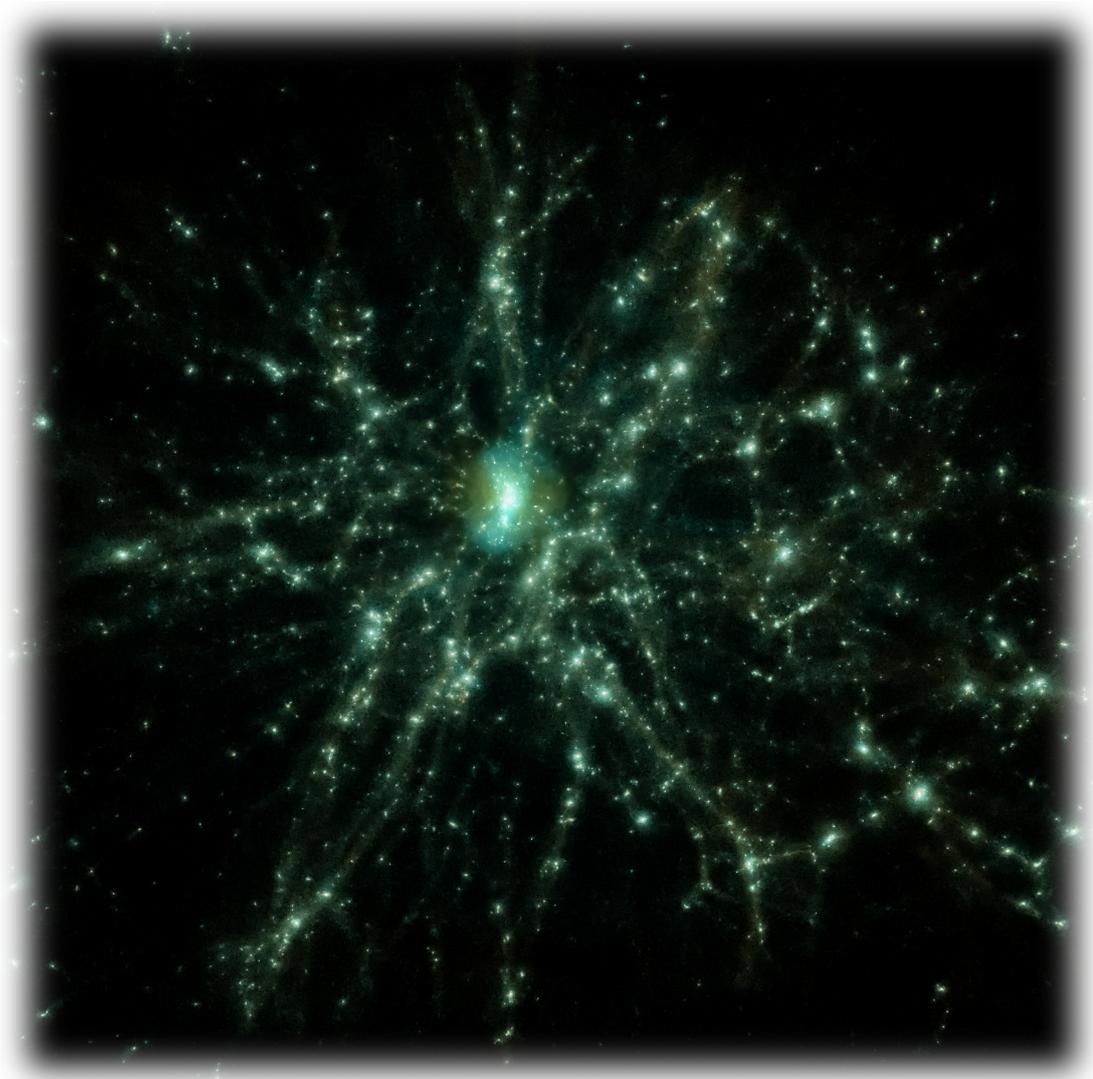




University of
Nottingham

UK | CHINA | MALAYSIA

**Searching for the Shadows of Giants:
The Absorption Signatures of
Protoclusters in the Lyman-alpha Forest**



Joel Sebastian Arthur Miller

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy, June 2021.

