Impact of inhomogeneous reionization on the Lyman-α forest

Paulo Montero-Camacho,1,2⋆ Christopher M. Hirata,1,2,3 Paul Martini1,3 and Klaus Honscheid1,2
1Center of Cosmology and Astroparticle Physics, The Ohio State University, 191 West Woodruff Lane, Columbus, Ohio 43210, USA.
2Department of Physics, The Ohio State University, 191 West Woodruff Lane, Columbus, OH 43210, USA
3Department of Astronomy, The Ohio State University, 140 West 18th Avenue, Columbus, OH 43210, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
The Lyman-α forest at high redshifts is a powerful probe of reionization. Modeling and observing this imprint comes with significant technical challenges: the memory fades away swiftly, hence one must focus on high redshifts. In addition, inhomogeneous reionization must be taken into account while simultaneously being able to resolve the web-like small-scale structure, which has gone fully non-linear prior to reionization. Here we use both small box simulations capable of handling the small-scale structure of the Lyman-α forest and semi-numerical large box simulations capable of representing the effects of inhomogeneous reionization. We find that inhomogeneous reionization could produce a measurable effect on the Lyman-α forest power spectrum. The deviation in the 3D power spectrum at \( z_{\text{obs}} = 4 \) and \( k = 0.14 \text{ Mpc}^{-1} \) ranges from 19 – 36\%, with a larger effect for later reionization. The corrections decrease to 2.0 – 4.1\% by \( z_{\text{obs}} = 2 \). The impact on the 1D power spectrum is smaller, and ranges from 3.3 – 6.5\% at \( z_{\text{obs}} = 4 \) to 0.35 – 0.75\% at \( z_{\text{obs}} = 2 \), values which are comparable to the statistical uncertainties in current and upcoming surveys. Furthermore, we study how can this systematic be constrained with the help of the quadrupole of the 21 cm power spectrum.

Key words: dark ages – first stars – intergalactic medium – reionization

1 INTRODUCTION
In the standard model of cosmology, the diffuse gas in the Universe has undergone a complex thermal history marked by several important transitions. At early times, the gas was hot, dense, ionized, and fully coupled to the thermal radiation in the early Universe. As cosmic recombination occurred at \( z \sim 1100 \), the gas became transparent; perturbations from this epoch are directly observable to us as fluctuations in the cosmic microwave background (CMB) temperature and polarization. During the subsequent epoch – the cosmic Dark Ages – perturbations in the gas and dark matter grew by gravitational instability, which eventually went non-linear and began to form collapsed objects. Some of these objects emitted ionizing radiation, triggering cosmic reionization. This transition, currently believed to have occurred at \( z_{\text{re}} = 7.7 \pm 0.8 \) (Planck Collaboration et al. 2018), also must have re-heated the intergalactic medium (IGM) to temperatures up to a few \( \times 10^4 \) K after the passing of an ionization front (McQuinn 2016; D’Aloisio et al. 2018). After reionization it is possible to probe the IGM using the absorption of neutral hydrogen in quasar spectra from overdense regions – the “Lyman-α forest”. Each quasar spectrum probes a 1-dimensional skewer through the Universe, with the fraction of flux transmitted at each observed wavelength being inversely related to the gas density at the corresponding redshift.

In addition to probing the astrophysics of the IGM, the Lyman-α forest has become an important tool for observational cosmology. It probes a redshift range (2 < \( z < 5 \) ) where conventional large-area galaxy surveys are very difficult due to the enormous luminosity distance and the redshifting of bright nebular emission lines into the infrared. In contrast, while bright background quasars are rare, each one provides information from the many structures along their line of sight. The Lyman-α forest is also in some ways simpler to model than galaxies: most of the forest consists of moderately overdense gas, where the physics is dominated by gravitational instability rather than feedback processes (Cen et al. 1994; Hernquist et al. 1996; Hui et al. 1997;...
for a recent review. The Lyman-\( \alpha \) forest has been used to study reionization of hydrogen and helium (Fan et al. 2002; Cen et al. 2009; McQuinn et al. 2009; Pritchard et al. 2010; Becker et al. 2011; Compostella et al. 2013; Greig et al. 2015; McGreer et al. 2015; Bouwens et al. 2015; Nasir et al. 2016; Walther et al. 2018a; Onorbe et al. 2018); the matter power spectrum (particularly on small scales that are too non-linear in galaxy surveys; Viel et al. 2008; Weinberg et al. 1998, 2003); constrain cosmological parameters and the late evolution of the universe (Hannestad et al. 2002; Seljak et al. 2005, 2006; Croft et al. 2006; Chabanier et al. 2018); and measure the neutrino mass via its effect on the growth of structure (Palanque-Delabrouille et al. 2015; Yèche et al. 2017).

Motivated by both cosmology and astrophysics, there has been a growing number of observed Lyman-\( \alpha \) forest sightlines in the range \( 2 < z < 5 \) (Bechtold 1994; Croft et al. 1998, 1999, 2002; Rauch 1998; McDonald et al. 2000, 2006; Lidz et al. 2010; Slosar et al. 2011, 2013; Eisenstein et al. 2011; Dawson et al. 2013; Font-Ribera et al. 2014; Delubac et al. 2015; du Mas des Bourboux et al. 2017; Bautista et al. 2017; Iríñez et al. 2017; Walther et al. 2018b; Lee et al. 2018). A range of statistical measures of the Lyman-\( \alpha \) forest have been used; for cosmology purposes, the most common has been the 1D correlation function or power spectrum (that is, the power spectrum of individual sightlines, treated as 1D random fields). With the Baryon Oscillation Spectroscopic Survey (BOSS), the number of sightlines became great enough to extract cosmological parameters from the 3D correlation function of the Lyman-\( \alpha \) forest (that is, using correlations between different lines of sight). BOSS also measured cross-correlations between Damped Lyman-\( \alpha \) systems (DLA) and the Lyman-\( \alpha \) forest, and measured Baryon Acoustic Oscillations (BAO) in the Lyman-\( \alpha \) forest. The number of observed Lyman-\( \alpha \) forest will soon be dramatically enhanced with the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2015).

