The effect of inhomogeneous reionization on the Lyman α forest power spectrum at redshift $z > 4$: implications for thermal parameter recovery

Margherita Molaro [ID],[1][*] Vid Iršič,[2] James S. Bolton [ID],[1] Laura C. Keating [ID],[3] Ewald Puchwein [ID],[3] Prakash Gaikwad [ID],[2] Martin G. Haehnelt,[2] Girish Kulkarni [ID][4] and 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]*Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, D-14482 Potsdam, Germany*
[4]*Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India*
[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*

ABSTRACT

We use the Sherwood–Relics suite of hybrid hydrodynamical and radiative transfer simulations to model the effect of inhomogeneous reionization on the 1D power spectrum of the Lyman α (Lyα) forest transmitted flux at redshifts $4.2 \leq z \leq 5$. Relative to models that assume a homogeneous ultraviolet background, reionization suppresses the power spectrum at small scales, $k \sim 0.1 \, \mathrm{km}^{-1} \, \mathrm{s}$, by $\sim$10 per cent because of spatial variations in the thermal broadening kernel and the divergent peculiar velocity field associated with overpressurized intergalactic gas. On larger scales, $k < 0.03 \, \mathrm{km}^{-1} \, \mathrm{s}$, the power spectrum is instead enhanced by 10–50 per cent by large-scale spatial variations in the neutral hydrogen fraction. The effect of inhomogeneous reionization must therefore be accounted for in analyses of forthcoming high precision measurements. We provide a correction for the Lyα forest power spectrum at $4.1 \leq z \leq 5.4$ that can be easily applied within other parameter inference frameworks using similar reionization models. We perform a Bayesian analysis of mock data to assess the extent of systematic biases that may arise in measurements of the intergalactic medium if ignoring this correction. At the scales probed by current high-resolution Lyα forest data at $z > 4$, $0.006 \leq k \leq 0.2 \, \mathrm{km}^{-1} \, \mathrm{s}$, we find inhomogeneous reionization does not introduce any significant bias in thermal parameter recovery for the current measurement uncertainties of $\sim$10 per cent. However, for 5 per cent uncertainties, $\sim 1\sigma$ shifts between the estimated and true parameters occur.

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

1 INTRODUCTION

The study of Lyman α (Lyα) absorption features in the spectra of bright, high-redshift quasars offers a valuable insight into the physical properties of the intergalactic medium (IGM) in the early Universe (Fan, Carilli & Keating 2006; Mortlock et al. 2011; Becker et al. 2015; Eilers et al. 2017; Bosman et al. 2018; Yang et al. 2020). Cumulatively referred to as the 'Lyα forest', these absorption features provide constraints on the ionization and thermal state of the IGM during and immediately after the final

stages of reionization at redshifts $z \simeq 5$–7 (Oñorbe et al. 2017b; Kulkarni et al. 2019; Walther et al. 2019; Gaikwad et al. 2020; Keating et al. 2020; Nasir & D'Aloisio 2020; Qin et al. 2021). Additionally, because the Lyα forest closely tracks the dark matter down to scales of $\sim$100 ckpc, it is also sensitive to the suppression of the matter power spectrum on small scales (see e.g. Seljak et al. 2006; Viel et al. 2008; Boyarsky et al. 2009). The power spectrum

of the Lyα forest transmitted flux thus provides one of the tightest lower limits on the mass of a putative warm dark matter thermal relic (Viel et al. 2013; Iršič et al. 2017a; Garzilli et al. 2019; Palanque-Delabrouille et al. 2020; Rogers & Peiris 2021).

Existing constraints on reionization and warm dark matter from the Lyα forest typically rely on high-quality quasar spectra combined with numerical models for the distribution of matter in the IGM (see e.g. McQuinn 2016). On the observational side, the number of known quasars at $z > 5$ has significantly increased in recent years (e.g. Bañados et al. 2016; Matsuoka et al. 2018; Reed et al. 2019; Wang et al. 2019), and the latest measurements of the Lyα forest power spectrum exhibit improved precision and extend toward smaller scales (Iršič et al. 2017c; Boera et al. 2019; Chabanier et al. 2019; Karaçaylı et al. 2021; Wilson, Iršič & McQuinn 2021). The computational demands on state-of-the-art simulations of the high-redshift Lyα forest are, however, still formidable. The simulations must capture the patchy thermal and ionization state of the IGM following reionization, have a large dynamic range that resolves gas at the Jeans scale (Theuns, Schaye & Haehnelt 2000; Bolton & Becker 2009) while simultaneously sampling a volume large enough

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

to follow the percolation of ionized bubbles (Iliev et al. 2014), and they must furthermore span a large and uncertain parameter space.

One of the most expensive physical processes to implement is the prescription for the radiative transfer (RT) of ultraviolet (UV) photons during inhomogeneous reionization (Gnedin 2000; Razoumov et al. 2002; Ciardi, Ferrara & White 2003; Mellema et al. 2006; Finlator et al. 2018; Molaro et al. 2019). A common approximation used in hydrodynamical simulations of the Lyα forest that sidesteps this requirement are pre-computed, spatially homogeneous ultraviolet background (UVB) models (Haardt & Madau 2012; Khaire & Srianand 2019; Puchwein et al. 2019; Faucher-Giguère 2020). While these 'UVB synthesis' models provide an excellent description of the IGM ionization state following reionization, when the mean free path for Lyman continuum photons is $\gtrsim$100 cMpc (see e.g. Lukić et al. 2015; Bolton et al. 2017; Rossi 2020; Villasenor et al. 2021), they neglect the large-scale fluctuations in the ionization and thermal state of the IGM that exist immediately following the completion of reionization (Becker et al. 2015; D'Aloisio, McQuinn & Trac 2015; Davies & Furlanetto 2016; Chardin, Puchwein & Haehnelt 2017; Kulkarni et al. 2019). It has been suggested that ignoring these fluctuations could weaken existing Lyα forest constraints on the free streaming length of dark matter, leaving the door firmly open for alternatives to cold dark matter (e.g. Hui et al. 2017).

Further progress in this area therefore necessitates the development of efficient and accurate numerical schemes that capture the effect of inhomogeneous hydrogen reionization on the Lyα forest at $z > 4$. Ideally, any such scheme should also be straightforward to incorporate into existing Lyα forest power spectrum parameter estimation frameworks. While significant progress has been made on this problem using hydrodynamically decoupled RT simulations (e.g. Cen et al. 2009; Keating, Puchwein & Haehnelt 2018; D'Aloisio et al. 2019) or seminumerical reionization models (e.g. Lidz & Malloy 2014; Montero-Camacho et al. 2019), these neglect the dynamical effect of patchy heating on the small-scale structure of the Lyα forest (i.e. the scales at wavenumbers $k \sim 0.1\,\mathrm{km}^{-1}\,\mathrm{s}$ that are most sensitive to the IGM thermal state and the coldness of dark matter at $z > 4$). Recent efforts toward addressing this deficiency have been presented by Oñorbe et al. (2019) and Wu et al. (2019), using independent approaches. Oñorbe et al. (2019) used a hybrid method that couples Eulerian hydrodynamical simulations performed with NYX(Almgren et al. 2013) with a seminumerical reionization model where energy is injected into the IGM by hand. This hybrid approach has the advantage of speed and efficiency, but at the cost of using an approximate treatment for the photoheating of the IGM. By contrast, Wu et al. (2019) used AREPO-RT (Kannan et al. 2019) to perform multifrequency radiation hydrodynamical (RHD) simulations. The RHD simulations self-consistently model photoheated gas temperatures, but at the expense of increased computational cost.

In this work, we complement these studies by adopting a third approach that is intermediate between seminumerical models and full RHD simulations. The simulations we use here are part of the Sherwood–Relics project (Puchwein et al., in preparation), a large-scale set of IGM simulations that directly builds upon our earlier Sherwood simulation project (Bolton et al. 2017). In Sherwood–Relics, we model the effect of inhomogeneous reionization on the high-redshift Lyα forest using a hybrid approach that combines RT calculations performed using ATON (Aubert & Teyssier 2008, 2010) with P-GADGET-3 cosmological hydrodynamical simulations (Springel 2005). By using a two-step approach, where ionization maps produced by empirically calibrated RT calculations are applied on-the-fly to the hydrodynamical simulations (see Section 2 for de-

tails), we are able to capture the patchy ionization and thermal state of the IGM *and* self-consistently model the hydrodynamical response of gas following inhomogeneous heating (see also Gaikwad et al. 2020; Šoltinský et al. 2021, for other recent applications of this approach). We then use the Sherwood–Relics simulations to construct and test a generalized 'patchy reionization' correction to the Lyα forest power spectrum at $z > 4$ predicted by (homogeneous UVB) hydrodynamical simulations. A key advantage of this approach is that the correction can be straightforwardly applied to existing grids of hydrodynamical simulations used in cosmological parameter inference frameworks (e.g. Bird et al. 2019; Boera et al. 2019; Walther et al. 2019; Rossi 2020), thus avoiding the need to perform large numbers of additional RHD simulations.

This paper is structured as follows. In Section 2, we describe the Sherwood–Relics simulations and introduce the reionization models used in this work. We examine the effect of inhomogeneous reionization on the Lyα forest power spectrum at $4.2 \leq z \leq 5$ in our fiducial reionization model in Section 3, and discuss the physical origin of the generic features we observe: an enhancement of power on large scales, $k < 0.03\,\mathrm{km}^{-1}\,\mathrm{s}$, and $\sim$10 per cent suppression of power on small scales, $k > 0.1\,\mathrm{km}^{-1}\,\mathrm{s}$ (cf. Oñorbe et al. 2019; Wu et al. 2019). In Section 4, we then expand our analysis to consider different reionization histories and construct a generalized inhomogeneous reionization correction to the Lyα forest power spectrum that can be applied to 'traditional' hydrodynamical simulations of the IGM (for readers wishing to skip the details, this correction is implemented using equation 3 and Table 2, and a simple PYTHON script to compute this correction is available at https://github.com/marghemolaro/RT_1dps_correction.git). In Section 5, we then assess the importance of any biases that may be introduced to measurements of the IGM thermal state from the Lyα forest power spectrum by applying this correction within our existing Markov chain Monte Carlo (MCMC) analysis framework (Viel et al. 2013; Iršič et al. 2017a). Finally, we conclude in Section 6. Supplementary information is provided in Appendix A.

2 SIMULATIONS OF THE LYMAN α FOREST

2.1 Hydrodynamical simulations of the IGM during reionization

We use simulations drawn from the Sherwood–Relics project (see Gaikwad et al. 2020; Šoltinský et al. 2021; Puchwein et al., in preparation). These are a series of high-resolution cosmological hydrodynamical simulations that use a customized version of P-GADGET-3 (see Springel 2005 for the original GADGET-2 reference). We use cosmological boxes of size 20 or 40 h^{-1} cMpc with 2×1024^3 or 2×2048^3 dark matter and gas particles. The box size and mass resolution have been chosen to adequately resolve the small-scale structure that contributes to the power spectrum of the Lyα forest transmitted flux at $z > 4$, while still retaining a relatively large cosmological volume (Bolton & Becker 2009; Lukić et al. 2015; Bolton et al. 2017). Note, however, that coherent ionized or neutral structures on scales larger than the box size will not be present in our simulations (Iliev et al. 2014; Kaur, Gillet & Mesinger 2020). In all models we use a simple, computationally efficient scheme whereby a gas particle is converted into a collisionless star particle if it reaches an overdensity $\Delta = 1 + \delta > 10^3$ and temperature $T < 10^5$ K (Viel, Haehnelt & Springel 2004). We assume a flat Λ cold dark matter (ΛCDM) cosmology with $\Omega_\Lambda = 0.692$, $\Omega_m = 0.308$, $\Omega_b = 0.0482$, $\sigma_8 = 0.829$, $n_s = 0.961$, $h = 0.678$ (Planck Collaboration XVI 2014), and a primordial helium fraction by mass of $Y_p = 0.24$ (Hsyu et al.

Table 1. List of simulations used in this work. From left to right, the columns list the simulation name, the box size in h^{-1} cMpc, the number of particles, the dark matter and gas particle mass in h^{-1} M$_\odot$, the redshift of reionization (defined as the redshift when the volume-averaged ionized fraction $1 - x_{\rm HI} \leq 10^{-3}$), the Thomson scattering optical depth $\tau_{\rm e}$, the gas temperature at the mean density, T_0, the cumulative energy input per proton mass at the mean density, u_0, for $4.6 \leq z \leq 13$ (cf. Boera et al. 2019), and the method used for modelling the photoionization of the IGM by UV photons. The upper section of the table lists the models in the first set of simulations – all performed using (variations of) the Puchwein et al. (2019) UVB synthesis model – that we use for our MCMC analysis (see text for details). The lower section of the table lists our second set of simulations, which includes hybrid-radiative transfer (RT) simulations and the corresponding 'paired' homogeneous models that are matched to the thermal and reionization histories in the hybrid-RT runs. We do not quote T_0 or u_0 for the hybrid-RT models, as this quantity will vary spatially and depend on when any given gas element is reionized. However, note that (by design) the *average* thermal history will be very similar to the paired homogeneous simulations.

| Name	$L_{\rm box}$ (h^{-1} cMpc)	$N_{\rm part}$	$M_{\rm dm}$ (h^{-1} M$_\odot$)	$M_{\rm gas}$ (h^{-1} M$_\odot$)	$z_{\rm R}$	$\tau_{\rm e}$	$T_0(z=4.6)$ (K)	$u_0(z=4.6)$ (eV $m_{\rm p}^{-1}$)	UVB model
20-1024	20.0	2×1024^3	5.37×10^5	9.97×10^4	6.00	0.062	10 066	7.7	P19
20-1024-zr54	20.0	2×1024^3	5.37×10^5	9.97×10^4	5.37	0.055	10 069	6.6	Rescaled P19
20-1024-zr67	20.0	2×1024^3	5.37×10^5	9.97×10^4	6.70	0.071	10 050	9.6	Rescaled P19
20-1024-zr74	20.0	2×1024^3	5.37×10^5	9.97×10^4	7.40	0.079	10 003	11.4	Rescaled P19
20-1024-cold	20.0	2×1024^3	5.37×10^5	9.97×10^4	5.98	0.062	6598	4.3	Rescaled P19
20-1024-zr54-cold	20.0	2×1024^3	5.37×10^5	9.97×10^4	5.35	0.055	6409	3.6	Rescaled P19
20-1024-zr67-cold	20.0	2×1024^3	5.37×10^5	9.97×10^4	6.69	0.070	6803	5.4	Rescaled P19
20-1024-zr74-cold	20.0	2×1024^3	5.37×10^5	9.97×10^4	7.39	0.079	6806	6.4	Rescaled P19
20-1024-hot	20.0	2×1024^3	5.37×10^5	9.97×10^4	6.01	0.063	13 957	14.4	Rescaled P19
20-1024-zr54-hot	20.0	2×1024^3	5.37×10^5	9.97×10^4	5.38	0.055	13 451	12.5	Rescaled P19
20-1024-zr67-hot	20.0	2×1024^3	5.37×10^5	9.97×10^4	6.71	0.071	14 369	17.8	Rescaled P19
20-1024-zr74-hot	20.0	2×1024^3	5.37×10^5	9.97×10^4	7.41	0.080	14 624	21.1	Rescaled P19
40-2048	40.0	2×2048^3	5.37×10^5	9.97×10^4	6.00	0.062	10 063	7.7	P19
RT-late	40.0	2×2048^3	5.37×10^5	9.97×10^4	5.30	0.056	–	–	Hybrid-RT
RT-mid	40.0	2×2048^3	5.37×10^5	9.97×10^4	5.99	0.057	–	–	Hybrid-RT
RT-early	40.0	2×2048^3	5.37×10^5	9.97×10^4	6.64	0.064	–	–	Hybrid-RT
Homog-late	40.0	2×2048^3	5.37×10^5	9.97×10^4	5.30	0.054	10 545	5.4	Paired homogeneous
Homog-mid	40.0	2×2048^3	5.37×10^5	9.97×10^4	5.99	0.055	10 431	5.5	Paired homogeneous
Homog-early	40.0	2×2048^3	5.37×10^5	9.97×10^4	6.64	0.062	10 245	6.2	Paired homogeneous

2020). The initial conditions for the simulations are identical to those used for our earlier Sherwood simulation project (Bolton et al. 2017).