Cover: Image showing all of the gas, star, black hole and dark matter particles at redshift $z = 2.4$ that will collapse to form an $M = 10^{14.12} M_{\odot}$ cluster at $z = 0$ in the Sherwood (80-1024-ps13) simulation. This object is also featured in Fig. 3.3 from Chapter 3. Gas particles are coloured in cyan, dark matter particles in orange, star particles in white and black hole particles in black. Two images were rendered using the PyQtGraph¹ Python package with different alpha values for the particles, before being combined and post-processed in Affinity Photo².

¹<http://www.pyqtgraph.org/>

²<https://affinity.serif.com/en-gb/photo/>

Contents

Abstract	v
Acknowledgements	vii
List of Tables	viii
List of Figures	ix
Abbreviations	xii
Chapter 1 Introduction	1
1.1 The Λ CDM Cosmological Model	1
1.2 Lyman- α Forest	5
1.2.1 Definition and Properties	5
1.2.2 Observational Properties	8
1.2.3 The Proximity Effect	10
1.2.4 The Fluctuating Gunn-Peterson Approximation	11
1.3 Protoclusters	12
1.3.1 Definition and Properties	12
1.3.2 Identification of Protoclusters	14
1.4 Cosmological Simulations	21
1.5 Summary	23
Chapter 2 The effect of local ionisation on the Lyman-α absorption signatures of protoclusters at redshift $z \sim 2.4$	25
2.1 Introduction	25
2.2 Simulating Ly- α absorption from protoclusters	27
2.2.1 Cosmological hydrodynamical simulations	27
2.2.2 Protocluster identification	28
2.2.3 Local ionisation models	28
2.2.4 Mock Ly- α absorption spectra	33
2.3 Ly- α absorption profiles around haloes	34
2.4 The effect of box size and mass resolution on Ly- α absorption around haloes	37
2.5 The effect of local ionisation on the Ly- α transmission around protoclusters	40

2.5.1	Smoothed Ly- α forest tomographic maps	40
2.5.2	Protocluster masses and the correlation between LAEs and Ly- α transmission in smoothed tomographic maps	46
2.6	Summary	51
Chapter 3	Characterising Protoclusters With Line of Sight Lyman-α Absorption	55
3.1	Introduction	55
3.2	Hydrodynamical simulations	56
3.2.1	Sherwood	58
3.2.2	EAGLE	59
3.2.3	Illustris	59
3.2.4	Generation of mock Ly- α absorption sight-lines	60
3.3	The HI column density distribution function	61
3.4	Simulated protoclusters	63
3.4.1	Protocluster definitions	63
3.4.2	The size and shape of simulated protoclusters	64
3.5	Characterising mass overdensities in Ly- α absorption on 15 h^{-1} cMpc scales	68
3.5.1	Relationship between mass and effective optical depth on 15 h^{-1} cMpc scales	68
3.5.2	CoSLAs: are they mass overdensities?	72
3.5.3	Effect of box size and mass resolution on CoSLAs	76
3.6	Characterising protoclusters in Ly- α absorption	80
3.7	Selecting protoclusters with line of sight Ly- α absorption	84
3.8	Summary	90
Chapter 4	Characterising Protoclusters With Line of Sight Lyman-α Absorption Redux - The Impact of Box Size and the Most Massive Protoclusters	93
4.1	Introduction	93
4.2	Hydrodynamical Simulations	94
4.2.1	Generation of mock Ly- α absorption sightlines	97
4.3	Protocluster Properties	97
4.4	Characterising mass overdensities in Ly- α absorption on 15 h^{-1} cMpc scales	99
4.5	The Fraction of CoSLAs in Protoclusters	105
4.6	Summary	105
Chapter 5	Conclusions and Outlook	108
5.1	Key Conclusions	108
5.2	Limitations of this Thesis	111

