Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **532,** 4876–4888 (2024) https://doi.org/10.1093/mnras/stae1830
Advance Access publication 2024 July 29

The effect of AGN feedback on the Lyman-α forest signature of galaxy protoclusters at $z \sim 2.3$

Chenze Dong [1,2]★ Khee-Gan Lee [1,2] Weiguang Cui [3,4] Romeel Davé [3,5] and Daniele Sorini [6]

[1] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*
[2] *Center for Data-Driven Discovery, Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*
[3] *Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[4] *Departamento de Física Teórica and CIAFF, Modulo 8 Universidad Autónoma de Madrid, E-28049 Madrid, Spain*
[5] *Department of Physics and Astronomy, University of the Western Cape, Bellville, Cape Town 7535, South Africa*
[6] *Institute for Computational Cosmology, Durham University, South Park Road, Durham DH1 3LE, UK*

Accepted 2024 July 25. Received 2024 June 22; in original form 2024 February 21

ABSTRACT

The intergalactic medium in the vicinity of galaxy protoclusters are interesting testbeds to study complex baryonic effects such as gravitational shocks and feedback. Here, we utilize hydrodynamical simulations from the SIMBA and The Three Hundred suites to study the mechanisms influencing large-scale Lyman-α transmission in $2 < z < 2.5$ protoclusters. We focus on the matter overdensity-Lyman-α transmission relation ($\delta_m - \delta_F$) on Megaparsec-scales in these protoclusters, which is hypothesized to be sensitive to the feedback implementations. The lower density regions represented by the SIMBA-100 cosmological volume trace the power-law $\delta_m - \delta_F$ relationship often known as the fluctuating Gunn–Peterson approximation. This trend is continued into higher density regions covered by simulations that implement stellar feedback only. Simulations with active galactic nucleus (AGN) thermal and AGN jet feedback, however, exhibit progressively more Lyman-α transmission at fixed matter overdensity. Compared with the seven protoclusters observed in the COSMOS field, only two display the excess absorption expected from protoclusters. The others exhibit deviations: four show some increased transparency suggested by AGN X-ray thermal feedback models while the highly transparent COSTCO-I protocluster appears to reflect intense jet feedback. Discrepancies with the stellar-feedback-only model suggests processes at play beyond gravitational heating and/or stellar feedback as the cause of the protocluster transparencies. Some form of AGN feedback is likely at play in the observed protoclusters, and possibly long-ranged AGN jets in the case of COSTCO-I. While more detailed and resolved simulations are required to move forward, our findings open new avenues for probing AGN feedback at Cosmic Noon.

Key words: methods: numerical – galaxies: clusters: general – intergalactic medium – large-scale structure of Universe.

1 INTRODUCTION

Galaxy clusters did not form in a day; these giants were still in their adolescence stage at redshift $z > 2$, in structures commonly known as ‘galaxy protoclusters’. At this cosmic noon epoch, they were still in an active stage of mass assembly, and their member galaxies may or may not end up in a present-day cluster (Chiang, Overzier & Gebhardt 2013). Protoclusters contribute a significant fraction to the cosmic star formation rate density (Chiang et al. 2017) at Cosmic Noon when cosmic star-formation activity was at its peak. Galaxy protoclusters do not necessarily exhibit very conspicuous galaxy overdensities, with most nascent protocluster cores accumulating masses of $M \lesssim 10^{14}\, M_\odot$ by this epoch, and the mass growth of these protoclusters is not necessarily monotonic (i.e. the most massive protocluster cores at $z \sim 2$ will not always grow into the most massive $z \sim 0$ clusters, for example, Cui et al. 2020; Remus, Dolag & Dannerbauer 2023).

Aside from their important role in star formation and massive galaxy assembly, protoclusters also show interesting interactions with the intergalactic medium (IGM), (for example Cai et al. 2017; Kikuta et al. 2019; Kooistra et al. 2022a; Emonts et al. 2023). After the end of cosmic reionization around $z \sim 6$, the free streaming photons and the expansion of the universe quickly erase the thermal memory of IGM towards reionization. The inverse Compton scattering of electrons off the Cosmic Microwave Background (CMB) is also an important cooling mechanism at high redshift, but its cooling efficiency is about one magnitude lower than the cooling caused by cosmic expansion at $2 < z < 3$ (McQuinn & Upton Sanderbeck 2016). The equilibrium between adiabatic expansion and photoionization heating leads to a power-law relation between the temperature and density in the IGM, $T = T_0(1+\delta)^{\gamma-1}$ (Hui & Gnedin 1997), where $\delta = \rho/\bar{\rho} - 1$ is the overdensity, T_0 is the typical temperature for IGM with an average density. The temperature-density equation-of-state index, $\gamma \approx 1.7$ (Lee et al. 2015; Hiss et al. 2018; Rorai et al. 2018), is a parameter arising from the balance between recombination and photoionization. This leads to a relation between hydrogen Lyman-alpha (Lyα) forest optical depth and IGM

★ E-mail: dong-chenze@g.ecc.u-tokyo.ac.jp

state, often known as the Fluctuating Gunn–Peterson Approximation (FGPA; Croft et al. 1998, 2002):

$$\tau_{\mathrm{Ly}\alpha} \propto \frac{T^{-0.7}}{\Gamma_{UV}}(1+\delta)^2 \propto (1+\delta)^{\beta}, \quad (1)$$

where Γ_{UV} corresponds to the background ultraviolet (UV) photoionization rate, and $\beta = 2 - 0.7(\gamma - 1)$ is another positive power-law index. By observing Lyman-α absorption, this relation facilitates the probing of overdensity δ along lines of sight to background objects, for example, CLAMATO (Lee et al. 2014b, 2018; Horowitz et al. 2022) and LATIS (Newman et al. 2020). Protoclusters are usually the sites with the highest overdensities and strongest Lyman-α absorption, so this relation further enables the detection of protoclusters at Cosmic Noon (Stark et al. 2015; Cai et al. 2016; Lee et al. 2016; Miller, Bolton & Hatch 2021; Qezlou et al. 2022).

Baryonic processes associated with galaxy formation in and around the protoclusters add great complexity to the connection between the IGM and galaxy protoclusters. Stellar and active galactic nucleus (AGN) feedback can inject energy into the IGM as they regulate the star formation, potentially causing deviations from the FGPA due to collisional heating and ionization. Furthermore, even shock heating purely driven by gravitational collapse can raise the temperature of the gas within high-mass haloes (Kereš et al. 2005), eventually driving the evolution into the complex and ‘missing’ phases of the IGM at late times (Davé et al. 2001; Cen & Ostriker 2006). Early works (Theuns, Mo & Schaye 2001) hinted at this more complex connection by inspecting individual quasar spectra. On the simulation side, various studies assessed the impact of feedback processes both on the summary statistics of the Lyman-α forest (for example Kollmeier et al. 2003; Christiansen et al. 2020; Nagamine et al. 2021; Burkhart et al. 2022; Khaire et al. 2023, 2024; Tillman et al. 2023a, b) and on Lyman-α absorption in the circumgalactic medium of galaxies and quasars (Kollmeier et al. 2006; Meiksin, Bolton & Tittley 2015; Meiksin, Bolton & Puchwein 2017; Turner et al. 2017; Sorini et al. 2018; Sorini, Davé & Anglés-Alcázar 2020; Appleby et al. 2021; Mallik et al. 2023; Maitra, Mallik & Srianand 2024). Kooistra et al. (2022a) also predicted an elevation of IGM temperature by the manual phenomenological implementation of AGN feedback.

Recently, Dong et al. (2023) found a particularly interesting case, the COSTCO-I protocluster, which potentially holds strong implications for the interplay between heating processes on the Lyman-α forest in protoclusters. This protocluster was first identified in density reconstructions of the COSMOS spectroscopic catalogue (Ata et al. 2021) and resulting constrained simulations (COSTCO project; Ata et al. 2022) as the progenitor of a late-time $M = (4.6 \pm 2.2) \times 10^{14}\, M_{\odot}$ galaxy cluster. However, a comparison with the Wiener-filtered tomographic IGM map from the CLAMATO survey (Horowitz et al. 2022) indicated that COSTCO-I is anomalously transparent to Lyman-α absorption, contrary to the significant absorption expected through FGPA given its galaxy overdensity. It is a special case among protoclusters residing in the overlapping volume of CLAMATO and COSTCO (hereafter CLAMATO×COSTCO protoclusters), as none of the other protoclusters have such transparent Lyman-α absorption at first glance. Moreover, the unexpectedly high transmission is still evident after averaging with a $r = 15$ cMpc h^{-1} tophat kernel, indicating that this phenomenon spans multiple physical Megaparsecs. Recent work by Miller et al. (2022) suggests that the UV radiation from protocluster AGN seems unlikely to be the cause of extended Lyman-α transmission to the extent that we see unless an ultraluminous quasar is present (Schmidt et al. 2018). Therefore, according to equation (1), such a Lyman-α transparency could be attributed to a high IGM temperature across several physical Mpc, given the uniform UV background and overdensity confirmed by COSTCO. We hypothesize that the extended gas around COSTCO-I is a possible site of large-scale heating processes other than UV background ionization, which breaks the power-law relation between the temperature and density of IGM. However, the mechanism that caused the heating still remains unclear, as it could be some combination of AGN and/or stellar feedback along with gravitational shock heating.

In this paper, we use the SIMBA cosmological galaxy formation simulation and The Three Hundred Cluster zoom simulation suite to investigate the impact of gravitational shock heating and various feedback processes on the Lyman-α forest within and around large haloes at $z = 2.5$. In Section 2, we introduce the data adopted in this study and describe how we derive the quantities of interest. In Section 3, we present our main findings based on the measurement from both observations and simulations. In Section 4, we discuss the results and summarize the study. Throughout this paper, we adopt the cosmology of $H_0 = 100h\ \mathrm{km\,s^{-1}\,Mpc^{-1}} = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.315$, and $\Omega_\Lambda = 0.685$ for the observational data to maintain consistency with the cosmology in Dong et al. (2023).