We consider two distinct sets of simulations in this work: (i) a large grid of 'traditional' hydrodynamical simulations of the IGM that use the UVB synthesis model from Puchwein et al. (2019), and (ii) a smaller set of simulations, half of which follow the reionization of the IGM using a hybrid-RT approach and include the hydrodynamical response of the gas to inhomogeneous heating, and half of which use the same spatially uniform UVB approach adopted in set (i), but where this has now been adjusted to match the average ionization and thermal histories of the hybrid-RT simulations. The models are summarized in Table 1, and the details of each set are discussed below.

The first set of simulations is constructed using modifications to the spatially uniform UVB synthesis model introduced by Puchwein et al. (2019). These simulations will be used to construct the grid of models we use for our MCMC analysis in Section 5, and are similar to models we have used in our earlier work on the Lyα forest power spectrum at high redshift (Viel et al. 2013; Nasir, Bolton & Becker 2016; Iršič et al. 2017a). The main improvements in this study are the larger dynamic range of the simulations, the use of a non-equilibrium thermochemistry solver (Puchwein et al. 2015) and an improved treatment of the IGM opacity that consistently captures the transition from a neutral to ionized IGM (see also Oñorbe, Hennawi & Lukić 2017a). In addition to running a model with the fiducial Puchwein et al. (2019) UVB, we also vary the photoheating rates to achieve models with different gas temperatures and/or end redshifts for reionization, following the approach described in Becker et al. (2011).

In the second set of simulations, half of the models include a hybrid-RT approach, in which we follow the RT of monochromatic

UV photons with the moment-based M1-closure RT code ATON (Aubert & Teyssier 2008). We construct the models as follows (see also Puchwein et al., in preparation): ATON is initially run on a base P-GADGET-3 simulation, using snapshots that are spaced at time intervals of $t = 40$ Myr. This produces three-dimensional maps of spatially varying H I photoionization rates, $\Gamma_{\rm HI}$, as a function of redshift. The luminosity of H I photoionizing sources is assumed to be proportional to the total halo mass, with a halo mass threshold of $M_{\rm h} > 10^9\ h^{-1}$ M$_\odot$ and a mean energy for the ionizing photons of 18.6 eV. Note, however, the ATON simulations do not make a prediction for the *absolute* ionizing luminosity associated with each halo. Instead, we take advantage of the efficiency of ATON to calibrate the ionizing emissivity, thus producing specific realizations of the reionization history. In this work, we construct three different reionization models that correspond to an 'early' ($z_{\rm R} \sim 5.3$), 'mid' ($z_{\rm R} \sim 6.0$), and 'late' ($z_{\rm R} \sim 6.6$) end redshift for reionization, $z_{\rm R}$, defined as the redshift when the volume-averaged H I fraction first falls below 10^{-3}. Adjusting the emissivity in this way is equivalent to treating the (uncertain) escape fraction from the ionizing sources as a free parameter. The resulting H I photoionization rate maps are then used as input in a rerun of the base P-GADGET-3 model, but now incorporating the *hydrodynamical* response of the gas to the inhomogeneous ionization and heating. The H I photoheating rate is obtained by multiplying through the photoionization maps by the excess energy per H I photoionization, 18.6 eV − 13.6 eV = 5.0 eV, yielding IGM gas temperatures consistent with recent observational constraints from Gaikwad et al. (2020). The He I photoionization rate is set equal to the H I photoionization rate, and the He I photoheating rate is 1.3 times that of the H I photoheating rate, matching the ratios from Puchwein et al. (2019). Finally, the He II photoionization and
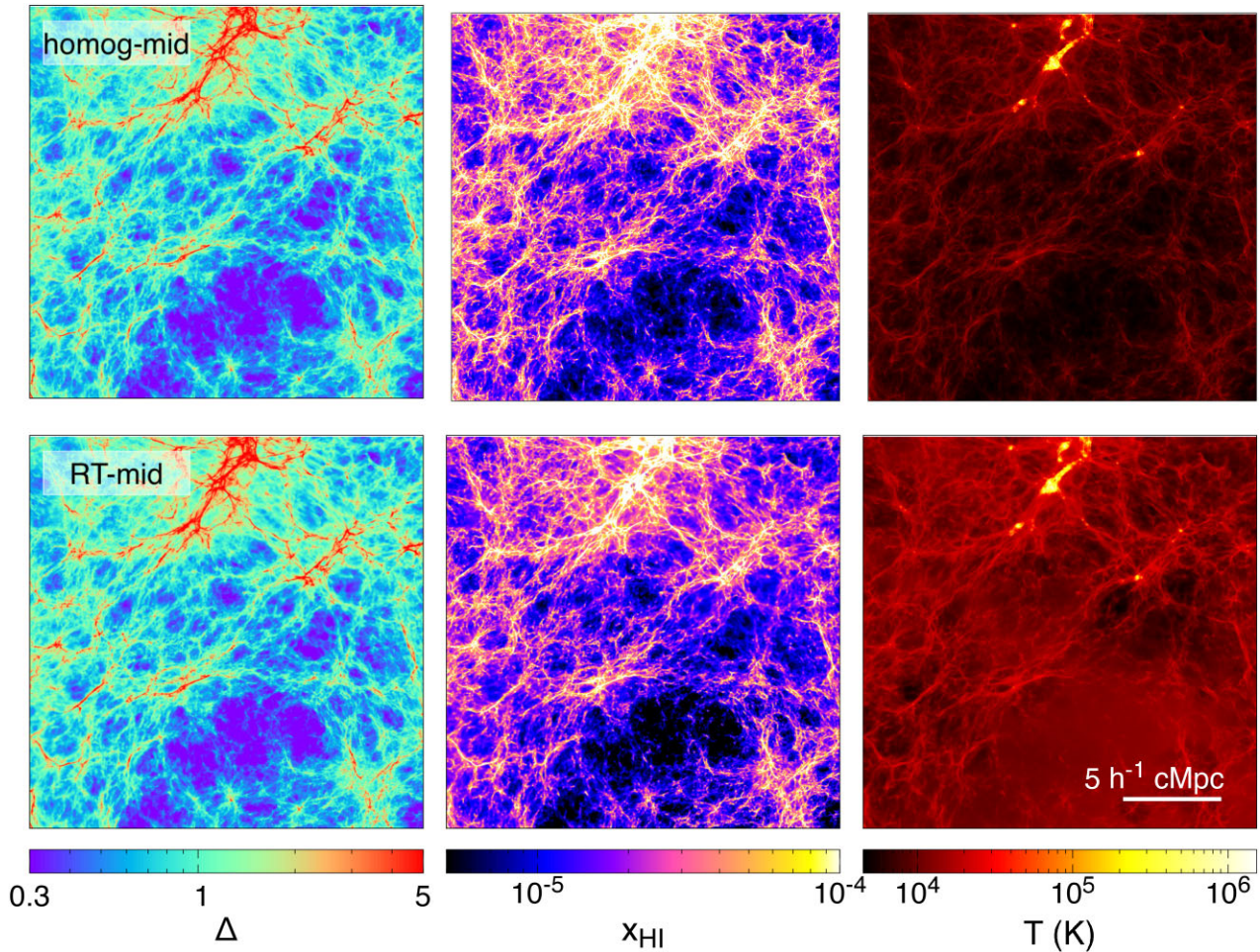
Figure 1. The gas density $\Delta = \rho/\langle\rho\rangle$ (left-hand column), neutral hydrogen fraction $x_{\mathrm{H\,I}}$ (middle column), and gas temperature T (right-hand column) in a $20\,h^{-2}\,\mathrm{cMpc}^2$ region of the Homog-mid (top row) and RT-mid (bottom row) simulation at $z = 5.4$. The slices have been projected over a slab of thickness $2\,h^{-1}\,\mathrm{cMpc}$. Differences arising from patchy reionization are particularly visible in the large void in the lower right-hand corner of the slice, where the later reionization and heating of the low-density gas lead to a higher temperatures and lower neutral fractions in the hybrid-RT model.

photoheating rates are assumed to be spatially homogeneous[1] and are adopted from Puchwein et al. (2019). The advantage of this multistep approach is that it consistently models the small-scale structure of the diffuse IGM within different constrained reionization histories. We can then directly contrast the results to simulations with spatially uniform ionization and heating by a homogeneous UVB (see also Oñorbe et al. 2019, for a related approach).

For the other half of the models in our second set of simulations, we therefore construct paired simulations that make use of the same homogeneous UVB approach as our grid of 'traditional' models, but where this has now been adjusted to reproduce the average reionization histories in the hybrid-RT models.[2] This allows us to

perform a direct comparison of the hybrid-RT simulations to models with a spatially uniform ionizing background. We ensure that – for each reionization history considered – the evolution of the average H I fraction and the gas temperature at mean density are consistent across the paired simulations. We achieve this by tuning the spatially uniform UVB model in the paired homogeneous simulations to match the volume-averaged H I fraction, the volume-averaged IGM temperature (at mean density) during reionization, and the median IGM temperature (at mean density) after reionization. The median temperature has the advantage that it is less affected by the shock heating of a small fraction of the gas to high temperatures following reionization. Further details can also be found in Puchwein et al. (in preparation). This allows us to isolate the effect of the 'patchiness' of reionization on the Lyα forest power spectrum from differences that otherwise arise from changes in the spatially averaged ionization and thermal history.

An illustration of the gas density, neutral hydrogen fraction, and temperature predicted by the hybrid-RT and paired homogeneous simulations (in this case for the RT-mid and Homog-mid models

---

[1]Note that if He II reionization begins at $z > 4$ (Bolton et al. 2012; Makan et al. 2021), our results may still somewhat underestimate the effect of large-scale temperature fluctuations on the high-redshift Lyα forest by missing the effect of inhomogeneous He II photoheating (see e.g. Meiksin & Tittley 2012; Greig, Bolton & Wyithe 2015).

[2]This approach is slightly different to that used by Oñorbe et al. (2019) and Wu et al. (2019), who both instead compared the results of their inhomogeneous reionization models to hydrodynamical simulations using 'flash' (i.e. very rapid) reionization histories. The flash models were constructed to have the same mid-point of reionization as the inhomogeneous reionization

simulations, rather than matching to the average ionization and thermal history as we do here.
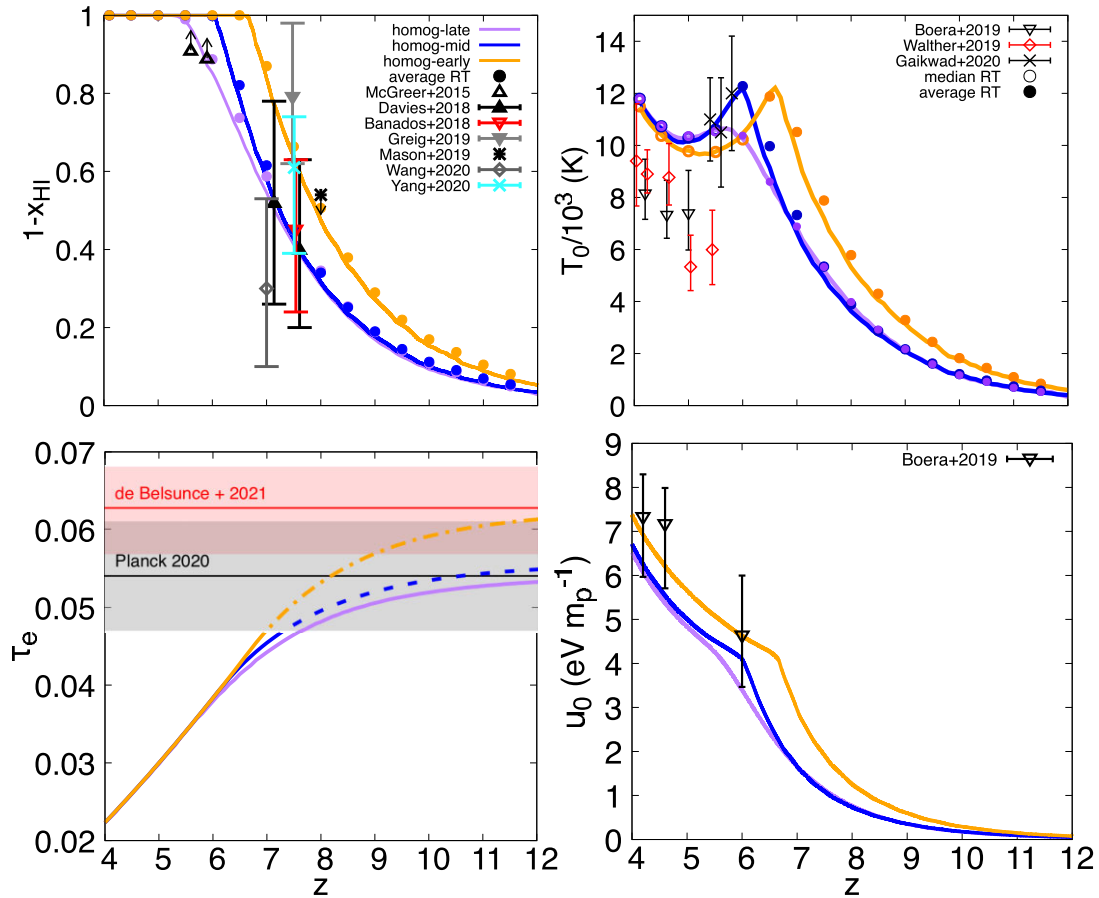
Figure 2. Summary of the hybrid-RT/paired homogeneous reionization models used in this work. The curves in each panel display the redshift evolution in the homogeneous simulations for three different reionization histories ending at $z_R = 5.30$ (late, solid violet curve), $z_R = 5.99$ (mid, dashed blue curve), and $z_R = 6.64$ (early, dot–dashed yellow curve). Top left: the volume-averaged ionized fraction evolution, $1 - x_{H_I}$, shown against recent observational measurements from McGreer, Mesinger & D'Odorico (2015), Bañados et al. (2018), Davies et al. (2018), Greig, Mesinger & Bañados (2019), Mason et al. (2019), Yang et al. (2020), and Wang et al. (2020). The higher redshift data points from Greig et al. (2019) and Davies et al. (2018) were shifted in redshift by -0.02 and $+0.02$, respectively, for clarity. The filled circles show the average ionized fraction in the corresponding RT simulations – note these match by design. Top right: the temperature at mean density, T_0, shown against recent measurements from Boera et al. (2019), Walther et al. (2019) (shifted by $+0.04$ in redshift for clarity), and Gaikwad et al. (2020). The filled (open) circles show the average (median) temperature from the corresponding RT simulations. Bottom left: the Thomson optical depth, τ_e, shown against the Planck Collaboration I (2020) and the de Belsunce et al. (2021) constraints. Bottom right: the cumulative energy per proton mass at the mean density, u_0, shown against the observational estimates from Boera et al. (2019).

at $z = 5.4$) is displayed in Fig. 1. Note in particular the higher temperatures and lower H I fractions in the hybrid-RT model within the prominent void in the lower half of each panel. This arises because the void has been reionized recently and is therefore hotter compared to the paired homogeneous model, as there is less time for subsequent cooling (see e.g. Trac, Cen & Loeb 2008; Cen et al. 2009; Furlanetto & Oh 2009; Raskutti et al. 2012; Lidz & Malloy 2014; D'Aloisio et al. 2015; Keating et al. 2018; Davies et al. 2019). Further discussion of the resulting differences in the $T–\Delta$ relation between the two simulations can be found in Puchwein et al. (in preparation) and Gaikwad et al. (2020).

2.2 Reionization histories used in the hybrid-RT simulations

In Fig. 2, we show the redshift evolution of the volume-averaged ionized hydrogen fraction, $1 - x_{H_I}$, the temperature at mean density, T_0, the Thomson scattering optical depth, τ_e, and the cumulative energy per proton mass deposited at the mean density, u_0, for each for the three reionization histories we use for the hybrid-RT and paired homogeneous simulations. The results from the homogeneous

models are displayed by the curves, while the corresponding $1 - x_{H_I}$ and T_0 values in the hybrid-RT runs are shown at redshift intervals of $\Delta z = 0.5$ by the open and filled circles. By design, these match very closely. Note also that in the case of the gas temperature, we show the median T_0 instead of the mean following reionization, as the median is less sensitive to high-temperature, shock-heated gas.