5.3 Outlook	112
Bibliography	116

Abstract

In this thesis, I use large scale state-of-the-art hydrodynamical cosmological simulations to investigate the Ly- α absorption properties of protocluster regions at redshift $z \sim 2.4$. Using the IllustrisTNG simulations, I investigate how the AGN proximity effect and hot, collisionally ionised gas arising from gravitational infall and black hole feedback changes the Ly- α absorption associated with the $z \simeq 2.4$ progenitors of $M \simeq 10^{14} M_{\odot}$ clusters. Within these protoclusters, I find that galaxy overdensities exhibit a weak anti-correlation with Ly- α transmission in tomographic maps of the intergalactic medium (IGM), but local H I ionisation enhancements due to hot $T > 10^6$ K gas or nearby AGN can disrupt this relationship within individual protoclusters. On average, however, I find that strong reductions in the IGM neutral fraction are limited to within $\lesssim 5h^{-1}$ cMpc of the dark matter haloes. Local ionisation enhancements will therefore have a limited 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 may result in systematically lower estimates of the $z = 0$ descendant mass of the most massive protoclusters.

I then use simulations from the Sherwood, EAGLE and Illustris projects to examine the Ly- α absorption signatures of $z \sim 2.4$ protoclusters with a descendant mass of $M \simeq 10^{14} M_{\odot}$. I find there is a weak correlation between the mass overdensity, δ_m , and the Ly- α effective optical depth relative to the mean, $\delta_{\tau_{\text{eff}}}$, averaged over $15 h^{-1}$ cMpc scales; however, the scatter in the δ_m - $\delta_{\tau_{\text{eff}}}$ plane means it is not possible to uniquely identify large scale overdensities with strong Ly- α absorption. Assuming perfect removal of contamination by Ly- α absorbers with damping wings, more than half of the remaining sightlines with $\delta_{\tau_{\text{eff}}} > 3.5$ (CoSLAs) trace protoclusters. It is furthermore possible to identify a model dependent $\delta_{\tau_{\text{eff}}}$ threshold that eliminates the contamination from filamentary structure and

Lyman-limit systems that are preferentially aligned along the line of sight. However, such regions are rare: excluding absorption caused by damped systems, less than 0.1% of sightlines that pass through a protocluster have $\delta_{\tau_{\text{eff}}} > 3.5$, meaning that any protocluster sample selected in this manner will also be highly incomplete.

Finally, I use the TNG300-1 simulation from the IllustrisTNG project to study the extent to which simulation box size—and the resulting presence of the most massive protoclusters—is responsible for the differences between the CoSLA populations observed in large N-body simulations and those in comparatively small hydrodynamical simulations. Using the better statistics afforded by the larger simulation volume, I confirm that CoSLAs are rare objects, finding a number of CoSLAs per unit redshift $dn/dz \sim 1 \times 10^{-2}$ at redshift $z = 2.444$. Furthermore, I find the $\delta_{\tau_{\text{eff}}}-\delta_{\text{m}}$ relation in a $205^3 h^{-3} \text{Mpc}^3$ volume is consistent with that for a smaller $75^3 h^{-3} \text{Mpc}^3$ box size, with systems with $\delta_{\text{m}} > 1$ existing over a wide range of $\delta_{\tau_{\text{eff}}}$ values. I find the CoSLAs in TNG300-1 to have a median $\delta_{\text{m}} = 0.45 \pm 0.05$, which remains somewhat lower than the $\delta_{\text{m}} = 0.64 \pm 0.38$ previously found in larger N-body simulations. These findings suggest that the differences between the CoSLA populations in hydrodynamical and N-body simulations are not solely the result of the differing simulation box sizes. Instead I suggest that these discrepancies are caused by the modelling of the hot, collisionally ionised, gas that surrounds massive haloes in protocluster regions.

Acknowledgements

First and foremost, I would like to thank my supervisors Jamie and Nina for their amazing support, advice and mentoring over the past 4 years. I am immensely grateful for all of the time and effort you have dedicated towards my education. Additional thanks to Nina for also being a fantastic undergraduate tutor for the 4 years prior to my PhD, as well as inviting me to go observing at the William Herschel Telescope for a week during the first year of my PhD. I'm sure you will be glad to be rid of me at last! I would also like to acknowledge STFC for funding my PhD. Thanks also go to the community over at Stack Overflow (as well as the LaTeX Stack Exchange) for providing the answers to countless coding questions. It would also be wrong of me to not acknowledge all of the IT support that Phil Parry has provided me and others with over the past 4 years, particularly as we were all switching to remote working at the beginning of the pandemic.