2 DATA

In order to quantify the heating of IGM around massive protoclusters in simulations and compare with the observed case with large-scale gas heating, we focus on the following quantities: overdensity δ_m and Lyman-alpha transmission δ_F. The definition and measurement methodology of these quantities for the CLAMATO×COSTCO observational data are described in Section 2.1. The simulation suites used in this study are introduced in Section 2.2, and in Section 2.4 we applied the same measurement to the simulations.

2.1 CLAMATO and COSTCO observations

This paper was motivated by the observational discovery of the COSTCO-I galaxy protocluster at $z = 2.30$ (Lee et al. 2016; Ata et al. 2022; Dong et al. 2023). The unusual nature of this protocluster was first hinted at by Lee et al. (2016), but the official discovery was made through constrained N-body simulations (Ata et al. 2022) based on the reconstructed matter density field of the COSMOS galaxy spectroscopic sample in the redshift range $2.0 < z < 2.5$ (Ata et al. 2021). To briefly summarize, the initial conditions (ICs) of the N-body simulations were first sampled using the `COSMIC-BIRTH` (Kitaura et al. 2020) algorithm, conditional to the observed spectroscopic galaxy redshift distribution at $2.00 < z_{\mathrm{obs}} < 2.52$ in the COSMOS field. Ata et al. (2022) then ran the `PKDGRAV3`N-body code (Potter, Stadel & Teyssier 2017) on the ICs derived from this analysis, with a total of 57 posterior realizations and produced simulation snapshots at $z = 2.3$ as well as at $z = 0$. This allows us to compute various quantities for all the CLAMATO×COSTCO protoclusters, including some that are not typically accessible through observations such as the matter density field. In addition, this allows us to estimate the uncertainties related to the derived quantities thanks to the posterior realizations.

Table 1 lists all the CLAMATO×COSTCO protoclusters that we analyse in this study. Among the entries, ZFIRE was first detect through medium-band imaging (Spitler et al. 2012) and later confirmed by Keck/MOSFIRE spectroscopy (Yuan et al. 2014). Hyperion (Casey et al. 2015; Chiang et al. 2015; Diener et al. 2015; Lee et al. 2016; Wang et al. 2016) is a complex of galaxy protoclusters identified at redshift $2.423 < z < 2.507$, where only Hyperion 1, 3,

Table 1. CLAMATO×COSTCO galaxy protoclusters analysed in this work. Note that Ata et al. (2022) only reported the final mass for Hyperion clusters as the sum of all member clusters (Hyperion 1–7), including Hyperion 2; Hyperion 6 and Hyperion 7 are outside the volume of the CLAMATO Lyman-α map.

Protocluster	RA	Dec	$z_{\rm obs}$	Final mass ($h^{-1}M_{\odot}$)
ZFIRE	150.094	2.251	2.095	$(1.2 \pm 0.3) \times 10^{15}$
Hyperion 1	150.093	2.404	2.468	$(2.5 \pm 0.5) \times 10^{15}$
Hyperion 3	149.999	2.253	2.444	–
Hyperion 4	150.255	2.342	2.469	–
Hyperion 5	150.229	2.338	2.507	–
COSTCO-I	150.11	2.161	2.298	$(4.6 \pm 2.2) \times 10^{14}$
COSTCO-III	150.129	2.275	2.16	$(5.3 \pm 2.6) \times 10^{14}$

4, and 5 are enclosed in the volume of CLAMATO. COSTCO-I and COSTCO-III are newly discovered galaxy protocluster reported by Ata et al. (2022) based on constrained-simulation analysis.

To derive the quantities of interest, we first measured the matter overdensity, δ_m around the protoclusters from the N-body outputs. To do this, we converted the particle data from the simulation output into a grid of $\delta_{m,\rm raw}$ and used the mean density of the universe to normalize the matter density,

$$\delta_{m,\rm raw} = \rho/\bar{\rho} - 1. \quad (2)$$

Then, we smoothed the $\delta_{m,\rm raw}$ field with a Gaussian kernel of 7.5 cMpc h^{-1} smoothing radius, and took the value of δ_m at the positions of protoclusters as the overdensity associated with the protocluster.

We measured the Lyman-α forest transmission fluctuation, $\delta_F = F/\bar{F} - 1$ (where $F = \exp(-\tau_{\rm Ly\alpha})$ is the Lyman-α transmission, and $\bar{F}$ is the mean Lyman-α transmission at $z = 2.5$ given by Becker et al. 2013) associated with the protoclusters based on the Wiener-filtered $\delta_F^{\rm WF}$ from the CLAMATO survey (Horowitz et al. 2022). To match the measurement of overdensity above, we simply smoothed the $\delta_F^{\rm WF}$ field with a Gaussian kernel of 7.5 cMpc h^{-1} and evaluated δ_F at the position of the protocluster, which was used to compare with the δ_F matter overdensity. In addition, Wiener filtering under irregular sightline sampling can lead to biased reconstructions, δ_F, compared to the underlying truth (see Appendix A). We corrected the bias by introducing a rescaling factor for the $\delta_F^{\rm WF}$ values of the CLAMATO×COSTCO protoclusters, calibrated from simulated mock sightlines.

2.2 Simulations

In order to explore a wide range of halo masses, larger simulation boxes are generally desired. However, this typically requires a trade-off between the simulation scale and the resolution given finite computational resources. In our work, we aim to achieve adequate resolution while retaining the ability to study massive haloes ($M_{\rm halo} > 10^{14}h^{-1}M_{\odot}$ at $z = 0$) when picking the runs in each simulation suite. We, therefore, use the large volume available in the `SIMBA` suite of cosmological hydrodynamical simulation, i.e. the $L = 100\,h^{-1}$ cMpc box. We also study `The Three Hundred`, a series of zoom-in simulations that focus on the evolution of the most massive galaxy clusters at $z = 0$. These simulations have been run with several different feedback prescriptions. We introduce the simulation suites below with a focus on their jet feedback implementation and summarize them in Table 2. Due to the limited snapshots available in `The Three Hundred`, we study the property of massive protoclusters with snapshots at $z = 2.5$, which correspond to the highest redshift in the CLAMATO volume. The choice of snapshot redshift can be justified for the following reasons. First, the median redshift of CLAMATO×COSTCO protoclusters is $z = 2.444$ (see Table 1), therefore adopting $z = 2.5$ is nominally the closest snapshot. Secondly, most clusters in `The Three Hundred` sample have larger final masses ($> 10^{15}h^{-1}M_{\odot}$) than CLAMATO×COSTCO sample, thus a slightly higher redshift for the simulated sample facilitates a better match with the evolutionary stage of the observed protoclusters.

2.2.1 SIMBA

The `SIMBA` simulation (Davé et al. 2019) is a cosmological simulation suite based on a forked version of the `Gizmo` (Hopkins 2015) Gravity+hydrodynamics solver. The meshless-finite-mass (MFM) scheme in `Gizmo` gets rid of artificial viscosity in its hydrodynamic implementation, which handles shocks in a more precise and realistic way. The `SIMBA` simulation includes a variety of input physics. The radiative cooling and photoionization heating are implemented using `Grackle-3.1` (Smith et al. 2017) with an on-the-fly, and self-consistent self-shielding treatment (Rahmati et al. 2013). `SIMBA` adopts the ionizing background in Haardt & Madau (2012) (HM12), which is further modified to account for self-shielding. The star formation formalism is similar to its precursor, `MUFASA` (Davé, Thompson & Hopkins 2016). The star formation rate (SFR) is calculated with an H_2-based Schmidt (1959) law, where the SFR is proportional to the H_2 density and the inverse of dynamic time. `SIMBA` incorporates the stellar feedback from Type Ia & Type II Supernovae (SNe) and Asymptotic Giant Branch (AGB) stars. The `SIMBA` applied the decoupled two-phase winds for stellar feedback, of which the effect is aligned with high-resolution FIRE simulation (Feedback In Realistic Environments; Hopkins et al. 2014).

The major improvement in `SIMBA` subgrid models relative to `MUFASA` is the black hole physics. For the black hole accretion, `SIMBA` employs a combination of Bondi–Hoyle–Lyttleton accretion (Bondi 1952) (from hot gas) and Torque-limited accretion (Hopkins & Quataert 2011) (from cold gas). The AGN feedback is modelled in a two-phase approach, featuring criteria on the black hole mass $M_{\rm BH}$ and Eddington ratio $f_{\rm Edd}$. The 'radiative' mode is activated for the black holes in high accretion state ($f_{\rm Edd} > 0.2$) or with low mass ($M_{\rm BH} < 10^{7.5}{\rm M}_{\odot}$), in which the black holes exert kinetic kick on surrounding gas to create 'AGN winds' with a speed of $\sim 10^3$ km s^{-1}. For more massive black holes with lower accretion rate, the feedback enters 'jet' mode, where the gas at black hole vicinity is heated to $\gtrsim 10^7$ K and launched with a speed of ~ 7000 km s^{-1}. The crucial feature of SIMBA's AGN feedback arises from the directionality of momentum ejection. For both wind and jet feedback, the ejected gas launches from the direction perpendicular to the inner disc around the SMBH with an opening angle of zero. Given the high velocity of the jet feedback, it leads to a highly collimated bipolar outflow, which has an impact on the scale of observed radio lobes. The authors argue the highly collimated implementation of feedback is later alleviated by other processes on longer time-scales, such as the precession of the inner disc and the interaction of outflow and ambient gas. For more detail, we refer the readers to the introductory paper of SIMBA (Davé et al. 2019).

`SIMBA` has simulation boxes with settings different in different box sizes and resolutions publicly available. In this study, we adopt the flagship run (hereafter SIMBA-100) with the box size of 100 cMpc h^{-1} and 2×1024^3 particles to increase the volume and hence halo mass coverage compared with the 50 h^{-1} cMpc versions.

Table 2. Hydrodynamical simulations used in this work. We specify the alias, box size, cosmology, resolution, and feedback model in each run.