The reionization histories are deliberately chosen to span a range of z_R, rather than calibrated in detail to reproduce observational data. For example, the large fluctuations in the Lyα forest opacity observed at $z = 5.5$ (Becker et al. 2015) will only be captured by the late-reionization model (cf. Kulkarni et al. 2019; Nasir & D'Aloisio 2020; Bosman et al. 2021; Choudhury, Paranjape & Bosman 2021; Qin et al. 2021). Nevertheless, there is good agreement between our simulations and existing constraints on the IGM ionization history. In the upper left-hand panel we compare to a selection of measurements from McGreer et al. (2015), based on dark gaps in the Lyα and Lyβ forests, from Bañados et al. (2018), Davies et al. (2018), Greig et al. (2019), Yang et al. (2020), and Wang et al. (2020), based on Lyα damping wings in high-redshift quasars, and from Mason et al. (2019), based on the visibility of Lyα emitting galaxies. All

three reionization histories are furthermore consistent with the Planck Collaboration I (2020) Thomson scattering optical depth, displayed in the lower left-hand panel, although only the Homog-early model is within 1σ of the slightly higher τ_e inferred recently by de Belsunce et al. (2021).

In the right-hand panels the simulations are compared to the gas temperature and cumulative energy per proton mass at the mean density measured by Boera et al. (2019) and Walther et al. (2019) from the Lyα forest power spectrum, and by Gaikwad et al. (2020) using Lyα transmission spike widths. There is again reasonable agreement between our simulations and the data, although note there is $\sim 2\sigma$–2.5σ difference between the Boera et al. (2019) and Walther et al. (2019) T_0 measurement from the Lyα forest power spectrum and the gas temperature evolution in the simulations. This may indicate the rise in the IGM temperature at $z < 5$ due to He II photoheating may occur too early in the simulations (but see Makan et al. 2021). Alternatively, there may also be systematic differences between the measurements due to a degeneracy between T_0 and the pressure smoothing scale, where models with increased smoothing can lead to a systematic decrease in the inferred T_0 from the power spectrum (cf. the 3σ discrepancy between the Walther et al. 2019 and Gaikwad et al. 2021 measurements at $z = 5.5$ – see section 5.4.3 in Gaikwad et al. 2021 for further discussion of this point). We will also return to this point later in Section 5.

2.3 Simulating the Lyman α forest power spectrum

We obtain the power spectrum of the Lyα forest transmitted flux using mock absorption spectra extracted from the simulations (e.g. Theuns et al. 1998). We extract 5000 lines of sight, each with 2048 pixels, drawn parallel to the cosmological box boundaries. The line-of-sight positions are the same for all the simulations, and furthermore all the models use initial conditions generated with the same random seed. This allows a direct, pixel-by-pixel comparison of the Lyα transmission across different models. As we are primarily interested in comparing different simulations, we do not add noise and instrumental broadening effects to the spectra in this analysis.

Once we have obtained the Lyα optical depth, τ_α, in each pixel, we rescale the transmitted flux $F = e^{-\tau_{Ly\alpha}}$ in each pixel to match the observed redshift evolution of the Lyα forest optical depth, $\tau_{eff} = -\ln\langle F \rangle$, where $\langle F \rangle$ is the mean observed transmission. Uncertainties in the IGM temperature and background photoionization rate mean a rescaling is commonly used to match the simulation to the data as closely as possible to the observed τ_{eff} (Bolton et al. 2005; Lukić et al. 2015). It furthermore conveniently allows us to vary the effective optical depth in our MCMC analysis without requiring additional simulations. Note, however, this scaling is only a good approximation following reionization, as it implicitly assumes the gas in the low-density IGM is in photoionization equilibrium, such that $\tau_\alpha \propto x_{HI} \propto \Gamma_{HI}^{-1}$. It is important to emphasize that, for this reason, we do not apply this optical depth rescaling to our hybrid-RT simulation outputs prior to the redshift at which reionization ends, z_R, in the simulation volumes. The redshift evolution for τ_{eff} we adopt is

$$\tau_{eff} = \begin{cases} -0.132 + 0.751[(1+z)/4.5]^{2.90} & \text{if } 2.2 \le z < 4.4, \\ 1.142[(1+z)/5.4]^{4.91} & \text{if } 4.4 \le z \le 5.5. \end{cases} \quad (1)$$

This is taken from Viel et al. (2013) and Becker et al. (2013) for the upper and lower redshift ranges, respectively.

Once this rescaling has been performed, we calculate the power spectrum of the transmitted flux, $P(k)$, using the estimator $\delta_F = F/\langle F \rangle - 1$. Since in this work we are primarily interested in analysing differences arising between the hybrid-RT and paired homogeneous

simulations, we will focus mainly on the ratio, $R(k, z)$, of the power spectra, that is

$$R(k, z) = \frac{P_{RT}(k, z)}{P_{homog}(k, z)}, \quad (2)$$

where $P_{RT}(k, z)$ and $P_{homog}(k, z)$ are the power spectrum from the hybrid-RT and paired homogeneous simulations, respectively. We will concentrate in particular on this ratio in three redshift bins, $z = 4.2$, 4.6, and 5.0, which correspond to those observed by Boera et al. (2019). We do not consider a redshift bin at $z = 5.4$ within our framework, despite the availability of observational constraints here (Viel et al. 2013). This is because – as already discussed above – our assumption of ionization equilibrium when rescaling τ_{eff} will break down for late-reionization models, invalidating our approach.

3 THE EFFECT OF INHOMOGENEOUS REIONIZATION ON THE LYα FOREST POWER SPECTRUM

We now turn to examine the effect of inhomogeneous reionization on the shape of the Lyα forest transmitted flux power spectrum. We first consider the intermediate reionization history with an end redshift for reionization at $z_R = 6.0$. We refer to the hybrid-RT and paired homogeneous simulations for this model as 'RT-mid' and 'Homog-mid', respectively (see Table 1). We will examine the effect of changing the end redshift of reionization on the power spectrum ratio, $R(k, z)$, in Section 4.

The black curves in Fig. 3 show the ratio $R(k, z)$ of the power spectrum compared to the 1σ uncertainty in the Boera et al. (2019) power spectrum measurements in three redshift bins ($z = 4.2$, 4.6, and 5.0). Note the Boera et al. (2019) data cover $-2.2 \le \log(k/\text{km}^{-1}\,\text{s}) \le -0.7$, whereas in the simulations we also consider larger scales up to $\log(k/\text{km}^{-1}\,\text{s}) = -2.9$. At small k values, the uncertainties in the power spectrum measurements are typically dominated by cosmic variance and the low number of large-scale modes. At high k values, on the other hand, the uncertainties are mainly due to noise, metals, and instrumental broadening corrections. In all three redshift bins the hybrid-RT prescription boosts the power spectrum at large scales (small k values) by ~ 10–40 per cent and suppresses it at small scales (large k values) by 10–15 per cent. The difference between the hybrid-RT and paired homogeneous simulations is comparable to or smaller than the 1σ uncertainties from the Boera et al. (2019) power spectrum measurements at scales $\log(k/\text{km}^{-1}\,\text{s}) > -2.2$. This already suggests that analyses of current Lyα forest power spectrum measurements from high-resolution ($R \sim 40\,000$) data at $z \le 5$ (e.g. Iršič et al. 2017b; Garzilli et al. 2021; Rogers & Peiris 2021) should not be strongly biased if ignoring the effect of inhomogeneous reionization (cf. Hui et al. 2017). This will change, however, for more precise measurements and/or data that extend to larger scales. We explore this further in Section 5.

An explanation for the change in the shape of the power spectrum in the hybrid-RT model may be obtained by successively isolating the physical quantities that influence $R(k, z)$ (i.e. the H I fraction, gas temperature, gas density, and peculiar velocity). The other curves in Fig. 3 display the results of this process, where we substitute quantities in the hybrid-RT simulation with those from the paired homogeneous run, and then reextract the Lyα forest spectra we use to calculate the power spectrum. In particular, in the hybrid-RT model we will replace the peculiar velocity (labelled 'homog v_{pec}' in Fig. 3) and the gas temperature ('homog T') with the values from the paired homogeneous run. Additionally, after the gas temperature in the hybrid-RT model has been substituted with the one from the homogeneous simulation,
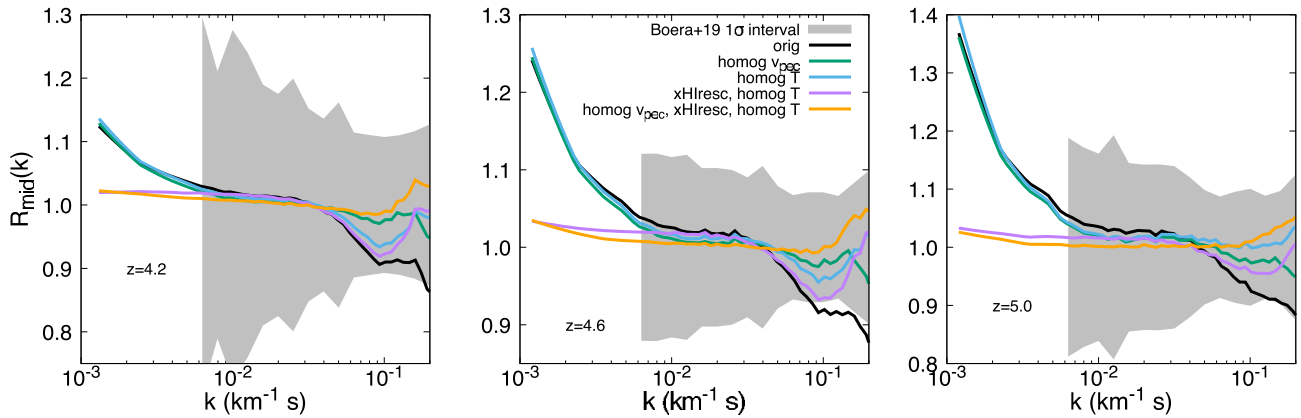
Figure 3. The ratio of Lyα forest transmitted flux power spectrum obtained from the RT-mid and Homog-mid simulations in three redshift bins (from left to right, $z = 4.2$, 4.6, and 5.0). At each redshift, all mock spectra have been rescaled to the same effective optical depth, given by equation (1). The ratio taken directly from the simulations is shown by the black curves, while the other curves are obtained by progressively replacing quantities in the RT-mid simulation with those from Homog-mid, and then recomputing the simulated Lyα forest spectra (see text for details). These are the peculiar velocity field ('homog $v_{\rm vpec}$'), the gas temperature ('homog T'), and ionized hydrogen fraction ('xHIresc', obtained by rescaling the ionization fraction to account for the temperature dependence of the recombination coefficient). The shaded grey region shows the 1σ uncertainties from the power spectrum measurements presented by Boera et al. (2019) using high-resolution Keck/High Resolution Echelle Spectrometer (HIRES) and Very Large Telescope (VLT)/Ultraviolet and Visual Echelle Spectrograph (UVES) data, centred around $R(k) = 1$. Note the different scale on the vertical axis of each panel.

we also rescale the neutral hydrogen fraction, $x_{\rm HI}$ ('xHIresc'), to account for differences in the temperature-dependent recombination rate ($\alpha_{\rm HII}(T) \propto T^{-0.7}$; e.g. McQuinn 2016). For further assistance with visualizing the changes this process causes in the mock Lyα forest spectra, in the supplementary Appendix A we show the transmitted flux along a single line of sight for each case considered in Fig. 3.

First, the orange curve in Fig. 3 shows the case where all quantities ($v_{\rm pec}$, T, and $x_{\rm HI}$) in the hybrid-RT model have been substituted with those from the homogeneous simulation, in the manner described above. The only remaining difference between the models is the underlying gas distribution, where we find the ratio, $R(k)$, varies by <5 per cent at all scales. Pressure (or Jeans) smoothing of the gas distribution – which, as discussed in Section 2, is implemented self-consistently in our hybrid-RT approach – will counter gravitational collapse and help push the baryons outward from peaks in the dark matter density (e.g. Gnedin & Hui 1998; Kulkarni et al. 2015; Rorai et al. 2017). This means that if the IGM has been reionized and heated more recently in the hybrid-RT model, gas will have had less time to dynamically respond to the change in pressure (D'Aloisio et al. 2019), which will increase power on small scales. Indeed, we find a small excess (<5 per cent) in the power spectrum ratio shown by the orange curve at $k > 0.1$ km^{-1} s. This implies that while changes in the pressure smoothing due to differences in the *average* thermal history of the IGM will alter the shape of the power spectrum following reionization (e.g. Nasir et al. 2016; Oñorbe et al. 2017b; Wu et al. 2019), variations around this average due to the 'patchiness' of pressure smoothing will have a much smaller effect.

The purple curve in Fig. 3 also replaces T and $x_{\rm HI}$ in the hybrid-RT model as previously described, however it now relies on the original hybrid-RT $v_{\rm pec}$. Comparison to the orange curve in Fig. 3 therefore demonstrates the effect that differences between the peculiar velocities in the hybrid-RT and homogeneous simulations have on the power spectrum. This suggests that the suppression of power by 5–10 per cent on scales of $k \simeq 0.1$ km^{-1} partly arises from changes in the IGM peculiar velocity field. This can also be seen by comparing the black and green curves, where in the latter case only $v_{\rm pec}$ has been

replaced in the hybrid-RT model using the homogeneous simulation values. As we demonstrate shortly, this suppression of power occurs because of the divergent peculiar velocity field associated with gas that has been recently heated, resulting in Lyα transmission that is more smoothed out in velocity space (see also e.g. Gaikwad et al. 2020).

Thermal broadening will also lead to a suppression of the power spectrum at small scales by smoothing the Lyα absorption in velocity space (Zaldarriaga, Hui & Tegmark 2001; Nasir et al. 2016; Walther et al. 2019). Indeed, this is what is shown by the blue curve in Fig. 3, where the gas temperature in the hybrid-RT model has been replaced with the values from the homogeneous simulation. The power at small scales is now increased relative to the black curve by up to 10 per cent at the smallest scales. The higher gas temperatures in the recently reionized regions in the hybrid-RT model therefore act (in combination with the peculiar velocity field) to suppress the power spectrum at small scales.

We may gain further insight into this behaviour in Fig. 3 by examining differences in the IGM properties in the hybrid-RT and paired homogeneous simulations. First, in Fig. 4, we show the fraction of pixels with gas density $\Delta < 3$ in the RT-mid simulation (i.e. the gas typically responsible for transmission in the Lyα forest at high redshift; see fig. 2 in Nasir et al. 2016) that have crossed different photoionization rate thresholds (in this case, $\Gamma_{\rm HI} > 10^{-15}$, 10^{-14}, 5×10^{-14}, and 10^{-13} s^{-1}) by a given redshift. Whereas this occurs by construction at a single redshift in the Homog-mid simulation (corresponding to $z = 7.03$ for $\Gamma_{\rm HI} > 10^{-14}$ s^{-1}), inhomogeneous reionization means this occurs at different redshifts in the RT-mid model. Note also that while the distributions for $\Gamma_{\rm HI} > 10^{-14}$ s^{-1} and $\Gamma_{\rm HI} > 10^{-15}$ s^{-1} are similar in Fig. 3, for the larger thresholds the distribution becomes more strongly peaked at $z \sim 6.2$, reflecting the fact that the amplitude of the ionizing background increases toward lower redshift.