The current paradigm for reionization is that it is an extended and inhomogeneous process, with “bubbles” of ionized gas forming in regions with more sources of ionizing photons, and the end of reionization is when these bubbles overlap and approach a filling factor of unity. Inhomogeneous reionization leaves its imprint in the Lyman-\( \alpha \) forest even after reionization ends, because the thermal state of the gas depends on when it was reionized (Miralda-Escudé & Rees 1994) and because the thermal history itself affects the distribution of the gas at the Jeans scale. Because the attractor evolution of the IGM temperature causes the memory of reionization to fade with time (see McQuinn & Upton Sanderbeck 2016, for a detailed explanation), as well as the complication of helium reionization turning on at later redshifts (\( z \sim 3.5 \)), most studies on the effect of reionization in the Lyman-\( \alpha \) forest have been at the highest redshifts that have a sufficient number of sightlines (Hui & Haiman 2003; Trac et al. 2008; Furlanetto & Oh 2009; Lidz & Malloy 2014).

The goal of this paper is to compute a first estimate of the effect of inhomogeneous reionization in the Lyman-\( \alpha \) forest at the lower redshifts (\( 2 < z < 4 \)) relevant to precision cosmology programs such as DESI. The remnants of hydrogen reionization are expected to be small, but in the DESI era even small effects in the Lyman-\( \alpha \) forest power spectrum will be important. Due to the enormous dynamic range in spatial scales, we use two types of simulations. Our small-scale simulations use full hydrodynamics to follow the dynamics of gas down to below the Jeans scale, but with an ionization history set by hand. These use the same modification of GADGET-2 (Springel 2005) that Hirata (2018) used to study the streaming velocity effect in the Lyman-\( \alpha \) forest. The large-scale boxes use a semi-numerical approach (21cmFAST; Mesinger & Furlanetto 2007; Mesinger et al. 2011) to model the ionized bubbles and their correlation with large-scale structure.

Besides estimating the change \( \Delta P_{\mathrm{Ly}\alpha}(k,z) \) in Lyman-\( \alpha \) forest power spectrum due to reionization effects, we consider how 21 cm observations could be used to predict \( \Delta P(k,z) \) and mitigate this systematic effect. The H\( i \) 21 cm hyperfine transition is a potentially powerful probe of reionization, and global measurements of the signal have already ruled out some models (Monsalve et al. 2017; Singh et al. 2017, 2018). We find that there is a quantitative relation between \( \Delta P_{\mathrm{Ly}\alpha}(k,z) \) and the full history of the cross-correlation of the matter and the ionization fraction. The latter is related to the redshift-space distortion in the 21 cm power spectrum (Barkana & Loeb 2005; Mao et al. 2012). We show that the simplest approach — using linear perturbation theory to map the \( \ell = 2 \) quadrupole of the 21 cm signal into a matter-ionization fraction cross-power spectrum and using this to predict \( \Delta P_{\mathrm{Ly}\alpha}(k,z) \) — is not accurate. We thus recommend that future work consider mitigation using models for the 21 cm redshift space distortion that go beyond linear theory (e.g., Mao et al. 2012).

We expect the effect studied here to be of importance for current and near-future Lyman-\( \alpha \) experiments. In particular, DESI will measure the three dimensional (3D) and one dimensional (1D) power spectrum of the Lyman-\( \alpha \) forest (DESI Collaboration et al. 2016), both of which will be altered by inhomogeneous reionization. The 1D power spectrum contains the correlation of the spatial structure of the neutral hydrogen regions along the line of sight of the same quasar (McDonald et al. 2005, 2006; Palanque-Delabrouille et al. 2013), i.e., the individual sightlines. In contrast, the 3D spectrum has correlations between different lines of sight (Blomqvist et al. 2015; Bautista et al. 2017; Font-Ribera et al. 2018). The fractional effect \( \Delta P(k,z)/P(k,z) \) of inhomogeneous reionization is larger on large scales (smaller \( k \)) than at smaller scales, and is larger in the 3D Lyman-\( \alpha \) forest power spectrum than in 1D (because large \( k \) in 3D can map into small \( k \) in 1D but not the other way around). The 1D power spectrum is affected by non-linearities, even for small \( k \)\( \parallel \), due to small-scale processes that govern the evolution of the IGM (Palanque-Delabrouille et al. 2013; Arinyo-i-Prats et al. 2015).

This paper is organized as follows. We introduce the relevant formalism in §2. We describe the two types of simulations we use in §3. Then we proceed to assess the possible contamination of the Lyman-\( \alpha \) flux in §4, and confirm that for this systematic the 3D Lyman-\( \alpha \) power spectrum is more affected than the 1D. In §5 we explore the use of 21 cm observables to address this contamination, and find that non-linear effects must be taken into account to appropriately utilize the link between the imprint of inhomogeneous reionization in the Lyman-\( \alpha \) forest and the 21 cm quadrupole. We
summarize our results and discuss directions for future work in §6.