Simulation	Box size (cMpc h^{-1})	Cosmology (h, σ_8, Ω_m, Ω_b)	Mass resolution ($h^{-1} M_\odot$)	Feedback model
SIMBA-100	100	(0.68, 0.82, 0.3, 0.048)	6.51×10^7	SIMBA full physics
300-SIMBA	zoom-in	(0.678, 0.823, 0.307, 0.048)	1.27×10^9	SIMBA full physics
300-GadgetX	zoom-in	(0.678, 0.823, 0.307, 0.048)	1.27×10^9	Stellar & AGN thermal feedback
300-GadgetMUSIC	zoom-in	(0.678, 0.823, 0.307, 0.048)	1.27×10^9	Stellar feedback

2.2.2 *The THREE HUNDRED*

`The Three Hundred` project (Cui et al. 2018, 2022) is a set of re-simulations of 324 large galaxy clusters. These clusters were first selected from the 1 cGpc h^{-1} MDPL2 dark-matter-only simulation (Klypin et al. 2016) based on their halo mass at $z = 0$ ($M_{h,z=0} > 6.42 \times 10^{14} M_\odot$). A spherical volume with a radius $r = 15$ cMpc h^{-1} was then selected as the re-simulation region. The ICs were generated by the `GINNUNGAGAP` code with the re-simulation region split into gas and dark matter particles, and the outer regions are resampled using low-resolution particles to save the computation cost. `The Three Hundred` resimulation set provides a mass-complete sample of the extremely massive haloes with flexibility in baryon model prescriptions. Our study includes the runs with GadgetX, GadgetMUSIC, and GIZMO-SIMBA models for all 324 clusters.

The detail of three prescriptions in `The Three Hundred` is listed as follows.

(i) `GadgetMUSIC` (hereafter 300-GadgetMUSIC): This prescription is based on Sembolini et al. (2012), running with `GADGET-3` Tree-PM code, an updated version of `GADGET-2` (Springel 2005) The classic entropy-conserving SPH formulation with a 40 neighbour M3 interpolation kernel is used in the hydrodynamic solver of 300-GadgetMUSIC. It adopts a metal-independent approach for radiative cooling and a homogeneous UV background (UVB) by Haardt & Madau (2001) (HM01) for photoionization heating. On star formation, it follows the hybrid multiphase model in Springel & Hernquist (2003) with the Salpeter (1955) initial mass function (IMF). The kinetic stellar feedback and thermal stellar feedback in 300-GadgetMUSIC are realized based on Springel & Hernquist (2003) and the 2-phase model in Yepes et al. (1997), respectively. There is neither black hole accretion nor black hole feedback implemented in 300-GadgetMUSIC.

(ii) `GadgetX` (hereafter 300-GadgetX): This prescription has a few differences compared with 300-GadgetMUSIC. It uses an updated SPH scheme (Beck et al. 2015) with improvement in artificial thermal diffusion, time-dependent artificial viscosity, high-order Wendland C4 interpolating kernel, and wake-up scheme. On the gas heating and cooling, 300-GadgetX adopts Haardt & Madau (1996) (HM96) UVB and a metal-dependent cooling calculation (Wiersma, Schaye & Smith 2009). It models star formation using Tornatore et al. (2007) and Chabrier (2003) IMF. 300-GadgetX also implemented kinetic stellar feedback following Springel & Hernquist (2003). Contrary to 300-GadgetMUSIC, 300-GadgetX suite has a black hole physics implementation by Steinborn et al. (2015), which includes a multiphase gas accretion and thermal AGN feedback.

(iii) `GIZMO-SIMBA` (hereafter 300-SIMBA): This prescription adopted a slightly modified model based on the SIMBA cosmological runs. Because the resolution of The Three Hundred is downgraded compared to SIMBA-100, Cui et al. (2022) re-calibrated the feedback model based on a series of criteria related to stellar mass and halo mass in their section 3.2. The modification is mainly motivated by the deficit of star formation in satellite galaxies. The star formation efficiency was then tuned higher to alleviate the deficit, which led to an overproduction of stars in central galaxies. In turn, the jet feedback was tuned more powerful (with a max speed of ∼ 15 000 km s^{-1}) to regulate stellar mass in the central galaxies. As a result, the jets in 300-SIMBA are the strongest among the simulation suites in this study, which results in a relatively flat gas density profile at z = 0 (Li et al. 2023).

Apart from the snapshot at $z = 2.5$, we also use the output of halo finder `AHF` (Knollmann & Knebe 2009) to determine the position of the most massive haloes in each re-simulation.

2.3 Mock Lyman-alpha skewers

As all the simulations used in this study utilized `GADGET` as their backbone, we use a light-weighted Python package pygad to handle the simulation output. The package integrates the modules for reading simulation data with automatically inferred units, producing maps of quantities with the given SPH kernel, and generating the Lyman-α optical depth. When generating the Lyman-α skewers, the package will use the cosmology defined in the simulation output (see Table 2). The cosmologies adopted in simulations have subtle per cent-level differences from the one used for the observational data reduction (Ata et al. 2022; Horowitz et al. 2022). In this paper, we are aiming at a largely qualitative comparison, so we do not consider these differences to be important. Future analyses should, however, take careful account of the assumed cosmology.

In addition, the cosmological simulations suites adopted in this study adopted different empirical UVBs, which may cause inconsistencies between the simulation and observed Lyman-α mean flux. To facilitate a fair comparison between simulations and observations we calibrated the transmission flux in simulations following the approach in Christiansen et al. (2020). The details can be found in Appendix B.

2.4 Simulation observables

For the simulations, we first identified the positions of massive haloes ($M_h > 10^{14} M_\odot$ at present day) in SIMBA-100 and the most massive halo in each of the `The Three Hundred` zoom-in simulation at $z = 2.5$, then calculated the same Lyman-α forest and mass proxies as for CLAMATO×COSTCO protoclusters: δ_m and δ_F. This involved computing the gridded matter density and Lyman-α transmission for each zoom-in simulation snapshot using `pygad`. We then applied the same smoothing kernels as described in 2.1, and finally evaluated the value at the position of corresponding haloes. Some of the central haloes in `The Three Hundred` were at the edge of the simulation zoom-in regions and were not properly simulated at full resolution, so we removed these haloes from our subsequent analysis.

3 RESULTS

Fig. 1 illustrates the distribution of gas particles on the temperature-density phase diagram. As a 'baseline' run for this study, SIMBA-100

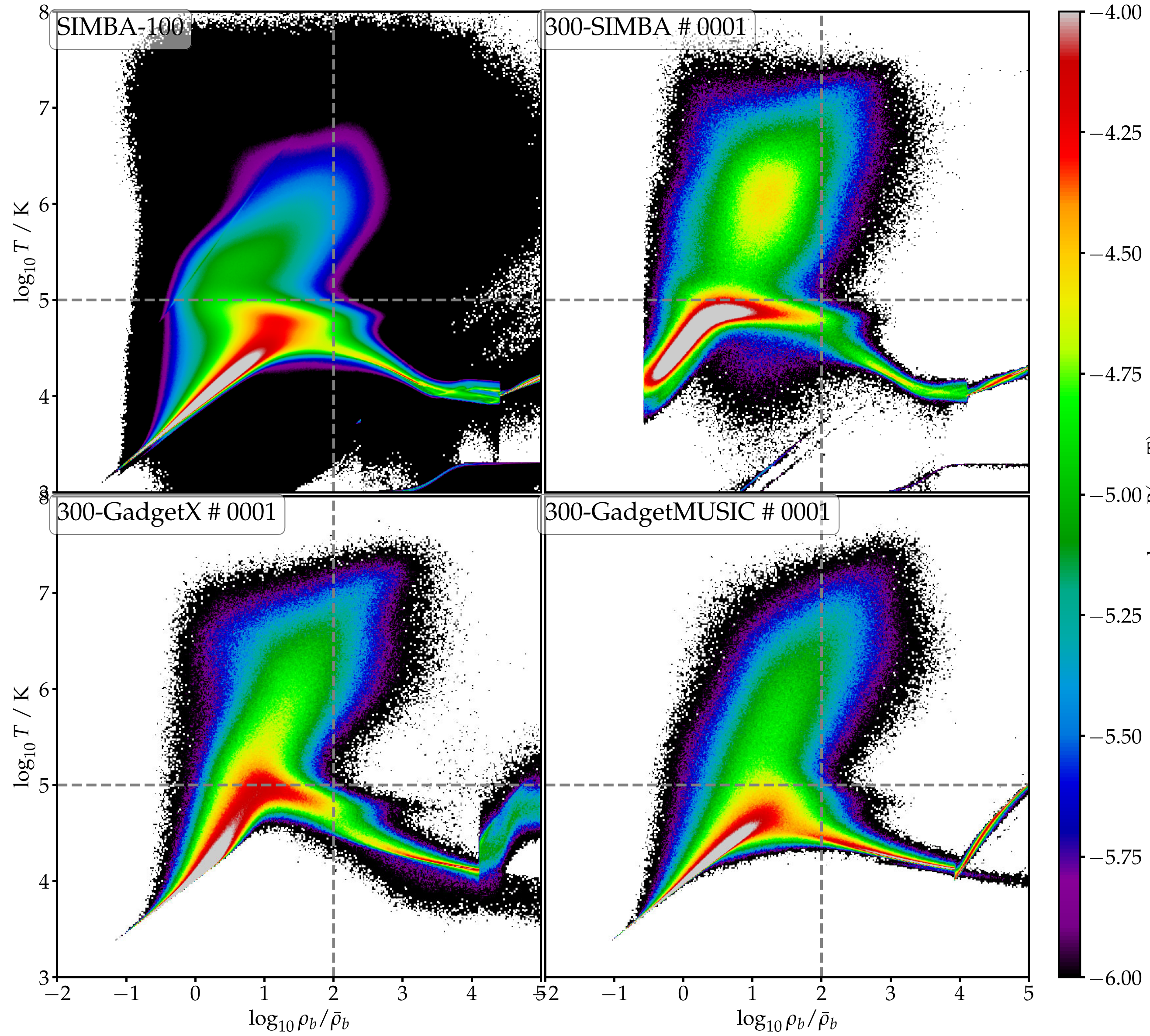

Figure 1. The mass-weighted temperature–density relation for SIMBA-100 (upper left) and three feedback prescriptions (rest panels) on `The Three Hundred` cluster #0001 at $z = 2.5$. In each panel, the vertical grey dash line represents $\log_{10} \rho_b/\bar{\rho}_b = 2$ while the horizontal line is $T = 10^5$K. The two dash lines divide the $\rho - T$ phase space into diffused cold gas (lower left), warm-hot intergalactic medium (WHIM, upper left), hot halo gas (upper right), and condensed gas (lower right). Among the `The Three Hundred` runs, the fraction of the WHIM component in each panel reflects the strength of feedback.

reproduces the power-law temperature–density relation in the diffuse cold gas regime, and shows a population of WHIM gas on $T > 10^5$ K. All `The Three Hundred` phase diagrams exhibit a more notable WHIM component since `The Three Hundred` runs are zoom-in simulations around extremely massive regions rather than a regular cosmological volume used in SIMBA-100. Among these phase diagrams, the WHIM fraction of 300-SIMBA run exceeds 300-GadgetX run, and 300-GadgetX outmatches 300-GadgetMUSIC run, which reflects the latter has relatively weak feedback.

