Searching for the shadows of giants – II. The effect of local ionization on the Ly α absorption signatures of protoclusters at redshift $z \sim 2.4$

Joel S. A. Miller [ID],⋆ James S. Bolton [ID] and Nina A. Hatch [ID]

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

ABSTRACT

Local variations in the intergalactic medium (IGM) neutral hydrogen fraction will affect the Ly α absorption signature of protoclusters identified in tomographic surveys. Using the IllustrisTNG simulations, we investigate how the AGN proximity effect and hot, collisionally ionized gas arising from gravitational infall and black hole feedback changes the Ly α absorption associated with $M_{z=0} \simeq 10^{14} \, M_\odot$ protoclusters at $z \simeq 2.4$. We find that protocluster galaxy overdensities exhibit a weak anticorrelation with Ly α transmission in IGM transmission maps, but local H I ionization enhancements due to hot $T > 10^6$ K gas or nearby AGN can disrupt this relationship within individual protoclusters. On average, however, we find that strong reductions in the IGM neutral fraction are limited to within $\lesssim 5h^{-1}$ cMpc of the dark matter haloes. Local ionization enhancements will therefore have a minimal impact on the completeness of protocluster identification in tomographic surveys if smoothing Ly α transmission maps over scales of $\sim 4h^{-1}$ cMpc, as is typically done in observations. However, if calibrating the relationship between the matter density and Ly α transmission in tomographic maps using simple analytical models for the Ly α forest opacity, the presence of hot gas around haloes can still result in systematically lower estimates of $M_{z=0}$ for the most massive protoclusters.

Key words: galaxies: clusters: general – intergalactic medium – quasars: absorption lines.

1 INTRODUCTION

A fundamental prediction of Lambda cold dark matter (ΛCDM) cosmogonies is that galaxy clusters are built from the assembly of lower mass haloes such as galaxy groups and isolated galaxies (e.g. White & Frenk 1991). At low redshifts, clusters are single dark matter haloes, filled with massive, evolved, early-type galaxies orbiting the brightest cluster galaxy. The progenitor of this halo at redshift $z > 2$ is a diffuse collection of smaller haloes spread over tens of comoving Mpc, all of which are rapidly growing and merging (Chiang, Overzier & Gebhardt 2013; Muldrew, Hatch & Cooke 2015). These ensembles of gravitationally bound but not yet virialized structures are known as protoclusters (Overzier 2016). They are the highest density regions in the early Universe and therefore are the most active sites of structure assembly.

Traditionally, protoclusters are located as galaxy overdensities in either photometric (Daddi et al. 2009; Chiang, Overzier & Gebhardt 2014) or spectroscopic (Steidel et al. 2005; Cucciati et al. 2014; Chiang et al. 2015; Lemaux et al. 2017; Harikane et al. 2019) surveys. Hundreds of protocluster candidates have now been identified as regions of high galaxy density on scales of a few to tens of arcminutes (Wylezalek et al. 2013; Toshikawa et al. 2018). A small but growing number of protoclusters have been selected not by their galaxy properties, however, but by their gaseous properties from the dominant baryonic component that lies in the intergalactic/intraprotocluster medium. X-ray and mm-wavelength observations – the latter via the Sunyaev–Zel'dovich effect (Sunyaev & Zeldovich 1972) – are

used to identify the $T \sim 10^7$ K intracluster medium within massive collapsed clusters and groups (Ebeling et al. 2010; Finoguenov et al. 2010; Bleem et al. 2015), leading to the detection of clusters up to redshifts of $z = 1.7$ (Strazzullo et al. 2019). Yet the majority of the gas by volume within protoclusters has not yet been shock heated to X-ray emitting temperatures and instead has $T \sim 10^4$ K (Miller, Bolton & Hatch 2019).

Fortunately, neutral hydrogen in the intergalactic medium (IGM) also traces the underlying dark matter structure closely on scales $\gtrsim 1h^{-1}$ cMpc (Croft et al. 2002; Viel, Haehnelt & Springel 2004). Protoclusters at redshift $z > 2$ have sizes of 10–$50 \, h^{-1}$ cMpc (Muldrew et al. 2015) so their large overdensities may be traced by neutral hydrogen. This neutral hydrogen can be detected in the spectra of background quasi-stellar objects (QSOs) as a series of Ly α absorption lines known as the Ly α forest (Rauch 1998). Coherent, large-scale decrements in the Ly α forest transmitted flux within individual QSO spectra may then correspond to intergalactic H I associated with significant mass overdensities. In particular, Cai et al. (2017) have used spectral regions with strong Ly α transmission decrements over a scale of $15h^{-1}$ cMpc – which they name Coherently Strong Ly α Absorption systems (CoSLAs) – to locate a protocluster at $z = 2.3$ (see also Cai et al. 2016; Shi et al. 2021; Zheng et al. 2021). However, using cosmological hydrodynamical simulations, Miller et al. (2019, hereafter Paper I) also showed that such CoSLAs are rare[1] and are not an exclusive probe of protoclusters. It is possible to adopt a rather strict (and model

---

[1]On analysing mock Ly α forest spectra drawn through the $(80h^{-1} \, \text{cMpc})^3$ Sherwood simulation volume (Bolton et al. 2017) with an average transverse

⋆ E-mail: joel.miller@nottingham.ac.uk

dependent) CoSLA detection threshold that removes contamination from coherent structures originating in the diffuse IGM, but any such protocluster sample is then incomplete.

An alternative technique for protocluster identification that also makes use of intergalactic Ly α absorption is IGM or Ly α forest tomography (Pichon et al. 2001; Caucci et al. 2008; Lee et al. 2014; Stark et al. 2015; Horowitz et al. 2019; Porqueres et al. 2020; Li, Horowitz & Cai 2021). If a sufficient number of individual Ly α forest sight-lines sample a given volume, a three-dimensional map of the Ly α transmission from the IGM may be reconstructed. A large sample of QSO sightlines can be used for this purpose (see e.g. Ravoux et al. 2020), but recent observational advances have also allowed spectra from background star-forming galaxies to be used, thus providing the high density of sight-lines needed to reconstruct the Ly α transmission on scales of a few comoving Mpc (Lee et al. 2016, 2018; Newman et al. 2020; Mukae et al. 2020a). This approach has been successfully used to locate dense structures in the early Universe – some of which are expected to be protoclusters. Furthermore, combining these new tomographic Ly α transmission maps with coeval galaxy surveys provides a powerful insight into the galaxy-IGM connection at $z > 2$, and can improve the accuracy of the tomographic reconstruction of the underlying density field (Mukae et al. 2017; Momose et al. 2021; Mukae et al. 2020b; Liang et al. 2021; Horowitz et al. 2021).

A key component in all of these recent IGM tomography studies are numerical simulations of the Ly α transmitted flux; these are used to translate the 3D reconstruction of the Ly α transmitted flux into the underlying matter density. The most common approach used to create simulated Ly α tomographic maps is to apply the fluctuating Gunn & Peterson (1965) approximation (FGPA) to the density field from large collisionless dark matter simulations (Stark et al. 2015; Lee et al. 2016; Newman et al. 2020). The FGPA assumes that the baryons trace the dark matter density field (modulo a correction for smoothing on the Jeans scale), that the neutral hydrogen is in photoionization equilibrium with a spatially uniform UV background, and there is a single, power-law relationship between the gas density and temperature (see e.g. Rauch 1998; Becker, Bolton & Lidz 2015). These assumptions are usually excellent ones when modelling the diffuse IGM at low densities, $\Delta = \rho/\langle\rho\rangle \lesssim 10$

It is well known, however, that neutral hydrogen (and hence also the Ly α transmission) is not a completely unbiased tracer of the underlying density field, particularly in highly overdense regions. More specifically, the FGPA will no longer hold for (i) gas that is hot and pre-dominantly collisionally ionized, either due to shocks from gravitational infall or energetic feedback from supernovae driven winds and/or black-hole accretion, (ii) high-density gas that is self-shielded to Lyman continuum photons, and (iii) local enhancements in the otherwise spatially uniform metagalactic UV background due to the presence of bright, rare sources (i.e. the proximity effect). Indeed, the presence of hot, highly ionized gas was suggested by Lee et al. (2016) as an explanation for the lack of a strong Ly α transmission decrement associated with a galaxy overdensity in their tomographic maps at $z \simeq 2.3$. Mukae et al. (2020a) also demonstrated a spatial offset of $\sim$3–5h^{-1} cMpc between Ly α emitting galaxies and the minimum Ly α transmission in their tomographic reconstruction around the MAMMOTH-1 nebula. These authors suggested that local fluctuations in the ionizing background may explain this offset, by changing the distribution of neutral hydrogen in the surrounding IGM

separation of $0.75h^{-1}$ cMpc, only $\sim$0.1 per cent of sight-lines exhibited a CoSLA (Paper I).

(see also Momose et al. 2021, for a similar result obtained from the cross-correlation of galaxies and Ly α tomographic maps).

In this paper, we investigate this issue of 'ionization bias' further using state-of-the-art hydrodynamical simulations from the IllustrisTNG project. We explore how local ionization variations in the IGM – either due to the presence of hot, collisionally ionized gas or the QSO proximity effect – impact on the detectability of protoclusters with Ly α forest tomography. Furthermore, we assess how these variations may affect the relationship between the Ly α transmission and the distribution of coeval galaxies, and how the assumption of the FPGA may bias constraints on protocluster mass. The goal of this work is not to test the efficacy of tomographic reconstruction techniques; this is already discussed in the literature in some detail (see e.g. Stark et al. 2015; Horowitz et al. 2019; Porqueres et al. 2020; Li et al. 2021). In this work, rather than using a full forward model, we instead create idealised, noiseless Ly α transmission maps by degrading our simulations to match the resolution of the tomographically reconstructed observations from Lee et al. (2018) and Newman et al. (2020). The advantage of this approach is that it allows us to isolate the effect of astrophysical systematics from any uncertainties associated with the reconstruction methodology.

In Section 2, we introduce the hydrodynamical simulations and local ionization models used throughout this work, and then examine the expected Ly α transmission profiles around dark matter haloes in Section 3. We discuss Ly α transmission maps of protoclusters and their relationship with coeval Ly α emitting galaxies in Section 4, and assess the role that local ionization variations may play in Ly α tomography measurements. Finally, we conclude in Section 5. Throughout this paper, we refer to comoving distance units using the prefix 'c' and to proper distance units using the prefix 'p'.

2 SIMULATING LY α ABSORPTION FROM PROTOCLUSTERS

2.1 Cosmological hydrodynamical simulations

In this work, we shall primarily use the publicly available TNG100-1 simulation from the IllustrisTNG collaboration (Nelson et al. 2019). IllustrisTNG has been performed using the moving-mesh hydrodynamics code AREPO (Springel 2010), and is described in detail in a series of five introductory papers (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018; Springel et al. 2018). We use three further IllustrisTNG models with different box sizes and mass resolutions (TNG100-2, TNG100-3, and TNG300-1) to assess the numerical convergence of our results (see the Appendix for further details).

In addition to the IllustrisTNG models, we also use the earlier Illustris-1 simulation (Vogelsberger et al. 2014; Nelson et al. 2015) to assess the effect of a different subgrid physics model on our results. The key differences between IllustrisTNG and Illustris-1 are summarized in table 2 of Nelson et al. (2019). These include changes to the stellar and AGN feedback implementations, and the addition of ideal magnetohydrodynamics in IllustrisTNG (Pakmor, Bauer & Springel 2011). There are also small differences in the ΛCDM cosmological parameters used in the two models. Importantly, however, the TNG100-1 initial conditions have the same random seed as Illustris-1, so we are able to directly compare the large-scale structure of intergalactic gas in these models.

All five of the simulations used in this work are summarized in Table 1. For each simulation, we use the snapshots and halo catalogues at $z = 2.44$ and $z = 0$.