2 CONVENTIONS AND FORMALISM

Throughout this paper we use the ΛCDM cosmological parameters from *Planck* 2015 TT+TE+EE+lowP-lensing+ext (Planck Collaboration et al. 2016), namely: $\Omega_b h^2 = 0.02230$, $\Omega_m h^2 = 0.14170$, $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\sigma_8 = 0.8159$ and $n_s = 0.9667$.

For convenience throughout this paper we use the dimensionless power spectrum when we refer to a given power spectrum, i.e. no Mpc unless explicitly stated. However, in the case of the 21 cm power spectrum the $k^3P(k)/2\pi^2$ have units of mK$^3$. Furthermore, we define all fluctuations as $\delta = p/p - 1$, with the only exception of the neutral hydrogen fraction fluctuations where $\delta_{\text{HI}} = \delta_H - \delta_m$.

The fluctuations of the Lyman-α forest transmitted flux are described by $\delta F = (F - \bar{F})/\bar{F}$, where $\bar{F}$ is the mean observed flux and the transmitted flux is normalized such that $0 \leq F \leq 1$. Then at linear order the flux fluctuations should be proportional to the matter fluctuations. Furthermore, taking into account redshift space distortions due to velocity gradients we have $\delta F = b_p (1 + \beta_p \mu^2) \delta_m + b_\perp \psi(z_{\text{re}})$.

In this work we explore the effect of inhomogeneous reionization in the Lyman-α transmission,

$$\delta F = b_p (1 + \beta_p \mu^2) \delta_m + b_\perp \psi(z_{\text{re}}),$$

where $b_p$ is the radiation bias defined by $b_p = \delta \ln \bar{F} / \delta \ln n_{\text{HI}}$ (Arinyo-i-Prats et al. 2015; Hirata 2018). Here $\tau_1$ is the optical depth that must be assigned to a patch of gas with mean density $\delta H = 1 + \delta_H = 1$ and temperature $T = 10^4 \text{ K}$ in order for the mean transmitted flux $\bar{F}$ to match observations. This bias parameter encodes the way large scale ionizing fluctuations affect the Lyman-α flux (Arinyo-i-Prats et al. 2015). In addition, we define

$$\psi(z_{\text{re}}) = \Delta \ln n_{\text{HI}}(z_{\text{re}}, \bar{z}_{\text{re}}) = \left[ \frac{\tau_1(z_{\text{re}})}{\tau_1(z_{\text{re}}, \bar{z}_{\text{re}})} \right],$$

which parametrizes the variations in the transparency for an opacity cube that suddenly reionizes at $z_{\text{re}}$ relative to an opacity cube that reionizes at $\bar{z}_{\text{re}}$. The sign convention is such that $\psi(z_{\text{re}})$ > 0 if gas reionized at $z_{\text{re}}$ is more transparent (has higher $F$) than gas reionized at $\bar{z}_{\text{re}}$.

The corresponding 3D Lyman-α flux power spectrum is given by

$$P_F^{3D}(k, \mu, z_{\text{obs}}) = \frac{b_p^2}{2} (1 + \beta_p \mu^2)^2 P_m(k, z_{\text{obs}}) + 2b_\perp b_p (1 + \beta_p \mu^2) P_m \psi(k, z_{\text{obs}}),$$

where $P_m$ is the matter power spectrum and $P_m \psi$ is the cross-power spectrum of matter and $\psi$. Note that we have neglected the third term in equation (3), the auto-power spectrum of $\psi$, because it is second order in $\psi$.

In principle the problem is now reduced to devising a way to model $P_m \psi(k, z)$ and, since it involves physics from different scales, we will divide and conquer. We will relate the cross-power spectrum of $\psi$ and matter by using the cross-power spectrum of matter and neutral fraction as our proxy to describe how matter and bubble spatial structure are correlated at a given redshift, i.e. how patchy reionization enters the fray. Furthermore, we directly introduce the change of the transparency of the IGM by means of its derivative with respect to redshift. We integrate from $z = 34.7$ to $z = 5.90$, such that in our default model we cover from the beginning of the Wouthuysen-Field coupling through the end of reionization. We then compute

$$P_m \psi(z_{\text{obs}}, k) = \int_{5.90}^{34.7} dz P_m \psi(z, k) \frac{D(z_{\text{obs}})}{D(z)},$$

where we have extrapolated to the observed redshift by including the ratio of growth functions. Also, note that the minus sign converts from neutral to ionized. We compute $P_m \psi$ by modifying the code available in 21cmFAST (Mesinger et al. 2011). It is important to highlight that in fact what gets computed from this procedure is the dimensionless power spectrum $\Delta_{m,HI}$ and hence $\Delta_m \psi$.

3 SIMULATIONS

To predict the effect of reionization on the Lyman-α forest, we need two types of simulations.

The “small box” simulation uses a box size that is smaller than the typical scale of a reionization bubble ($L = 2.55 \text{ Mpc}$), and can resolve structures down to below the Jeans scale. Its purpose is to determine how the transmitted flux of the Lyman-α forest depends on the reionization redshift $z_{\text{re}}$, i.e., to determine the function $\psi(z_{\text{re}})$. The small boxes are assumed to reionize instantaneously, and to determine the functional form $\psi(z_{\text{re}})$ we run a sequence of boxes with $z_{\text{re}}$ stepped over the interesting range (here 6–12). The small boxes have high resolution, since the thermal state of the IGM depends on the way in which small-scale structures are disrupted following reheating (Hirata 2018). Aside from triggering reionization at a particular redshift, these boxes do not need additional information on ionizing sources.

In contrast, the “large box” simulations are large compared to the scale of reionization bubbles. We choose $L = 300 \text{ Mpc}$ such that the simulations have enough statistical power. Their purpose is to predict the spatial structure of the bubbles and their correlation with matter. The simulations have a prescription for placing ionizing sources, and a fast approximation scheme to track the fate of ionizing photons.