In Fig. 2 we display the projected matter density, gas temperature, and Lyman-α transmission maps of The Three Hundred cluster #0001. The matter density distribution is almost identical among the different feedback prescriptions, which supports the idea that even strong feedback does not significantly alter the overall matter distribution (Sunseri, Li & Liu 2023). The gas temperature maps, however, show distinct differences. The 300-SIMBA AGN jet-feedback runs show the highest overall IGM temperature in `The Three Hundred`, which is centred on hot ($> 10^7$ K) bubbles at the nodes of the cosmic web and elevated gas temperature ($\sim 10^5$ K) that extends over several Megaparsecs into regions with milder overdensity. As a result, the Lyman-α forest absorption is almost completely suppressed in this 300-SIMBA protocluster (top-right panel of 2), and the protocluster appears as a roughly mean-density region when viewed through the Lyman-α forest. A search for protoclusters using only Lyman-α forest data (for example, Stark et al. 2015; Cai et al. 2016; Lee et al. 2016) would probably miss protoclusters like this.

Hot bubbles are also seen in the 300-GadgetX thermal AGN run, but they do not extend into underdensities and the gas temperature in those regions remain cool ($\sim 10^4$ K). As a result, these regions

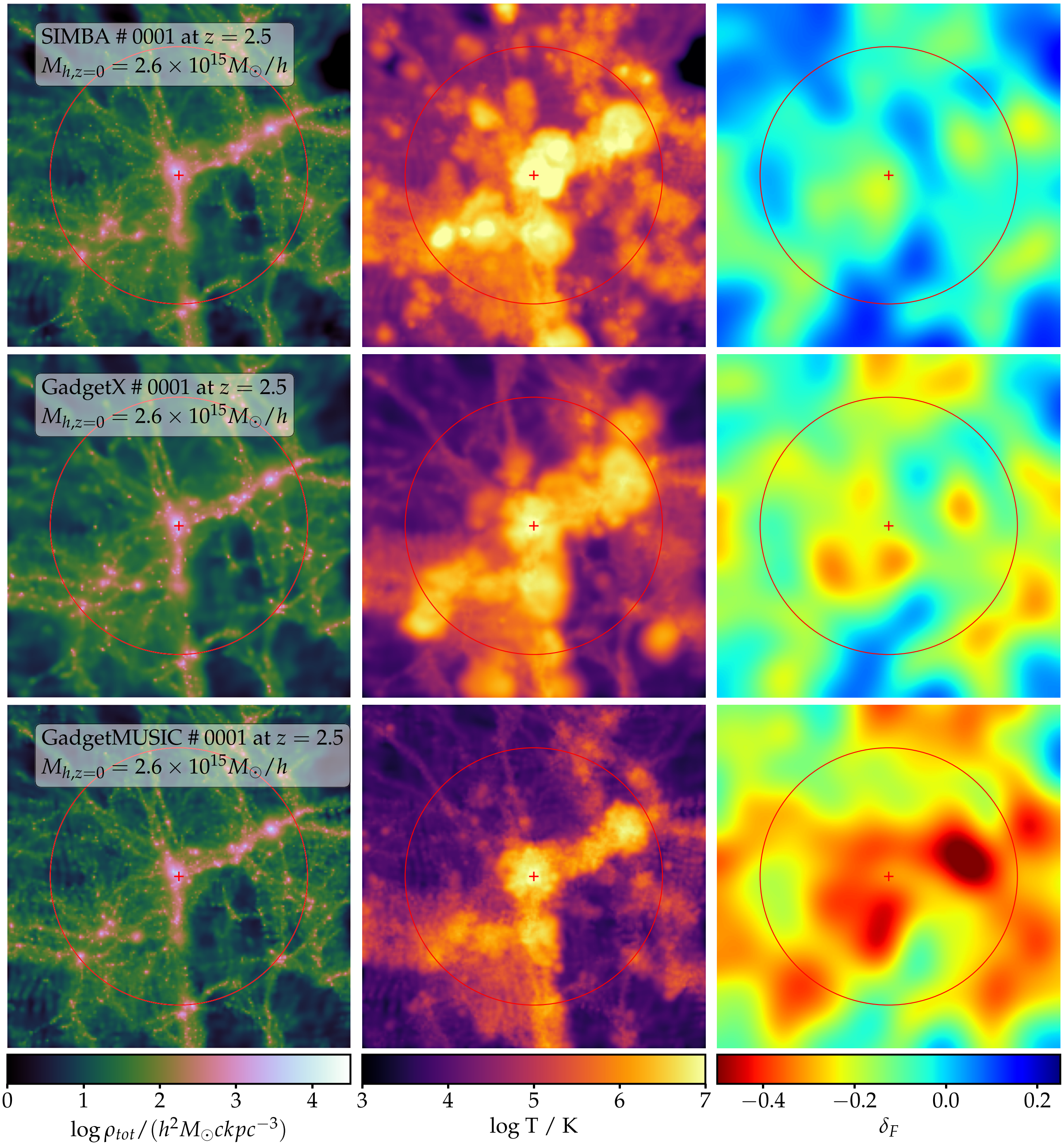

Figure 2. $20 \times 20 \times 2$ cMpc h^{-1} slice of The Three Hundred # 0001. The left column shows matter density maps, the middle panels are the gas temperature maps, and the right panel are the maps of Lyman-α transmission fluctuation δ_F. The δ_F map is smoothed with a Gaussian kernel with $R = 1$ cMpc h^{-1}. From top to bottom, the maps are from GIZMO-SIMBA, GadgetX, and GadgetMUSIC prescriptions, respectively. The centres of haloes are marked with red plus signs; the red circles illustrate the 7.5 cMpc h^{-1} scale used for Gaussian kernels of smoothing. The jet feedback of the 300-SIMBA run significantly heats up the surrounding IGM, leading to a high overall Lyman-α transmission even with the massive galaxy protocluster overdensity.

exhibit some Lyman-α forest absorption such that the 300-GadgetX case is not as transparent as in 300-SIMBA.

In the 300-GadgetMUSIC runs, high temperature gas is seen only in the direct vicinity of filaments within $\lesssim$ 1Mpc from galaxies, and the Lyman-α transmission value is therefore the lowest among the different feedback prescriptions of The Three Hundred. The Lyman-α absorption of this protocluster is clearly evident and would be easily detected in a search of the Lyman-α forest.

Fig. 3 shows the smoothed Lyman-α forest transmission, δ_F, plotted against the correspondingly smoothed matter overdensity, δ_m, measured for the ensemble of simulated protoclusters at $z = 2.5$. We find a correlation between the two quantities within the individual simulation suites: δ_F generally decreases when a halo has higher δ_m. The 300-SIMBA suite has the highest overall Lyman-α transmission among The Three Hundred runs, while the δ_F value of 300-GadgetX protoclusters shows an offset of $\Delta\delta_F \sim -0.1$, i.e. more Lyman-α opacity. The 300-GadgetMUSIC suite exhibits the least transparent Lyman-α forest, which is probably due to the lack of AGN feedback in its feedback prescription. For SIMBA-100, we show the matter-transmission relationship based on the fluctuating