Taking $\Gamma_{\rm HI} > 10^{-14}$ s^{-1} as a proxy when a pixel is first ionized, in Fig. 5 we show the probability distribution function (PDF) for the gas density, temperature, and the derivative of the line-of-sight peculiar velocity $\partial v_{\rm pec}/\partial v_{\rm H}$ (where $v_{\rm H}$ is the Hubble velocity) at redshift $z = 5$ in the RT-mid and Homog-mid simulations. For the
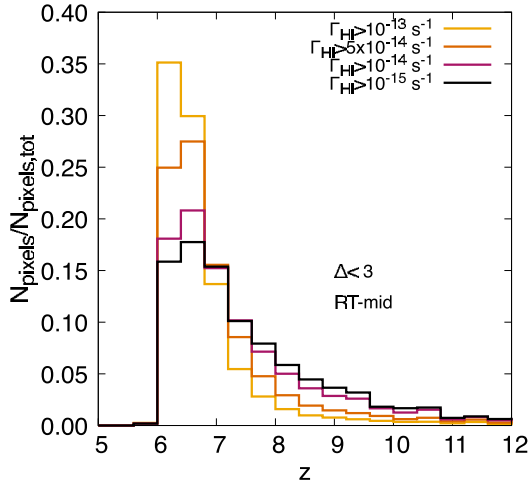
Figure 4. The distribution of redshifts at which pixels with gas densities $\Delta < 3$ along the lines of sight drawn from the RT-mid simulation first pass a fixed threshold in the hydrogen photoionization rate, $\Gamma_{\mathrm{H\,I}}$. These are $\Gamma_{\mathrm{H\,I}} > 10^{-15}\,\mathrm{s}^{-1}$ (black), $\Gamma_{\mathrm{H\,I}} > 10^{-14}\,\mathrm{s}^{-1}$ (red), $\Gamma_{\mathrm{H\,I}} > 5 \times 10^{-14}\,\mathrm{s}^{-1}$ (orange), and $\Gamma_{\mathrm{H\,I}} > 10^{-13}\,\mathrm{s}^{-1}$ (yellow). The distributions are normalized to the total number of pixels, $N_{\mathrm{pixels,\,tot}}$, in the simulated lines of sight. Note that in the paired homogeneous simulation, Homog-mid, by construction all pixels exceed a given $\Gamma_{\mathrm{H\,I}}$ threshold at the same redshift of $z = 7.03$.

latter quantity, positive (negative) values correspond to a divergent (convergent) peculiar velocity field. We show the PDFs for all pixels in the mock spectra (upper panels) and, in the case of RT-mid only, broken down by the redshift at which the pixels first cross the $\Gamma_{\mathrm{H\,I}} > 10^{-14}\,\mathrm{s}^{-1}$ threshold (lower panels). The gas density PDF for all pixels appears similar in both models, consistent with the rather small differences we observe for $R(k)$ when comparing the black and orange curves in Fig. 3. However, the temperature and peculiar velocity gradient PDFs show larger differences. In particular there is a broader distribution of gas temperatures in the hybrid-RT model, with a peak in the distribution that is $\sim$4000 K hotter compared to the homogeneous simulation. The fraction of pixels with $\partial v_{\mathrm{pec}}/\partial v_{\mathrm{H}} > 1$ is also larger in the hybrid-RT run, due to newly expanding gas that has recently been reionized and heated.

The RT-mid PDF for all pixels (red curve in the upper panels) is an average of the distributions shown in the lower panels, weighted by the relative amplitude of each redshift bin in Fig. 4. First, note that for gas that is reionized earlier, the gas-density PDF in the lower left-hand panel of Fig. 5 shifts toward larger values of Δ, with a high-density tail that becomes shallower. This is consistent with the highest density regions in the simulation reionizing first. Second, we observe for the T and $\partial v_{\mathrm{pec}}/\partial v_{\mathrm{H}}$ PDFs that higher gas temperatures and divergent peculiar velocities arise from gas that has been recently ionized. As discussed for Fig. 3, we argue here that the suppression of power in the RT-mid simulation on small scales, $k > 0.05\,\mathrm{km}^{-1}\,\mathrm{s}^{-1}$, relative to the Homog-mid model is primarily due to these differences acting together in redshift space, rather than a physical smoothing of the gas density.

Finally, we note the results displayed in Fig. 3 have also been discussed by a number of other studies using different numerical methods (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). All these studies find that large-scale fluctuations in the post-reionization IGM temperature will increase the power spectrum on large scales; our results are in good agreement with this canonical

expectation. Of particular note, however, is the recent study by Wu et al. (2019). These authors used fully coupled radiation–hydrodynamical simulations to model the effect of reionization on the power spectrum at $z \sim 5$, finding that power was enhanced up to $\sim$30–40 per cent at $k \sim 0.002\,\mathrm{km}^{-1}\,\mathrm{s}^{-1}$ and suppressed by up to $\sim$10 per cent on scales $k > 0.1\,\mathrm{km}^{-1}\,\mathrm{s}$ (see their fig. 5). Our results appear to be remarkably consistent, suggesting that the hybrid approach we adopt is well suited for efficiently modelling the power spectrum.

4 A CORRECTION FOR THE EFFECT OF LARGE-SCALE TEMPERATURE FLUCTUATIONS ON THE LYα FOREST POWER SPECTRUM

We now assess how different patchy reionization models affect the Lyα forest power spectrum. Alongside the RT-mid and Homog-mid models, we now use the simulations with the two additional reionization histories displayed in Fig. 2: a 'late' model with reionization[3] ending at $z_{\mathrm{R}} \simeq 5.3$, and an 'early' model with $z_{\mathrm{R}} \simeq 6.6$. In Fig. 6, the power spectrum ratio, $R(k)$, is displayed for all three models in redshift bins at $z = 4.2$, 4.6, and 5.0. The grey regions once again show the 1σ uncertainties on the power spectrum measurements from Boera et al. (2019). At large scales, $k \leq 0.03\,\mathrm{km}^{-1}\,\mathrm{s}$, the extra power due to large-scale temperature fluctuations is enhanced for a later end redshift to reionization in all three redshift bins. This is because the temperature fluctuations will begin to fade once the reionization process has completed and the low-density IGM adiabatically cools towards the thermal asymptote (e.g. Theuns et al. 2002; Hui & Haiman 2003; McQuinn & Upton Sanderbeck 2016). The suppression of power at the smallest scales, on the other hand, shows no strong trend across the three redshift bins, although the early-reionization model has the smallest suppression in the small-scale power. This is consistent with the early-reionization model being slightly colder than the other two models (see Fig. 2). As discussed in Section 3, differences between the hybrid-RT and homogeneous simulations at small scales are primarily due to thermal broadening (i.e. differences in the instantaneous gas temperature) and peculiar velocities associated with recently heated, expanding gas. At large scales, the enhanced power is instead due to variations in the ionized hydrogen fraction associated with large-scale temperature fluctuations.

Next, we consider whether the large-scale enhancement of power at $k \leq 0.03\,\mathrm{km}^{-1}\,\mathrm{s}$ in the different reionization histories can be emulated, simply by varying the redshift at which the power spectrum is measured from a *single* reionization model. Since the large-scale temperature fluctuations fade predictably as the IGM cools following reionization, we make the *ansatz* that different reionization histories should be equivalent in terms of their large-scale power enhancement at similar time intervals following the completion of reionization. We test this by comparing the power spectrum ratio, $R_{\mathrm{early}}(k)$, from the paired early-reionization models to the ratio from the mid-reionization models, $R_{\mathrm{mid}}(k)$. First, we compute $R_{\mathrm{early}}(k)$ on a redshift grid, using simulation outputs at intervals of $\Delta z = 0.1$. We linearly interpolate in redshift to obtain $R_{\mathrm{early}}(k)$ between these intervals. Next, we find the redshift, z', at which $R_{\mathrm{early}}(k)$ best matches $R_{\mathrm{mid}}(k)$ on scales $k \leq 0.03\,\mathrm{km}^{-1}\,\mathrm{s}$ at redshift z, where we use $z = 4.2$, 4.6,

---

[3]Recall that we formally define the end redshift of reionization, z_{R}, in the models as the redshift when the volume-averaged H I fraction first falls below $x_{\mathrm{H\,I}} = 10^{-3}$.
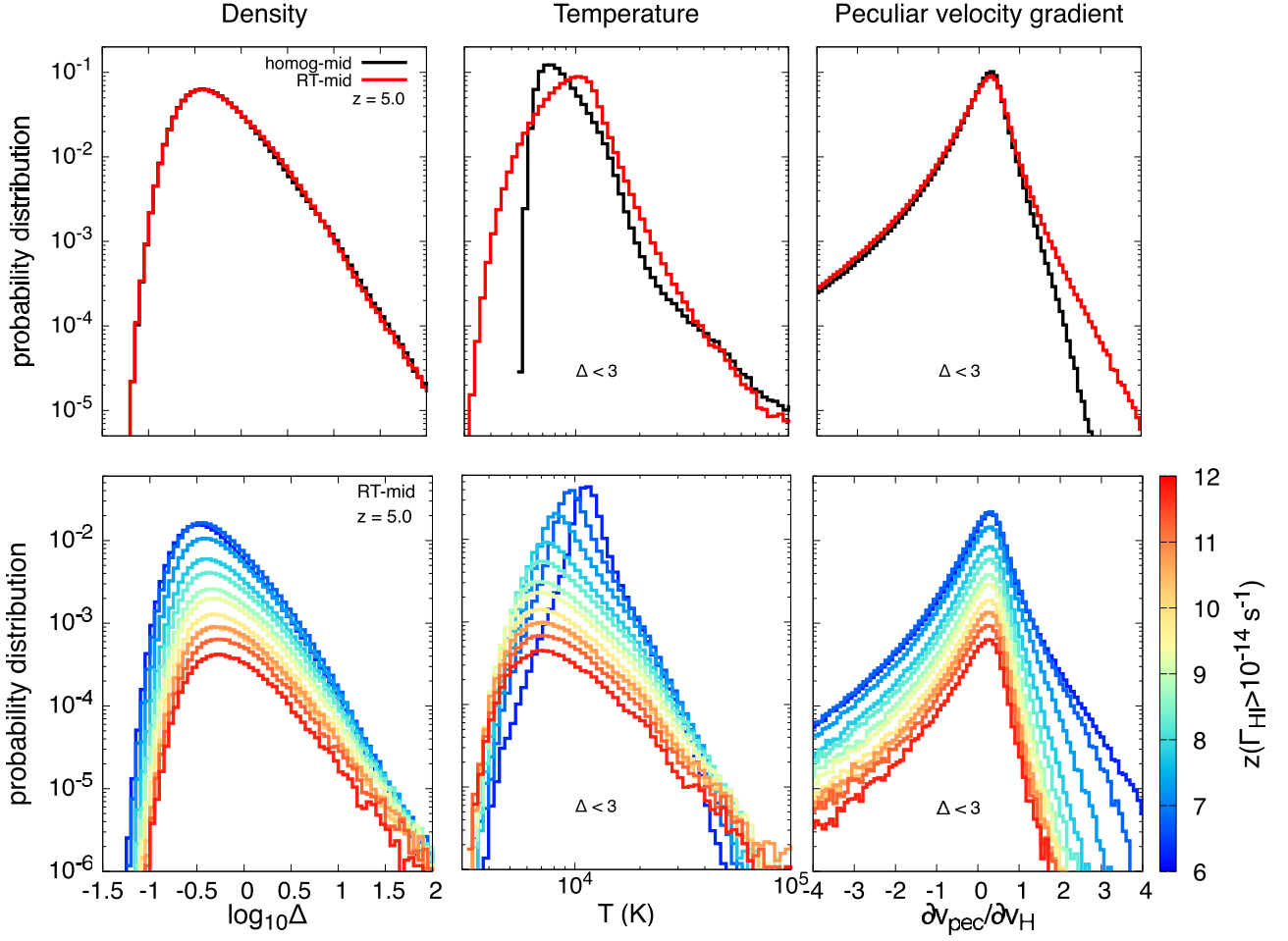
Figure 5. Top row: comparison between the gas density (left-hand column), temperature (middle column), and peculiar velocity gradient $\partial v_{pec}/\partial v_H$ (right-hand column) PDFs at redshift $z = 5$ in the Homog-mid (black curves) and RT-mid (red curves) simulations. The temperature and $\partial v_{pec}/\partial v_H$ distributions are for gas with $\Delta < 3$, to exclude the higher density gas that is not visible in transmission at high redshift. Bottom row: PDFs from the RT-mid simulation at redshift $z = 5$, but now shown only for pixels in which the photoionization rate, Γ_{HI}, first exceeds 10^{-14} s^{-1} at a selection of different redshifts. The colour axis shows the central value of the redshift bins we use here, which have width $\Delta z = 0.5$. Note that the normalization of the PDF in each bin reflects the fraction of the volume that exceeds the Γ_{HI} threshold at that redshift (cf. Fig. 4), highlighting how the different redshift bins contribute to the shape of the overall distribution shown in the top row.
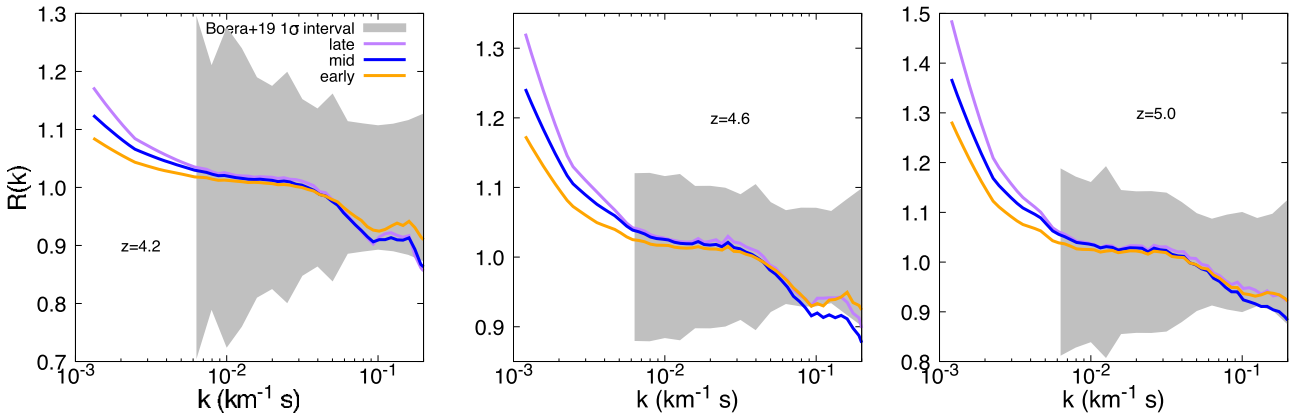
Figure 6. The ratio of the Lyα forest power spectrum in the hybrid-RT and paired homogeneous simulations, $R(k)$, at redshifts $z = 4.2$, 4.6, and 5.0, for the three reionization histories shown in Fig. 2. At each redshift, all mock spectra have been rescaled to the same effective optical depth, given by equation (1). The shaded grey region shows the 1σ uncertainties on the power spectrum measurements presented by Boera et al. (2019). Note the different scale on the vertical axis of each panel.
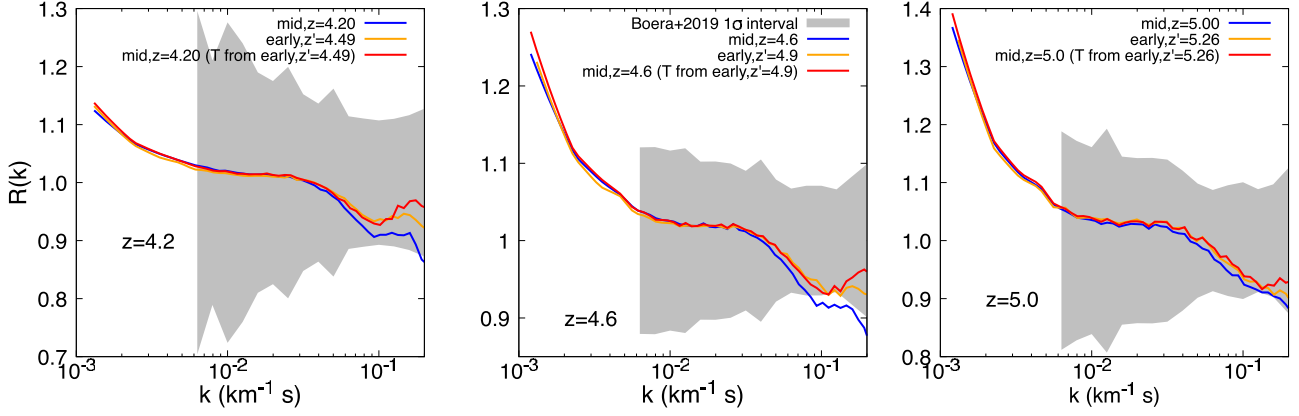
Figure 7. The ratio of the Lyα forest power spectrum from the RT-mid and Homog-mid simulations, $R_{\rm mid}(k, z)$ (blue curves), in three redshift bins ($z = 4.2$, 4.6, and 5.0), compared to $R_{\rm early}(k, z')$ (orange lines) at *different* redshifts z'. The redshift z' has been found by minimizing the difference between $R_{\rm mid}$ and $R_{\rm early}$ at $k \leq 0.03\,{\rm km^{-1}\,s}$. This results in an excellent agreement between the large-scale power predicted by the two reionization models, with less than 1 per cent difference between the two ratios at $k \leq 0.03\,{\rm km^{-1}\,s}$. This good agreement can be extended up to $k \sim 0.1\,{\rm km^{-1}\,s}$ by substituting the gas temperatures in the RT-mid and Homog-mid simulations with those in the RT-early and Homog-early simulations, respectively (red lines), while leaving the original ionization fraction unchanged. This corrects for the remaining differences in the power spectra due to thermal broadening at small scales. The shaded grey region shows the 1σ uncertainties on the power spectrum measurements presented by Boera et al. (2019). Note the different scale on the vertical axis of each panel.
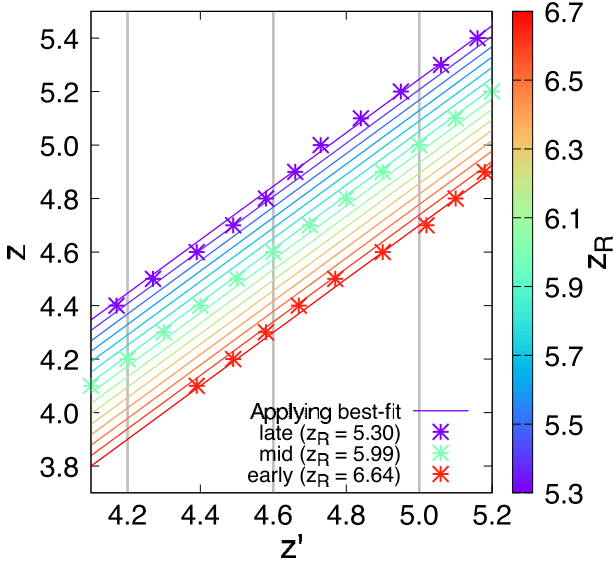
Figure 8. The redshifts, z', at which $R_{\rm early}(k, z')$ (orange asterisks) and $R_{\rm late}(k, z')$ (purple asterisks) most closely match $R_{\rm mid}(k, z)$ at redshift z at $k \leq 0.03\,{\rm km^{-1}\,s}$ (see text for details, as well as Fig. 7). The one-to-one identity relation for $R_{\rm mid}(k, z)$ is also displayed (cyan asterisks). The solid lines correspond to equation (3), and show the best-fitting linear relation for z as a function of both z' and the end redshift of reionization, $z_{\rm R}$, where the colour of each line corresponds to the $z_{\rm R}$ shown in the colour bar (in $\Delta z_{\rm R} = 0.1$ steps, starting at $z_{\rm R} = 5.3$). The vertical grey lines show the redshift bins that correspond to the power spectrum measurements presented by Boera et al. (2019).

and 5.0 as before. In Fig. 7, we show the resulting 'match' of $R_{\rm early}(k, z')$ (orange curves) to $R_{\rm mid}(k, z)$ (blue curves) in the three redshift bins, where we find $z' = 4.49$, 4.90, and 5.26. The ratios now show excellent agreement at $k \leq 0.03\,{\rm km^{-1}\,s}$, with less than 1 per cent difference between the two ratios.