It is only with the combination of the two types of boxes that we can compute all of the ingredients in §2. Note that these simulations incorporate different physical ingredients and hence we will use two completely separate codes.

3.1 Small boxes

We follow Hirata (2018) in the methodology of the small box simulations, but we only use the biggest box size stud-

---

1 See Arinyo-i-Prats et al. (2015) for a description of the linear biasing coefficients and their range of validity.

2 This is the same usage as in Hirata (2018).
ied in their work. In particular, we use a modified version of GADGET-2 (Springel et al. 2001; Springel 2005), which is a smoothed particle hydrodynamics (SPH) code. In this modified GADGET-2 we include the most relevant heating and cooling processes for the gas. We add Compton heating/cooling for neutral gas (with residual ionization) and ionized gas; see Eq. (17) in Hirata (2018). Also for ionized gas we include recombination cooling, photonization heating, free-free cooling, and He\textsuperscript{II} line cooling; see Eqs. (19-23) in Hirata (2018). Following the default treatment in Hirata (2018), reionization is treated with an uniform post-reionization temperature of $T_{\text{re}} = 2$ (or $T_{\text{re}} = 2 \times 10^4$ K) everywhere.

Each simulation we used in GADGET-2 has the same box size, $L = 1728 h^{-1}$ = 2551 kpc and the same number of particles, $2 \times (384)^3$. The dark matter particle mass in the simulations is $9.72 \times 10^3 M_\odot$, while the gas particle mass is $1.81 \times 10^2 M_\odot$. The only difference among them is when reionization turns on. We simulate eight realizations in order to reduce the statistical error due to the limited box size by a factor of $\sqrt{8}$.

As in Hirata (2018) we start all the simulations at recombination, $z_{\text{dec}} = 1059$ with a modified version of N-GEN-IC (the default initial condition generator in GADGET-2) to enable streaming velocities between the baryons and dark matter (Tseliakhovich & Hirata 2010). The boxes are evolved with neutral gas physics until reionization, at which point the temperature is reset to $2 \times 10^4$ K, and the box is evolved further with singly ionized (H\textsuperscript{II}/He\textsuperscript{II}) primordial gas physics. There is no He\textsuperscript{II} reionization in these simulations. These simulations were tested for convergence in §5.3 of Hirata (2018).

All of the small box simulations were done on the Ruby cluster at the Ohio Supercomputer Center (Ohio Supercomputer Center 2015).

### 3.2 Large Boxes

We include the physics of reionization-bubble-scale by using semi-numerical simulations, specifically we use 21cmFAST (Mesinger et al. 2011). We utilize all the default parameters in 21cmFAST with the exceptions of a few parameters that we change for different alternatives scenarios in order to examine different reionization histories, and the use of our chosen background cosmology. The box size for each one of these simulations is $L = 300$ Mpc. We run eight realizations for each of the three different reionization scenarios with the same physical setup but with different random seeds to diminish the simulation variance. Furthermore, we generate snapshots that track – among other parameters – the density, neutral fraction of hydrogen and the 21 cm temperature fields for $34.70 \geq z \geq 5.90$, with a step size of 2%, i.e., $z_{\text{c}} = (z_{\text{s}} + 1)/1.02 - 1$. Therefore, in total we have 84 snapshots (each with 3 dependent variables) for the computation of the cross-power spectrum of matter and neutral fraction, and the 21 cm quadrupole per realization. In order to compute the cross-power spectra we only modify “delta.ps” in 21cmFAST, such that it can take both a density and neutral fraction snapshot. Furthermore, to be able to compute the 21 cm quadrupole we include the ability to compute $P_T(k, \mu)$ instead of $P_T(k)$.

To explore different reionization histories we change the number of ionizing photons escaping into the IGM per baryon in collapsed structures, i.e. $H\text{II}$\textunderscore EFF\textunderscore FACTOR in 21cmFAST, which represents the ionizing efficiency

$$\zeta = N_{H\text{II}} f_{\text{esc}} f_b,$$

where $f_b$ is the baryon fraction of a halo in units of the cosmic mean value $\Omega_b/\Omega_m$, $f_s$ is the fraction of baryons from the halos that form stars, $N_{H\text{II}}$ corresponds to the number of ionizing photons produced per stellar baryon, and $f_{\text{esc}}$ is the fraction of produced ionizing photons which escape into the IGM.

We define a default model with optical depth to reionization equal to the current best fit value of the recent final results of Planck 2018 $TT+TE+EE+lowE+lensing$ (Planck Collaboration et al. 2018), i.e. $\tau = 0.054$, which corresponds to an ionizing efficiency of $\zeta = 25.3$. In our default model this optical depth corresponds to a redshift of reionization of $z_{\text{re}} = 7.61$, which corresponds to when the mean neutral hydrogen fraction is 0.5. For comparison purposes we chose alternative models with later ("model A") and earlier ("model B") reionization than our default model. We select these alternative scenarios so that they encompass the range of models consistent with the new Planck results and other reionization probes at $z \geq 5.9$ (see Bouwens et al. 2015 for a compact list of these probes). We choose our model A such that the volume weighted neutral fraction at $z = 5.9$ approximately matches the 1σ upper limit extracted from the dark pixel measurements reported in McGreer et al. (2015), $x_{\text{HI}} \leq 0.11$. This corresponds to roughly $\zeta = 20.9$, and by construction the reionization process has not been completed by the end of our large simulations.