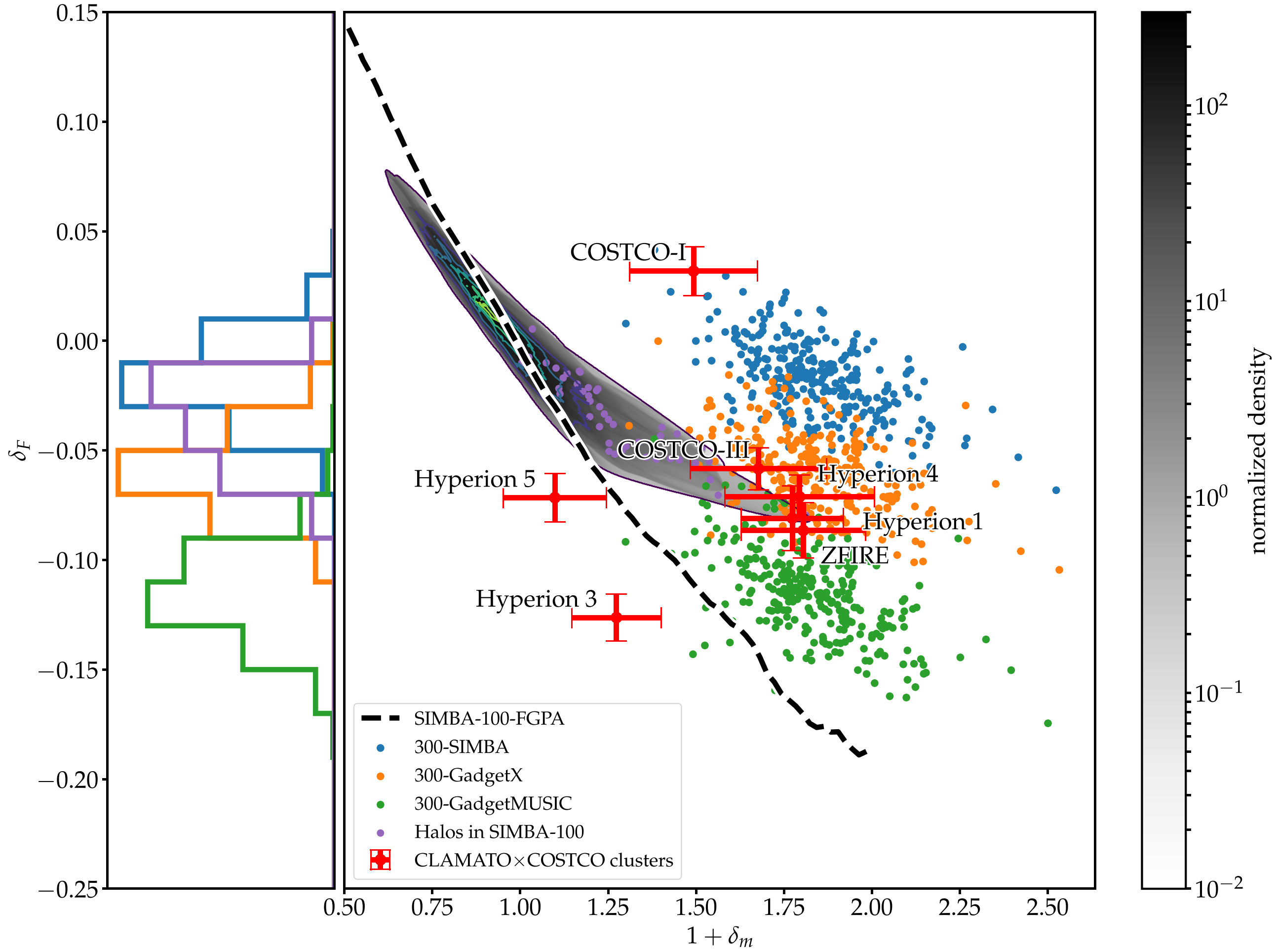

Figure 3. The relation between δ_m and δ_F in simulations as well as data. The scatter plots with blue, orange, and green dots are measured from 300-SIMBA, 300-GadgetX, and 300-GadgetMUSIC, respectively. The black dashed line depicts the FGPA relationship in SIMBA-100, where $\beta = 1.6$. The shadow region and contours show the global relation of the two quantities in SIMBA-100 simulation box, where the purple points are the values of massive haloes in SIMBA-100. The red points with error bars are the measurements from the CLAMATO×COSTCO protocluster. The left histogram shows the distribution of δ_F measured from massive protoclusters in our simulation samples, with the same colour code as in the right panel. Most of the CLAMATO×COSTCO protoclusters appear to be deviating from the standard transmission-density relationship, while COSTCO-I shows a good match with the 300-SIMBA protoclusters that are implemented with jet feedback.

Gunn–Peterson relationship over the entire box as the black dashed line in Fig. 3. In addition, we generate `pygad` Lyman-α skewers from the simulated H I gas density and show the skewer-based matter-transmission relationship over the entire box as a grey-scale contour. The corresponding values centred on individual protoclusters within SIMBA-100 are also indicated separately as purple points. The FGPA of SIMBA-100 appears to flatten out in δ_F with increasing δ_m, but interestingly the points in 300-GadgetMUSIC seem to continue following the FGPA relation traced out by the lower density SIMBA-100 points. This is in qualitative agreement with Viel, Schaye & Booth (2013b), who argued that stellar feedback does not have a significant effect on the power spectrum (i.e. fluctuations) of the Lyman-α forest.

The most massive haloes in the SIMBA-100 box are generally not as massive as those in `The Three Hundred`, but they do not appear to follow a continuous relationship with those in 300-SIMBA. This is suspicious as one would expect the haloes in SIMBA-100 to have similar properties as haloes in 300-SIMBA due to the similar feedback prescriptions; we believe this to be due to the coarser resolution and the somewhat strengthened feedback 300-SIMBA, which we will discuss further in Section 4. The data point for COSTCO-I closest to the trend of 300-SIMBA simulations: its overdensity is comparable with the less massive haloes in The Three Hundred, while its transmission falls in the range of δ_F of 300-SIMBA.

We have also overplotted in Fig. 3 the measured δ_F from the CLAMATO Wiener-filtered maps and the δ_m derived from the COSTCO constrained simulations, for the observed protoclusters listed in Table 1. The error bars were estimated from the ensemble of posterior realizations in the COSTCO simulation suite.

The δ_m values of the observed CLAMATO×COSTCO protoclusters are generally $\delta_m > 1$ as expected for overdensities, although those for Hyperion 3 and Hyperion 5 are only mildly overdense. This is not in contradiction with their identification as protoclusters, because protoclusters are often only mildly overdense over scales of multiple Megaparsecs (Chiang et al. 2013). Hyperion 3 and Hyperion 5 are components of the Hyperion proto-supercluster (Cucciati et al. 2018), which was shown by Ata et al. (2022) to be the progenitor of an eventual $L \sim 100$ Mpc long filamentary supercluster with a total mass of $M_{h,z=0} \approx 2.5 \times 10^{15}\, h^{-1}\, M_{\odot}$. However, when compared with SIMBA simulation, these two protoclusters fall slightly below the FGPA δ_m-δ_F relationship, i.e. they are more Lyman-α opaque than expected. However, this discrepancy is at the $\sim 1 - 2\,\sigma$ level and is likely due to random scatter caused by observational uncertainties.

On the other hand, one sees clear deviations from SIMBA-100 FGPA in the other CLAMATO×COSTCO structures. The majority of the protoclusters (COSTCO-III, ZFIRE, Hyperions 1 and 4) overlaps with the δ_m-δ_F points from the 300-GadgetX cluster simulations, which incorporates the effect of thermal AGN feedback. It indicates a clear departure from the global FGPA-like power-law relationship for these protoclusters, in that the amount of observed Lyman-α absorption is not commensurate with the matter overdensity traced by the COSMOS galaxies.

COSTCO-I exhibits an even more significant departure from the FGPA relationship, with a slightly positive δ_F value more usually indicative of underdensities. Its δ_m-δ_F relationship does not appear consistent with those of the 300-GadgetX protoclusters, and lies among those of the 300-SIMBA simulations where AGN jet feedback has been implemented.

4 DISCUSSION

4.1 Gravitational collapse as the origin of protocluster heating?

In this work, we made a comparison of the large-scale Lyman-α forest transmission of several galaxy protoclusters observed within the COSMOS field (Ata et al. 2022; Horowitz et al. 2022) with several hydrodynamical simulation suites that have varied galaxy and AGN feedback prescriptions. In particular, we aimed to investigate the anomalously low amount of absorption seen in the COSTCO-I protocluster at $z = 2.30$ (Dong et al. 2023).

Interestingly, the different feedback prescriptions lead to clearly distinct Lyman-α transmission values δ_F in the simulated protoclusters, especially if the controlling variable of underlying matter overdensity δ_m is taken into account. This indicates that measuring the FGPA-like relationship between Lyman-α flux and matter overdensity within galaxy protoclusters can be a sensitive probe of galaxy feedback. In a previous paper, Kooistra, Lee & Horowitz (2022b) showed that hydrodynamical simulations with different feedback distributions exhibit different effective slopes in their global $\delta_F - \delta_m$ relationships, but this is a subtle effect that would probably require next-generation telescopes with $\sim$ 30m diameter apertures to obtain sufficient signal to beat down the observational uncertainties. The results of Fig. 3, on the other hand, suggests that focusing on galaxy protoclusters could provide clearer constraints on the feedback mechanisms.

Only two of the observed protoclusters within the CLAMATO×COSTCO volume appear to follow the global power-law FGPA relationship, while most of the protoclusters exhibit clear deviations from FGPA. The majority of our protoclusters appear to follow the $\delta_F - \delta_m$ relationship traced by the 300-GadgetX simulations, which incorporate the effect of AGN thermal X-ray feedback. The COSTCO-I protocluster, on the other hand, exhibits an even more extreme deviation from the FGPA, which places it within the relationship traced by the AGN jet-feedback model of the 300-SIMBA runs. Indeed, COSTCO-I is at a $\sim 5\sigma$ discrepancy with the simulated 300-GadgetMUSIC protoclusters, which model only stellar/supernova feedback as well as the generic hydrodynamical effect of gravitational shock-heating.

In Dong et al. (2023), we discussed the possibility that the transparent Lyman-α forest of COSTCO-I might simply be due to shock-heating of the cosmic web driven by non-linear gravitational collapse within the overdense protocluster region. Given that stellar feedback appears to have only a minor effect on the FGPA-like photoionization equilibrium of the Lyman-α forest in the 300-GadgetMUSIC protoclusters (Viel et al. 2013a), the 300-GadgetMUSIC protoclusters set an upper limit to the impact of gravitational shock heating in protoclusters at this redshift. While there is a possibility that `The Three Hundred` has not fully converged for the Lyman-α forest (see Appendix C, also the discussion on simulation resolution in Bourne, Zubovas & Nayakshin 2015), the clear deviations from FGPA in the majority of the observed protoclusters suggest that effects beyond simple gravitational collapse are likely at play. Upcoming data sets from `The Three Hundred` with higher resolutions (Cui et al., in preparation) would help us to distinguish the effect of the unresolved baryon process in the case. The fact also appears to be true for the COSTCO-III, Hyperion 1, Hyperion 4, and ZFIRE protoclusters: although their δ_F are not as high as COSTCO-I, they appear to deviate clearly from the adiabatic δ_m-δ_F relationship. Even for the most massive simulated protocluster progenitors ($M_{h,z=0} > 10^{15} M_{\odot}$) in the weak feedback model of 300-GadgetMUSIC, the Megaparsec-scale overdensities have not yet collapsed sufficiently at Cosmic Noon to drive enough gravitational shock heating to suppress the Lyman-α forest signal to the level seen in the majority of the CLAMATO×COSTCO protoclusters. This should also be true for lower mass, COSTCO-I-like haloes $M_{h,z=0} = (4.6 \pm 2.2) \times 10^{14}\, h^{-1}\, M_{\odot}$, as they are even further from virialization.