Thank you to my family for supporting me throughout my lengthy education, without whom I would not be where I am today. On the topic of the pandemic, special thanks to my mum for providing me with a place to hide out the worst of it and to my brother for equipping it with a gym.

I also want to thank everyone at UoNU for making my time in Nottingham a truly life-changing experience. To Liam and Ben for being instrumental in getting me hooked on ultimate, such that it became a core component of my life; to George, my UoN teammate for as many as eight consecutive seasons, and to everyone else who has come and gone since September 2013. Furthermore, I'd like to thank all my teammates from East Midlands Open and GB Men (and indeed numerous players from other teams) for many memorable experiences both on and off the pitch during the course of my studies.

I've been fortunate enough to have made many great friends while in Nottingham (some of whom didn't even play ultimate). Whilst I don't have space to thank you all for your support, I would like to thank Vicky, Tomáš, Liam and Maya for assisting me with the proofreading of this thesis, as well as my various housemates over the course of my PhD for the countless hours of entertainment. Particular thanks also go to my best friend Emma, who has provided me with a level of emotional support over the years that I doubt I'll ever be able to fully repay.

Last but not least, a huge thank you to everyone in CAPT for making it a fantastic place to work over the past 4 years! I am grateful to everyone who has helped me with science/programming questions and I will miss the lunchtime and office discussions greatly.

List of Tables

2.1	Hydrodynamical simulations used in Chapter 2	27
2.2	Fraction of protoclusters from the TNG100-1 simulation that would be selected according to the thresholds defined by Lee et al. (2016) and Newman et al. (2020)	48
3.1	Hydrodynamical simulations used in Chapter 3	57
3.2	Hydrodynamical simulations used to test the effect of box size and mass resolution on CoSLAs	77
3.3	All CoSLAs in the Sherwood simulation that are associated with protoclusters	90
4.1	Hydrodynamical simulations used in Chapter 4	95

List of Figures

1.1	Illustration depicting the history of the Universe	2
1.2	The CMB temperature power spectrum	3
1.3	Formation of an $M \sim 10^{15} M_{\odot}$ cluster from the Millennium-II simulation	4
1.4	An example Ly- α forest spectra	6
1.5	The onset of Gunn-Peterson troughs in QSO spectra	7
1.6	Two DLAs in the spectrum of QSO PSS0209+0517	9
1.7	2-D projection of a tomographic IGM map from Mukae et al. (2020a)	10
1.8	Renderings of the IGM opacity in the LATIS fields	19
1.9	The BOSS1441 galaxy overdensity at redshift $z = 2.32 \pm 0.02$ from Cai et al. (2017b)	20
2.1	2-D projections of an $M_{z=0} = 10^{14.46} M_{\odot}$ protocluster in the TNG100-1 simulation at redshift $z = 2.44$	29
2.2	Radial Ly- α transmission profiles around haloes in TNG100-1 at redshift $z = 2.44$	35
2.3	Comparison of radial Ly- α transmission profiles around haloes at redshift $z = 2.44$ between the TNG100-1, TNG100-2 and TNG100-3 simulations	38
2.4	Comparison of radial Ly- α transmission profiles around haloes at redshift $z = 2.44$ between the TNG100-2 and TNG300-1 simulations	39
2.5	Smoothed tomographic maps showing δ_{F}/σ for three protoclusters in the TNG100-1 simulation at redshift $z = 2.44$	42
2.6	The Ly- α luminosity function for all galaxies in the TNG100-1 simulation at $z = 2.44$	43
2.7	The two-point correlation function for LAEs in the TNG100-1 simulation at $z = 2.44$	44
2.8	The relationship between minimum relative Ly- α transmission and $z = 0$ cluster mass	47
2.9	The relationship between relative transmission and LAE overdensity	50
3.1	The H I column density distribution function at $z \simeq 2.4$	62