### Table 1: Summary of the different reionization models

| Reionization model | $\tau$ | $z_{\text{re}}$ | $\zeta$ |
|--------------------|--------|-----------------|--------|
| A: later reionization | 0.0512 | 7.22 | 20.9 |
| Default: Planck 2018 reionization | 0.0548 | 7.61 | 25 |
| B: earlier reionization | 0.0615 | 8.34 | 35 |

The main difference between the Planck 2015 and 2018 results is the value of the optical depth which in 2015 was $\tau = 0.066$. We keep all other cosmological parameters equal to the cosmology from Planck 2015 used throughout this paper.

We note that our model A might be considered as conservative in light of recent high redshift constraints of the hydrogen neutral fraction, e.g., $x_{\text{HI}} = 0.88$ at $z = 7.6$ (Hoag et al. 2019) and $x_{\text{HI}} > 0.76$ at $z = 8$ (Mason et al. 2019).

The lower $\tau$ from Planck was reported too late to change the small-box simulations to run to lower $z$.  

---

3 The main difference between the Planck 2015 and 2018 results is the value of the optical depth which in 2015 was $\tau = 0.066$. We keep all other cosmological parameters equal to the cosmology from Planck 2015 used throughout this paper.

4 We note that our model A might be considered as conservative in light of recent high redshift constraints of the hydrogen neutral fraction, e.g., $x_{\text{HI}} = 0.88$ at $z = 7.6$ (Hoag et al. 2019) and $x_{\text{HI}} > 0.76$ at $z = 8$ (Mason et al. 2019).

5 The lower $\tau$ from Planck was reported too late to change the small-box simulations to run to lower $z$.  

---

MNras 000, 1-10 (2017)
4 ASSESSMENT OF CONTAMINATION

The goal of this work is to quantify the effect of inhomogeneous reionization in the Lyman-α flux. In principle to estimate the contamination to the signal we simply require both terms in the RHS of equation (3), i.e. the linear theory term and the non-linear extension.

4.1 Linear power spectrum (3D)

We start by computing the linear term. We simplify this process by assuming that \( \mu = 0 \), i.e. perpendicular to the line of sight so that one can ignore redshift space distortions. We compute the matter power spectrum from the density boxes generated in 21cmFAST with our chosen cosmology. For the sake of illustrating the contamination we choose a redshift of observation of 2.5 and a wavenumber of 0.14 Mpc\(^{-1} \), both values are typical of Lyman-α measurements. Then, we have

\[
P_{F}^{3D}(z_{\text{obs}} = 2.5, k = 0.14 \text{ Mpc}^{-1}) = 995.3 \beta_{F}(z_{\text{obs}}) \text{ Mpc}^{3}.
\]  

We extract the Lyman-α flux bias from Table 1 in McQuinn & White (2011). We summarize the relevant values in Table 2. Since we are using the results from McQuinn & White (2011) we will be consistent with their assumption that the redshift distortion parameter is equal to unity, i.e. \( \beta_{F} = 1 \) throughout our work.

4.2 Computation of \( \psi(z_{\text{re}}) \)

We use our smaller simulations to calculate how the transparency of the Lyman-α forest depends on when reionization occurs. We obtain \( \psi \) from Eq. (2) and present its redshift dependence in Table 3. In what follows, we linearly interpolate \( \psi \) between the redshifts in the table.

4.3 Cross-power spectrum of matter and \( \psi \)

In Eq. (4) we accounted for the inhomogeneous nature of reionization by including the cross-power spectrum of matter and hydrogen neutral fraction. We obtain this cross-power spectrum from our modified version of 21cmFAST. We plot the dimensionless cross-power spectrum of matter and neutral fraction as a function of reionization for our different models of reionization history. All cross-power spectra have been evaluated at wavenumber \( k = 0.14 \text{ Mpc}^{-1} \).

We determine the impact of inhomogeneous reionization on the Lyman-α flux from both our calculations of how the transparency of the IGM depends on the redshift of reionization §4.2, and the cross-correlation of matter and the neutral hydrogen fraction. Using Eq. (4) evaluated at \( k = 0.14 \text{ Mpc}^{-1} \) and observing at redshift 2.5 for our default model, we obtain \( P_{m,\psi} = -26.7 \beta_{F}(z_{\text{obs}}) \beta_{\psi}(z_{\text{obs}}) \text{ Mpc}^{3} \). Therefore we see that the contamination to the linear term in Eq. (6), ignoring the bias factors, is approximately 5.36%. We ignored the bias ratio to quantify this change, however as seen in Table 2 both bias factors can significantly affect the deviation. We tabulated the deviation of the Lyman-α forest power spectra, taking into account the role of the bias factors, for wavenumber 0.14 Mpc\(^{-1} \) and different redshifts of observation in Table 4. We obtained the radiation bias from our GADGET 2 simulations\(^{6} \) (see Table 2).

In Fig. 3 we plot a comparison of the strength of the effect by looking at the ratio of the 3D dimensionless cross-

---

\( \psi \) is calculated as the redshift distortion parameter.

---

Note that the bias factors obtained from our simulations seem to be slightly higher than the ones reported in Hirata (2018); nevertheless, we consider these good enough for a first estimation of the effects of inhomogeneous reionization in the Lyman-α flux.

