma$. This demonstrates that IGM parameter recovery is significantly more sensitive to inhomogeneous reionization when including power spectrum data on large scales.

- We perform a first analysis of the Ly α forest power spectrum measurements presented recently by Karaçaylı et al. (2022), which are based on the combined observational data from the Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ) (O’Meara et al. 2017), the Spectral Quasar Absorption Database (SQUAD) (Murphy et al. 2019) and the XQ-100 survey (López et al. 2016). These data cover four redshift bins at $z = 4.0, 4.2, 4.4$, and 4.6 , and extend to large scales, $\log_{10}(k/\text{km}^{-1} \text{s}) \approx -3$ and have typical relative uncertainties of 10–20 per cent in each wavenumber bin. We find a preference at the 2.7σ level for an enhancement in the Karaçaylı et al. (2022) power spectrum at large scales relative to models that assume a spatially uniform background. This additional power may be due to patchy reionization, although we caution that systematic effects (e.g. variations in the shape of the continuum placement, or the damping wings of high column density systems) could still contribute.

- The precision of the Karaçaylı et al. (2022) power spectrum data at large scales is insufficient for recovering the signature of inhomogeneous reionization at high significance. However, forthcoming surveys such as the Dark Energy Spectroscopic Instrument (DESI) survey (Vargas-Magana et al. 2019) and William Herschel Telescope Enhanced Area Velocity Explorer QSO (WEAVE-QSO) survey (Pieri et al. 2016) will also extend to large scales at $2 \leq z \leq 4.5$, and will measure the 1D Ly α forest power spectrum to a precision of a few per cent. Any enhancement in the power spectrum on large scales should then be recoverable. We anticipate that combining the 1D power spectrum on large scales in a joint analysis with the transmitted flux distribution (e.g. Bosman et al. 2022) should provide a powerful constraint on the timing of reionization.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous referee for very helpful comments. We thank Steven Bamford for very useful discussions. The simulations used in this work were performed using the Joliot Curie supercomputer at the Très Grand Centre de Calcul (TGCC) and the Cambridge Service for Data Driven Discovery (CSD3), part of which is operated by the University of Cambridge Research Computing on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). We acknowledge the Partnership for Advanced Computing in Europe (PRACE) for awarding us time on Joliot Curie in the 16th call. The DiRAC component of CSD3 was funded by BEIS capital funding via STFC capital grants ST/P002307/1 and ST/R002452/1 and STFC operations grant ST/R00689X/1. This work also used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility. The

equipment was funded by BEIS capital funding via STFC capital grants ST/P002293/1 and ST/R002371/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. MM and JSB are supported by STFC consolidated grant ST/T000171/1. VI is supported by the Kavli foundation. MGH acknowledges support from the UKRI STFC (grant numbers ST/N000927/1 and ST/S000623/1). LCK was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 885990. We thank Volker Springel for making P-GADGET-3 available. We also thank Dominique Aubert for sharing the ATON code. We acknowledge use of the keras2c library in this work (Conlin et al. 2021). For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

DATA AVAILABILITY

All data and analysis code used in this work are available from the first author on reasonable request.