At smaller scales, however, larger differences at the ∼5 per cent level between the two ratios remain. We expect these differences to be mainly due to spatial variations in the thermal broadening

kernel,[4] as the models will not have exactly the same instantaneous gas temperatures at z' and z. To verify this, we once again substitute the gas temperatures with those from other models, but this time substituting the Homog-early and RT-early temperatures at redshift z' into the Homog-mid and RT-mid simulations at redshift z. The result is shown by the red curves in Fig. 7, where we find that the good agreement now extends up to $k \sim 0.1\,{\rm km^{-1}\,s}$. We have also verified that a similar result holds on repeating this procedure for the late-reionization model ratio, $R_{\rm late}(k)$. This implies that, in general, we may find $R(k, z') \simeq R_{\rm mid}(k, z)$ for any $z_{\rm R}$ within the range our simulations presently cover. In this way we may emulate a 'patchy' correction to the power spectrum simply by finding a best-fitting relation for z' and z as a function of $z_{\rm R}$. This property is particularly useful for implementing a correction for inhomogeneous reionization within existing grids of hydrodynamical simulations that use a spatially uniform ionizing background.

We implement this correction within a Bayesian parameter inference framework in Section 5 as follows. In Fig. 8, we show all the z'–z pairs we have obtained from matching $R_{\rm early}(k)$ (green asterisks, $z_{\rm R} = 5.3$) and $R_{\rm late}(k)$ (purple asterisks, $z_{\rm R} = 6.6$) to $R_{\rm mid}(k)$, as well the one-to-one identity relation for $R_{\rm mid}(k)$ (cyan asterisks, $z_{\rm R} = 6.0$). Note that we limit ourselves to $z' \leq 5.2$, in order to only consider redshifts where reionization has completed in all models. Furthermore, all our simulations currently finish at $z = 4.1$, explaining why there are no data points at $z' < 4.4$ for the late-reionization model. We then find the best-fitting linear relation in the z'–z plane for *each reionization model*. Furthermore, we assume there is an approximately linear increase in the best-fitting z'–z relation as a function of $z_{\rm R}$, over the range $5.3 \leq z_{\rm R} \leq 6.7$. The resulting linear relations are shown as the solid lines in Fig. 8, where the best fit for the redshift, z, at which $R_{\rm mid}(k)$ matches the power spectrum ratio for a model at z' with an

---

[4]In principle, any differences between $R(k, z')$ and $R(k, z)$ on small scales due to the *hydrodynamical* response of the gas to heating (e.g. divergent peculiar velocities) should also be minimized at z'. The IGM in different reionization models experience similar amounts of pressure smoothing when considered at approximately the same time interval after the initial heating (see e.g. D'Aloisio et al. 2019, for further discussion of this point).

Table 2. Tabulated values of $R_{\mathrm{mid}}(z, k)$ (see e.g. the black curves in Fig. 3). In combination with equation (3), this table can be used to find an approximate correction to the Lyα forest power spectrum for the effects of inhomogeneous reionization at redshift z_R. The corrected power spectrum is then $P_{\mathrm{RT}}(k, z') = R_{\mathrm{zr}}(k, z')P_{\mathrm{homog}}(k, z')$, where $R_{\mathrm{zr}}(k, z') = R_{\mathrm{mid}}(k, z)$. We have verified that linear interpolation between the tabulated redshift bins is sufficient for recovering the ratio for different reionization histories to 1 per cent accuracy for wavenumbers $k \leq 0.03$ km^{-1} s (or equivalently, $\log(k/\mathrm{km}^{-1}$ s$) \leq -1.5$) over the z_R range we consider. However, at smaller scales, $k \geq 0.03$ km^{-1} s, we expect the correction will only be accurate to within ~5 per cent due to differences associated with thermal broadening, although this is still well within the current measurement uncertainties. Note that the k bins are spaced following Boera et al. (2019).

| $\log(k/\mathrm{km}^{-1}$ s$)$ | | | | | | | z | | | | | | | |
	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0	5.1	5.2	5.3	5.4
−2.9	1.109	1.126	1.147	1.167	1.191	1.233	1.264	1.294	1.324	1.354	1.394	1.431	1.466	1.507
−2.8	1.092	1.105	1.122	1.138	1.157	1.182	1.205	1.228	1.250	1.272	1.300	1.328	1.353	1.383
−2.7	1.073	1.084	1.096	1.108	1.121	1.139	1.155	1.172	1.187	1.203	1.222	1.242	1.260	1.281
−2.6	1.057	1.065	1.074	1.082	1.092	1.104	1.116	1.128	1.140	1.152	1.167	1.183	1.198	1.216
−2.5	1.047	1.055	1.061	1.068	1.075	1.085	1.095	1.103	1.112	1.122	1.135	1.149	1.163	1.180
−2.4	1.039	1.045	1.050	1.055	1.061	1.069	1.077	1.084	1.092	1.101	1.113	1.126	1.139	1.153
−2.3	1.032	1.037	1.041	1.044	1.049	1.054	1.060	1.066	1.073	1.079	1.088	1.097	1.105	1.117
−2.2	1.027	1.029	1.031	1.033	1.035	1.038	1.041	1.045	1.050	1.054	1.062	1.071	1.081	1.093
−2.1	1.021	1.023	1.025	1.027	1.028	1.029	1.032	1.035	1.038	1.040	1.045	1.051	1.060	1.071
−2.0	1.019	1.020	1.021	1.023	1.025	1.025	1.028	1.030	1.033	1.036	1.040	1.047	1.055	1.064
−1.9	1.015	1.016	1.017	1.018	1.019	1.020	1.023	1.024	1.027	1.031	1.036	1.043	1.050	1.057
−1.8	1.012	1.014	1.015	1.017	1.018	1.021	1.023	1.024	1.024	1.024	1.027	1.033	1.041	1.052
−1.7	1.012	1.012	1.014	1.016	1.017	1.017	1.018	1.020	1.024	1.028	1.032	1.038	1.048	1.062
−1.6	1.009	1.011	1.013	1.014	1.017	1.019	1.020	1.021	1.022	1.024	1.026	1.031	1.038	1.047
−1.5	1.003	1.004	1.005	1.007	1.010	1.012	1.014	1.015	1.018	1.023	1.028	1.035	1.042	1.050
−1.4	0.989	0.990	0.993	0.996	0.998	1.001	1.003	1.003	1.007	1.013	1.018	1.024	1.031	1.036
−1.3	0.975	0.975	0.976	0.978	0.979	0.980	0.983	0.987	0.991	0.993	0.995	0.999	1.003	1.011
−1.2	0.946	0.949	0.954	0.955	0.956	0.960	0.964	0.968	0.968	0.972	0.976	0.978	0.983	0.992
−1.1	0.924	0.923	0.924	0.923	0.930	0.935	0.938	0.940	0.941	0.946	0.950	0.954	0.960	0.966
−1.0	0.909	0.909	0.907	0.911	0.924	0.919	0.921	0.921	0.929	0.928	0.927	0.924	0.931	0.935
−0.9	0.908	0.913	0.915	0.916	0.922	0.917	0.919	0.915	0.915	0.914	0.906	0.909	0.914	0.914
−0.8	0.900	0.913	0.912	0.909	0.913	0.913	0.904	0.903	0.909	0.905	0.900	0.886	0.883	0.885
−0.7	0.869	0.862	0.876	0.870	0.877	0.879	0.882	0.886	0.883	0.883	0.882	0.877	0.869	0.864

end redshift of reionization z_R is

$$z(z', z_R) = z' + \xi z_R + \zeta, \tag{3}$$

where $\xi = -0.390 \pm 0.007$ and $\zeta = 2.31 \pm 0.04$ for 1σ bootstrapped uncertainties on the best-fitting parameters. The inhomogeneous reionization correction to the power spectrum is then $P_{\mathrm{RT}}(k, z') = R_{\mathrm{zr}}(k, z')P_{\mathrm{homog}}(k, z')$, where $R_{\mathrm{zr}}(k, z') = R_{\mathrm{mid}}(k, z)$, $R_{\mathrm{mid}}(k, z)$ is obtained by linearly interpolating the data in Table 2, and $P_{\mathrm{RT}}(k, z')$ and $P_{\mathrm{homog}}(k, z')$ are, respectively, the corrected power spectrum and the original power spectrum from a simulation with a spatially uniform ionizing background.

In Fig. 9, we perform a consistency check on this procedure. We use equation (3) to obtain the redshift $z(z', z_R)$, and hence $R_{\mathrm{mid}}(k, z)$ for all three reionization histories at $z' = 4.6$. For wavenumbers $k \leq 0.03$ km^{-1} s, the agreement is excellent, with a less than 1 per cent difference between the emulated (thick curves) and true (thin curves) ratios. The differences at the smallest scales are up to 5 per cent, but as already discussed, this is explained primarily by spatial fluctuations in the thermal broadening in the models. These differences are furthermore well within the uncertainties on existing measurements of the Lyα forest power spectrum (see Fig. 7).

In summary, Table 2 combined with equation (3) provides an approximate correction for inhomogeneous reionization that can be applied to the Lyα forest power spectrum predicted by hydrodynamical simulations with a spatially uniform ionizing background. A simple PYTHON script to calculate this correction is accessible at https://github.com/marghemolaro/RT_1dps_correction.git. It is important to emphasize, however, that this correction is model dependent. While we explore reionization models that are broadly consistent with current observational constraints, our approach will

not be applicable to reionization histories that differ significantly from those displayed in Fig. 2, for reionization models that finish well outside the range we have modelled, $5.3 \leq z_R \leq 6.7$. Furthermore, for simulation volumes that are much larger than we consider here, the accuracy of the power spectrum modelling may also benefit from an additional box size correction. Furthermore, note that our correction factor was obtained by comparing hybrid-RT and homogeneous UVB simulations that use a non-equilibrium thermochemistry solver. The application of our correction to grids of homogeneous UVB simulations that use an equilibrium solver (as is often the case in the literature) will therefore implicitly assume an initial spread in the temperature–density relation that differs from that predicted by equilibrium models, although such differences are expected to have a small effect on the 1D power spectrum (see e.g. Puchwein et al. 2015; Gaikwad et al. 2019).

5 RECOVERY OF THERMAL PARAMETERS FROM MOCK OBSERVATIONS

A commonly used method for physical parameter recovery from Lyα forest power spectrum observations is Bayesian inference, where the data are compared with a grid of hydrodynamical simulations that span the relevant parameter space (Viel et al. 2013; Iršič et al. 2017a; Yèche et al. 2017; Boera et al. 2019; Palanque-Delabrouille et al. 2020; Rogers & Peiris 2021). So far, aside from Oñorbe et al. (2019) and Wu et al. (2019), the parameter grids used in these studies have typically relied on simulations that assume homogeneous UVBs. In Section 3, however, we have shown that replacing a homogeneous UVB with a more realistic model for 'patchy' reionization can lead to significant differences in the resulting power spectrum.
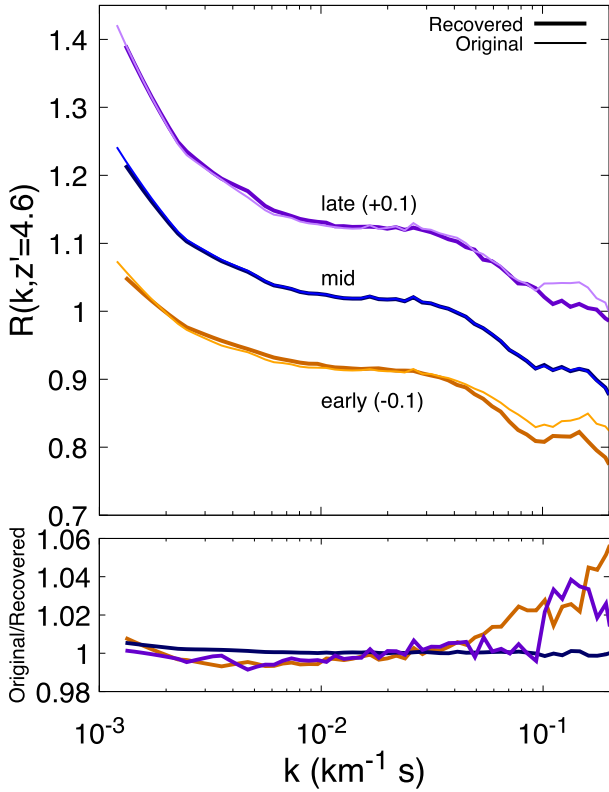
Figure 9. Upper panel: consistency test of the procedure we use for emulating a patchy reionization correction to the Lyα forest power spectrum. The test shows the recovery of $R_{zr}(k, z')$ at $z' = 4.6$ for the late (orange curves), mid (blue curves), and early (purple curves) reionization histories using only the power spectrum ratio from the mid model (see Table 2) and the best-fitting relation in equation (3) to determine $z(z_R, z')$. Thick curves show the emulated ratio, while the thin curves show the original ratio from the simulations. For clarity, the early (late) models are shifted vertically by -0.1 ($+0.1$). Lower panel: The ratio of the original to recovered ratios shown in the upper panel.

Given that any observations may include such patchy effects, it is therefore important that we ask whether a grid of simulations using a homogeneous UVB is still able to accurately recover the underlying physical parameters from observations, or whether patchy reionization effects on the power spectrum could significantly bias any constraints (see e.g. Hui et al. 2017, for a discussion of this point).

In order to answer this question, we proceed as follows. First, we generate a set of mock Lyα forest data (D_{mock}) based on one of our patchy simulations. We then construct two grids of simulations for interpolating between the physical parameters of interest: one, G_{homog}, relying on homogeneous UVB simulations, and the other, G_{patchy}, additionally including the patchy correction described in equation (3). We then perform a Bayesian inference analysis to assess if either model grid allows us to accurately recover the thermal parameters used to generate the mock data.