---

\( ^{6} \) Note that the bias factors obtained from our simulations seem to be slightly higher than the ones reported in Hirata (2018); nevertheless, we consider these good enough for a first estimation of the effects of inhomogeneous reionization in the Lyman-α flux.
Table 3. Results for the small-scale simulations. Transparency variations in the IGM, $\psi = \Delta \ln \tau_1(\text{Sim. B})/\tau_1(\text{Sim. A})$. The number in square brackets is the redshift at which reionization turns on. Note that $\Delta \ln \tau_1$ is negative if simulation A is more transparent than simulation B.

| Box size [ckpc] | Sim. A | Sim. B | $z = 2.0$ | $z = 2.5$ | $10^5 \times \Delta \ln \tau_1$ |
|----------------|--------|--------|-----------|-----------|------------------|
| 2551           | [8]    | [6]    | 6064 ± 553| 6910 ± 895| 9772 ± 973       |
|                 | [7]    |        | 2112 ± 235| 2399 ± 367| 3612 ± 436       |
| 2551           | [8]    | [9]    | -878.8 ± 166| -649.6 ± 252| -1083 ± 315  |
|                 | [10]   |        | -1300 ± 303| -407.5 ± 406| -847.8 ± 560  |
| 2551           | [11]   |        | -1425 ± 413| -52.51 ± 582| -318.9 ± 724  |
| 2551           | [12]   |        | -1552 ± 482| 199.7 ± 674 | 283.5 ± 850   |

Table 4. Percentage deviation of the 3D and 1D Lyman-α power spectrum due to patchy reionization for the different reionization models considered. Here we have used $k = 0.14$ Mpc$^{-1}$, and we included the bias ratio.

| Simulation | Model A 3D | Default 3D | Model B 3D | Model A 1D | Default 1D | Model B 1D |
|------------|------------|------------|------------|------------|------------|------------|
| $z_{\text{obs}} = 2.0$ | 4.09 ± 0.47 | 5.56 ± 0.98 | 1.96 ± 0.32 | 0.75 ± 0.08 | 0.61 ± 0.08 | 0.35 ± 0.06 |
| $z_{\text{obs}} = 2.5$ | 1.96 ± 1.45 | 1.64 ± 1.69 | 1.36 ± 0.26 | 1.81 ± 0.28 | 2.92 ± 0.31 | 0.52 ± 0.21 |
| $z_{\text{obs}} = 3.0$ | 19.6 ± 1.87 | 31.0 ± 2.31 | 19.3 ± 1.73 | 3.63 ± 0.34 | 5.45 ± 0.37 | 1.52 ± 0.26 |
| $z_{\text{obs}} = 3.5$ | 35.9 ± 2.63 | 31.0 ± 2.31 | 19.3 ± 1.73 | 6.53 ± 0.42 | 5.45 ± 0.37 | 3.27 ± 0.29 |
| $z_{\text{obs}} = 4.0$ | 39.1 ± 2.63 | 34.1 ± 2.31 | 19.9 ± 1.73 | 6.83 ± 0.42 | 5.45 ± 0.37 | 3.27 ± 0.29 |

4.4 Linear power spectrum (1D)

Once we have the 3D power spectrum (the cross-correlation between different skewers) we can compute the 1D Lyman-α power spectrum (the cross-correlation between pixels of the same skewer) which is given by averaging over the perpendicular direction to the line of sight, i.e.

\[
P_{1D}^k(k, z_{\text{obs}}) = \int_0^\infty \frac{dk}{2\pi} k^2 P_{1D}^k(k, z_{\text{obs}}) = b_P^2 \int_0^\infty \frac{dk}{2\pi} k^2 (1 + \mu^2)^2 P_{m, \psi}^k(k, z_{\text{obs}}) + 2 b_P b_F \Gamma \int_0^\infty \frac{dk}{2\pi} k^2 (1 + \mu^2) P_{m, \psi}^k(k, z_{\text{obs}}).
\]

Since our 21cmFAST simulations use a $k_{\text{max}} \approx 3.20$ Mpc$^{-1}$, we set this wavenumber as the upper limit of the integrals over the perpendicular direction.

We report the percentage of the deviation, i.e. the ratio of $2(b_P/b_F)P_{m, \psi}^k/P_{m, \psi}^k \times 100\%$ from our simulations in Table 4. Furthermore, we plot the ratio of $P_{m, \psi}^k/P_{1D}^k$ as a function of wavenumber for the different redshifts of observation that we explored in Fig. 4. We use the 1D Lyman-α power spectrum from BOSS (Palanque-Delabrouille et al. 2013). From Fig. 4 and Fig 3 we confirm that the effect of reionization is stronger for the Lyman-α 3D power spectrum than for the Lyman-α 1D power spectrum, as one could have expected due to the integration smoothing the deviation. Moreover, the 1D ratio is also consistent with the redshift hierarchical structure (within the error bars). The error bars shown in Fig. 4 have been computed in the same way we described in §4.3, and hence we have ignored the error bars from the BOSS.
Reionization effect on the Lyman-α forest

Figure 3. Comparison for the different reionization models between the dimensionless cross-power spectrum of ψ and matter, and the dimensionless power spectrum of matter as a function of wavenumber for fixed redshift of observation.

Figure 4. Comparison between the 1D dimensionless cross-power spectrum of ψ and matter, and the 1D dimensionless power spectrum of matter as a function of wavenumber for fixed redshift of observation.

5 CONSTRAINING THE EFFECT WITH 21 CM COSMOLOGY

In the previous section we described a potentially very significant systematic for Lyman-α forest measurements, one that is particularly important for later reionization scenarios. In this section we will illustrate how this systematic could become an interesting link between Lyman-α and 21 cm cosmology.

In this section we show how non-parametric mitigation of the thermal imprint of reionization in the Lyman-α forest is possible through the use of the “linear µk-decomposition” scheme (Barkana & Loeb 2005). Unfortunately, as we will
illustrate, non-linear effects are significant and hence in future work we will continue our efforts of non-parametric mitigation by employing the “quasi-linear $\mu_k$-decomposition” (Mao et al. 2012).