Table 1. Hydrodynamical simulations used in this work. The columns list, from left to right: the simulation name, the box size in h^{-1} cMpc, the total number of gas cells and dark matter particles, and the typical dark matter particle and gas cell masses. The IllustrisTNG simulations assume a Planck Collaboration et al. (2015) consistent cosmology, with $\Omega_m = 0.3089$, $\Omega_\Lambda = 0.6911$, $\Omega_b = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$ and $h = 0.6774$. The cosmological parameters used in the Illustris-1 simulation instead take values consistent with WMAP-9 (Hinshaw et al. 2013), giving $\Omega_m = 0.2726$, $\Omega_\Lambda = 0.7274$, $\Omega_b = 0.0456$, $\sigma_8 = 0.809$, $n_s = 0.963$ and $h = 0.704$.

Name	Box size (h^{-1}cMpc)	N_{gas+DM}	M_{dm} ($M_\odot$)	M_{gas} ($M_\odot$)
TNG100-1	75	2×1820^3	7.50×10^6	1.40×10^6
Illustris-1	75	2×1820^3	6.26×10^6	1.26×10^6
TNG100-2	75	2×910^3	5.97×10^7	1.12×10^7
TNG100-3	75	2×455^3	4.78×10^8	8.92×10^7
TNG300-1	205	2×2500^3	5.90×10^7	1.10×10^7

2.2 Protocluster identification

Throughout this paper, we take advantage of the ability of simulations to connect physical structures at different instances in time. Following Paper I, we define protoclusters in the TNG100-1 model as the structures that form clusters with $M_{z=0} \geq 10^{14}$ $M_\odot$ at redshift $z = 0$. We identify the simulation resolution elements that belong to these protoclusters as being all those within friends-of-friends haloes with $M_{z=0} \geq 10^{14}$ $M_\odot$ at redshift $z = 0$. We then find these resolution elements at redshift $z = 2.44$ and use their positions to compute the centre of mass of each protocluster and the radial extent around the centre of mass, R_{95}, that contains 95 per cent of the protocluster's $z = 0$ mass. This procedure yields a total of 22 protoclusters in the TNG100-1 volume at $z = 2.44$.

An example of one such protocluster from the TNG100-1 simulation with $z = 0$ mass $M_{z=0} = 10^{14.46}$ $M_\odot$ is displayed in the upper left and central panels of Fig. 1. A 2D projection of the logarithm of the normalized gas density, $\Delta = \rho/\langle\rho\rangle$, and logarithm of the gas temperature, T, are shown within $\pm 1\,h^{-1}$ cMpc of the protocluster centre of mass. The white dashed circle in each panel shows the radial extent of the protocluster, R_{95}. As discussed in detail in Paper I, a wide range of protocluster morphologies are expected using our protocluster definition, where typically $R_{95} = 5-10 h^{-1}$ cMpc. On average, the gas in protoclusters will exhibit slightly higher densities, temperatures, and neutral hydrogen fractions compared to the surrounding IGM.

2.3 Local ionization models

The primary focus of this work is assessing the impact that local variations in the IGM ionization state may have on the identification of protoclusters using Ly α absorption. We now turn to describing the three different IGM ionization models we use for this purpose.

In our fiducial ionization model, we adopt a similar approach to Paper I and assume a spatially uniform UV background using the Faucher-Giguère (2020) synthesis model. For reference, the Faucher-Giguère (2020) model has an H I photoionization rate $\Gamma_{HI} = 9.76 \times 10^{-13}$ s^{-1} at $z = 2.44$, which is consistent with independent constraints on Γ_{HI} from the Ly α forest opacity (Becker & Bolton 2013). We calculate neutral hydrogen fractions in each cell of the simulation under the assumption of ionization equilibrium by using the coupled equations given by Katz, Weinberg & Hernquist (1996), after updating the recombination and collisional ionisation rates to match those used by Bolton et al. (2017). We also use the Rahmati et al. (2013) prescription for self-shielding to obtain the correct

incidence of absorbers that are optically thick to Lyman continuum photons (i.e. for $N_{HI} \geq 10^{17.2}$ cm^{-2}). We have already verified in Paper I (see fig. 1 in that work) that this procedure reproduces the shape of the observed H I column density distribution over the range 10^{12} cm$^{-2} \leq N_{HI} \leq 10^{22}$ cm^{-2} very well.

In addition to our fiducial model, we investigate two further, alternative ionization models. In the first we assume a spatially uniform UV background, but now ignore the effects of collisional ionization and self-shielding on the neutral hydrogen fraction. We achieve this by setting the collisional ionisation rates to zero and neglecting the Rahmati et al. (2013) correction when calculating the H I fractions in each gas cell of the hydrodynamical simulations. Collisional ionization will be particularly important for the ionization state of gas around haloes, where gas is heated to $T > 10^6$ K by gravitational infall and AGN or supernovae feedback (see e.g. the protocluster in the upper central panel of Fig. 1). Neglecting collisional ionization in these hot, dense regions will result in an overestimate of the H I fraction, and hence, an overestimate of the Ly α optical depth associated with the gas. By contrast, ignoring self-shielding will instead result in an underestimate of the number of rare, high-column density absorption systems with $N_{HI} \geq 10^{17.2}$ cm^{-2} that arise from cool, dense gas. We refer to this model as 'No Collisional' – shortened to *NoCol* – throughout this paper. The *NoCol* model is chosen to be similar (but not identical) to the fluctuating Gunn–Peterson approximation (FGPA) that has been commonly used in the recent literature to link the Ly α optical depth to the underlying gas or dark matter density (e.g. Stark et al. 2015; Newman et al. 2020). The FGPA assumes photoionization equilibrium in an IGM which follows a power-law temperature density relation, $T = T_0 \Delta^{\gamma-1}$, which is a good approximation only for gas with $\Delta \leq 10$ at $z \simeq 2$ (see e.g. Rauch 1998).

Our second alternative ionization model includes the effect of local enhancements in the IGM ionization state due to quasars and active galactic nuclei (AGNs; i.e. the proximity effect, Murdoch et al. 1986; Bajtlik, Duncan & Ostriker 1988). At the redshift we consider in this work, $z = 2.44$, the mean free path of Lyman continuum photons is ~ 300 pMpc (Worseck et al. 2014); on smaller scales the UV background is to a good approximation spatially uniform. However, the presence of AGNs in close proximity to protoclusters could mean the background photoionization rate is significantly enhanced on scales up to a few proper Mpc in the vicinity of the AGN.

We model the effect of a local enhancement in the ionization level of neutral hydrogen following the simple model described in Bolton & Viel (2011). We populate the TNG100-1 simulation with AGN at $z = 2.44$ by requiring the number of AGN in a comoving volume, V, satisfies

$$N_{AGN} = V \int_{L_{min}}^{\infty} \phi(L_{1450}) dL_{1450}, \tag{1}$$

where $\phi(L_{1450})$ is the AGN luminosity function from Kulkarni, Worseck & Hennawi (2019) at $z = 2.44$. We assume a minimum luminosity of $L_{min} = 10^{43.2}$ erg s^{-1}, corresponding to an absolute AB magnitude $M_{1450} = -18$. We assign a luminosity, L_{1450}, to each AGN by Monte Carlo sampling the luminosity function from Kulkarni et al. (2019), and then populate the simulation by assigning AGN to haloes in a one-to-one rank order fashion, such that the most luminous AGN resides in the most massive halo (i.e. we effectively assume an AGN duty cycle of one). This yields 281 AGN within the TNG100-1 volume, with a median $M_{1450} = -18.9$ and a minimum of $M_{1450} = -24.7$.
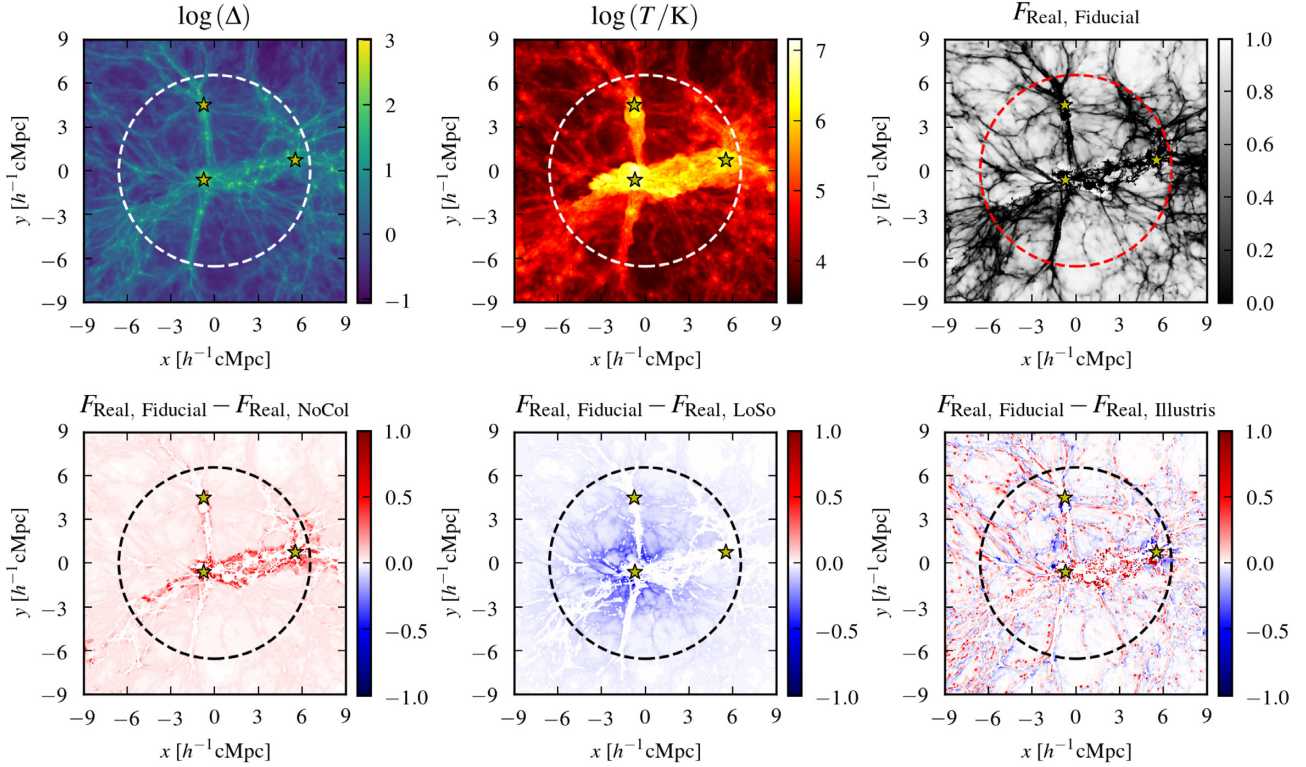
Figure 1. Top: A series of 2D projections of the gas overdensity (left), gas temperature (centre), and the real space Ly α forest transmission $F_{\rm real}$ (right, see the text for details) for a protocluster with $z = 0$ mass $M_{z=0} = 10^{14.46}$ M$_\odot$ in the TNG100-1 simulation at $z = 2.44$. The slices are projected over a distance of $2\,h^{-1}$ cMpc and are centred on the protocluster centre of mass. The dashed circle represents R_{95} for the protocluster, while the yellow stars denote the locations of haloes within the slice that are populated with AGN in our local source ionization model (see Section 2.3 for details). Bottom: Slices showing the difference in $F_{\rm real}$ between our fiducial model and a model with no collisional ionization or self-shielding (left), a model with local enhancements in the IGM ionization state due to the proximity effect from AGN (centre), and the Illustris-1 simulation which uses a different subgrid physics implementation compared to TNG100-1. Here, red represents a larger Ly α transmission (less absorption) in the fiducial model, while blue represents a smaller transmission (more absorption).