3.2	The size and shape of simulated protoclusters in Sherwood, EAGLE and Illustris	64
3.3	Projected maps of protoclusters in the Sherwood, EAGLE and Illustris simulations	66
3.4	Gas properties for protocluster and field regions in the Sherwood, EAGLE and Illustris simulations	67
3.5	The relationship between δ_m and $\delta_{\tau_{\text{eff}}}$ on $15 h^{-1}$ cMpc scales	69
3.6	Absorber classification fraction as a function of δ_m	71
3.7	Mass PDF for CoSLAs in the Sherwood simulation	73
3.8	2-D projections of CoSLAs from an overdense and an underdense region in the Sherwood simulation	74
3.9	Comparison between CoSLAs in the fiducial and QLy- α Sherwood models	76
3.10	Comparison of CoSLA mass PDFs in an $80 h^{-1}$ cMpc and $160 h^{-1}$ cMpc side length simulation box	78
3.11	Comparison of CoSLA mass PDFs at three different mass resolutions	79
3.12	Simulated sightline segment of a CoSLA at two different mass resolutions	80
3.13	Fraction of absorption systems associated with protoclusters and their classification as a function of $\delta_{\tau_{\text{eff}}}$	81
3.14	Example of four different $15 h^{-1}$ cMpc sightline segments that pass through protoclusters in the Sherwood simulation	83
3.15	Mass PDF of CoSLAs in protoclusters vs those in the field from the Sherwood simulation	85
3.16	Contamination and “completeness” of a protocluster sample selected from all segments above a $\delta_{\tau_{\text{eff}}}$ threshold	86
3.17	Reverse cumulative distribution functions for all $15 h^{-1}$ cMpc associated with protoclusters in the Sherwood simulation	87
3.18	Example of four different CoSLAs that pass through protoclusters in Sherwood	89
4.1	The H I column density distribution function at $z = 2.444$	96
4.2	The size and shape of protoclusters in TNG300-1	98
4.3	The δ_m - $\delta_{\tau_{\text{eff}}}$ relation for TNG100-1, TNG100-2 and TNG300-1	100
4.4	The relationship between 3-D mass and 1-D effective optical depth from Cai et al. (2016)	101
4.5	Mass PDFs of CoSLAs from TNG100-1, TNG100-2, TNG300-1 and Chapter 3	102

4.6	Mass PDFs of CoSLAs in protocluster vs the field from TNG100-1, TNG100-2 and TNG300-1	104
-----	--	-----

Abbreviations

Λ CDM Lambda Cold Dark Matter.

Ly- α Lyman-Alpha.

AGN Active Galactic Nuclei.

CDDF Column Density Distribution Function.

CGM Circumgalactic Medium.

CMB Cosmic Microwave Background.

CoM Centre of Mass.

CoSLA Coherently Strong Lyman-alpha Absorption System.

DLA Damped Lyman-Alpha system.

DSFG Dusty Star Forming Galaxy.

FGPA Fluctuating Gunn-Peterson Approximation.

FoF Friends-of-Friends.

IGM Intergalactic Medium.

IMF Initial Mass Function.

LAB Lyman-Alpha Blob.

LAE Lyman-Alpha Emitter.

LLS Lyman-Limit System.

PDF Probability Distribution Function.

QSO Quasi-Stellar Object.

SFR Star Formation Rate.

SLLS Super Lyman-Limit System.

SPH Smoothed Particle Hydrodynamics.

UV Ultraviolet.

Chapter 1

Introduction

“Given enough time, hydrogen starts to wonder where it came from, and where it is going.”

—Edward R. Harrison

1.1 The Λ CDM Cosmological Model

The current understanding of cosmology points towards a flat Λ CDM Universe (Planck Collaboration et al., 2020), within which structure formation proceeds in a hierarchical fashion. The evolution of the Universe from the Big Bang at $t = 0$ (left) to the present epoch (right) is depicted in Fig. 1.1. The very left-hand side of Fig. 1.1 shows the epoch of cosmic inflation (Guth, 1981; Linde, 1982). During this epoch a rapid expansion of spacetime takes place between 10^{-36} seconds and 10^{-32} seconds after the Big Bang. This process amplifies the primordial quantum fluctuations to macroscopic scales, resulting in small density fluctuations throughout the underlying matter distribution of the Universe (Mukhanov and Chibisov, 1981).