Under the linear $\mu_k$-decomposition the 3D 21 cm power spectrum is given by (Mao et al. 2012)

$$ P_T^{3D}(k) = P_{\mu_0}(k) + P_{\mu^2}(k) P_{\mu}(k) P_{\mu^4}(k), \quad \text{(10)} $$

where each $P_{\mu}(k) = \langle P_{\mu}(k) \rangle$ is angle-averaged over constant k-shells. Moreover, in the limit of $T_k \gg T_{\text{CMB}}$, which corresponds to the range where $P_{m,HI}$ is important (see Fig. 2), Eq. (10) becomes

$$ P_{\mu_0}(k, z) = (\delta T_b(z))^2 P_{\delta m}(k, z), \quad \text{(11)} $$

$$ P_{\mu^2}(k, z) = 2 (\delta T_b(z))^2 P_{\delta m}^2 P_{\delta HI}(k, z), \quad \text{(12)} $$

$$ P_{\mu}(k, z) = (\delta T_b(z))^2 P_{\delta HI}(k, z), \quad \text{(13)} $$

where $\delta_{m,HI} = \delta_m + \delta_{\text{HI}} + \delta_{\text{HI,HI}}$ and $\delta T_b$ is the mean of the brightness temperature. In the limit of $T_k \gg T_{\text{CMB}}$ we have

$$ \delta T_b(z) = \frac{3c^3 A_{HI} T_{b,HI}(z)}{32\pi v_{HI}^2 (1 + z) T_{HI}(z)} \approx 26.6 T_{b,HI}(z) \left( \frac{\Omega_m h^2}{0.0223} \right) \left( \frac{0.1417}{10} \right)^{1/2} \text{mK}, \quad \text{(14)} $$

where $A_{HI}$ is the Einstein coefficient of the hyperfine transition, $T_{b,HI}$ is the 21 cm hyperfine transition in temperature units and $v_{HI}$ is the velocity of the HI.

Rewriting Eq. (10) into $\ell$-multipoles we have

$$ P_T^{3D}(k, \ell) = P_{\mu_0} + P_{\mu^2} + P_{\mu^4} + \frac{2}{3} \left( P_{\mu^2} + \frac{6}{7} P_{\mu^4} \right) L_2(\mu) + \frac{8}{35} P_{\mu^4} L_4(\mu), \quad \text{(15)} $$

where $L_\ell$ are the Legendre polynomials. Note that in our notation the quadrupole term can be expanded as

$$ P_T^{\ell=2} = \frac{2}{3} (\delta T_b)^2 \left( \frac{2}{3} P_{m,HI} + \frac{20}{7} P_m \right), \quad \text{(16)} $$

where we have taken advantage of the fact that hydrogen traces the matter distribution, and the extra factor of neutral hydrogen fraction comes from the way we have defined our perturbations on the neutral fraction, i.e. $\delta_{\text{HI}} = x_{\text{HI}} - 3\delta_{\text{HI}}$.

Hence the cross-power spectrum of matter and fraction of neutral hydrogen atoms is related through Eq. (16) to the $\ell = 2$ multipole component of the 21 cm power spectrum. Therefore with a measurement of the 21 cm power spectrum one can in principle constrain the cross-power responsible for the systematic imprint in the Lyman-\(\alpha\) forest.

The first step for our mitigation scheme to be successful is to be able to reproduce the 21 cm power spectrum. As was quantified in Fig. 10 of Mao et al. (2012), the error in the linear method can easily reach thirty percent or more for the smallest scales. Thus one should expect significant errors in the quadrupole of the 21 cm power spectrum computed with the linear approximation.

We test the accuracy of the linear decomposition, i.e. we compare the right-hand side of Eq. (16) with the output from the 21cmFAST simulations. We show the failure of the linear $\mu_k$-decomposition in Fig. 5 for the default model at redshift 7. The linear decomposition fails similarly for the other models.

6 DISCUSSION

Hydrogen reionization is one of the defining events in the thermal and dynamical history of the IGM. It heats up the IGM to $> 10^4$ K, and by increasing the Jeans mass, it disrupts pre-existing small-scale structure. We have seen that the thermal and dynamical effects persist for cosmological timescales, and that the transmission of the Lyman-\(\alpha\) forest even at $z_{\text{obs}} < 3$ is sensitive to the reionization model. Since reionization is believed to have occurred inhomogeneously – by expanding and overlapping ionized “bubbles” correlated with large-scale structure – this dependence on reionization redshift $z_{\text{obs}}$ translates into a spatial modulation of the Lyman-\(\alpha\) forest and hence a correction to the power spectrum.
Reionization effect on the Lyman-α forest

The magnitude of the effect is largest in the 3D power spectrum at the highest observed redshifts and on large scales (which are better matched to the scale of reionization bubbles). For example, in the 3D power spectrum at \( z_{\text{obs}} = 4 \) and \( k = 0.14 \text{ Mpc}^{-1} \), we find corrections of 19–36\% depending on the reionization model chosen. At lower \( z_{\text{obs}} \), the effect of reionization is reduced, declining to 2.0–4.1\% at \( z_{\text{obs}} = 2 \). For the 1D Lyman-\( \alpha \) power spectrum the deviation is significantly smaller: again at \( k = 0.14 \text{ Mpc}^{-1} \) we estimate 3.3–6.5\% at \( z_{\text{obs}} = 4 \), declining to 0.35–0.75\% at \( z_{\text{obs}} = 2 \). The corrections due to reionization are small, but we should remember that the 1D power spectrum is already measured at very high signal-to-noise ratio: for example at \( z_{\text{obs}} \approx 2.2 \) and \( k = 0.116 \text{ Mpc}^{-1} \), BOSS-eBOSS have a statistical error of 1.2\% per bin\(^7\) (Chabanier et al. 2018). These statistical errors will shrink further in the DESI era.