In the literature, the importance of gravitational heating has mostly been discussed in the context of star formation quenching in galaxies. The pioneering work by Dekel & Birnboim (2006) suggests that the occurrence of gravitational shocks leads to the bimodality of galaxies, as the gravitational shocks shut down the continuous gas supply via cold streams. However, such a scenario is largely restricted by the redshift evolution of halo mass and shock heating mass (c.f. fig. 7 of Dekel & Birnboim 2006), limiting the effect of gravitational shocks to $z < 2$ galaxies. More recent works, for example, Liu & Cen (2017) and Zhu & Feng (2017), also set the scope of such shock-heating-induced quenching to $z < 2$. In addition to the redshift restriction of gravitational-induced shocks, the shock radii of these $z \sim 2.5$ haloes also suggest that the shock scenario is not preferable. As a rough estimation, the most massive halo in the $z = 2.5$ snapshot of SIMBA-100 has a total mass of $8.53 \times 10^{13} M_{\odot}\, h^{-1}$, which translates into $r_{200c} = 946$ ckpc h^{-1}. The central galaxy group in COSTCO-I ($M_h \approx 6 \times 10^{13}\, h^{-1}\, M_{\odot}$; Dong et al. 2023) has a similar halo mass, thus the extent of the central core halo spans a significantly smaller region than the observed scale of enhanced Lyman-α transparency in COSTCO-I, which is a few comoving Megaparsecs. These scales would need to encompass the filamentary structures within the protocluster region, which are at lower densities than haloes and therefore expected to collapse and shock-heat later at $z \sim 1$ (for example, Martizzi et al. 2019).

4.2 Can AGN jet feedback explain COSTCO-I?

Among the feedback prescriptions implemented in The Three Hundred, we find that the 300-SIMBA AGN jet feedback model yields the best match with the observed Lyman-α transmission of COSTCO-I. We also find a discrepancy between 300-SIMBA zoom-in runs and the SIMBA-100 cosmological volume, which ostensibly have similar feedback prescriptions.

The massive cluster environment and the modified parameters from the original SIMBA model to 300-SIMBA contribute to the difference in transmission, but the simulation resolution might also be an issue. In Appendix C we compare the SIMBA-100 run with two additional tests with the 300-SIMBA prescription. We find that the stronger jet feedback gives rise to the WHIM gas with $T \sim 10^7$ K by $z = 2.5$, whereas the downgraded resolution changes the power-law $\rho - T$ relation in the diffuse cold gas regime. As stated in Section 4.1, the fact suggests that the 300-SIMBA runs have yet to reach resolution convergence, but this will not undermine the robustness of comparison between the different models implemented for `The Three Hundred`, nor does it affect our discovery that the majority of our observed protoclusters are in disagreement with FGPA. Understanding the impact of resolution better necessitates zoom-in simulation suites with improved resolution (for example, TNG-Cluster; Nelson et al. 2023).

We conclude that extreme transmission in the 300-SIMBA sample stems from the collective effect of a massive cluster environment, stronger feedback, and low resolution, thus it is premature to claim that this particular AGN jet feedback model is necessarily an accurate depiction of the Lyman-α forest observed in COSTCO-I. Nevertheless, 300-SIMBA results are still useful for qualitative investigations into mechanisms that might facilitate large-scale heating of the IGM. In particular, it confirms the highly collimated jet feedback can boost the large-scale heating of surrounding gas, and make massive protoclusters transparent in the Lyman-α forest. This is broadly consistent with results from previous works on the effect of the different feedback prescriptions in the SIMBA suite of simulations on the distribution and thermal state of gas in the IGM and CGM (Bradley et al. 2022; Sorini et al. 2022; Khrykin et al. 2024; Yang et al. 2024).

Within the next few years, the IGM tomography component of the Subaru Prime Focus Spectrograph (PFS; Takada et al. 2014; Greene et al. 2022) will discover $\gtrsim 100$ galaxy protoclusters at $2.2 < z < 2.7$ to enable a statistical sample to compare with the hydrodynamical feedback models. By then, we aim to have resolved suites of simulations to compare with the observations.

5 SUMMARY AND CONCLUSIONS

In this paper, we examined the effect of various galaxy and AGN feedback mechanisms on the large-scale (~Mpc) Lyman-α forest transmission of galaxy protoclusters at Cosmic Noon ($z = 2.5$). This was motivated by the recent observational discovery of a $z = 2.30$ galaxy protocluster, dubbed COSTCO-I (Dong et al. 2023), which exhibits excess Lyman-α transmission (i.e. too little absorption) given the significant matter overdensity that it represents.

In addition to the $100\,h^{-1}$ cMpc SIMBA-100 cosmological hydrodynamical simulation that features multiple stellar- and AGN-feedback mechanisms, we also analysed a sample of cluster zoom-in simulations from `The Three Hundred` suite, in which each cluster was also simulated with three different feedback prescriptions (stellar feedback, AGN thermal feedback, and AGN jet feedback). We focused on the redshift snapshot at $z = 2.50$ when most of the clusters were still in the protocluster stage. The simulated protocluster sample from `The Three Hundred` suite covers a range of cluster masses up to $3 \times 10^{15} M_{\odot}\,h^{-1}$ at $z = 0$, providing a good statistical sample to study COSTCO-I.

The observable we focused on was the smoothed Lyman-α transmission, δ_F, as a function of the similarly smoothed underlying matter overdensity δ_m. While these quantities can be straightforwardly calculated in the simulations, on the observational side we can do direct comparisons thanks to the combination of the 3D Lyman-α forest tomographic maps from the CLAMATO survey, as well as the COSTCO constrained simulations covering the same volume, that were based on the observed spectroscopic galaxy distribution in the COSMOS field.

We summarize our findings as follows:

(i) When averaged over apertures of several Megaparsecs around individual galaxy protoclusters, we find clear differences in the Lyman-α transmission depending on the feedback model. There is a weak dependence of Lyman-α transmission on the underlying matter density, but the different feedback models studied here all occupy distinct populations in the matter density transmission ($\delta_m - \delta_F$) plane. This establishes that the Lyman-α forest around protoclusters is a potentially powerful probe of feedback mechanisms at Cosmic Noon ($z \sim 2 - 3$).

(ii) From the $\delta_m - \delta_F$ relation for several of the observed CLAMATO×COSTCO protoclusters, we find deviations from the global power-law relationship that is expected to hold for the photoionized IGM at these epochs, i.e. there is a deficit of Lyman-α absorption in these structures. These protoclusters are in marginal disagreement with the simulated protoclusters that incorporate stellar feedback only (300-GadgetMUSIC model). Since this model also sets an upper limit on the effect of shock heating from gravitational collapse, we argue that processes beyond gravitational shock-heating or stellar feedback are possibly at play in enhancing the transparency of the Lyman-α forest seen in these protoclusters.

(iii) The COSTCO-I protocluster, in particular, exhibits an even stronger deviation ($\sim 5\sigma$) from the FGPA relationship and 300-GadgetMUSIC. At face value, COSTCO-I is best matched to the protocluster population simulated with the AGN jet feedback model (300-SIMBA). However, lower mass protoclusters in the SIMBA-100 cosmological hydrodynamical box, which nominally has a similar feedback model, do not form a continuous relation with the 300-SIMBA population in the $\delta_m - \delta_F$ plane. This is probably due to the large-scale extreme environment within the 300-SIMBA regions even in early times, but there is a possibility that `The Three Hundred` is not fully converged concerning the Lyman-α forest. Future analyses will need to be careful to ensure simulation convergence.

Strong AGN feedback in galaxy protoclusters has potential implications towards the use of the Lyman-α forest as a cosmological probe. Cosmological analyses of the Lyman-α forest power spectrum typically make the implicit assumption that the IGM at Cosmic Noon follows the FGPA, which stellar feedback largely does not perturb (Viel et al. 2013b). Galaxy protoclusters represent regions of the Universe with the strongest density fluctuations at the $z \gtrsim 2$ epoch, therefore widespread AGN feedback in these structures could potentially smooth out the corresponding fluctuations that manifest into the Lyman-α forest. While Chabanier et al. (2020) analysed the effect of AGN jet feedback on the 1D Lyman-α forest power spectrum in the context of the HorizonAGN simulations, they did not check the specific case of galaxy protoclusters. It is therefore unclear whether their AGN model would reproduce the protocluster effects seen in this paper. A recent paper by Rogers & Poulin (2023) found an anomalous small-scale suppression in the Lyman-α forest power spectrum in tension with Planck, but our findings indicate that the effect of AGN feedback needs to be taken into account in such analyses.

In the near future, larger samples of Cosmic Noon galaxy protoclusters, such as those to be observed by the Subaru PFS Galaxy Evolution survey, in combination with converged hydrodynamical simulations, will provide strong constraints on the AGN feedback mechanisms at play in galaxy protoclusters.

ACKNOWLEDGEMENTS

The authors thank the anonymous reviewer for the constructive feedback helping us to enhance the quality of this paper. The authors also thank Drew Newman for useful discussions on the correction of Wiener-filter effects in the observational data. Kavli IPMU is supported by World Premier International Research Center Initiative (WPI), MEXT, Japan. This work was performed in part at the Center for Data-Driven Discovery, Kavli IPMU (WPI). This research was supported by Forefront Physics and Mathematics Program to Drive Transformation (FoPM), a World-leading Innovative Graduate Study (WINGS) Program, the University of Tokyo. KGL acknowledges support from JSPS Kakenhi grant Nos. JP18H05868 and JP19K 14755. WC is supported by the STFC AGP Grant ST/V000594/1, the Atracción de Talento Contract no. 2020-T1/TIC-19882 granted by the Comunidad de Madrid in Spain and the science research grants from the China Manned Space Project. He also thanks the Ministerio de Ciencia e Innovación (Spain) for financial support under Project grant PID2021-122603NB-C21, ERC: HORIZON-TMA-MSCA-SE for supporting the LACEGAL-III project with grant number 101086388, and the science research grants from the China Manned Space Project with No. CMS-CSST-2021-A01, and CMS-CSST-2021-A03.