Next, for each AGN we assume the spectral energy distribution used by Kulkarni et al. (2019),

$$L(\nu) \propto \begin{cases} \nu^{-0.61} & (912\,\text{Å} < \lambda \leq 1450\,\text{Å}), \\ \nu^{-1.70} & (\lambda \leq 912\,\text{Å}). \end{cases} \tag{2}$$

We then compute the specific intensity, $J(\boldsymbol{r}, \nu)$, of the ionizing emission from the AGN on a 256^3 grid, assuming each AGN emits isotropically and that the IGM is optically thin within the periodic simulation volume. Hence,

$$J(\boldsymbol{r}, \nu) = \frac{1}{4\pi} \sum_{i=1}^{N_{\rm AGN}} \frac{L_i(\boldsymbol{r}, \nu)}{4\pi |\boldsymbol{r}_i - \boldsymbol{r}|^2}, \tag{3}$$

where $|\boldsymbol{r}_i - \boldsymbol{r}|$ is the distance of the ith AGN from $\boldsymbol{r}$. Finally, we compute the spatially varying photoionization rate from the AGN by evaluating

$$\Gamma_{\rm H\,I}(\boldsymbol{r}) = \int_{\nu_{\rm H\,I}}^{4\nu_{\rm H\,I}} \frac{4\pi J(\boldsymbol{r}, \nu)}{h_{\rm P}\nu} \sigma_{\rm H\,I}(\nu)\,{\rm d}\nu, \tag{4}$$

where $\sigma_{\rm H\,I}(\nu)$ is the photoionization cross-section from Verner et al. (1996) and $\nu_{\rm H\,I}$ is the frequency at the hydrogen Lyman limit. The photoionization rate for each gas cell is then obtained by trilinear interpolation of the nearest 256^3 grid points to the Voronoi cell centre. If the photoionization rate from equation (4) exceeds the value from the Faucher-Giguère (2020) synthesis model at $z = 2.44$ in any given cell, we use the former to calculate the ionization fraction.

Throughout this work, we shall refer to this as our 'local sources' – shortened to *LoSo* – model.

In Fig. 1, we perform an initial assessment of the effect of these ionization models on the average Ly α forest transmission. We consider a region of width $\Delta R = 2\,h^{-1}$ cMpc centred around the protocluster, and obtain an estimate of the real-space transmitted flux, $F_{\rm real}$, from the column density, $N_{\rm H\,I}$, in each pixel following a similar approach to Kulkarni et al. (2015), where

$$F_{\rm real} = \exp\left(-\frac{3\lambda_{\rm Ly\alpha}^3 \gamma_{\rm Ly\alpha}}{8\pi H(z)} \frac{N_{\rm H\,I}}{\Delta R} \right). \tag{5}$$

Here, $\gamma_{\rm Ly\alpha} = 6.265 \times 10^8\,{\rm s}^{-1}$ is the Ly α damping constant and $\lambda_{\rm Ly\alpha} = 1216$ Å. This approximation ignores the effect of peculiar velocities and thermal broadening on the Ly α opacity, and as a consequence it does not provide an accurate value for the average Ly α transmission along a given line of sight. However, it provides a convenient illustration of the relative effect of our ionization models on the Ly α opacity within protoclusters.

In the top right-hand panel of Fig. 1, we show $F_{\rm real}$ for the example TNG100-1 protocluster at $z = 2.44$ in the fiducial ionization model, whilst in the lower panels we show the difference in $F_{\rm real}$ between the fiducial model and the *NoCol* model (left), *LoSo* model (centre), and the same region in the Illustris-1 simulation (right). In the *NoCol* model, there is a decrease in the transmission from the filaments within the protocluster, leading to relatively more transmission in the fiducial model. The filaments (which are most apparent in the
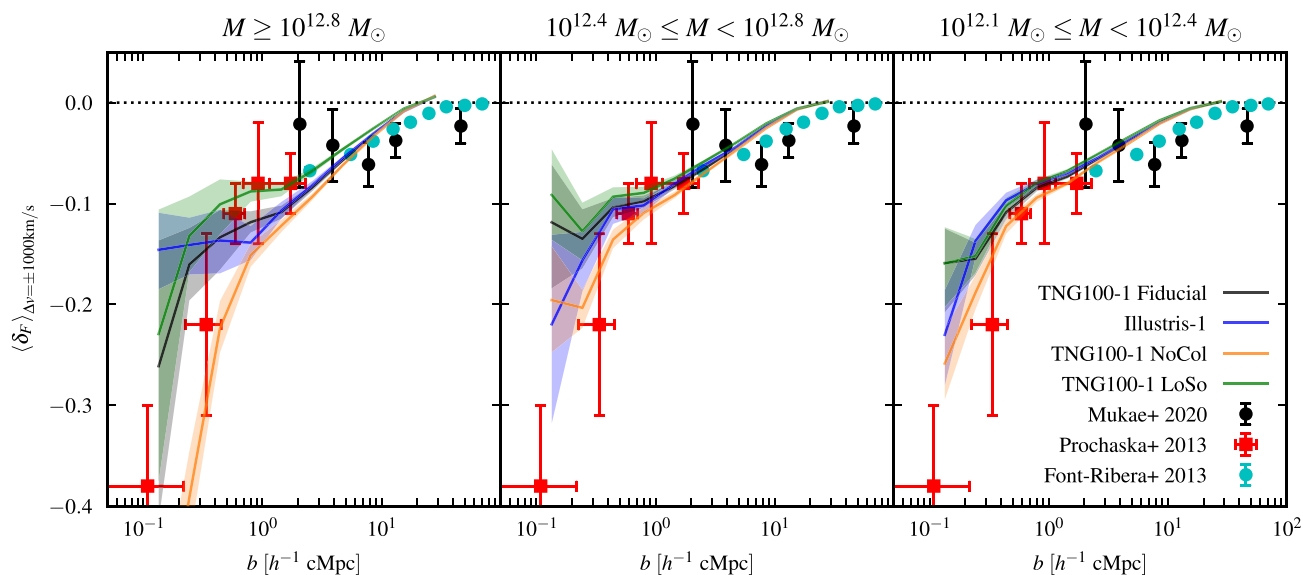
Figure 2. The median of the relative Ly α transmission, δ_F, averaged over a velocity window of width $\Delta v = 2000 \, \mathrm{km \, s^{-1}}$, as a function of impact parameter, b, around haloes in three mass bins at redshift $z = 2.44$: $M \geq 10^{12.8} \, \mathrm{M_\odot}$ (left), $10^{12.4} \, \mathrm{M_\odot} \leq M < 10^{12.8} \, \mathrm{M_\odot}$ (centre) and $10^{12.1} \, \mathrm{M_\odot} \leq M < 10^{12.4} \, \mathrm{M_\odot}$ (right). A negative (positive) value of δ_F corresponds to a decrease (increase) in the Ly α transmission relative to the mean transmitted flux, $\langle F \rangle$, of the IGM. The spatially uniform UV background model from the TNG100-1 simulation (black curves) is compared to the case ignoring collisional ionization and self-shielding (orange curves, *NoCol* model) and with the addition of a proximity effect due to AGN (green curves, *LoSo* model). The corresponding haloes in the Illustris-1 simulation are shown by the blue curves. The shaded regions bound 68 per cent of the distribution around the median, and are obtained by 10^3 bootstrap samples of the transmission profiles around each halo. For comparison, the data points and 1σ error bars display observational constraints on the Ly α transmission around QSO host haloes from Font-Ribera et al. (2013) (cyan circles), Prochaska et al. (2013) (red squares), and Mukae et al. (2020a) (black circles).

upper panels of Fig. 1) span up to $\sim 10 \, h^{-1} \, \mathrm{cMpc}$ in length, with widths on the order of $\sim 100 \, h^{-1} \, \mathrm{ckpc}$, and consist of overdense gas ($\Delta \sim 10$–100) at high temperatures ($T \sim 10^5$–$10^7 \, \mathrm{K}$). This corresponds to gas that has been heated by shocks and outflows and is therefore collisionally ionized in the fiducial model. Hence, we expect that ignoring hot, collisionally ionized gas will underestimate the Ly α transmission from the gas in protoclusters. By contrast, in the *LoSo* model there is an increase in the Ly α transmission in the protocluster relative to the fiducial model, with a magnitude that decreases radially and is generally more pronounced in cooler, less dense regions where photoionization dominates. Finally, comparing the different subgrid physics implementation used in Illustris-1 to the fiducial TNG100-1 model (see Section 2.1 for further details), we find the variation in transmission along the filaments of the cosmic web is more complex. Once again, these differences are driven primarily by changes in the thermal and ionization state of the hydrogen gas.[2] Note the transmission from the low-density IGM with $\Delta \lesssim 1$ remains unchanged; however, as the gas in voids is largely unaffected by shocks, AGN or supernovae driven winds at $z = 2.44$ (e.g. Theuns et al. 2002; Viel, Schaye & Booth 2013).

2.4 Mock Ly α absorption spectra

In the remainder of this work, we will analyse the Ly α absorption associated with protoclusters using simulated Ly α forest spectra. We again follow the procedure described in Paper I, which we briefly repeat here. Mock Ly α absorption spectra are extracted from the simulations by assigning each Voronoi cell a smoothing length, h_i,

based on the cell volume, V_i, such that

$$h_i = \left(\frac{3 N_{\mathrm{sph}} V_i}{4\pi} \right)^{1/3}. \tag{6}$$

We assume $N_{\mathrm{sph}} = 64$ for the number of smoothing neighbours. The interpolation scheme described by Theuns et al. (1998) is then used to extract Ly α optical depths using the Voigt profile approximation from Tepper-García (2006). Unless otherwise stated, we also rescale the optical depths of each pixel in our mock spectra by a constant to match observational constraints on the Ly α forest effective optical depth, $\tau_{\mathrm{eff}} = -\ln \langle F \rangle = 0.20$ at $z = 2.4$, from Becker et al. (2013).

The transmitted flux in each pixel is then given by $F = e^{-\tau}$, and we define the transmitted flux contrast, δ_F as the relative transmission – averaged over some velocity window of width Δv – around the IGM mean value

$$\delta_F = \frac{\langle F \rangle_{\Delta v}}{\langle F \rangle} - 1. \tag{7}$$

A negative (positive) value of δ_F thus represents a decrease (increase) in the Ly α transmission relative to the mean transmitted flux, $\langle F \rangle$, of the IGM.

3 LY α ABSORPTION PROFILES AROUND HALOES

We perform a consistency test of our mock Ly α absorption spectra in Fig. 2, where we show the transmitted flux contrast for our different ionization models around haloes in three mass bins: $M \geq 10^{12.8} \, \mathrm{M_\odot}$ (left), $10^{12.4} \, \mathrm{M_\odot} \leq M < 10^{12.8} \, \mathrm{M_\odot}$ (centre) and $10^{12.1} \, \mathrm{M_\odot} \leq M < 10^{12.4} \, \mathrm{M_\odot}$ (right). We select the mock spectra using a grid of sight-lines running the length of the simulation box in all three cardinal directions, with a mean transverse separation of $1.96 \, h^{-1} \, \mathrm{cMpc}$. We then calculate the mean transmission within

[2]This comparison is not exact, however, due to the slightly different cosmological parameters used in Illustris-1 and TNG-1 (see Table 1).

a velocity window of 2000 km s^{-1}, and bin the transmission in terms of the halo impact parameter, b. The results are compared to observational measurements of δ_F around QSOs from Mukae et al. (2020a), Prochaska et al. (2013), and Font-Ribera et al. (2013). Note that we display the Mukae et al. (2020a) MAMMOTH1-QSO measurements only on scales above the resolution limit of their Ly α tomographic maps. The Font-Ribera et al. (2013) data correspond to the Baryon Oscillation Spectroscopic Survey (BOSS) QSO-Lyα cross-correlation measurement, and have been converted to δ_F by Sorini et al. (2018).