To perform our Bayesian inference analysis, we sample the parameter space using a MCMC sampler, based on the Metropolis–Hastings algorithm (e.g. Iršič et al. 2017a). In Bayesian inference, the resulting distribution in the parameter space is directly proportional to the likelihood of the data given the input parameters, i.e. the so-called Bayes' theorem,

$$p(\theta|D) \propto \mathcal{L}(D|\theta)\pi(\theta),$$ (4)

where $\mathcal{L}(D|\theta)$ is the likelihood, $\pi(\theta)$ contains the prior information on the parameters, and $p(\theta|D)$ is the posterior probability of the parameters obtained as the end result in our analysis. In order to sample the posterior distribution one needs to evaluate the likelihood and priors at each point in the parameter space. For the mock measurements of the Lyα forest power spectrum we use a Gaussian likelihood with the covariance matrix evaluated from the Homog-late simulation using a bootstrap method. We have also tested whether using either the homogeneous or hybrid-RT simulations to generate the covariance matrix affects our results, and have verified that the resulting differences are negligible compared to the width of the posteriors. Different redshift bins in the analysis are treated as independent, and as such the Bayesian analysis at each of the three redshift bins we consider ($z = 4.2$, 4.6, and 5.0) is independent.

5.1 Thermal parameters and their priors

We now discuss the thermal parameters and priors we consider in our parameter recovery analysis. Each of the (homogeneous UVB) simulations used for constructing our parameter grid features different heating rates at the onset of hydrogen reionization (see the upper section of Table 1). These different heating rates are partly characterized by the cumulative energy per unit mass deposited into gas at the mean background density (see e.g. Fig. 2). This quantity, u_0, is highly correlated with the suppression of the Lyα forest power spectrum on small scales (Nasir et al. 2016) and it acts as a proxy for the amount of pressure (or Jeans) smoothing induced by the integrated thermal history. Furthermore, Boera et al. (2019) (see their appendix J) have demonstrated that u_0 correlates very well with the physical scale used to characterize the physical extent of the pressure smoothing (e.g. Gnedin & Hui 1998; Kulkarni et al. 2015), with typical values of $u_0 \sim 6$ eV m_p^{-1} corresponding to a smoothing scale $\lambda_p \sim 60$ ckpc at $z = 4.2$. It has also been shown that the small-scale power in the Lyα forest at the three redshifts we consider here, $z = 4.2$, 4.6, and 5.0, tightly correlates with u_0 integrated over a redshift range of $z = [4.2–12.0]$, $[4.6–12.0]$, and $[6.0–13.0]$, respectively (Boera et al. 2019). Here we similarly adopt this definition, where the u_0 range covered by our parameter grid is $u_0 = [4.02, 21.1]$, $[3.65, 21.1]$, and $[2.46, 18.7]$ eV m_p^{-1} at $z = 4.2$, 4.6, and 5.0, respectively (see Table 3).

However, u_0 alone is insufficient for describing the full range of possible IGM thermal histories. For instance, the same u_0 could be obtained if reionization started early but the *instantaneous* gas temperature of the IGM remained low (e.g. if reionization were driven by ionizing sources with soft spectra), or where reionization occurred late but was driven by hard ionizing sources that heated the IGM to much higher temperatures. Our choice of thermal parameters must therefore capture both the duration of the reionization process and the instantaneous temperature. To decouple this information we therefore use two further parameters that describe the temperature–density relation of the post-reionization IGM in the homogeneous UVB models[5] – the temperature at mean density, T_0, and the power-law index of the temperature–density relation, γ. The relation between the (instantaneous) temperature and the density of the gas can thus

---

[5]Note that a well-defined power-law temperature–density relation will not always apply in our patchy reionization simulations, due to spatial fluctuations in the IGM temperature (see e.g. Keating et al. 2018). As already discussed earlier, however, the patchy correction given by equation (3) effectively corrects for the effect that any departure from this power-law relation has on the Lyα forest power spectrum.

Table 3. Summary of the astrophysical parameters and priors used in our analysis of the Lyα forest power spectrum.

Parameter	Prior (flat)
$\tau_{\rm eff}(z = 4.2)$	$[0.3\text{–}1.8] \times 1.01$
$\tau_{\rm eff}(z = 4.6)$	$[0.3\text{–}1.8] \times 1.37$
$\tau_{\rm eff}(z = 5.0)$	$[0.3\text{–}1.8] \times 1.92$
$T_0(z = 4.2)$	$[0.5\text{–}1.5] \times 10^4 \,{\rm K}$
$T_0(z = 4.6)$	$[0.5\text{–}1.5] \times 10^4 \,{\rm K}$
$T_0(z = 5.0)$	$[0.5\text{–}1.5] \times 10^4 \,{\rm K}$
$\gamma(z = 4.2)$	$[1.0\text{–}1.7]$
$\gamma(z = 4.6)$	$[1.0\text{–}1.7]$
$\gamma(z = 5.0)$	$[1.0\text{–}1.7]$
$u_0^{z=4.2}(4.2\text{–}12.0)$	$[4.02\text{–}21.1]\,{\rm eV}\,m_{\rm p}^{-1}$
$u_0^{z=4.6}(4.6\text{–}12.0)$	$[3.65\text{–}21.1]\,{\rm eV}\,m_{\rm p}^{-1}$
$u_0^{z=5.0}(6.0\text{–}13.0)$	$[2.46\text{–}18.7]\,{\rm eV}\,m_{\rm p}^{-1}$

be written as $T = T_0\Delta^{\gamma - 1}$, where $\Delta = \rho/\langle\rho\rangle$ (Hui & Gnedin 1997; McQuinn 2016). As for u_0, our assumed priors for T_0 and γ are listed in Table 3; the prior ranges are chosen to encompass a physically plausible range of values. Furthermore, since the goal of this analysis is to investigate potential shifts and biases in the recovery of these thermal parameters due to inhomogeneous reionization, in the subsequent MCMC analysis we adopt agnostic flat priors on all three of these thermal parameters, T_0, γ, and u_0.

Lastly, the transmission in the Lyα forest also depends on the H I photoionization rate, $\Gamma_{\rm H I}$. We therefore adopt the standard approach (e.g. Viel et al. 2004) of varying the effective optical depth in the simulated spectra, $\tau_{\rm eff}$, around the observed value for each of the redshift bins (see also the discussion in Section 2.3).

5.2 Emulator for the Lyα forest power spectrum

To evaluate the likelihood function used in our Bayesian inference analysis, we build an emulator that predicts the power spectrum of the Lyα forest transmitted flux at any given point in the parameter space. Our emulator is based on the linear interpolation of a grid of simulated flux power spectra. This is a method that – despite its simplicity – works remarkably well and has been extensively utilized in earlier Lyα forest studies (Viel et al. 2013; Iršič et al. 2017a; Yèche et al. 2017; Palanque-Delabrouille et al. 2020) and adapted in more sophisticated Gaussian process emulators (Bird et al. 2019; Pedersen et al. 2021).

We start by first constructing a grid of simulated flux power spectra from the suite of 12 simulations in $20\,h^{-1}$ cMpc boxes listed in the upper section of Table 1 (for further details see Section 2). In order to construct a sufficiently well-sampled grid of models spanning the entire parameter range, we then post-process the 12 initial simulations to achieve different parameter combinations for u_0, T_0, γ, and $\tau_{\rm eff}$. We refer to this interpolated grid as $\boldsymbol{G}_{\rm homog}$. In order to interpolate the T_0–γ plane, we follow the method described in Boera et al. (2019) and Gaikwad et al. (2020). Briefly, we rotate and translate the gas particles in the temperature–density plane to obtain models with different T_0 and γ. This preserves the temperature–density cross-correlation coefficient, allowing one to inexpensively construct models with different thermal parameters on a finely spaced grid. We also include a small correction for the smaller box size of these runs relative to our hybrid-RT models using the 40-2048 model in the upper part of Table 1. The size of the correction is at most 4 per cent at large scales, and below 2 per cent on average, over the scales of interest for this work.

Using this method, we construct a $15 \times 10 \times 7 = 1050$ grid of parameter values on top of *each* of the twelve $20\,h^{-1}$ cMpc simulations in the upper section of Table 1. The grid ($\boldsymbol{G}_{\rm homog}$) consists of 10 values of T_0 spanning the range from 5000 to 15 000 K in steps of 1000 K, seven values of γ from 1.0 to 1.7 in steps of 0.1, and 15 different values of $\tau_{\rm eff}$ ranging from 0.3 to 1.8 times the value given by equation (1), in multiplicative steps of 0.1. For 12 simulations with different u_0 values, this gives a total of $12 \times 1050 = 12\,600$ combinations on our parameter grid.

We then adopt the patchy correction from equation (3) in conjunction with the values in Table 2 to construct a second 'recovered' grid of models ($\boldsymbol{G}_{\rm patchy}$) that, while relying on the same homogeneous UVB simulations adopted in $\boldsymbol{G}_{\rm homog}$, now also corrects for the effect of patchy reionization on the transmitted flux power spectra. The choice of $z_{\rm R}$ is set by the end point of reionization in each of the homogeneous models. More specifically, when 'correcting' the simulations in the top section of Table 1 to obtain $\boldsymbol{G}_{\rm patchy}$, we adopt the $z_{\rm R}$ values listed in column 6, and then use equation (3) to modify the power spectrum appropriately. This allows us to construct, using the same method described above, an emulator that includes the effect of a patchy UVB. Note, however, that for the models with $z_{\rm R} \simeq 7.4$, this requires the uncertain extrapolation of the data in Table 2 to $z' < 4.1$. We have therefore imposed a floor to equation (3), such that we always assume $z' = 4.1$ if using a $z_{\rm R}$ values that give $z' < 4.1$ in equation (3). This may slightly overestimate the expected patchy correction for the $z_{\rm R} = 7.4$ simulations, although note the patchy reionization correction is already modest at $z' = 4.1$ and will become progressively smaller toward lower redshift. As a check, we have also compared the results obtained from this approach to those obtained for a linear extrapolation of the data below $z' = 4.1$ in Table 2. We find the differences are very small, and we do not expect this choice to impact on our conclusions.

5.3 Generation of mock Lyα forest power spectrum data

Our mock Lyα forest power spectrum data were constructed from our RT-late simulation. This is the reionization model in which the effects of inhomogeneous reionization on the power spectrum of the transmitted flux will be most prominent. The range of scales where we perform the power spectrum analysis is chosen to match Boera et al. (2019), $-2.2 \leq \log(k/{\rm km}^{-1}\,{\rm s}) \leq -0.7$, where the patchy correction is at the 5–10 per cent level at scales of $k \simeq 0.1\,{\rm km}^{-1}\,{\rm s}$, depending on the redshift and the reionization model used. We do not add any instrumental effects (e.g. noise or line broadening

due to spectral resolution) to the mock data, although any detailed comparison to observed data will ultimately require this.

We construct two sets of data from RT-late, one with $\sim$10 per cent relative errors ($D_{\rm mock}^{10\,{\rm per\,cent}}$), and one with $\sim$5 per cent relative errors ($D_{\rm mock}^{5\,{\rm per\,cent}}$). The Boera et al. (2019) measurements typically have a scale-dependent uncertainty between 10 and 25 per cent (see e.g. Fig. 7) with a redshift path-length of $dz = 2.26$, 5.63, and 2.41 at $z = 4.2$, 4.6, and 5.0, respectively. The covariance matrices for $D_{\rm mock}^{10\,{\rm per\,cent}}$ ($D_{\rm mock}^{5\,{\rm per\,cent}}$) were obtained by bootstrapping the mock data using single redshift snapshots with path-lengths $\times 2$, $\times 1$, and $\times 1.6$ ($\times 9.5$, $\times 4.3$, and $\times 7.3$) compared to those considered by Boera et al. (2019) in the $z = 4.2$, 4.6, and 5.0 redshift bins, respectively. These path-lengths were selected to achieve a mean uncertainty of 5 ± -1 per cent and 10 ± 1 per cent, respectively, for the k-bins under consideration. Hence, in the case of $D_{\rm mock}^{10\,{\rm per\,cent}}$ we consider mocks that are slightly more precise than current observed data sets from high-resolution data. On the other hand, $D_{\rm mock}^{5\,{\rm per\,cent}}$ tests the significantly improved precision that might be achieved by future measurements. This is likely to be achievable in the near future with a combination of the current generation of large-scale surveys at low spectral resolution ($R \leq 5000$) such as the Dark Energy Spectroscopic Instrument (DESI; Vargas-Magana et al. 2019) and the WHT Enhanced Area Velocity Explorer-quasi-stellar object (WEAVE-QSO; Pieri et al. 2016), and the increasing availability of homogeneous samples of high-resolution quasar spectra ($R \geq 20\,000$) such as the Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ; O'Meara et al. 2021), the Spectral Quasar Absorption Database (SQUAD; Murphy et al. 2019), and the XQR-30 (Bosman et al. 2021).

5.4 Results of MCMC analysis

The main results of our MCMC analysis are summarized in Fig. 10. The three panels show the redshift bins $z = 4.2$ (top left), 4.6 (top right), and 5.0 (bottom). Each panel consists of 4×4 subpanels showing the marginalized two-dimensional or one-dimensional posterior distributions for each of the four parameters, $\tau_{\rm eff}$, T_0, γ, and u_0. The subpanels in the upper triangular section of each panel show the posteriors in the case where the relative errors on the mock Lyα forest power spectrum are $\sim$10 per cent ($D_{\rm mock}^{10\,{\rm per\,cent}}$). The parameter estimation using either $G_{\rm homog}$ (green) or $G_{\rm patchy}$ (red) is almost identical, with the posteriors largely overlapping. The purple stars (and vertical dashed lines) in the subpanels indicate the true parameter values used to construct the mocks. The true parameters are well recovered and are always within 1σ–2σ of the posterior distribution. The best-fitting χ^2 values are 44.2 and 46.1 for $G_{\rm patchy}$ and $G_{\rm homog}$, respectively, with 36 degrees of freedom. Similarly, the best-fitting parameter values, indicated as a green or red cross for $G_{\rm homog}$ and $G_{\rm patchy}$, respectively, are well captured within the posterior distributions. Overall, this suggests that for $\sim$10 per cent uncertainty on the flux power spectrum measurements – which is similar or slightly better than the uncertainty on current high-resolution ($R \sim 40\,000$) data at $z > 4$ (Boera et al. 2019) – the effects of patchy reionization do not affect the recovery of the thermal state of the IGM (see also Wu et al. 2019).

Similar behaviour is observed in the lower section of each panel in Fig. 10, where the results using $G_{\rm homog}$ (blue) and $G_{\rm patchy}$ (orange) and the mock data with 5 per cent relative errors ($D_{\rm mock}^{5\,{\rm per\,cent}}$) are displayed. As expected, the posteriors are tighter compared to the $D_{\rm mock}^{10\,{\rm per\,cent}}$ case. The true parameter values are again indicated as purple stars, showing that even when the uncertainty on the power

spectrum is reduced by a factor of 2, the parameters are still well recovered for $G_{\rm patchy}$ (orange contours). The best-fitting $G_{\rm patchy}$ model has the χ^2 value of 30.2 for 36 degrees of freedom. Note, however, the best-fitting values are sometimes shifted along a degeneracy axis with respect to the true parameter values. While these correlations indicate that the parameters are not completely independent, they are none the less informative. There are three obvious correlations between $\tau_{\rm eff}$–γ, T_0–γ, and T_0–u_0. The first arises because the response of the power spectrum to $\tau_{\rm eff}$ or γ is nearly scale independent, due to the power spectrum being sensitive to gas close to mean density at $z \simeq 4$–5. The second parameter correlation, T_0–γ, arises because the Lyα forest is sensitive to gas over a narrow range of densities (e.g. Becker et al. 2011). This correlation therefore disappears when the typical overdensity probed by the Lyα forest becomes close to the mean density of the Universe. Finally, the last parameter correlation is between u_0 and T_0, which is also the most pronounced in our analysis. This anticorrelation results from the small-scale power suppression in the Lyα forest by thermal broadening traced by T_0, and the pressure smoothing traced by u_0 (Nasir et al. 2016; Garzilli et al. 2019; Wu et al. 2019)

For comparison, the best-fitting model for $G_{\rm homog}$ (blue contours) has the χ^2 value of 32.0 for 36 degrees of freedom, but there are now up to $\sim$1σ shifts between the estimated and true parameters for $G_{\rm homog}$ (blue contours). This implies that patchy reionization effects should start to become important for Lyα forest power spectrum analyses at small scales, $k \sim 0.1\,{\rm km^{-1}\,s}$, if measurement uncertainties are at the $\sim$5 per cent level. At this level, however, several other effects may also become important in the future data sets, such as the accuracy of the power spectrum emulator (Bird et al. 2019; Pedersen et al. 2021), simulation initial conditions (Bird et al. 2020), and hydrodynamical methods (Regan, Haehnelt & Viel 2007; Walther et al. 2021). Furthermore, the accuracy of the approximation for the patchy correction starts to approach (at most) the 5 per cent level at the smallest scales (see Fig. 9). As an example, in the left-hand panel of Fig. 11, we show the best-fitting power spectra obtained using $G_{\rm patchy}$ (solid) and $G_{\rm homog}$ (dashed) models for $D_{\rm mock}^{5\,{\rm per\,cent}}$ (i.e. the orange and blue contours in Fig. 10). These agree with the mock data very well (recall their respective χ^2 are 30.2 and 32.0 for $G_{\rm patchy}$ and $G_{\rm homog}$, respectively, for 36 degrees of freedom). However the differences between the *true power spectrum* and these two models are at around the same level as the difference between the true power spectrum and the emulator predicted power at the *true parameter values* (the right-hand panel of Fig. 11 shows the ratio of the true power spectrum and emulated power spectrum for the true parameter values). This implies that the emulator accuracy would need to be improved to analyse data with much less than a 5 per cent relative error on the power spectrum.