This work has been made possible by the 'The Three Hundred' collaboration.[1] The 300 simulations were performed at the MareNostrum Supercomputer of the BSC-CNS through The Red Española de Supercomputación grants (AECT-2022-3-0027, AECT-2023-1-0013), and at the DIaL–DiRAC machines at the University of Leicester through the RAC15 grant: Seedcorn/ACTP317. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

DATA AVAILABILITY

The results shown in this work use data from The Three Hundred galaxy clusters sample. These data are available on request following the guidelines of The Three Hundred collaboration, at https://www.the300-project.org. The data specifically shown in this paper will be shared upon request to the authors. The SIMBA-100 cosmological simulation is publicly available at http://simba.roe.ac.uk/.

REFERENCES

Appleby S., Davé R., Sorini D., Storey-Fisher K., Smith B., 2021, MNRAS, 507, 2383
Ata M., Kitaura F.-S., Lee K.-G., Lemaux B. C., Kashino D., Cucciati O., Hernández-Sánchez M., Le Fèvre O., 2021, MNRAS, 500, 3194
Ata M., Lee K.-G., Vecchia C. D., Kitaura F.-S., Cucciati O., Lemaux B. C., Kashino D., Müller T., 2022, Nat. Astron., 6, 857
Beck A. M. et al., 2015, MNRAS, 455, 2110
Becker G. D., Hewett P. C., Worseck G., Prochaska J. X., 2013, MNRAS, 430, 2067
Bondi H., 1952, MNRAS, 112, 195
Bourne M. A., Zubovas K., Nayakshin S., 2015, MNRAS, 453, 1829
Bradley L., Davé R., Cui W., Smith B., Sorini D., 2022, preprint (arXiv:2203.15055)
Burkhart B., Tillman M., Gurvich A. B., Bird S., Tonnesen S., Bryan G. L., Hernquist L., Somerville R. S., 2022, ApJ, 933, L46
Cai Z. et al., 2016, ApJ, 833, 135
Cai Z. et al., 2017, ApJ, 837, 71
Casey C. M. et al., 2015, ApJ, 808, L33
Cen R., Ostriker J. P., 2006, ApJ, 650, 560
Chabanier S., Bournaud F., Dubois Y., Palanque-Delabrouille N., Yèche C., Armengaud E., Peirani S., Beckmann R., 2020, MNRAS, 495, 1825
Chabrier G., 2003, PASP, 115, 763
Chiang Y.-K., Overzier R., Gebhardt K., 2013, ApJ, 779, 127
Chiang Y.-K. et al., 2015, ApJ, 808, 37
Chiang Y.-K., Overzier R. A., Gebhardt K., Henriques B., 2017, ApJ, 844, L23
Christiansen J. F., Davé R., Sorini D., Anglés-Alcázar D., 2020, MNRAS, 499, 2617
Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, ApJ, 495, 44
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., 2018, A&A, 619, A49
Cui W. et al., 2018, MNRAS, 480, 2898
Cui W., Qiao J., Davé R., Knebe A., Peacock J. A., Yepes G., 2020, MNRAS, 497, 5220
Cui W. et al., 2022, MNRAS, 514, 977
Davé R. et al., 2001, ApJ, 552, 473
Davé R., Thompson R., Hopkins P. F., 2016, MNRAS, 462, 3265
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Diener C. et al., 2015, ApJ, 802, 31
Dong C., Lee K.-G., Ata M., Horowitz B., Momose R., 2023, ApJ, 945, L28
Emonts B. H. C. et al., 2023, Science, 379, 1323
Greene J., Bezanson R., Ouchi M., Silverman J., the PFS Galaxy Evolution Working Group, 2022, preprint (arXiv:2206.14908)
Haardt F., Madau P., 1996, ApJ, 461, 20
Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays, Savoie, France, preprint (arXiv:astro-ph/0106018)
Haardt F., Madau P., 2012, ApJ, 746, 125
Hiss H., Walther M., Hennawi J. F., Oñorbe J., O'Meara J. M., Rorai A., Lukić Z., 2018, ApJ, 865, 42
Hopkins P. F., 2015, MNRAS, 450, 53
Hopkins P. F., Quataert E., 2011, MNRAS, 415, 1027
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguére C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Horowitz B. et al., 2022, ApJS, 263, 27
Hui L., Gnedin N. Y., 1997, MNRAS, 292, 27
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Khaire V., Hu T., Hennawi J. F., Burchett J. N., Walther M., Davies F., 2023, preprint (arXiv:2311.08470)
Khaire V., Hu T., Hennawi J. F., Walther M., Davies F., 2024, MNRAS, 527, 4545
Khrykin I. S., Sorini D., Lee K.-G., Davé R., 2024, MNRAS, 529, 537
Kikuta S. et al., 2019, PASJ, 71, L2
Kitaura F.-S., Ata M., Rodríguez-Torres S. A., Hernández-Sánchez M., Balaguera-Antolínez A., Yepes G., 2020, MNRAS, 502, 3456
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kollmeier J. A., Weinberg D. H., Davé R., Katz N., 2003, ApJ, 594, 75
Kollmeier J. A., Miralda-Escudé J., Cen R., Ostriker J. P., 2006, ApJ, 638, 52
Kooistra R., Inoue S., Lee K.-G., Cen R., Yoshida N., 2022a, ApJ, 927, 53
Kooistra R., Lee K.-G., Horowitz B., 2022b, ApJ, 938, 123
Lee K.-G., Hennawi J. F., White M., Croft R. A. C., Ozbek M., 2014a, ApJ, 788, 49
Lee K.-G. et al., 2014b, ApJ, 795, L12
Lee K.-G. et al., 2015, ApJ, 799, 196
Lee K.-G. et al., 2016, ApJ, 817, 160
Lee K.-G. et al., 2018, ApJS, 237, 31

---

[1] https://www.the300-project.org

Li Q. et al., 2023, MNRAS, 523, 1228
Liu J., Cen R., 2017, preprint (arXiv:1701.00866)
Lukić Z., Stark C. W., Nugent P., White M., Meiksin A. A., Almgren A., 2014, MNRAS, 446, 3697
Maitra S., Mallik S., Srianand R., 2024, MNRAS, 530, 3013
Mallik S., Srianand R., Maitra S., Gaikwad P., Khandai N., 2023, MNRAS, 523, 2296
Martizzi D. et al., 2019, MNRAS, 486, 3766
McQuinn M., Upton Sanderbeck P. R., 2016, MNRAS, 456, 47
Meiksin A., Bolton J. S., Tittley E. R., 2015, MNRAS, 453, 899
Meiksin A., Bolton J. S., Puchwein E., 2017, MNRAS, 468, 1893
Miller J. S. A., Bolton J. S., Hatch N. A., 2021, MNRAS, 506, 6001
Miller T. B. et al., 2022, ApJ, 941, L37
Nagamine K. et al., 2021, ApJ, 914, 66
Nelson D., Pillepich A., Ayromlou M., Lee W., Lehle K., Rohr E., Truong N., 2024, A&A, 686, A157
Newman A. B. et al., 2020, ApJ, 891, 147
Potter D., Stadel J., Teyssier R., 2017, Comput. Astrophys. Cosmol., 4, 2
Qezlou M., Newman A. B., Rudie G. C., Bird S., 2022, ApJ, 930, 109
Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, MNRAS, 430, 2427
Remus R.-S., Dolag K., Dannerbauer H., 2023, ApJ, 950, 191
Rogers K. K., Poulin V., 2023, preprint (arXiv:2311.16377)
Rorai A., Carswell R. F., Haehnelt M. G., Becker G. D., Bolton J. S., Murphy M. T., 2018, MNRAS, 474, 2871
Salpeter E. E., 1955, ApJ, 121, 161
Schmidt M., 1959, ApJ, 129, 243
Schmidt T. M., Hennawi J. F., Worseck G., Davies F. B., Lukić Z., Oñorbe J., 2018, ApJ, 861, 122
Sembolini F., Yepes G., De Petris M., Gottlöber S., Lamagna L., Comis B., 2012, MNRAS, 429, 323
Smith B. D. et al., 2017, MNRAS, 466, 2217
Sorini D., Oñorbe J., Hennawi J. F., Lukić Z., 2018, ApJ, 859, 125
Sorini D., Davé R., Anglés-Alcázar D., 2020, MNRAS, 499, 2760
Sorini D., Davé R., Cui W., Appleby S., 2022, MNRAS, 516, 883
Spitler L. R. et al., 2012, ApJ, 748, L21
Springel V., 2005, MNRAS, 364, 1105
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Stark C. W., White M., Lee K.-G., Hennawi J. F., 2015, MNRAS, 453, 311
Steinborn L. K., Dolag K., Hirschmann M., Prieto M. A., Remus R.-S., 2015, MNRAS, 448, 1504
Sunseri J., Li Z., Liu J., 2023, Phys. Rev. D, 107, 023514
Takada M. et al., 2014, PASJ, 66, R1
Theuns T., Mo H. J., Schaye J., 2001, MNRAS, 321, 450
Tillman M. T. et al., 2023a, AJ, 166, 228
Tillman M. T., Burkhart B., Tonnesen S., Bird S., Bryan G. L., Anglés-Alcázar D., Davé R., Genel S., 2023b, ApJ, 945, L17
Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050
Turner M. L., Schaye J., Crain R. A., Rudie G., Steidel C. C., Strom A., Theuns T., 2017, MNRAS, 471, 690
Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013a, Phys. Rev. D, 88, 043502
Viel M., Schaye J., Booth C. M., 2013b, MNRAS, 429, 1734
Wang T. et al., 2016, ApJ, 828, 56
Wiersma R. P. C., Schaye J., Smith B. D., 2009, MNRAS, 393, 99
Yang T., Davé R., Cui W., Cai Y.-C., Peacock J. A., Sorini D., 2024, MNRAS, 527, 1612
Yepes G., Kates R., Khokhlov A., Klypin A., 1997, MNRAS, 284, 235
Yuan T. et al., 2014, ApJ, 795, L20
Zhu W., Feng L.-L., 2017, ApJ, 847, 17

APPENDIX A: CORRECTION FOR OBSERVATIONAL WIENER-FILTER EFFECTS

The Wiener filter – which was applied to the CLAMATO Lyman-α forest absorption data to create 3D flux maps – has a prior of reconstructed transmission fluctuation $\delta_F^{\rm rec} = 0$ (i.e. the mean transmission) for voxels in the reconstruction field. This introduces a bias towards the mean flux value in regions without sufficient line-of-sight sampling. According to fig. 7 of Lee et al. (2014a), the Wiener filter will introduce a tilt to the relationship between reconstructed flux $\delta_F^{\rm rec}$ and underlying smoothed matter density, in cases where the 'true' flux $\delta_F^{\rm true}$ is known from numerical simulations. In this work, we correct this effect by fitting a relation between COSTCO-FGPA transmission data and the mock reconstruction result (see Dong et al. 2023) based on N-body simulations, assuming the same sightline sampling as in CLAMATO. To account for the effect of smoothing in this work, we applied the same smoothing scale 7.5 cMpc h^{-1} for the data. In Fig. A1, we find applying a correction factor of 0.735 yields the best match between the reconstruction value and the underlying truth.