Several earlier studies have already discussed the level of agreement between hydrodynamical simulations and observations of the neutral hydrogen distribution around QSOs (e.g. Fumagalli et al. 2014; Rahmati et al. 2015; Faucher-Giguère et al. 2016; Meiksin, Bolton & Puchwein 2017; Sorini, Davé & Anglés-Alcázar 2020; Nagamine et al. 2021). In general, differences in stellar and AGN feedback implementations, halo mass, and numerical resolution all play an important role. The differences we find here are consistent with earlier work, where the relative transmission at small scales, $b < 0.5h^{-1}$ cMpc, in Illustris-1 (blue curves) and TNG100-1 (black curves) is larger than the observed relative transmission. The relative difference between these two models, particularly in the $M \geq 10^{12.8}$ M$_\odot$ bin, is most likely associated with the more aggressive AGN feedback implementation within Illustris-1, which leads to more hot, collisionally ionized gas.

Recently, however, Sorini et al. (2020) have found very good agreement between the Prochaska et al. (2013) data and SIMBA simulations (Davé et al. 2019) on small scales, suggesting that the choice of stellar feedback model plays a key role in reproducing the observations (see also Faucher-Giguère et al. 2016). On larger scales ($b > 1h^{-1}$ cMpc), the level of agreement we find with the Font-Ribera et al. (2013) data is similar to Sorini et al. (2020), who suggest the overprediction of the relative transmission may be due to the small box size ($50h^{-1}$ cMpc) of the SIMBA simulation. We test this hypothesis by analysing the TNG-300-1 simulation in the Appendix, where we indeed find improved agreement with the Font-Ribera et al. (2013) observations for a larger box size of $300h^{-1}$ cMpc.

The main focus of this study are the differences caused by the various ionization models. In Fig. 2, we find these differences are largest for the highest mass haloes with $M \geq 10^{12.8}$ M$_\odot$. In general, the *NoCol* and *LoSo* models show less and more Ly α transmission relative to the fiducial TNG100-1 model, respectively. In the *NoCol* model (orange curves), this is due to neglecting collisional ionization from hot circumgalactic gas, where in general, the temperature and physical extent of the hot gas increases with halo mass. Interestingly, the *NoCol* model predicts too little transmission in the highest mass bin relative to the Prochaska et al. (2013) measurements, suggesting that collisional ionization (and hence gas temperature) plays an important role in setting the Ly α transmission at $b < 1h^{-1}$ cMpc (see also Sorini et al. 2018).

The increased transmission in the *LoSo* model (green curves) due to enhanced ionization by the proximity effect is also most pronounced in the $M \geq 10^{12.8}$ M$_\odot$ bin, as these haloes are populated with the highest luminosity AGN in our model.[3] Note, however,

that in contrast to the *NoCol* case (orange curves), the differences between the *LoSo* (green curves) and the fiducial model (black curves) are largest at $1h^{-1}$ cMpc $\leq b \leq 5h^{-1}$ cMpc. This is because the enhanced photoionization rate only begins to dominate over collisional ionization at $b \gtrsim 1h^{-1}$ cMpc. By contrast, the haloes in the lower two mass bins host either fainter AGN, or are unoccupied. As a result, the local source model does not have a significant effect on the Ly α transmission profiles for haloes with masses $M \lesssim 10^{12.8}$ M$_\odot$. Note, however, that we have deliberately adopted a model that maximises the proximity effect around the most massive haloes, and adopting a duty cycle $f_{duty} < 1$ (e.g. Shankar, Weinberg & Shen 2010) would push these AGN into lower mass hosts. Finally, in the $M \geq 10^{12.8}$ M$_\odot$ bin the *LoSo* model is in slightly better agreement with the Mukae et al. (2020a) data at $b < 5h^{-1}$ cMpc, although due to the large error bars the significance is not high. This appears to be consistent with the interpretation advanced by Mukae et al. (2020a) that the MAMMOTH1-QSO tomographic map exhibits a QSO proxmity zone.
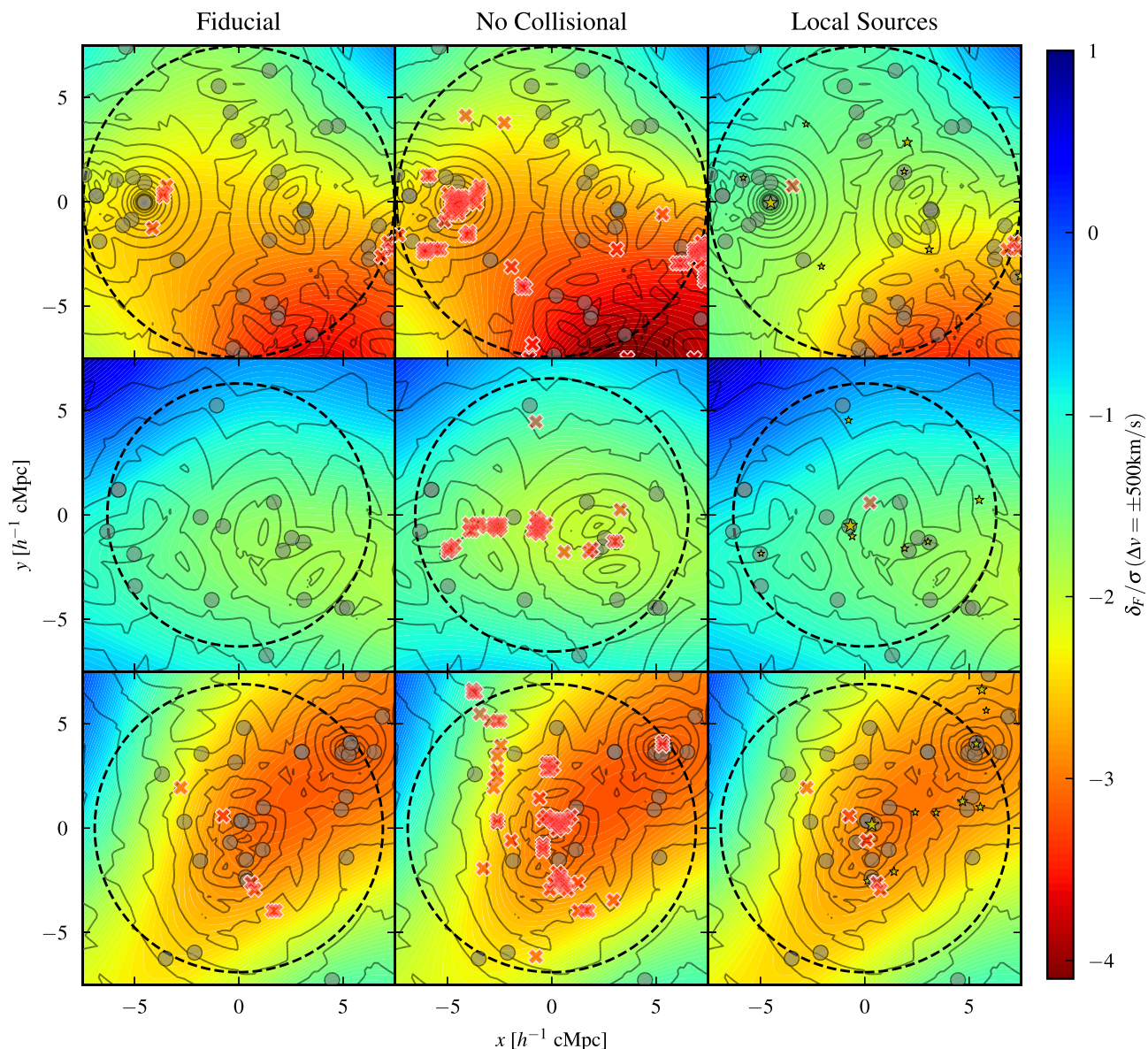
Since the *LoSo* and *NoCol* models effectively bracket the plausible range in the Ly α transmission profiles, we proceed to investigate the effect these models have on the Ly α transmission associated with protoclusters in TNG100-1. The different subgrid physics implementation in Illustris-1 sits between the extremes explored by these models, and so we do not investigate it further.

4 THE EFFECT OF LOCAL IONIZATION ON THE LY α TRANSMISSION AROUND PROTOCLUSTERS

4.1 Smoothed Ly α forest transmission maps

We now turn to investigate how the Ly α transmission around protoclusters is altered by changes in the local ionization state of the IGM. As already discussed, we do not create the Ly α transmission maps by forward modelling the observational data (e.g. Stark et al. 2015). Instead, we use the noiseless spectra drawn from the simulations to create idealized maps of the relative transmission, δ_F, around each of the 22 protoclusters in the TNG100-1 volume. We then degrade these maps to match the final resolution of the observational data presented by Lee et al. (2018) and Newman et al. (2020) by smoothing with a Gaussian filter. Our results will therefore not capture the effect of any systematic uncertainties associated with the accuracy of tomographic reconstruction techniques, or the signal-to-noise properties of the data.

We first extract spectra in a 90×90 grid in a 15 h^{-2} cMpc^2 area centred on each protocluster's centre of mass, following the procedure described in Section 2.4. We then construct Ly α transmission maps by obtaining the average Ly α transmission over velocity windows, $\Delta v = 1000$ km s^{-1} and then smoothing the relative transmission, δ_F, in the transverse direction using a Gaussian with standard deviation 4 h^{-1} cMpc. The velocity window, Δv, is chosen to match the full width at half maximum of the Gaussian filter at $z = 2.44$. This choice matches the transverse smoothing scale applied in the Ly α Tomography IMACS Survey (LATIS; Newman et al. 2020) and COSMOS Ly α Mapping and Observations survey (CLAMATO; Lee et al. 2018) tomographic surveys. Finally, we normalize each transmission map by the standard deviation of δ_F obtained from the full simulation volume. We obtain a standard deviation of $\sigma = 0.076$, $\sigma = 0.081$, and $\sigma = 0.070$ for the fiducial, *NoCol* and *LoSo* models, respectively. The standard deviation is slightly increased in the *NoCol* model with respect to fiducial, because ignoring collisional ionization decreases the neutral hydrogen fraction in

---

[3]If the AGN emission is preferentially beamed along the line of sight rather than in the transverse direction, our isotropic emission model will overestimate the impact of the proximity effect on the transmission profile. Similarly, non-equilibrium photoionization and light traveltime effects due to flickering AGN emission may also result in gas that is less highly ionized (e.g. Oppenheimer & Schaye 2013; Schmidt et al. 2019).

Figure 3. Smoothed transmission maps showing δ_F/σ for three protoclusters in the TNG100-1 simulation at redshift $z = 2.44$. Positive (negative) values of δ_F/σ correspond to Ly α transmission that is larger (smaller) relative to the average for the IGM. The maps are obtained by averaging the Ly α forest over a $\Delta v = 1000\,\mathrm{km\,s^{-1}}$ window, smoothing in the transverse direction with a Gaussian filter with standard deviation $4h^{-1}$ cMpc, and then centring on velocity windows $\Delta v = \pm 500\,\mathrm{km\,s^{-1}}$ where δ_F/σ is minimized within R_{95} for each protocluster. Each row shows a different protocluster for the fiducial (left), no collisional (centre), and local sources (right) ionization models. On top of each map, we display the locations of coeval Ly α emitting galaxies that satisfy the criteria $L_{\mathrm{Ly}\alpha} > 10^{41.5}\,\mathrm{erg\,s^{-1}}$ and $\mathrm{EW}_{\mathrm{Ly}\alpha} > 15$ Å (grey filled circles), using a simple empirical model (see the text for details). The grey contours show the logarithm of the LAE overdensity, $\log(1 + \delta_{\mathrm{LAE}}) = \log(\rho_{\mathrm{LAE}}/\langle \rho_{\mathrm{LAE}}\rangle)$, obtained from the distance to the fifth nearest neighbour in increments of 0.2 dex. In the right column, the yellow stars show the locations of AGN in the local sources model, with sizes scaled according to their luminosity. Red crosses display the locations of coherently strong Ly α absorption systems (CoSLAs). The dashed black circle shows the 2D cross-section of the sphere of radius R_{95} – the radius that contains 95 per cent of the $z = 0$ mass, $M_{z=0}$ – that intersects the velocity window for each protocluster. From top to bottom, the selected protoclusters have $M_{z=0} = 10^{14.18}\,\mathrm{M_\odot}$, $M_{z=0} = 10^{14.46}\,\mathrm{M_\odot}$, and $M_{z=0} = 10^{14.43}\,\mathrm{M_\odot}$. Note the protocluster in the middle row is also shown in Fig. 1.

dense, hot gas, thus increasing the amount of strong Ly α absorption. Conversely, the standard deviation is reduced relative to fiducial in our *LoSo* model, as the ionizing sources (AGN) are placed into high-density regions, thus reducing the incidence of strong Ly α absorption.