One second after the Big Bang, the first baryons formed creating the coupled photon-baryon soup of the early Universe. Within this opaque plasma, fluctuations in the matter density field propagated at the sound speed as baryon acoustic oscillations (Peebles and Yu, 1970; Sunyaev and Zeldovich, 1970). Some 380,000 years after the Big Bang the Universe cooled to a temperature that allowed the first atoms to form, coinciding with the decoupling photons and baryons from

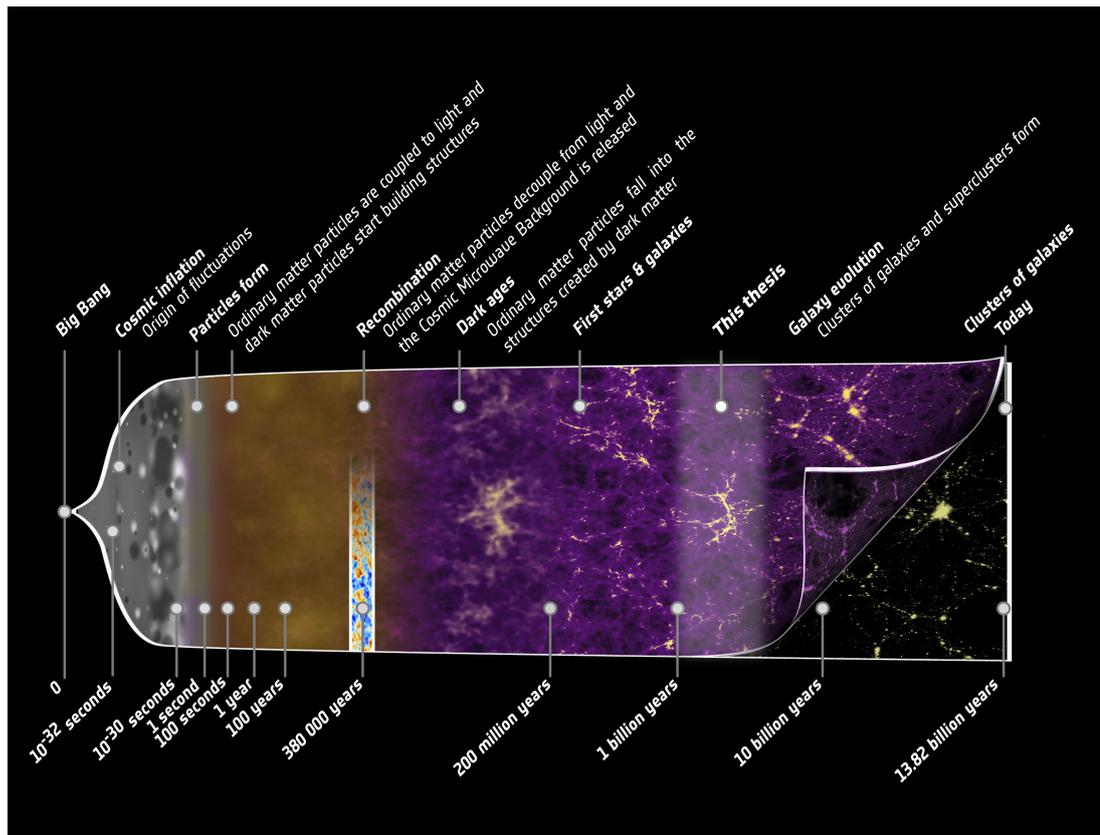


Figure 1.1: Illustration depicting the history of cosmic structure formation. Key events are indicated along the top with the age of the Universe at various points indicated along the bottom. The highlighted region from ~ 1 – 5 billion years represents approximately where the work contained within this thesis sits. Credit: ESA and the Planck Collaboration¹

one another. This process caused the Universe to become transparent, ‘freezing’ the density perturbations into the dark matter distribution and leaving behind the cosmic microwave background (CMB). This epoch is indicated on Fig. 1.1 by the signature red-green-blue contours around one third of the way from the left-hand side to the right-hand side. The CMB is the isotropic and homogeneous (on large scales) radiation field left over from the Big Bang. It was first discovered serendipitously by [Penzias and Wilson \(1965\)](#) and has since been shown to be a near perfect blackbody emitter with a temperature of $T \sim 2.7$ K at the present day ([Fixsen, 2009](#)). The CMB can be characterised by its power spectrum, shown in Fig. 1.2, which encodes the contribution to the anisotropies as a function of angular scale. The CMB power spectrum has provided cosmologists with a key observable of the early Universe that facilitates measurements of cosmological parameters. For example the locations at which the various peaks occur can be used to determine the curvature of the Universe ([Hu and White, 1997](#)).