6 CONCLUSIONS

In this work, we use simulations drawn from the Sherwood–Relics project (Puchwein et al., in preparation) to model the effect of inhomogeneous reionization on the 1D power spectrum of the Lyα forest transmitted flux in the post-reionization IGM at $4.2 \leq z \leq 5$. We use a novel hybrid approach for combining cosmological RT with hydrodynamical simulations that captures the patchy ionization and thermal state of the IGM and self-consistently models the hydrodynamical response of the gas. Our primary conclusions are summarized as follows.

(i) Inhomogeneous reionization suppresses the power spectrum at small scales, $k \sim 0.1\,{\rm km^{-1}\,s}$, by as much as 10–15 per cent. This is broadly consistent with Wu et al. (2019), who used fully
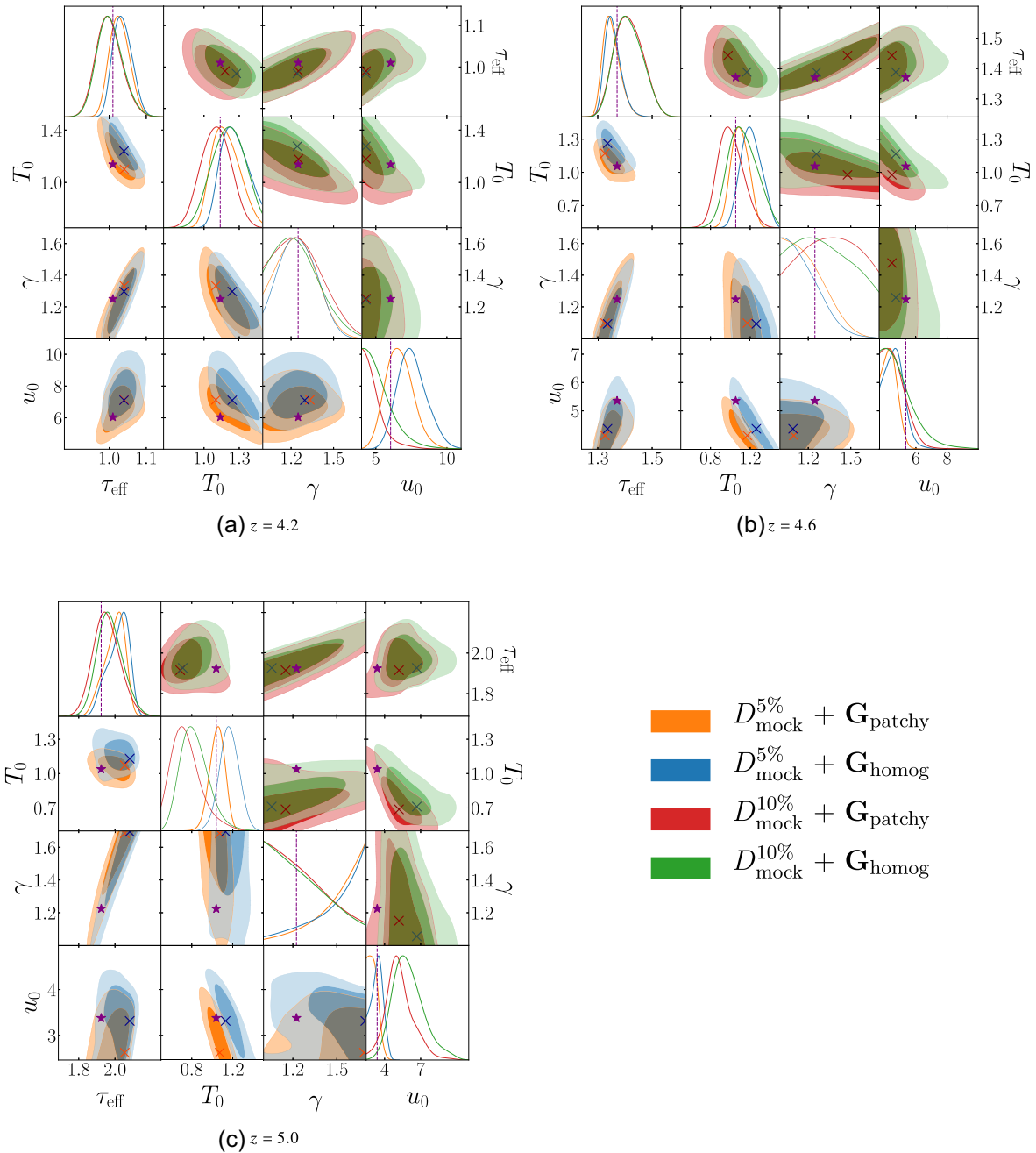
Figure 10. The one- and two-dimensional posterior distributions for τ_{eff}, T_0, γ, and u_0 from our analysis of mock Lyα forest power spectrum data drawn from the RT-late simulation at $z = 4.2$ (upper left), $z = 4.6$ (upper right), and $z = 5.0$ (lower left). The results are obtained using a grid of homogeneous UVB models (G_{homog}, blue and green contours) and when the patchy correction from equation (3) was applied to each of the grid models (G_{patchy}, orange and red contours). These two analyses were performed on mock power spectrum data with a relative error of ∼10 per cent ($D_{\mathrm{mock}}^{10\,\mathrm{per\,cent}}$, upper part of each panel), and ∼5 per cent ($D_{\mathrm{mock}}^{5\,\mathrm{per\,cent}}$, lower part of each panel). The purple stars and vertical dashed lines correspond to the true parameter values. The crosses correspond to the best-fitting parameters from the MCMC analysis with (G_{patchy}, orange/red) and without (G_{homog}, blue/green) the patchy correction from equation (3) applied to the parameter grid.

self-consistent radiation hydrodynamics simulations. We note, however, that our hybrid-RT approach allows us to provide a factor of ∼13 improvement in mass resolution and a factor of ∼4 in volume relative to Wu et al. (2019). We find this small-scale suppression is primarily driven by thermal broadening associated with the higher gas temperatures in the recently reionized regions, combined with the divergent peculiar velocity field that arises as the IGM dynamically responds to the increase in gas pressure.

(ii) On larger scales, $k < 0.03\,\mathrm{km}^{-1}$ s, we also recover the (now well established) result that inhomogeneous reionization boosts the power spectrum by 10–50 per cent (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). This arises from the large-scale variations in the IGM neutral fraction due to spatial fluctuations in the gas temperature, where gas that has been reionized recently is hotter and has less time to cool. The large-scale variations in
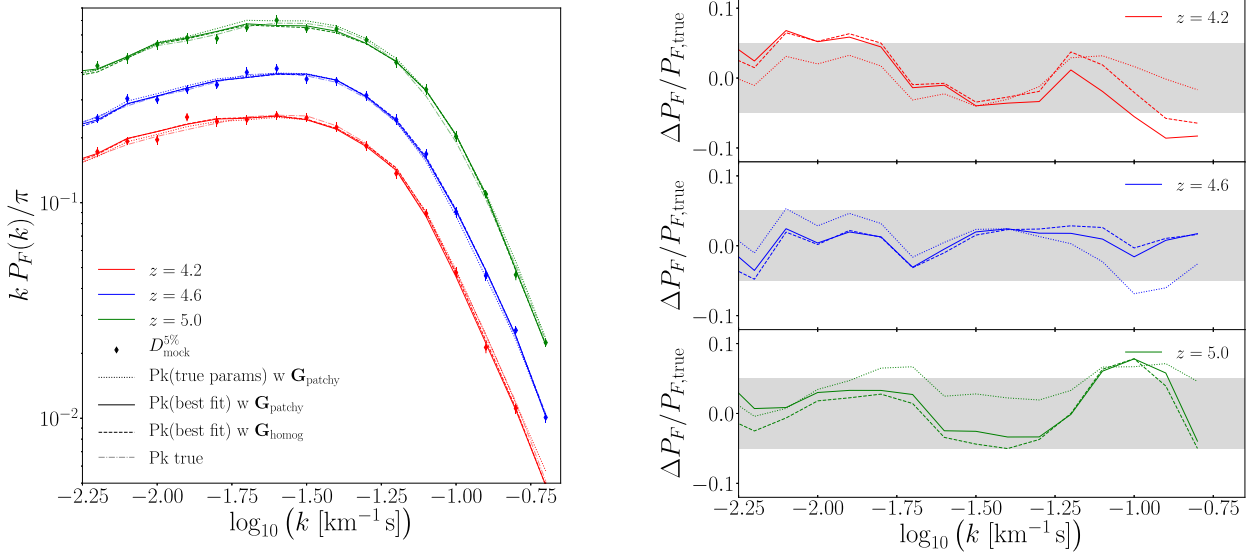
Figure 11. Left-hand panel: the best-fitting $P_F(k)$ models for mocks with 5 per cent relative errors. The points with error bars show the $D_{\mathrm{mock}}^{5\ \mathrm{per\ cent}}$ mock data, while the curves show the true underlying power from the RT-late simulation (dot–dashed) and the best-fitting models with ($\mathbf{G}_{\mathrm{patchy}}$, solid) and without ($\mathbf{G}_{\mathrm{homog}}$, dashed) the patchy correction applied to the models. The best-fitting χ^2/degrees of freedom values are 30.2/36 and 32.0/36 for $\mathbf{G}_{\mathrm{patchy}}$ and $\mathbf{G}_{\mathrm{homog}}$, respectively. For comparison, we also show the emulator prediction at the true parameter values (dotted). Right-hand panel: the residuals with respect to the true underlying model. The grey shading shows the range corresponding to a 5 per cent uncertainty.

the neutral fraction, $x_{\mathrm{H\,I}}$, then arise because of the temperature dependence of the H II recombination coefficient, where $x_{\mathrm{H\,I}} \propto T^{-0.72}$ for gas in ionization equilibrium. The boost to the large-scale power increases with increasing redshift and/or for a later end redshift for reionization. This is consistent with temperature fluctuations that gradually fade once reionization is complete, since the low-density IGM will adiabatically cool towards the thermal asymptote.

(iii) We use our hybrid-RT simulations to provide a 'patchy reionization' correction to the Lyα forest power spectrum on scales $-2.9 \leq \log(k/\ \mathrm{km^{-1}\,s}) \leq -0.7$. We provide a lookup table that enables this correction to be straightforwardly applied to existing grids of hydrodynamical simulations used in cosmological parameter inference frameworks at $4.1 \leq z \leq 5.4$ (see Table 2 and equation 3). This correction has been derived from simulations that assume $5.3 \leq z_R \leq 6.7$. We therefore caution against applying this correction to reionization histories that differ significantly from those considered here. This correction reproduces our simulation results to within 1 per cent (5 per cent) at scales $k \leq 0.03\,\mathrm{km^{-1}\,s}$ ($k \geq 0.03\,\mathrm{km^{-1}\,s}$).

(iv) We use Bayesian parameter inference to assess the importance of any biases due to inhomogeneous reionization for measurements of the IGM thermal state from the Lyα forest power spectrum. We apply our patchy reionization correction to a parameter grid of simulations with a homogeneous UVB using an MCMC analysis. Over the scales probed by the Boera et al. (2019) power spectrum measurements, $-2.2 \leq \log(k/\mathrm{km^{-1}\,s}) \leq -0.7$, inhomogeneous reionization does not introduce any significant bias in the inferred IGM parameters for mock data with a relative measurement uncertainty of $\sim$10 per cent. However, for relative uncertainties of $\sim$5 per cent – as might be achieved by future observations – there are $\sim$1σ shifts between the estimated and true parameters due to the effect of inhomogeneous reionization on the Lyα forest power spectrum. Note, however, that several other effects will also become important at the $\sim$5 per cent level, including the accuracy of the power spectrum emulator, simulation initial conditions, or the choice of hydrodynamics solver.

In summary, this study demonstrates that the effect of inhomogeneous reionization should be accounted for in future analyses of

high precision measurements of the Lyα forest power spectrum that extends to large scales at $z > 4$. For this purpose, we have provided a lookup table that gives our (model dependent) reionization correction for the Lyα forest power spectrum at $4.1 \leq z \leq 5.4$ in a form that can be easily applied within other parameter inference frameworks that use similar models for reionization. Fortunately, we also find that hydrodynamical simulations that use a homogeneous UVB are still well suited to predicting the Lyα forest power spectrum at $4.2 \leq z \leq 5$ on scales of $-2.2 \leq \log(k/\mathrm{km^{-1}\,s}) \leq -0.7$, given the *current* uncertainties ($\sim$10 per cent) on high-resolution ($R \sim 40\,000$) observational data (Boera et al. 2019). This is in good agreement with the earlier work of Wu et al. (2019), who noted that inhomogeneous reionization effects should not strongly bias existing constraints on alternative dark matter scenarios from the Lyα forest. However, this picture will change if the uncertainties on the power spectrum measurements approach the $\sim$5 per cent level at the scales that are most sensitive to the IGM thermal history or coldness of the dark matter, $k \sim 0.1\,\mathrm{km^{-1}\,s}$. This will also be the case for analyses of the Lyα forest power spectrum on larger scales, $k \leq 0.003\,\mathrm{km^{-1}\,s}$. Large-scale spectroscopic surveys that will obtain huge numbers of low-resolution spectra ($R \leq 5000$) such as DESI (Vargas-Magana et al. 2019) or WEAVE-QSO (Pieri et al. 2016) will push the power spectrum measurements to larger scales at $z \simeq 3$–4. These data will be sensitive to the remnant patchy heating associated with hydrogen reionization (e.g. Cen et al. 2009; D'Aloisio et al. 2018; Oñorbe et al. 2019; Montero-Camacho & Mao 2020). Our results provide a first step toward developing a simple template for incorporating this into the parameter inference frameworks that will be used in forthcoming analyses of the Lyα forest power spectrum.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

All data and analysis code used in this work are available from the first author on reasonable request. An open access preprint of the manuscript will be made available at arXiv.org.