The resulting Ly α transmission maps for three different protoclusters are shown in each row of Fig. 3. The different local ionization models for the protoclusters are displayed in each column, with the yellow stars in the right column showing the location of coeval AGN

in the local sources model. The three protoclusters have been selected to show the region containing the most massive halo (and hence also the brightest AGN) in the TNG100-1 simulation (upper row, $M_{z=0} = 10^{14.18}\,\mathrm{M_\odot}$), a region where the Ly α transmission within R_{95} is higher than average (middle row, $M_{z=0} = 10^{14.46}\,\mathrm{M_\odot}$), and a region that is representative of the average Ly α transmission associated with a protocluster in TNG100-1 (lower row, $M_{z=0} = 10^{14.43}\,\mathrm{M_\odot}$). Note the protocluster in the middle row is also displayed in Fig. 1. The position of the maps are selected using the velocity window within

R_{95}, where δ_F/σ is minimized, similar to how these structures are identified within observed tomographic maps.

In each map, we also mark the locations of individual sightlines that contain coherently strong Ly α absorption systems (CoSLAs) using red crosses. Following Cai et al. (2017), CoSLAs are defined as sightlines that exhibit a fluctuation in the Ly α forest effective optical depth, $\delta_{\tau_{eff}} > 3.5$, over a scale of 15 h^{-1} cMpc, after excluding any Ly α absorbers with damping wings, $N_{H I} \geq 10^{19}$ cm^{-2} (see also Paper I for further details). This allows us to assess how CoSLAs are distributed relative to the Ly α transmission maps.

Lastly, we also use a simple model based on empirically derived scaling relations to display the locations of coeval Ly α emitting galaxies (grey circles). The Ly α luminosities and equivalent widths for the galaxies were estimated using the stellar mass and star formation rate (SFR) for each subhalo in TNG100-1. We convert the instantaneous SFR into an luminosity at 1216 Å using the relation from Dijkstra (2017) for a Salpeter (1955) initial mass function,

$$L_{Ly\alpha} = 1.0 \times 10^{42} \text{ erg s}^{-1} f_{esc}^{Ly\alpha} \left(\frac{\text{SFR}}{\text{M}_\odot \text{ yr}^{-1}} \right). \tag{8}$$

We have implicitly assumed a Lyman continuum escape fraction $f_{esc}^{LyC} \simeq 0$ in equation (8), and $f_{esc}^{Ly\alpha}$ is the volume averaged effective Lyα escape fraction inferred by Hayes et al. (2011),

$$f_{esc}^{Ly\alpha} = C_{Ly\alpha} 10^{-0.4 A_{Ly\alpha}}, \tag{9}$$

where $C_{Ly\alpha} = 0.445$ (Hayes et al. 2011). The quantity $A_{Ly\alpha}$ is derived using the relation between extinction at Hα wavelengths and stellar mass derived by Garn & Best (2010), and then converting to $A_{Ly\alpha}$ using the Calzetti et al. (2000) dust law. We also estimate the rest frame equivalent width in Angstroms, $EW_{Ly\alpha}$, for each Ly α emitter (LAE) using the relation $EW_{Ly\alpha} = f_{esc}^{Ly\alpha}/0.0048$ from Sobral & Matthee (2019). The LAEs displayed in the maps are selected by requiring $L_{Ly\alpha} > 10^{41.5}$ erg s^{-1} and $EW_{Ly\alpha} > 15$ Å (e.g. Shimakawa et al. 2017). The grey dashed contours correspond to the logarithm of the LAE overdensity, $\log(1 + \delta_{LAE}) = \log(\rho_{LAE}/\langle\rho_{LAE}\rangle)$, determined by a fifth nearest neighbour algorithm. We note, however, that this simple model does not include a self-consistent coupling between the visibility of the Ly α emission line and the Ly α opacity of the intervening circumgalactic medium (CGM) or IGM in the TNG100-1 simulation. It furthermore does not follow the complex Ly α radiative transfer within the interstellar medium of the galaxies (e.g. Laursen, Sommer-Larsen & Razoumov 2011; Gurung-López et al. 2020). As such, while the model is consistent with average LAE properties by design, it may still underestimate the variation in $L_{Ly\alpha}$ for a given stellar mass.

We first consider the fiducial ionization model, displayed in the left column of Fig. 3. There is an anticorrelation between the LAE density and δ_F/σ for all three protoclusters, and any CoSLAs are typically situated where the galaxy clustering is strongest. The maps generally exhibit a smaller δ_F/σ (less Ly α transmission), where the LAE density is largest. However, for the protocluster displayed in the top row of Fig. 3 there is an offset between where the LAEs are most strongly clustered around a massive halo with $M = 10^{13.5}$ M$_\odot$ at $(x, y) \simeq (-4, 0)h^{-1}$ cMpc and the largest Ly α transmission decrement at $(x, y) \simeq (3, -6)h^{-1}$ cMpc. This is qualitatively similar to the observation from Lee et al. (2016), where no strong Ly α transmission decrement was detected around a galaxy overdensity in their CLAMATO tomographic maps. These authors speculated that higher gas temperatures due to shocks or feedback may play a role in ionising gas and hence suppressing Ly α absorption in the vicinity of galaxy overdensities. This is indeed the case for the example here; the gas around the massive halo at $(x, y) \simeq (-4, 0)h^{-1}$ cMpc has

been heated to $T > 10^6$ K and is therefore highly ionised, whereas the IGM associated with the Ly α transmission decrement in the lower right of the map is significantly cooler, with $T < 10^5$ K. The protocluster displayed in the middle row (see also the same object in Fig. 1) exhibits more Ly α transmission compared to the other protoclusters for a similar reason; in addition to the presence of larger underdensities within R_{95} in this protocluster, there is an extended region of $T > 10^6$ K gas around the protocluster centre of mass that further increases the Ly α transmission.

In the central column of Fig. 3, we show the smoothed Ly α transmission maps for the same three protoclusters, but now using the *NoCol* ionization model. As expected, all three protoclusters exhibit smaller values of δ_F/σ where the LAE density is largest, but there is no significant change in δ_F/σ where the LAE density is lower. This is because hot, collisionally ionized gas is found around massive haloes and filaments, and this is the environment where most of the LAEs reside in our model. Another striking feature of the *NoCol* ionization models is that they contain a much higher incidence of CoSLAs (red crosses). There are two reasons for this. The first is that ignoring collisional ionization produces larger H I fractions, and hence stronger Ly α absorption. However, in the *NoCol* model we also neglect the effect of self-shielding to Lyman continuum photons on the Ly α absorption. This means that strong Ly α absorbers with column densities $N_{H I} > 10^{19}$ cm^{-2} in the fiducial model (i.e. damped systems) are overionized and have *lower* column densities in the *NoCol* model. Hence, damped absorption systems that are excised when selecting the CoSLA sample in the fiducial model are erroneously classified as lower column density CoSLAs in the *NoCol* model. As discussed in Paper I, this highlights the importance of correctly modelling high column density absorbers when simulating the incidence of coherent Ly α systems.

Finally, in the right column of Fig. 3 we show the smoothed Ly α forest transmission maps for the *LoSo* ionization model. The AGN positions are marked in the maps with star symbols. Due to the proximity effect, all three protoclusters exhibit larger δ_F/σ (more transmission) in comparison to the fiducial and *NoCol* models. The greatest increase in δ_F/σ occurs where the LAE density is largest, but there is also a small increase in δ_F/σ at lower densities. This can be further understood from the halo profiles in Fig. 2, where the brightest AGN in the model can ionize their surroundings up to $\sim 5 h^{-1}$ cMpc from the centre of their host halo. The largest proximity zone in the simulation volume is shown in the upper left-hand panel of Fig. 3, where the massive halo at $(x, y) \simeq (-4, 0)h^{-1}$ cMpc hosts an AGN with $M_{1450} = -24.7$. This further enhances the existing spatial offset between the largest LAE density and the weakest Ly α transmission/strongest Ly α absorption. A qualitatively similar observational result, but on much larger scales of $40h^{-1}$ cMpc, has been reported by Mukae et al. (2020a), who find an H I underdensity in the CLAMATO tomographic maps associated with a LAE overdensity. These authors suggest this is due to the enhanced ionization of the IGM by multiple nearby QSO proxmity regions.

Due to the increased level of ionization around massive haloes, the *LoSo* model also has a slightly reduced incidence of CoSLAs in comparison to the fiducial model. Interestingly, however, there are a few cases in which a CoSLA is present in the *LoSo* model but missing in the fiducial model. An example of this can be seen in the central region of the protocluster in the middle row of Fig. 3. This is the result of absorption systems that are classified as damped ($N_{H I} > 10^{19}$ cm^{-2}) in the fiducial model and are thus rejected when selecting CoSLAs, but instead correspond to lower column density absorbers in the *LoSo* model. The column densities of the damped
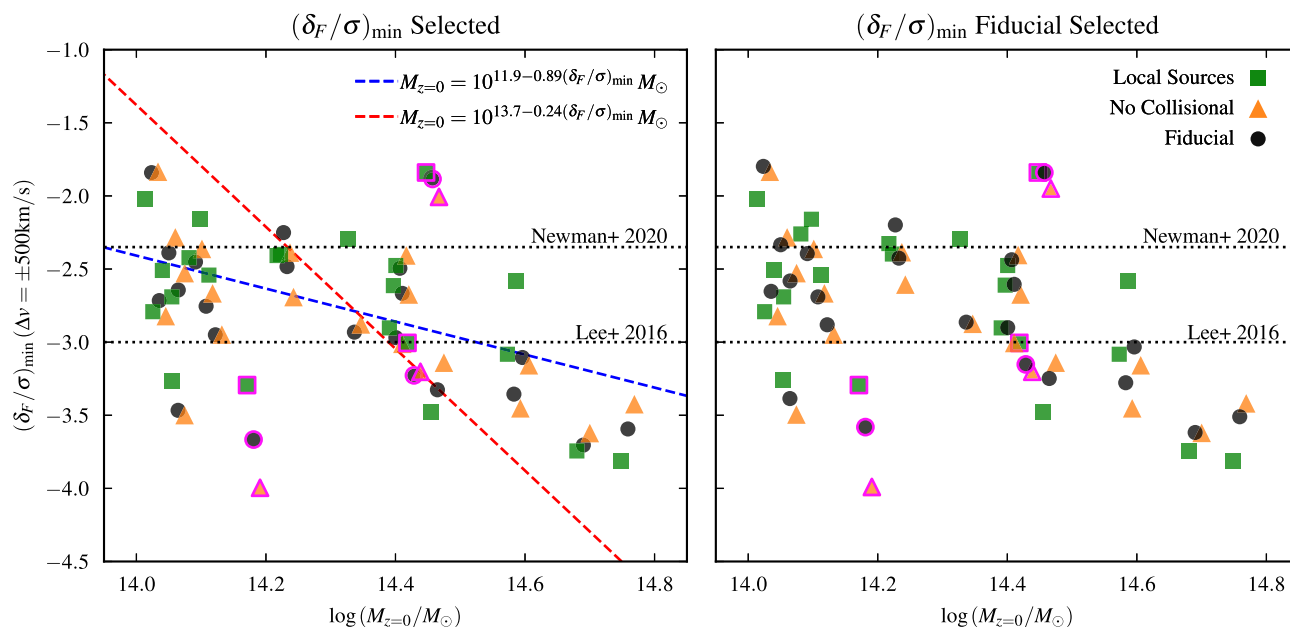
Figure 4. Scatter plot showing the relationship between the minimum relative Ly α transmission, $(\delta_F/\sigma)_{min}$, and the $z = 0$ cluster mass, $M_{z=0}$, in the smoothed transmission maps for the 22 protoclusters in the TNG100-1 volume at redshift $z = 2.44$. We show the results for our fiducial (black circles), *NoCol* (orange triangles), and *LoSo* (green squares) ionization models, where the *NoCol* and *LoSo* data points for each protocluster have been slightly offset on the horizontal axis for presentation purposes. The dotted lines show the thresholds used by Newman et al. (2020) and Lee et al. (2016) to select protoclusters from observed tomographic maps. These are at $(\delta_F/\sigma)_{min} = -2.35$ and -3, respectively. The data points outlined in fuchsia correspond to the three protoclusters shown in Fig. 3. The left-hand panel shows the $(\delta_F/\sigma)_{min}$ identified within R_{95} for each model. The right-hand panel instead shows $(\delta_F/\sigma)_{min}$ obtained when centring at the same location as $(\delta_F/\sigma)_{min}$ in the fiducial model, where we have also fixed $\sigma = \sigma_{fid}$ for all the ionization models. The red dashed line displays the relationship between $(\delta_F/\sigma)_{min}$ and $M_{z=0}$ obtained by Lee et al. (2016) from collisionless cosmological simulations post-processed with the fluctuating Gunn–Peterson approximation. The blue dashed line shows a linear best fit to the data from our fiducial model.

absorbers in the fiducial model are reduced due to the proximity effect, and these regions are then classified as CoSLAs.