REFERENCES

- Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, *MNRAS*, **447**, 3402
- Bird S., Rogers K. K., Peiris H. V., Verde L., Font-Ribera A., Pontzen A., 2019, *J. Cosmology Astropart. Phys.*, **2019**, 050
- 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., Haehnelt M. G., Viel M., Springel V., 2005, *MNRAS*, **357**, 1178
- Bosman S. E. I., Āurovčřková D., Davies F. B., Eilers A.-C., 2021, *MNRAS*, **503**, 2077
- Bosman S. E. I., et al., 2022, *MNRAS*, **514**, 55
- Cen R., McDonald P., Trac H., Loeb A., 2009, *The Astrophysical Journal*, **706**, L164
- Chabanier S., et al., 2019, *J. Cosmology Astropart. Phys.*, **2019**, 017
- Chabanier S., Bornaud F., Dubois Y., Palanque-Delabrouille N., Yęche C., Armengaud E., Peirani S., Beckmann R., 2020, *MNRAS*, **495**, 1825
- Conlin R., Erickson K., Abbate J., Kolemen E., 2021, *Engineering Applications of Artificial Intelligence*, **100**
- Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, *ApJ*, **581**, 20
- 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, *The Astrophysical Journal*, **874**, 154
- Eilers A.-C., Davies F. B., Hennawi J. F., 2018, *ApJ*, **864**, 53
- Fan X., Carilli C. L., Keating B., 2006, *ARA&A*, **44**, 415
- Fernandez M. A., Ho M.-F., Bird S., 2022, arXiv e-prints, p. arXiv:2207.06445
- Font-Ribera A., Miralda-Escudę J., 2012, *J. Cosmology Astropart. Phys.*, **2012**, 028
- Francis P. J., Hewett P. C., Foltz C. B., Chaffee F. H., 1992, *ApJ*, **398**, 476
- Fukushima K., 1975, *Biological cybernetics*, **20**, 121
- Gaikwad P., et al., 2020, *MNRAS*, **494**, 5091
- Gaikwad P., Srianand R., Haehnelt M. G., Choudhury T. R., 2021, *MNRAS*, **506**, 4389
- Garzilli A., Magalich A., Theuns T., Frenk C. S., Weniger C., Ruchayskiy O., Boyarsky A., 2019, *MNRAS*, **489**, 3456
- Gurvich A., Burkhart B., Bird S., 2017, *ApJ*, **835**, 175
- Hiss H., Walthor M., Hennawi J. F., Oñorbe J., O'Meara J. M., Rorai A., Lukić Z., 2018, *ApJ*, **865**, 42
- Hornik K., Stinchcombe M., White H., 1989, *Neural networks*, **2**, 359
- Hsu T., Cooke R. J., Prochaska J. X., Bolte M., 2020, *ApJ*, **896**, 77
- Hui L., Gnedin N. Y., 1997, *MNRAS*, **292**, 27
- Hui L., Haïman Z., 2003, *ApJ*, **596**, 9
- Iršič V., et al., 2017a, *Phys. Rev. D*, **96**, 023522
- Iršič V., et al., 2017b, *MNRAS*, **466**, 4332
- Karaçaylı N. G., Font-Ribera A., Padmanabhan N., 2020, *MNRAS*, **497**, 4742
- Karaçaylı N. G., et al., 2022, *MNRAS*, **509**, 2842
- Keating L. C., Puchwein E., Haehnelt M. G., 2018, *MNRAS*, **477**, 5501
- La Plante P., Trac H., Croft R., Cen R., 2018, *ApJ*, **868**, 106
- LeCun Y. A., Bottou L., Orr G. B., Müller K.-R., 2012, in , *Neural networks: Tricks of the trade*. Springer, pp 9–48
- López S., et al., 2016, *A&A*, **594**, A91
- McDonald P., Miralda-Escudę J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, *ApJ*, **543**, 1
- McDonald P., Seljak U., Cen R., Bode P., Ostriker J. P., 2005, *MNRAS*, **360**, 1471
- McGreer I. D., Mesinger A., D'Odorico V., 2015, *MNRAS*, **447**, 499
- McQuinn M., 2016, *ARA&A*, **54**, 313
- McQuinn M., Hernquist L., Lidz A., Zaldarriaga M., 2011, *MNRAS*, **415**, 977
- Mishra N., Gnedin N. Y., 2022, *ApJ*, **928**, 174
- Molaro M., et al., 2022, *MNRAS*, **509**, 6119
- Montero-Camacho P., Mao Y., 2020, *MNRAS*, **499**, 1640
- Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2019, *MNRAS*, **482**, 3458
- Nasir F., Bolton J. S., Becker G. D., 2016, *MNRAS*, **463**, 2335
- Oñorbe J., Davies F. B., Lukić Z., Hennawi J. F., Sorini D., 2019, *MNRAS*, **486**, 4075
- O'Meara J. M., Lehner N., Howk J. C., Prochaska J. X., Fox A. J., Peeples M. S., Tumlinson J., O'Shea B. W., 2017, *AJ*, **154**, 114
- Palanque-Delabrouille N., et al., 2013, *A&A*, **559**, A85
- Palanque-Delabrouille N., Yęche C., Schöneberg N., Lesgourgues J., Walthor M., Chabanier S., Armengaud E., 2020, *J. Cosmology Astropart. Phys.*, **2020**, 038
- Pan S. J., Yang Q., 2009, *IEEE Transactions on knowledge and data engineering*, **22**, 1345
- Pedersen C., Font-Ribera A., Rogers K. K., McDonald P., Peiris H. V., Pontzen A., Slosar A., 2021, *J. Cosmology Astropart. Phys.*, **2021**, 033
- Pedersen C., Font-Ribera A., Gnedin N. Y., 2022, arXiv e-prints, p. arXiv:2209.09895
- Pieri M. M., et al., 2016, in Reylę C., Richard J., Cambręsy L., Deleuil M., Pęcontal E., Tresse L., Vauglin I., eds, *SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*. pp 259–266 (arXiv:1611.09388)
- Planck Collaboration et al., 2014, *A&A*, **571**, A16
- Puchwein E., Haardt F., Haehnelt M. G., Madau P., 2019, *MNRAS*, **485**, 47
- Puchwein E., et al., 2022, arXiv e-prints, p. arXiv:2207.13098
- Rauch M., 1998, *ARA&A*, **36**, 267
- Rogers K. K., Peiris H. V., 2021, *Phys. Rev. Lett.*, **126**, 071302
- Rogers K. K., Bird S., Peiris H. V., Pontzen A., Font-Ribera A., Leistedt B., 2018, *MNRAS*, **474**, 3032
- Schaye J., Theuns T., Rauch M., Efstathiou G., Sargent W. L. W., 2000, *MNRAS*, **318**, 817
- Seljak U., Makarov A., McDonald P., Trac H., 2006, *Phys. Rev. Lett.*, **97**, 191303
- Stone M., 1974, *Journal of the royal statistical society: Series B (Methodological)*, **36**, 111
- Suzuki N., Tytler D., Kirkman D., O'Meara J. M., Lubin D., 2005, *ApJ*, **618**, 592
- Theuns T., Schaye J., Zaroubi S., Kim T.-S., Tzanavaris P., Carswell B., 2002a, *ApJ*, **567**, L103
- Theuns T., Viel M., Kay S., Schaye J., Carswell R. F., Tzanavaris P., 2002b, *ApJ*, **578**, L5
- Upton Sanderbeck P., Bird S., 2020, *MNRAS*, **496**, 4372
- Vargas-Magana M., Brooks D. D., Levi M. M., Tarle G. G., 2019, arXiv e-prints, p. arXiv:1901.01581
- Viel M., Haehnelt M. G., 2006, *MNRAS*, **365**, 231

- Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013a, *Phys. Rev. D*, **88**, 043502
- Viel M., Schaye J., Booth C. M., 2013b, *MNRAS*, **429**, 1734
- Villasenor B., Robertson B., Madau P., Schneider E., 2022, arXiv e-prints, p. [arXiv:2209.14220](https://arxiv.org/abs/2209.14220)
- Walther M., Hennawi J. F., Hiss H., Oñorbe J., Lee K.-G., Rorai A., O'Meara J., 2018, *ApJ*, **852**, 22
- Walther M., Oñorbe J., Hennawi J. F., Lukić Z., 2019, *ApJ*, **872**, 13
- Worseck G., Prochaska J. X., Hennawi J. F., McQuinn M., 2016, *ApJ*, **825**, 144
- Wu X., McQuinn M., Kannan R., D'Aloisio A., Bird S., Marinacci F., Davé R., Hernquist L., 2019, *MNRAS*, **490**, 3177
- Yèche C., Palanque-Delabrouille N., Baur J., du Mas des Bourboux H., 2017, *J. Cosmology Astropart. Phys.*, **2017**, 047
- Đurovčiková D., Katz H., Bosman S. E. I., Davies F. B., Devriendt J., Slyz A., 2020, *MNRAS*, **493**, 4256

ORIGINAL UNEDITED MANUSCRIPT