REFERENCES

Almgren A. S., Bell J. B., Lijewski M. J., Lukić Z., Van Andel E., 2013, ApJ, 765, 39

Aubert D., Teyssier R., 2008, MNRAS, 387, 295

Aubert D., Teyssier R., 2010, ApJ, 724, 244

Bañados E. et al., 2016, ApJS, 227, 11

Bañados E. et al., 2018, Nature, 553, 473

Becker G. D., Bolton J. S., Haehnelt M. G., Sargent W. L. W., 2011, MNRAS, 410, 1096

Becker G. D., Hewett P. C., Worseck G., Prochaska J. X., 2013, MNRAS, 430, 2067

Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, MNRAS, 447, 3402

Bird S., Rogers K. K., Peiris H. V., Verde L., Font-Ribera A., Pontzen A., 2019, J. Cosmol. Astropart. Phys., 02, 050

Bird S., Feng Y., Pedersen C., Font-Ribera A., 2020, J. Cosmol. Astropart. Phys., 06, 002

Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, ApJ, 872, 101

Bolton J. S., Becker G. D., 2009, MNRAS, 398, L26

Bolton J. S., Haehnelt M. G., Viel M., Springel V., 2005, MNRAS, 357, 1178

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., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897

Bosman S. E. I., Fan X., Jiang L., Reed S., Matsuoka Y., Becker G., Haehnelt M., 2018, MNRAS, 479, 1055

Bosman S. E. I. et al., 2021, preprint (arXiv:2108.03699)

Boyarsky A., Lesgourgues J., Ruchayskiy O., Viel M., 2009, J. Cosmol. Astropart. Phys., 05, 012

Cen R., McDonald P., Trac H., Loeb A., 2009, ApJ, 706, L164

Chabanier S. et al., 2019, J. Cosmol. Astropart. Phys., 07, 017

Chardin J., Puchwein E., Haehnelt M. G., 2017, MNRAS, 465, 3429

Choudhury T. R., Paranjape A., Bosman S. E. I., 2021, MNRAS, 501, 5782

Ciardi B., Ferrara A., White S. D. M., 2003, MNRAS, 344, L7

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

Davies F. B., Furlanetto S. R., 2016, MNRAS, 460, 1328

Davies F. B. et al., 2018, ApJ, 864, 142

Davies J. E., Mutch S. J., Qin Y., Mesinger A., Poole G. B., Wyithe J. S. B., 2019, MNRAS, 489, 977

de Belsunce R., Gratton S., Coulton W., Efstathiou G., 2021, MNRAS, 507, 1072

Eilers A.-C., Davies F. B., Hennawi J. F., Prochaska J. X., Lukić Z., Mazzucchelli C., 2017, ApJ, 840, 24

Fan X., Carilli C. L., Keating B., 2006, ARA&A, 44, 415

Faucher-Giguère C.-A., 2020, MNRAS, 493, 1614

Finlator K., Keating L., Oppenheimer B. D., Davé R., Zackrisson E., 2018, MNRAS, 480, 2628

Furlanetto S. R., Oh S. P., 2009, ApJ, 701, 94

Gaikwad P., Srianand R., Khaire V., Choudhury T. R., 2019, MNRAS, 490, 1588

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

Garzilli A., Magalich A., Ruchayskiy O., Boyarsky A., 2021, MNRAS, 502, 2356

Gnedin N. Y., 2000, ApJ, 542, 535

Gnedin N. Y., Hui L., 1998, MNRAS, 296, 44

Greig B., Bolton J. S., Wyithe J. S. B., 2015, MNRAS, 447, 2503

Greig B., Mesinger A., Bañados E., 2019, MNRAS, 484, 5094

Haardt F., Madau P., 2012, ApJ, 746, 125

Hsyu T., Cooke R. J., Prochaska J. X., Bolte M., 2020, ApJ, 896, 77

Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27

Hui L., Haiman Z., 2003, ApJ, 596, 9

Hui L., Ostriker J. P., Tremaine S., Witten E., 2017, Phys. Rev. D, 95, 043541

Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, MNRAS, 439, 725

Iršič V. et al., 2017a, Phys. Rev. D, 96, 023522

Iršič V., Viel M., Haehnelt M. G., Bolton J. S., Becker G. D., 2017b, Phys. Rev. Lett., 119, 031302

Iršič V. et al., 2017c, MNRAS, 466, 4332

Kannan R., Vogelsberger M., Marinacci F., McKinnon R., Pakmor R., Springel V., 2019, MNRAS, 485, 117

Karaçaylı N. G. et al., 2021, MNRAS, 509, 2842

Kaur H. D., Gillet N., Mesinger A., 2020, MNRAS, 495, 2354

Keating L. C., Puchwein E., Haehnelt M. G., 2018, MNRAS, 477, 5501

Keating L. C., Weinberger L. H., Kulkarni G., Haehnelt M. G., Chardin J., Aubert D., 2020, MNRAS, 491, 1736

Khaire V., Srianand R., 2019, MNRAS, 484, 4174

Kulkarni G., Hennawi J. F., Oñorbe J., Rorai A., Springel V., 2015, ApJ, 812, 30

Kulkarni G., Keating L. C., Haehnelt M. G., Bosman S. E. I., Puchwein E., Chardin J., Aubert D., 2019, MNRAS, 485, L24

Lidz A., Malloy M., 2014, ApJ, 788, 175

Lukić Z., Stark C. W., Nugent P., White M., Meiksin A. A., Almgren A., 2015, MNRAS, 446, 3697

McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499

McQuinn M., 2016, ARA&A, 54, 313

McQuinn M., Upton Sanderbeck P. R., 2016, MNRAS, 456, 47

Makan K., Worseck G., Davies F. B., Hennawi J. F., Prochaska J. X., Richter P., 2021, ApJ, 912, 38

Mason C. A. et al., 2019, MNRAS, 485, 3947

Matsuoka Y. et al., 2018, ApJ, 869, 150

Meiksin A., Tittley E. R., 2012, MNRAS, 423, 7

Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006, New Astron., 11, 374

Molaro M., Davé R., Hassan S., Santos M. G., Finlator K., 2019, MNRAS, 489, 5594

Montero-Camacho P., Mao Y., 2020, MNRAS, 499, 1640

Montero-Camacho P., Hirata C. M., Martini P., Honscheid K., 2019, MNRAS, 487, 1047

Mortlock D. J. et al., 2011, Nature, 474, 616

Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2019, MNRAS, 482, 3458

Nasir F., D'Aloisio A., 2020, MNRAS, 494, 3080

Nasir F., Bolton J. S., Becker G. D., 2016, MNRAS, 463, 2335

O'Meara J. M., Lehner N., Howk J. C., Prochaska J. X., 2021, AJ, 161, 45

Oñorbe J., Hennawi J. F., Lukić Z., 2017a, ApJ, 837, 106

Oñorbe J., Hennawi J. F., Lukić Z., Walther M., 2017b, ApJ, 847, 63

Oñorbe J., Davies F. B., Lukić Z., Hennawi J. F., Sorini D., 2019, MNRAS, 486, 4075

Palanque-Delabrouille N., Yèche C., Schöneberg N., Lesgourgues J., Walther M., Chabanier S., Armengaud E., 2020, J. Cosmol. Astropart. Phys., 04, 038

Pedersen C., Font-Ribera A., Rogers K. K., McDonald P., Peiris H. V., Pontzen A., Slosar A., 2021, J. Cosmol. Astropart. Phys., 05, 033

Pieri M. M. et al., 2016, in Reylé C., Richard J., Cambrésy L., Deleuil M., Pécontal E., Tresse L., Vauglin I., eds, SF2A-2016: Proceedings of the Annual Meeting of the French Society of Astronomy and Astrophysics. French Society of Astronomy and Astrophysics, Lyon, France, p. 259

Planck Collaboration XVI, 2014, A&A, 571, A16

Planck Collaboration I, 2020, A&A, 641, A1

Puchwein E., Bolton J. S., Haehnelt M. G., Madau P., Becker G. D., Haardt F., 2015, MNRAS, 450, 4081

Puchwein E., Haardt F., Haehnelt M. G., Madau P., 2019, MNRAS, 485, 47

Qin Y., Mesinger A., Bosman S. E. I., Viel M., 2021, MNRAS, 506, 2390

Raskutti S., Bolton J. S., Wyithe J. S. B., Becker G. D., 2012, MNRAS, 421, 1969

Razoumov A. O., Norman M. L., Abel T., Scott D., 2002, ApJ, 572, 695

Reed S. L. et al., 2019, MNRAS, 487, 1874

Regan J. A., Haehnelt M. G., Viel M., 2007, MNRAS, 374, 196

Rogers K. K., Peiris H. V., 2021, Phys. Rev. Lett., 126, 071302

Rorai A. et al., 2017, Science, 356, 418

Rossi G., 2020, ApJS, 249, 19

Seljak U., Makarov A., McDonald P., Trac H., 2006, Phys. Rev. Lett., 97, 191303

Šoltinský T. et al., 2021, MNRAS, 506, 5818

Springel V., 2005, MNRAS, 364, 1105

Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478

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

Trac H., Cen R., Loeb A., 2008, ApJ, 689, L81

Vargas-Magana M., Brooks D. D., Levi M. M., Tarle G. G., 2019, preprint (arXiv:1901.01581)

Viel M., Haehnelt M. G., Springel V., 2004, MNRAS, 354, 684

Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., Rauch M., Sargent W. L. W., 2008, Phys. Rev. Lett., 100, 041304

Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, Phys. Rev. D, 88, 043502

Villasenor B., Robertson B., Madau P., Schneider E., 2021, ApJ, 912, 138

Walther M., Oñorbe J., Hennawi J. F., Lukić Z., 2019, ApJ, 872, 13

Walther M., Armengaud E., Ravoux C., Palanque-Delabrouille N., Yèche C., Lukić Z., 2021, J. Cosmol. Astropart. Phys., 04, 059

Wang F. et al., 2019, ApJ, 884, 30

Wang F. et al., 2020, ApJ, 896, 23

Wilson B., Iršič V., McQuinn M., 2021, MNRAS, 509, 2423

Wu X., McQuinn M., Kannan R., D'Aloisio A., Bird S., Marinacci F., Davé R., Hernquist L., 2019, MNRAS, 490, 3177

Yang J. et al., 2020, ApJ, 897, L14

Yèche C., Palanque-Delabrouille N., Baur J., du Mas des Bourboux H., 2017, J. Cosmol. Astropart. Phys., 06, 047

Zaldarriaga M., Hui L., Tegmark M., 2001, ApJ, 557, 519

## APPENDIX A: ILLUSTRATION OF THE EFFECT OF PHYSICAL QUANTITIES ON LYα ABSORPTION FEATURES

In Section 3, we explored the effect that different physical quantities play in setting the shape of the ratio $R(k, z) = P_{RT}(k, z)/P_{homog}(k, z)$ by progressively replacing quantities in the RT-mid simulation with values from the paired Homog-mid simulation (see Fig. 3). Here we illustrate how this procedure affects the transmitted Lyα flux along a single line of sight.

In the uppermost panel of Fig. A1, we show the Lyα transmission along a 1000 km s^{-1} section of a single line of sight at redshift $z = 5$ for the cases considered in Fig. 3. The second panel shows the difference between each case and the original spectrum from the simulation (i.e. the black curve in the upper panel), where the line styles match those used in Fig. 3. We also show the corresponding normalized density Δ, gas temperature T, neutral hydrogen fraction x_{HI}, peculiar velocity derivative $\partial v_{pec}/\partial v_H$, and peculiar velocity v_{pec}, for both the RT-mid (red curves) and Homog-mid (black curves) models. Note that, as was the case for Fig. 3, the mean transmission for all 5000 lines of sight drawn from the simulations is rescaled to match the effective optical depth given by equation (1). The mean transmission along *individual* lines of sight, however, can still vary around the mean of the ensemble. Hence, the spectra displayed in Fig. A1 can have a different mean transmission over the 1000 km s^{-1} range displayed.

In a few locations along the line of sight (e.g. from 2400 to 2900 km s^{-1}), there is a coherent decrease in the transmission for the cases that include a rescaling of the neutral hydrogen fraction, x_{HI}, to account for differences in the temperature-dependent recombination rate between the RT-mid and Homog-mid models (i.e. the purple and orange curves in Fig. A1). As can be seen in the lower panels, these correspond to regions where the IGM is hotter – and therefore more highly ionized – in the RT-mid simulation. When rescaling the x_{HI} to correspond to the lower temperature in the Homog-mid simulation, the substitution produces an increase in x_{HI} since $x_{HI} \propto T^{-0.7}$, leading to more absorption. The extended nature of these hot regions in the hybrid-RT runs, and the resulting boost to the transmission due to the more highly ionized hydrogen within them, leads to the increased power at large scales (low-k values) in the hybrid-RT simulations (see also Keating et al. 2018; Oñorbe et al. 2019; Wu et al. 2019).

On the other hand, when substituting the peculiar velocities (i.e. the green and purple curves) the largest visible differences in the transmission occur at ∼2700 km s^{-1}, corresponding to a region where there is an enhancement in the divergence of the peculiar velocity field the RT-mid simulation. This arises in regions that have been recently heated and are expanding (see Fig. 5). The change in the peculiar velocities smooths the absorption and produces a localized decrease in the Lyα transmission. This effect contributes towards the suppression of the power spectrum at small scales (high-k) seen Fig. 3.
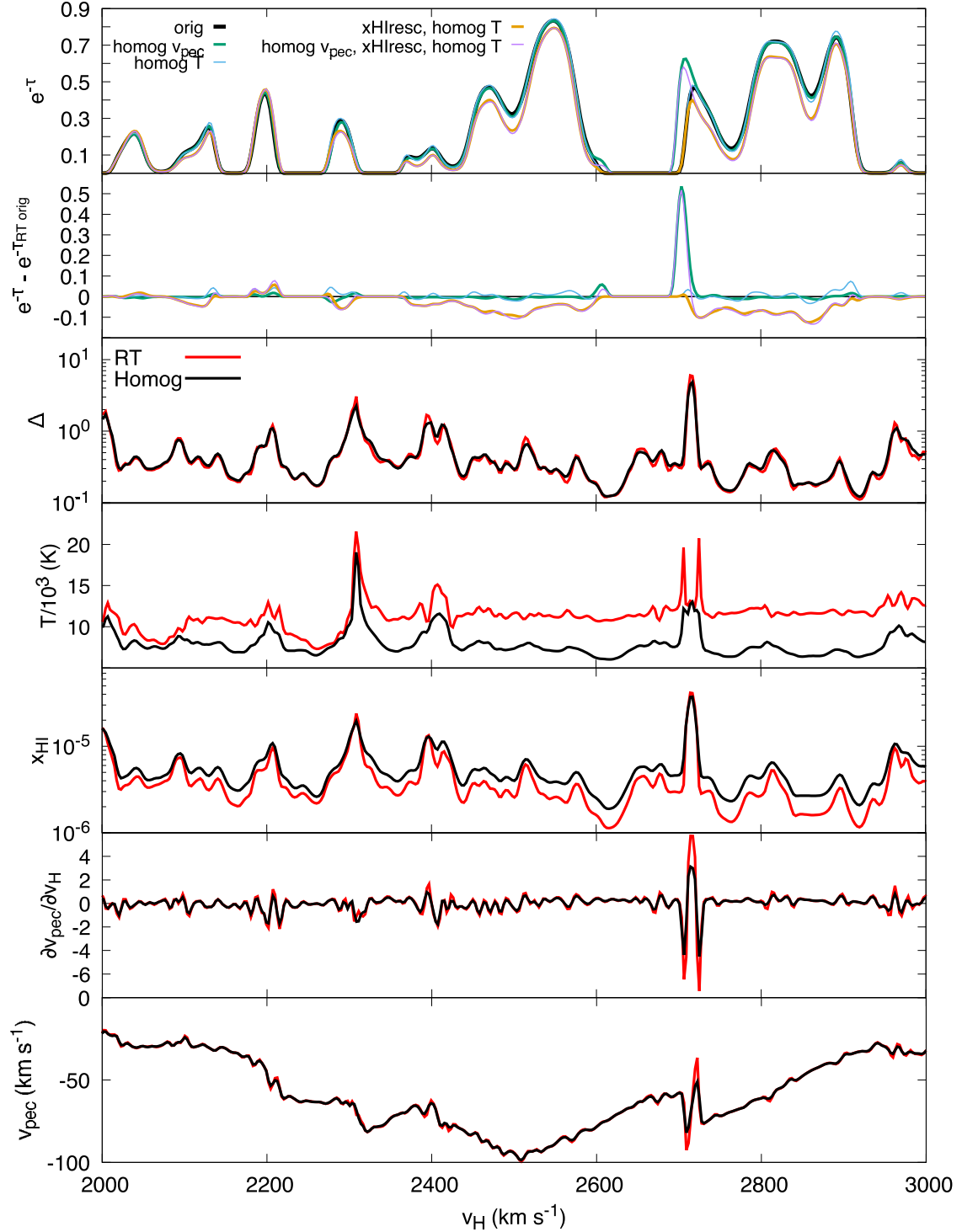
Figure A1. Top panel: the Lyα transmission along a section of a single line of sight for the cases shown in Fig. 3 at redshift $z = 5$ (note the line styles used here also match those used in Fig. 3). The original spectrum from the RT-mid simulation is shown by the black curve. Second panel: the difference in the Lyα transmission between these cases and the original RT-mid spectrum for the same line of sight. The remaining panels show the normalized gas density Δ, gas temperature T, neutral fraction x_{HI}, peculiar velocity derivative $\partial v_{pec}/\partial v_H$, and peculiar velocity v_{pec} profiles along the same segment for the RT-mid (red curves) and Homog-mid (black curves) simulations.

Finally, the effect of differences in the thermal broadening kernel on the Lyα transmission is shown by the blue curve in Fig. A1. The colder gas temperatures in the Homog-mid simulation results in less thermal broadening and hence slightly sharper absorption features compared to the original RT-mid spectrum (black curve). Hence, the hot gas associated with large-scale temperature fluctuations can also suppress power on small scales, although to a lesser extent than the peculiar velocity field. Note that the differences in the gas density, Δ, due to spatial variations in the pressure smoothing scale are typically small.

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