### 4.2 Protocluster masses and the correlation between LAEs and Ly α transmission in smoothed maps

It is apparent from our qualitative discussion of the IGM transmission maps in Section 4.1 that local ionization plays an important role in the correlation between Ly α transmission, coeval galaxies and the distribution of coherent Ly α absorption systems within *individual* protoclusters. We now consider how local ionization variations impact on two quantities derived from tomographic maps: estimates of the $z = 0$ protocluster mass, $M_{z=0}$ (Lee et al. 2016; Newman et al. 2020) and the correlation between LAE overdensity and Ly α transmission (Mukae et al. 2017, 2020a; Liang et al. 2021).

We first examine the relationship between the minimum relative transmission within the protocluster, $(\delta_F/\sigma)_{min}$, and the $z = 0$ mass of the clusters, $M_{z=0}$, in Fig. 4. Candidate protoclusters are identified in both the CLAMATO and LATIS tomographic surveys by applying (δ_F/σ) thresholds to the smoothed Ly α transmission maps. Newman et al. (2020) define their matter overdensity/protocluster candidates as regions with $\delta_F/\sigma < -2.35$ in the LATIS survey, whilst for the CLAMATO survey Lee et al. (2016) use a more conservative value of $\delta_F/\sigma < -3$.

The left-hand panel of Fig. 4 shows $(\delta_F/\sigma)_{min}$ obtained from our smoothed transmission maps, centred on the velocity window containing $(\delta_F/\sigma)_{min}$ within R_{95} for each protocluster. We find the wide variety of protocluster morphologies (see also Paper I) means the minimum transmission can be located anywhere up to $\sim 10h^{-1}$ cMpc from the true protocluster centre of mass. For comparison, the right-hand panel shows $(\delta_F/\sigma)_{min}$ when the maps

for all models are instead centred at the location of $(\delta_F/\sigma)_{min}$ in the fiducial model. Additionally, in this case we use $\sigma = \sigma_{fid}$ for all three models. This allows us to focus on how the ionization models affect δ_F in the same physical region, as opposed to selecting δ_F/σ_{min} in a way that mimics the observations. On average, the changes in $(\delta_F/\sigma)_{min}$ for the different local ionization models are small in the left-hand panel of Fig. 4, with the largest differences occurring for the protocluster with $M_{z=0} = 10^{14.22} \, M_\odot$ that harbours the brightest AGN/most massive halo in the TNG100-1 volume (see the upper panels of Fig. 3). This suggests that smoothing Ly α tomographic maps on $\sim 4h^{-1}$ cMpc scales should help mitigate for the possible bias in inferred cluster masses due to local ionisation variations, as well as optimising protocluster detectability (Stark et al. 2015). This is furthermore consistent with our earlier finding that the largest differences in the local ionization models occur on scales $< 1h^{-1}$ cMpc (see Fig. 2). For comparison, we find that when selecting the same physical locations in the transmission maps and fixing $\sigma = \sigma_{fid}$ (right-hand panel), in almost all cases $(\delta_F/\sigma)_{min}$ is largest in the *LoSo* model and lowest in the *NoCol*, as one would naively expect. However, the differences between the ionization models are again modest for most protoclusters.

From the left-hand panel of Fig. 4, the protocluster completeness for the selection thresholds $(\delta_F/\sigma) < -2.35 \, (< -3)$ in the fiducial model is 86 (36) per cent, and this remains similar at 86 (41) per cent and 82 (32) per cent for both the *NoCol* and *LoSo* models, respectively. However, we find the best fit linear relation between $(\delta_F/\sigma)_{min}$ and $M_{z=0}$ is slightly shallower compared to the relationship obtained by Lee et al. (2016) from collisionless cosmological simulations (red dashed line in Fig. 4). Our best-fitting relation to the fiducial model is $M_{z=0} = 10^{11.9-0.89(\delta_F/\sigma)_{min}} \, M_\odot$, shown by the blue dashed line in the left-hand panel of Fig. 4. This implies that, for a given (δ_F/σ), our
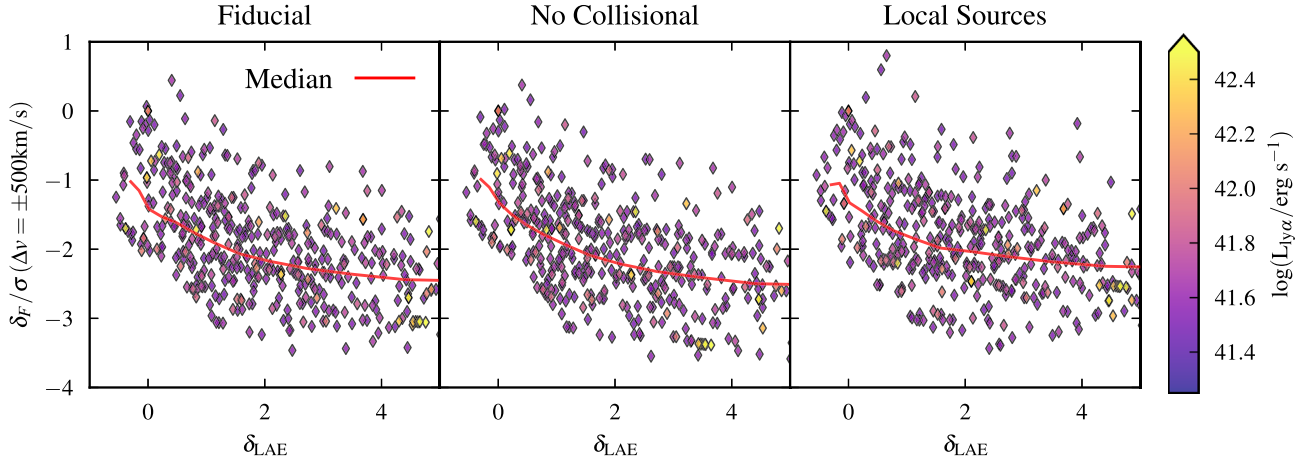
Figure 5. Scatter plot showing the relationship between the relative transmission, δ_F/σ, and the LAE overdensity, δ_{LAE}, in our smoothed transmission maps at redshift $z = 2.44$. From left to right, we show the results for our fiducial, no collisional and local source ionization models. The coloured diamonds correspond to LAEs with $L_{Ly\alpha} > 10^{41.5}$ erg s^{-1} and EW$_{Ly\alpha} > 15$ Å within R_{95} of all 22 protoclusters in the TNG100-1 volume. The colour scale shows the Ly α luminosity of the LAEs, and the red curve shows the median trend obtained by randomly sampling within R_{95} of the smoothed protocluster maps. There is a weak anti-correlation between δ_F and δ_{LAE} for $\delta_{LAE} \leq 1.5$.

fiducial model will favour larger $z = 0$ masses for the most massive candidate protoclusters compared to the Lee et al. (2016) calibration, possibly as a result of including hot gas with $T > 10^6$ K from shocks and AGN feedback [see also fig. 6 in Lee et al. (2016) and the related discussion]. We caution, however, that the relatively small TNG100-1 box means we also have a much smaller sample of $M_{z=0} > 10^{14}$ M$_\odot$ protoclusters compared to Lee et al. (2016), who use a collisionless dark matter simulation with box size $256h^{-1}$ cMpc. Note also that in this work we analyse idealized transmission maps, and we have not performed a tomographic reconstruction of the Ly α forest transmission using noisy data.

In the smoothed transmission maps in Fig. 3 we also observed an anticorrelation between the LAE overdensity, δ_{LAE}, and the relative transmission δ_F/σ. In Fig. 5, we examine this further by showing the relationship between δ_F/σ and δ_{LAE} for all 22 protoclusters in the TNG100-1 volume. The three different ionization models are shown in the individual panels. The filled diamonds correspond to the LAEs in the transmission maps centred on $(\delta_F/\sigma)_{min}$, while the red curve shows the median relation obtained by randomly sampling the maps. In all three models there is significant scatter in δ_F/σ at fixed δ_{LAE}, but the median trend shows decreasing δ_F/σ with increasing δ_{LAE} for $\delta_{LAE} \lesssim 1.5$. The Spearman's rank correlation coefficient for the LAEs with $\delta_{LAE} < 1.5$ is -0.4 in all three models, consistent with a weak anticorrelation. This is followed by a flattening at $\delta_{LAE} \gtrsim 1.5$ due to the $4h^{-1}$ cMpc Gaussian smoothing we apply to the transmission maps; we have verified that adopting a smaller smoothing scale reduces this apparent flattening and extends the anticorrelation to larger values of δ_{LAE}. As was the case in Fig. 4, the relative transmission in the *NoCol* and *LoSo* models typically decreases and increases, respectively, compared to the fiducial model. However, any changes remain very small compared to the scatter in the $\delta_F/\sigma-\delta_{LAE}$ plane, and are unimportant for the shape of the median trend. The colours of each point in Fig. 5 show the Ly α luminosity of the LAEs, which are selected using the criteria $L_{Ly\alpha} > 10^{41.5}$ erg s^{-1} and EW$_{Ly\alpha} > 15$ Å. There is no correlation (Spearman's rank coefficient -0.04 in all three models) apparent between δ_F/σ and the LAE luminosity, $L_{Ly\alpha}$ in our maps.

We may also compare the results in Fig. 5 to recent observational determinations of the relationship between δ_F and δ_{LAE} from Liang et al. (2021) (see also Mukae et al. 2017, 2020a; Momose et al.