In principle, measurements of diffuse 21 cm radiation from the epoch of reionization can constrain the reionization model, one of the key inputs in calculating the correction to the Lyman-\( \alpha \) power spectrum. We are particularly interested in the 21 cm quadrupole \( P_T^{2}(k, z) \), since it is sensitive to the specific power spectrum \( P_{\text{in}}(k, z) \) that we need. Unfortunately, the simplest implementation of this idea – the “linear µ\( z \) decomposition” theory – is not accurate in the range of parameter space we need. In future work we will investigate other correction schemes, including models with corrections to linear theory (building on past work, e.g., Mao et al. 2012), and schemes where reionization models are parameterized (e.g., source ionizing efficiency, minimum mass, IGM clumping parameters) and then 21 cm observations are used to constrain the parameters rather than directly infer power spectra.

Another avenue for future work is to incorporate some of the other physical effects that may interact with inhomogeneous hydrogen reionization. One of the most important may be He\( ii \) reionization, which is believed to have occurred around \( z \sim 3 \) and resulted in an additional energy injection into the IGM. This has likely been seen in the thermal evolution of the IGM inferred from the Lyman-\( \alpha \) forest (e.g., Becker et al. 2011; Walther et al. 2018a). This additional energy injection can reduce the sensitivity of the low-redshift IGM to its initial thermal state, e.g., it can reduce \( \partial \ln T(z = z_{\text{obs}})/\partial \ln T(z = z_{\text{rec}}) \) (see, e.g., the discussion in Hirata 2018 in the context of streaming velocities), though the change depends on the timeline and the relative contributions of EUV and X-ray radiation. Previous studies have also found that He\( ii \) reionization introduces its own imprints on the Lyman-\( \alpha \) forest (McQuinn et al. 2009; Compostella et al. 2013; Greig et al. 2015). In any case, it appears that at the level of precision of interest for modern cosmological Lyman-\( \alpha \) forest studies, the IGM may not have relaxed from inhomogeneous hydrogen reionization before helium reionization takes place.

A second issue is that we have taken only a simplified model for hydrogen reionization itself: we have ignored X-ray heating prior to hydrogen reionization, and we have neglected variations in the reheat temperature (e.g., due to spatial variation of the ionization parameter). The choice of modeling of these issues led to only minor changes in the simulations by Hirata (2018), so we did not consider them in this paper, but only a few alternative models were tested and more should be explored.

In conclusion, we have found that inhomogeneous hydrogen reionization results in an imprint on the Lyman-\( \alpha \) forest power spectrum, even at “low” redshifts \( 2 < z_{\text{obs}} < 4 \). The effect is present despite the “attractor” nature of the IGM temperature-density relation, because of the finite relaxation time and the low redshift of reionization favored by Planck. It can range from \( \lesssim 1\% \) at small scales and low redshifts, up to tens of percents in the large-scale 3D power spectrum at \( z_{\text{obs}} \gtrsim 3.5 \). While we have not yet developed a robust mitigation strategy, there are several clear paths forward on both the theory/simulation front, and with additional observations to help constrain reionization.

ACKNOWLEDGEMENTS

We thank Hy Trac, Tzu-Ching Chang, Jordi Miralda-Escudé, Lluís Mas-Ribas, Benjamin Buckman and Xiao Fang for useful discussions. PMC is grateful to Andrei Mesinger for fruitful suggestions with 21cmFAST. PMC and CMH are supported by the Simons Foundation, the US Department of Energy, the Packard Foundation, the NSF, and NASA. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC-0011726.

REFERENCES

Arinyo-i-Prats A., Miralda-Escudé J., Viel M., Cen R., 2015, J. Cosmology Astropart. Phys., 12, 017
Barkana R., Loeb A., 2005, ApJ, 624, L65
Bautista J. E., et al., 2017, A&A, 603, A12
Bechtold J., 1994, ApJS, 91, 1
Becker G. D., Bolton J. S., Haehnelt M. G., Sargent W. L. W., 2011, MNRAS, 410, 1096
Blomqvist M., et al., 2015, J. Cosmology Astropart. Phys., 11, 034
Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015, ApJ, 811, 140
Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M., 1994, ApJ, 437, L9
Cen R., McDonald P., Trac H., Loeb A., 2009, ApJ, 706, L164
Chabanier S., et al., 2018, arXiv e-prints, p. arXiv:1812.03554
Compostella M., Cantalupo S., Porciani C., 2013, MNRAS, 435, 3169
Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, ApJ, 495, 44
Croft R. A. C., Weinberg D. H., Pettini M., Hernquist L., Katz N., 1999, ApJ, 520, 1
Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, ApJ, 581, 20
Croft R. A. C., Banday A. J., Hernquist L., 2006, MNRAS, 369, 1090
D’Aloisio A., McQuinn M., Maupin O., Davies F. B., Trac H., Fuller S., Upton Sandebeck P. R., 2018, preprint, (arXiv:1807.00992)
DESI Collaboration et al., 2016, preprint, (arXiv:1511.00036)
Dawson K. S., et al., 2013, AJ, 145, 10
Delubac T., et al., 2015, A&A, 574, A59
Eisenstein D. J., et al., 2011, AJ, 142, 72

---

\(^7\) The bin size is \( \Delta z = 0.2 \) and \( \Delta k/k = 0.03 \), and corresponds to the second row of Table 4 in Chabanier et al. (2018).