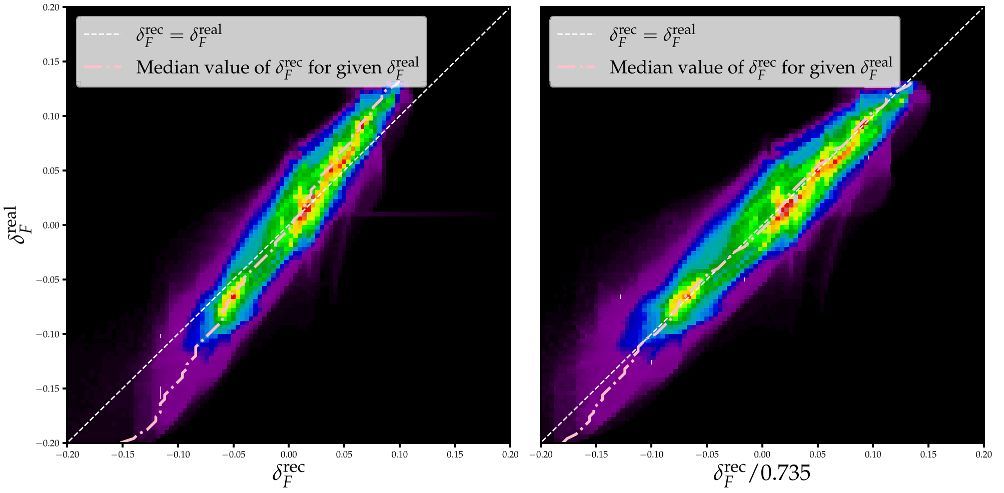

Figure A1. Left: the relation between the Wiener filter-reconstructed flux $\delta_F^{\rm rec}$ and the underlying true value $\delta_F^{\rm real}$ in one COSTCO-FGPA realization, smoothed with a 7.5 cMpc h^{-1} Gaussian kernel. Right: the same relation after applying a correction factor of 0.735 to the reconstructed values. After the correction, $\delta_F^{\rm rec}$ and $\delta_F^{\rm real}$ are in good agreement.

Table B1. A summary table on the UVB adopted, the global neutral Hydrogen abundance, un-calibrated mean Lyman-α flux and calibration factor in each simulation. The calibration factor $A_{\rm UVB}$ is the multiplier such that $\langle \exp(-A_{\rm UVB}\cdot\tau_{\rm raw})\rangle = \langle F\rangle_{z=2.5,\rm obs} = 0.79$, where $\tau_{\rm raw}$ is the optical depth before the calibration.

Simulation	UVB	neutral Hydrogen abundance	$\langle F\rangle_{\rm raw}$	$A_{\rm UVB}$
SIMBA-100	HM12	1.48×10^{-2}	0.79	0.91
300-GadgetMUSIC	HM01	5.97×10^{-3}	0.77	0.87
300-GadgetX	HM96	3.14×10^{-3}	0.78	0.93
300-SIMBA	HM12	4.20×10^{-3}	0.83	1.38

APPENDIX B: CALIBRATION OF LYMAN-α TRANSMISSION AT $z = 2.5$ IN SIMULATIONS

According to Christiansen et al. (2020), for the optically thin case, $\tau \propto 1/\Gamma_{HI}$ is a good approximation. Practically, this implies that changing the UV background is effectively multiplying the H I optical depth by a calibration factor. This allows us to derive the factor $A_{\rm UVB}$ for each simulation and match the observation via rescaling the 'raw' output of skewer generator.

For the SIMBA-100 simulation, we first obtained the 'raw' Lyman-α optical depth $\tau_{\rm raw}$ grid in SIMBA-100 with `pygad`, and then derived a calibration factor $A_{\rm UVB}$ such that $\langle A_{\rm UVB}\cdot\tau_{\rm raw}\rangle = \langle F\rangle_{z=2.5,\rm obs} = 0.79$, where the observational mean flux value is taken from Becker et al. (2013).

The Three Hundred simulations are centred on highly overdense regions that are not expected to yield the mean flux $\langle F\rangle$ when averaging over their Lyman-α sightlines. Therefore, we run complementary cosmological simulations to carry out the $A_{\rm UVB}$ calibration. These complementary runs adopted the same cosmology, numerical resolution, and feedback models as shown in 2. However, instead of zoom-in simulations of individual clusters, these complementary cosmological simulations are run with random initial conditions in a $(33\,{\rm cMpc\,h^{-1}})^3$ cosmological box. The choice of IC avoided the extremely massive protoclusters in `The Three Hundred` sample, thus providing a better calibration.

We present the calibration factors together with other information related to the UVB in Table B1.

APPENDIX C: CONSISTENCY TEST BETWEEN SIMBA-100 AND 300-SIMBA

In this appendix, we show our observations on consistency between SIMBA-100 and 300-SIMBA. To achieve this, we run two test simulation boxes with randomly generated ICs (i.e. not in the environment of massive clusters) but keep the feedback prescription identical as in The Three Hundred. One simulation box has the identical resolution as SIMBA-100 and a box size of 25 cMpc h^{-1}, while another box has a degraded resolution as in `The Three Hundred`, and a box size of 33 cMpc h^{-1}.

Fig. C1 illustrates the mass-weighted temperature–density relation of gas in the simulation (TDR; e.g. Cen & Ostriker 2006; Martizzi et al. 2019) in SIMBA-100 and two test runs at $z = 2.5$. For the connivance of discussion, we adopted the IGM phase classification of Lukić et al. (2014). The IGM gas is divided into four phases according to their temperature and density: diffuse cold gas ($T< 10^5K$, $\rho_b/\bar{\rho}_b < 100$), warm-hot intergalactic medium (WHIM; $T> 10^5K$, $\rho_b/\bar{\rho}_b < 100$), hot halo gas ($T> 10^5K$, $\rho_b/\bar{\rho}_b > 100$), and condensed gas ($T< 10^5K$, $\rho_b/\bar{\rho}_b > 100$). We find, in the diffuse regime, all the three simulation boxes reproduced a power-law ρ–T relation $T = T_0(1+\delta)^{\gamma-1}$, but the low-resolution run has a larger power-law index γ, which may be relevant to the high overall Lyman-α transmission shown in Section 3. The pattern of gas distribution in the WHIM regime is different from simulation to simulation: most of WHIM in SIMBA-100 has $T \le 10^6$ K, though there is a small fraction of WHIM reaching a temperature of $10^7 - 10^8$ K, which can be attributed to the high resolution and large box size of SIMBA-100. In the 300-SIMBA high-resolution test run, we find a branch of WHIM gas spanning from $(\log_{10}\rho_b/\bar{\rho}_b, \log_{10}T) = (1, 6)$ to $(\log_{10}\rho_b/\bar{\rho}_b, \log_{10}T) = (0, 7)$, implying the effect of stronger jet feedback in 300-SIMBA. The branch, however, only contributes a negligible portion of the total gas mass. A similar trend is found in the 300-SIMBA low-resolution test run, but the limitation on resolution makes the pattern less notable. We conclude that the modification of feedback strength mainly leads to a population of WHIM gas with $T > 10^7$ K at redshift $z = 2.5$, and a lower resolution is likely to have an impact on the diffuse cold gas.

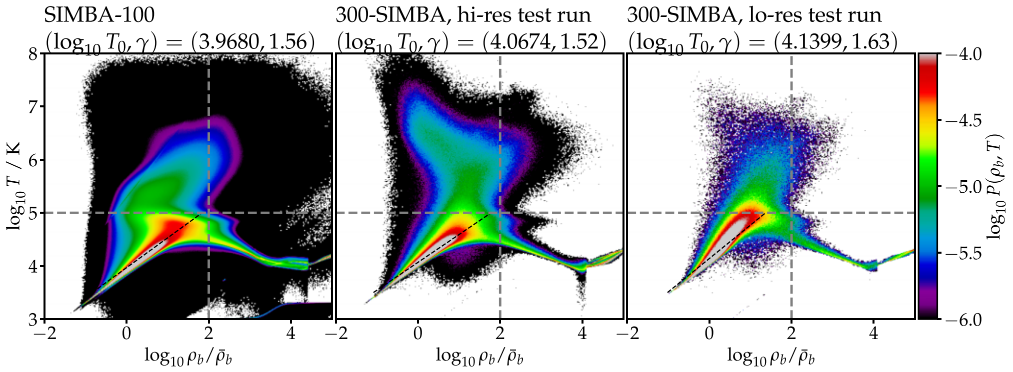

Figure C1. From left to right: the mass-weighted $z = 2.5$ temperature–density relation for SIMBA-100 and two (high resolution and low resolution) 300-SIMBA runs with ICs that are roughly at mean cosmic density. In each panel, the vertical grey dash line represents $\log_{10} \rho_b/\bar{\rho}_b = 2$ while the horizontal line is $T = 10^5$K. The two dash lines divide the $\rho - T$ phase space into diffused cold gas (lower left), warm-hot intergalactic medium (WHIM, upper left), hot halo gas (upper right), and condensed gas (lower right). In each panel, we add a reference dashed line to emphasize the linear $\rho - T$ relation of the diffuse cold gas regime and list the best-fitting parameters, $\log_{10} T_0$ (in Kelvin) and γ (defined in Section 1). All three panels show a similar pattern in the WHIM regime, while the low-resolution run has a distinct $\rho - T$ relation in the diffuse cold gas regime.

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