2021, for closely related work), as well as the results from other cosmological hydrodynamical simulations (Nagamine et al. 2021). Note that these different studies do not calculate δ_{LAE} in the same way as this work, so a direct comparison with the results we present here is not possible. Nevertheless, we may still gain some insight from a qualitative comparison. Liang et al. (2021) identify LAEs at $z \sim 2.2$ from Subaru/Hyper Suprime-Cam data and compare the LAE overdensity to nearby Ly α absorbers in the Extended-BOSS database (Dawson et al. 2016). These authors do not use a tomographic reconstruction of the Ly-α forest, and instead compute δ_{LAE} and δ_F within cylindrical apertures. Assuming a best-fitting relation of $\delta_F = m\delta_{LAE} + C$, these authors find an anticorrelation with $m = -0.116^{+0.018}_{-0.022}$ and $C = -0.248^{+0.082}_{-0.093}$. Similarly, Nagamine et al. (2021) use the GADGET3-Osaka simulations to find a shallower relation with $m = -0.0664 \pm 0.00476$ and $C = -0.100 \pm 0.0006$, also obtained using a cylindrical aperture matched to the Liang et al. (2021) measurement. Although the slopes and normalization of the linear fits from these two studies differ, the relationship between δ_F and δ_{LAE} is qualitatively similar to the weak anti-correlation we observe at $\delta_{LAE} \lesssim 1.5$. This is consistent with the interpretation that LAEs are preferentially located in regions with increased Ly α absorption and hence larger H I densities at $z \simeq 2$–3.

Finally, although the relationship between δ_F and δ_{LAE} is not significantly altered in our different ionization models, we note that the visibility of Ly α emission lines and variations in the IGM/circumgalactic medium (CGM) Ly α transmission are closely coupled. As discussed previously, the volume averaged effective escape fraction we use, $f_{esc}^{Ly\alpha}$, does not self-consistently capture the effect of this coupling on δ_{LAE} in the TNG100-1 simulation. Detailed Ly α radiative transfer models that include the effect of both inflows and outflows in the CGM will be required to investigate the relationship between δ_F and δ_{LAE} further (e.g. Barnes et al. 2011; Laursen et al. 2011; Gurung-López et al. 2020).

5 CONCLUSIONS

In this work, we have investigated the effect that local ionization variations in the IGM, due the proximity effect from AGN and hot, $T > 10^6$ K gas from shocks and AGN feedback, have on the Ly α

absorption signature of protoclusters in the IllustrisTNG simulations at redshift $z \simeq 2.4$. We consider three different local ionization models in our analysis: a fiducial model with a spatially uniform UV background model, a second model that ignores the effect of collisional ionization and self-shielding on the H I fraction in the IGM, and final model where we incorporate spatial variations in the UV background due to the proxmity effect from AGN. The impact of 'ionization bias' on the Ly α transmission profiles around massive haloes and Ly α transmission maps is then investigated. We quantify this by computing the relative Ly α transmission averaged over a velocity window Δv, $\delta_F = (\langle F \rangle_{\Delta v}/\langle F \rangle) - 1$, where a negative (positive) value of δ_F represents a decrease (increase) in the Ly α transmission relative to the mean IGM transmitted flux, $\langle F \rangle$. We furthermore examine the relationship between the relative Ly α transmission in the smoothed transmission maps and the distribution of coeval Ly α emitting galaxies (LAEs) and coherently strong Ly α absorption systems (CoSLAs). Our main conclusions are as follows:

(i) We find local ionization effects have a significant impact on Ly α absorption in the vicinity of massive dark matter haloes with $M \geq 10^{12.8}\,\mathrm{M_\odot}$ for impact parameters $b \lesssim 1\,h^{-1}$ cMpc (see also Sorini et al. 2018). In particular, the presence of hot ($T > 10^6$ K) collisionally ionized gas will strongly increase δ_F within $\sim 1\,h^{-1}$ cMpc of dark matter haloes. We furthermore find that the proximity effect associated with AGN (which have absolute magnitudes in the range $-24.7 \leq M_{1450} \leq -18.9$ in our model) results in a modest increase in δ_F for impact parameters, $1\,h^{-1}$ cMpc $\leq b \leq 5\,h^{-1}$ cMpc, corresponding to distances where the photoionization of the IGM begins to dominate over collisional ionization. However, both of these effects become less important on larger scales, $b \gtrsim 5\,h^{-1}$ cMpc, and around less massive haloes with $M < 10^{12.8}\,\mathrm{M_\odot}$ in our model.

(ii) We construct idealized mock Ly α transmission maps around the 22 protoclusters with $M_{z=0} \geq 10^{14}\,\mathrm{M_\odot}$ in the TNG100-1 volume (cf. Lee et al. 2018; Newman et al. 2020). We find that local ionization effects can play an important role in the correlation between the Ly α transmission, coeval galaxies, and CoSLAs within a small number of individual protoclusters. In particular, we find a spatial offset of $\sim 9\,h^{-1}$ cMpc between a LAE overdensity around a massive halo and the largest Ly α flux decrement in a protocluster with $M_{z=0} = 10^{14.18}\,\mathrm{M_\odot}$. This offset is due to collisionally ionised gas with temperature $T > 10^6$ K surrounding the halo associated with the LAE density peak. This is qualitatively similar to the galaxy–Ly α absorption offset observed by Lee et al. (2016) and Mukae et al. (2020a) in CLAMATO tomographic maps. The transmission contrast of this spatial offset is further enhanced by the proximity effect associated with a $M_{1450} = -24.7$ AGN hosted within the halo. We furthermore find that the incidence of CoSLAs within protoclusters is sensitive to changes in our local ionization models, largely as a result of changes in the number of self-shielded, damped Ly α absorbers with $N_{\mathrm{H\,I}} \geq 10^{19}$ cm^{-2} (see also Miller et al. 2019).

(iii) After smoothing the simulated Ly α transmission maps with a Gaussian of standard deviation $4\,h^{-1}$ cMpc (Lee et al. 2018; Newman et al. 2020), we find that local ionization effects have a rather limited impact on the completeness of protocluster identification if using a fixed identification threshold of $(\delta_F/\sigma)_{\min} \leq -2.35$ (Newman et al. 2020) or $(\delta_F/\sigma)_{\min} \leq -3.00$ (Lee et al. 2016). For an ensemble of 22 protoclusters drawn from the TNG100-1 volume, we obtain a completeness of 82–86 per cent and 32–41 per cent, respectively for these thresholds if applied across all three of our ionization models. These results suggest that, in addition to optimizing protocluster detection (Stark et al. 2015), smoothing the Ly α tomographic maps on $4\,h^{-1}$ cMpc scales may also help mitigate for a possible 'ionization bias' in the completeness of a statistical sample of protoclusters.

Within our model, this is because the largest differences in the Ly α forest transmission typically occur on scales $< 4\,h^{-1}$ cMpc around dark matter haloes. However, we also find the presence of hot gas around haloes may still result in systematically lower estimates of $M_{z=0}$ for the most massive protoclusters if calibrating against mock tomographic Ly α maps created using the fluctuating Gunn–Peterson approximation. We find $M_{z=0} = 10^{11.9-0.89(\delta_F/\sigma)_{\min}}\,\mathrm{M_\odot}$ for the 22 protoclusters in our fiducial model.

(iv) A simple model that uses empirically derived scaling relations for the volume averaged effective Ly-α escape fraction (Hayes et al. 2011) and the Ly α rest-frame equivalent width (Sobral & Matthee 2019) is used to populate the IGM transmission maps with Ly α emitting galaxies. In qualitative agreement with recent results from observations (Mukae et al. 2017; Liang et al. 2021) and cosmological hydrodynamical simulations (Nagamine et al. 2021), we observe a modest anticorrelation (Spearman's rank correlation coefficient of ~ -0.4) between δ_F and the LAE emitter overdensity, δ_{LAE} for all three of our ionization models at $\delta_{\mathrm{LAE}} \lesssim 1.5$. This is consistent with these galaxies being preferentially located in overdense regions which exhibit smaller δ_F (i.e. stronger Ly α absorption) at $z \simeq 2.4$ compared to the average IGM value.

There remains plenty of scope for improving upon the numerical modelling in this work. In particular, the dynamic range of the hydrodynamical simulations should ideally be larger. A mass resolution of $M_{\mathrm{gas}} \sim 10^6\,M_\odot$ is required to resolve Ly α absorption from the IGM at $z \simeq 2$ (Bolton & Becker 2009; Miller et al. 2019). While the Ly α forest in the TNG100-1 simulation is therefore well resolved, the statistics are somewhat limited with only 22 protoclusters with $M_{z=0} \geq 10^{14}\,\mathrm{M_\odot}$. Additionally, the lack of any $M_{z=0} \geq 10^{15}\,\mathrm{M_\odot}$ clusters in the TNG100-1 simulation means that we are unable to study the impact of local ionization effects on the most massive structures. In our local sources model we have furthermore assumed isotropic AGN emission, and have ignored non-equilibrium ionization and light traveltime effects (e.g. Schmidt et al. 2019). Finally, we have adopted a simple model for the distribution of LAEs in the transmission maps that does not include the effect of the local IGM opacity on Ly α emitter visibility. Simulations of Ly α radiative transfer through the interstellar and circumgalactic/intergalactic medium will be required to address this question further (e.g. Barnes et al. 2011; Laursen et al. 2011; Gurung-López et al. 2020).

In summary, models that incorporate all of these physical effects will be important for fully unravelling the relationship between H I gas density and galaxies from Ly α tomographic surveys at $z \gtrsim 2$. Encouragingly, however, our results confirm that the identification and completeness of $M_{z=0} \simeq 10^{14}\,\mathrm{M_\odot}$ protoclusters identified from Ly α forest transmission maps smoothed on scales $\gtrsim 4\,h^{-1}$ cMpc should not be strongly affected by the variations in the local ionization state of the IGM at $z \simeq 2.4$.

DATA AVAILABILITY

All data and analysis code used in this work are available from the first author on reasonable request. An open access pre-print of the manuscript is available at arXiv.org

REFERENCES

Bajtlik S., Duncan R. C., Ostriker J. P., 1988, ApJ, 327, 570
Barnes L. A., Haehnelt M. G., Tescari E., Viel M., 2011, MNRAS, 416, 1723
Becker G. D., Bolton J. S., 2013, MNRAS, 436, 1023
Becker G. D., Hewett P. C., Worseck G., Prochaska J. X., 2013, MNRAS, 430, 2067
Becker G. D., Bolton J. S., Lidz A., 2015, PASA, 32, e045
Bleem L. E. et al., 2015, ApJS, 216, 27
Bolton J. S., Becker G. D., 2009, MNRAS, 398, L26
Bolton J. S., Viel M., 2011, MNRAS, 414, 241
Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T. S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897
Cai Z. et al., 2016, ApJ, 833, 135
Cai Z. et al., 2017, ApJ, 839, 131
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Caucci S., Colombi S., Pichon C., Rollinde E., Petitjean P., Sousbie T., 2008, MNRAS, 386, 211
Chiang Y. K. et al., 2015, ApJ, 808, 37
Chiang Y. K., Overzier R., Gebhardt K., 2013, ApJ, 779, 127
Chiang Y. K., Overzier R., Gebhardt K., 2014, ApJ, 782, L3
Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, ApJ, 581, 20
Cucciati O. et al., 2014, A&A, 570, A16
Daddi E. et al., 2009, ApJS, 694, 1517
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Dawson K. S. et al., 2016, AJ, 151, 44
Dijkstra M., 2017, Lyman-alpha as an Astrophysical and Cosmological Tool, Springer-Verlag, Berlin Heidelberg, p. 1
Ebeling H., Edge A. C., Mantz A., Barrett E., Henry J. P., Ma C. J., Van Speybroeck L., 2010, MNRAS, 407, 83
Faucher-Giguère C.-A., 2020, MNRAS, 493, 1614
Faucher-Giguère C.-A., Feldmann R., Quataert E., Kereš D., Hopkins P. F., Murray N., 2016, MNRAS, 461, L32
Finoguenov A. et al., 2010, MNRAS, 403, 2063
Font-Ribera A. et al., 2013, J. Cosmol. Astropart. Phys., 2013, 018
Fumagalli M., Hennawi J. F., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J., 2014, ApJ, 780, 74
Garn T., Best P. N., 2010, MNRAS, 409, 421
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Gurung-López S., Orsi Á. A., Bonoli S., Padilla N., Lacey C. G., Baugh C. M., 2020, MNRAS, 491, 3266
Harikane Y. et al., 2019, ApJ, 883, 142
Hayes M., Schaerer D., Stlin G., Mas-Hesse J. M., Atek H., Kunth D., 2011, ApJS, 730, 8
Hinshaw G. et al., 2013, ApJS, 208, 19
Horowitz B., Lee K.-G., White M., Krolewski A., Ata M., 2019, ApJ, 887, 61
Horowitz B., Zhang B., Lee K.-G., Kooistra R., 2021, ApJ, 906, 110
Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19
Kulkarni G., Hennawi J. F., Oñorbe J., Rorai A., Springel V., 2015, ApJ, 812, 30
Kulkarni G., Worseck G., Hennawi J. F., 2019, MNRAS, 488, 1035
Laursen P., Sommer-Larsen J., Razoumov A. O., 2011, ApJ, 728, 52
Lee K. G., Hennawi J. F., White M., Croft R. A., Ozbek M., 2014, ApJS, 788, 49
Lee K.-G. et al., 2016, ApJ, 817, 160
Lee K.-G. et al., 2018, ApJS, 238

Lemaux B. C. et al., 2017, A&A, 615, 77
Li Z., Horowitz B., Cai Z., 2021, ApJ, 916, 20
Liang Y. et al., 2021, ApJ, 907, 3
Marinacci F. et al., 2018, MNRAS, 480, 5113
Meiksin A., Bolton J. S., Puchwein E., 2017, MNRAS, 468, 1893
Miller J. S., Bolton J. S., Hatch N., 2019, MNRAS, 489, 5381 (Paper I)
Momose R. et al., 2021, ApJ, 909, 117
Mukae S. et al., 2017, ApJ, 835, 281
Mukae S. et al., 2020a, ApJ, 896, 45
Mukae S. et al., 2020b, ApJ, 903, 24
Muldrew S. I., Hatch N. A., Cooke E. A., 2015, MNRAS, 452, 2528
Murdoch H. S., Hunstead R. W., Pettini M., Blades J. C., 1986, ApJ, 309, 19
Nagamine K. et al., 2021, ApJ, 914, 66
Naiman J. P. et al., 2018, MNRAS, 477, 1206
Nelson D. et al., 2015, Astron. Comput., 13, 12
Nelson D. et al., 2018, MNRAS, 475, 624
Nelson D. et al., 2019, Comput. Astrophys.Cosmol., 6, 2
Newman A. B. et al., 2020, ApJ, 891, 147
Oppenheimer B. D., Schaye J., 2013, MNRAS, 434, 1063
Overzier R. A., 2016, A&A Rev., 24, 14
Pakmor R., Bauer A., Springel V., 2011, MNRAS, 418, 1392
Pichon C., Vergely J. L., Rollinde E., Colombi S., Petitjean P., 2001, MNRAS, 326, 597
Pillepich A. et al., 2018, MNRAS, 475, 648
Planck Collaboration et al., 2015, A&A, 594, A13
Porqueres N., Hahn O., Jasche J., Lavaux G., 2020, A&A, 642, A139
Prochaska J. X. et al., 2013, ApJS, 776, 136
Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, MNRAS, 430, 2427
Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, MNRAS, 452, 2034
Rauch M., 1998, ARA&A, 36, 267
Ravoux C. et al., 2020, J. Cosmol. Astropart. Phys., 2020, 010
Salpeter E. E., 1955, ApJ, 121, 161
Schmidt T. M., Hennawi J. F., Lee K.-G., Lukić Z., Oñorbe J., White M., 2019, ApJ, 882, 165
Shankar F., Weinberg D. H., Shen Y., 2010, MNRAS, 406, 1959
Shi D., Cai Z., Fan X., Zheng X., Huang Y.-H., Xu J., 2021, ApJ, 915, 32
Shimakawa R. et al., 2017, MNRAS, 468, 1123
Sobral D., Matthee J., 2019, A&A, 623, A157
Sorini D., Oñorbe J., Hennawi J. F., Lukić Z., 2018, ApJ, 859, 125
Sorini D., Davé R., Anglés-Alcázar D., 2020, MNRAS, 499, 2760
Springel V. et al., 2018, MNRAS, 475, 676
Springel V., 2010, MNRAS, 401, 791
Stark C. W., White M., Lee K. G., Hennawi J. F., 2015, MNRAS, 453, 311
Steidel C. C., Adelberger K. L., Shapley A. E., Erb D. K., Reddy N. A., Pettini M., 2005, ApJ, 626, 44
Strazzullo V. et al., 2019, A&A, 622, A117
Sunyaev R. A., Zeldovich Y. B., 1972, Comments Astrophys. Space Phys., 4, 173
Tepper-García T., 2006, MNRAS, 369, 2025
Theuns T., Leonard A., Efstathiou G., Pearce F. R., Thomas P. A., 1998, MNRAS, 301, 478
Theuns T., Viel M., Kay S., Schaye J., Carswell R. F., Tzanavaris P., 2002, ApJ, 578, L5
Toshikawa J. et al., 2018, PASJ, 70, 12
Verner D. A., Ferland G. J., Korista K. T., Yakovlev D. G., 1996, ApJ, 465, 487
Viel M., Haehnelt M. G., Springel V., 2004, MNRAS, 354, 684
Viel M., Schaye J., Booth C. M., 2013, MNRAS, 429, 1734
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Worseck G. et al., 2014, MNRAS, 445, 1745
Wylezalek D. et al., 2013, ApJ, 769, 79
Zheng X. Z., Cai Z., An F. X., Fan X., Shi D. D., 2021, MNRAS, 500, 4354

APPENDIX: THE EFFECT OF BOX SIZE AND MASS RESOLUTION ON LY α ABSORPTION AROUND HALOES

Following on from Fig. 2 and the associated discussion in Section 3, the effect of simulation mass resolution and box size on the transmission profiles around haloes are shown in Figs A1 and A2. Here, in addition to the fidicial TNG100-1 model, we use the publicly available TNG100-2, TNG100-3, and TNG300-1 simulations. The properties of these additional simulations are outlined in Table 1. There is generally very good agreement between the different simulations in Fig. A1, suggesting that our results should be sufficiently converged with respect to mass resolution. However, for the simulations with varying box sizes in Fig. A2, we observe

larger differences. In the case of the highest mass bin (left-hand panel), there is a divergence between the two simulations at impact parameters $b < 1\,h^{-1}$ cMpc. This is caused by the larger number of massive haloes with $M \geq 10^{12.8}\,M_\odot$ present in the TNG300-1 simulation, many of which are surrounded by hot $T > 10^6$ K gas with correspondingly low H I fractions. This explanation is consistent with the fact that in both of the lower mass bins we observe a very good agreement between the two simulations. On larger scales ($b > 1 h^{-1}$ cMpc), the level of agreement with the Font-Ribera et al. (2013) data is improved for the TNG300-1 model, particularly for the largest halo mass bin. This appears to be consistent with the suggestion by Sorini et al. (2020) that smaller volumes lacking the most massive haloes may predict transmission that is systematically above the Font-Ribera et al. (2013) measurements.
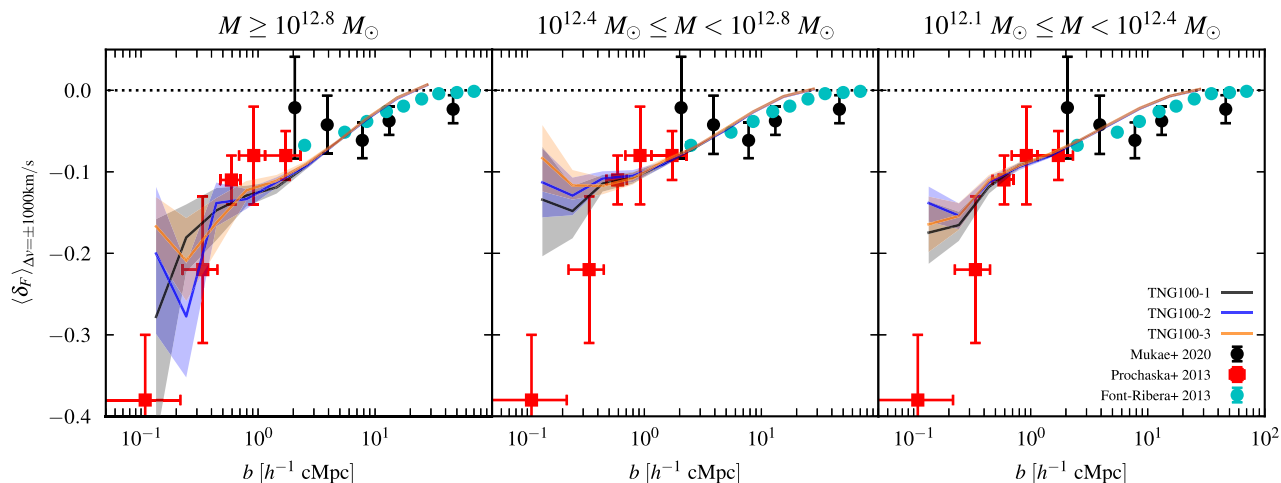
Figure A1. As for Fig. 2, but comparing simulations with fixed box size and different mass resolutions. The fidicial TNG100-1 simulation (black curves) is compared to the TNG100-2 (blue curves) and TNG100-3 (orange curves) simulations. These have dark matter particle masses a factor of 8 and 64 times larger than the TNG100-1 simulation, respectively. The transmission profiles are consistent within the 68 per cent scatter around the median, shown by the shaded regions.
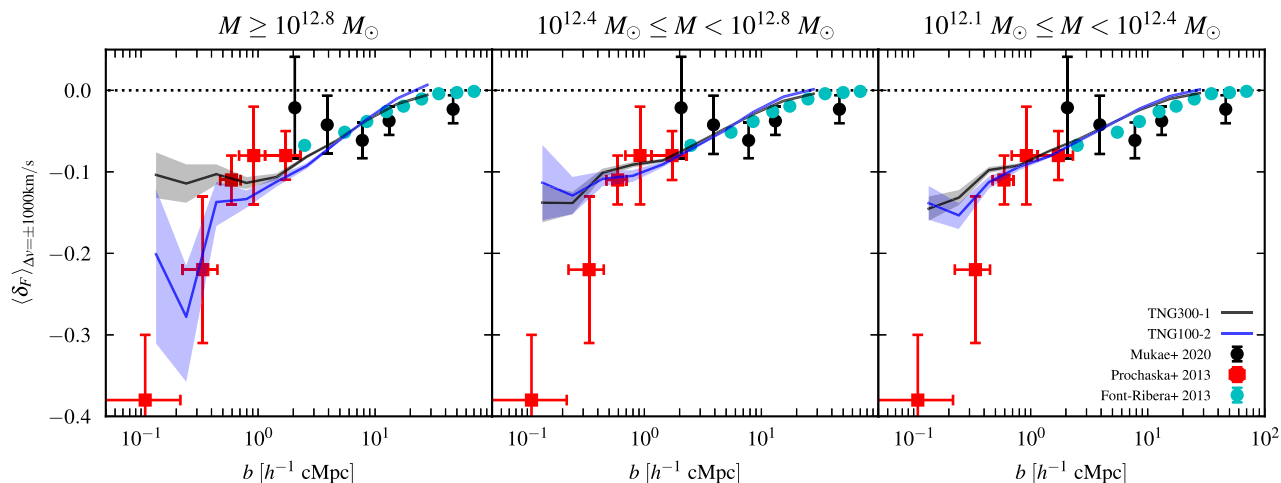
Figure A2. As for Fig. 2, but comparing simulations with different box sizes and very similar mass resolutions. The TNG100-2 simulation (black curves) uses the same box size as the fidicial TNG100-1 model, whereas the TNG300-1 simulation (blue curves) has a volume 27 times larger. The TNG300-1 simulation is in slightly better agreement with the observational measurements from Font-Ribera et al. (2013) on large scales, particularly for the highest halo mass bin in the left-hand panel. The TNG300-1 model also exhibits more transmission at impact parameters $b < 1h^{-1}$ cMpc around the most massive haloes.

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