¹https://www.esa.int/Science_Exploration/Space_Science/Planck/History_of_cosmic_structure_formation

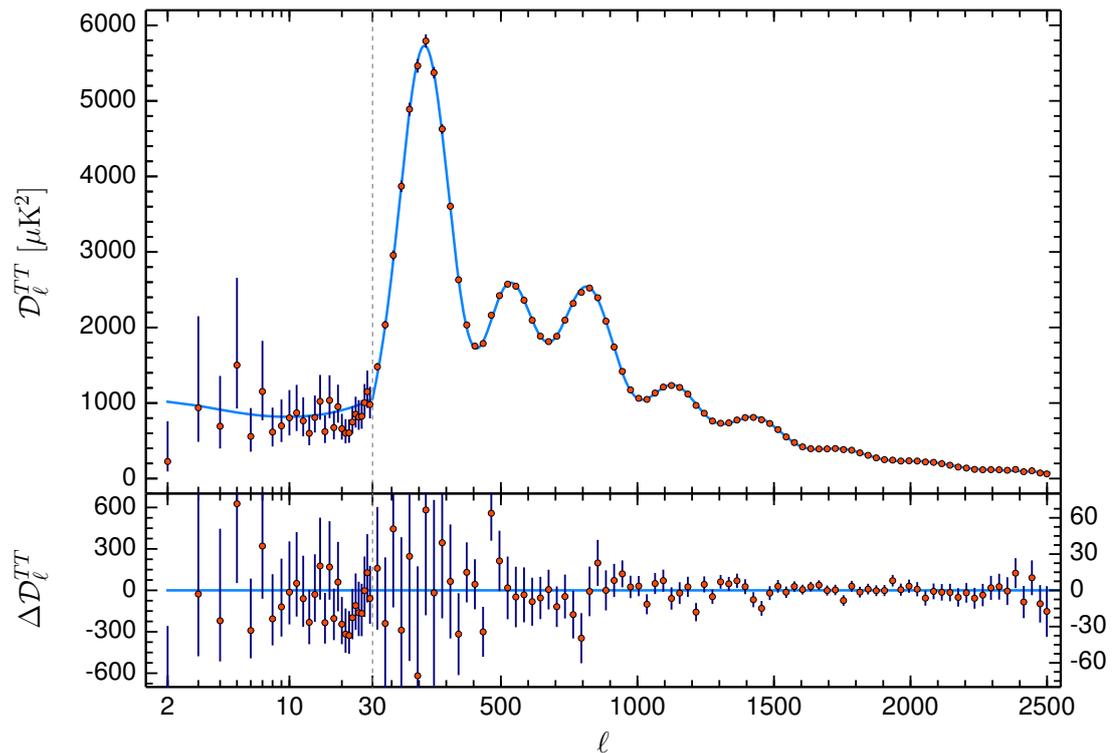


Figure 1.2: Top: The CMB temperature power spectrum from [Planck Collaboration et al. \(2020\)](#). The solid blue line shows the best fit base- Λ CDM theoretical spectrum to the Planck TT,TE,EE+lowE+lensing likelihoods. Bottom: The residuals with respect to the best fit model.

Moving further towards the right-hand side of Fig. 1.1, density perturbations grow with the expansion of the Universe until they reach a maximum radius of R_{ta} at t_{ta} (the turnaround epoch). At this point, they begin to collapse to form virialised haloes with $R_f = R_{\text{ta}}/2$ at $t = t_f = 2t_{\text{ta}}$ ([Kravtsov and Borgani, 2012](#)). This collapse takes place over a timescale, $t_f - t_{\text{ta}}$, which is on the order of the free-fall time, t_{ff} . For a uniform sphere of radius R_{ta} and mass M , the free-fall time is $t_{\text{ff}} = \pi\sqrt{R_{\text{ta}}^3/8GM} = \pi\sqrt{R_f^3/GM}$, showing that less dense regions will collapse over a longer time period than more dense ones. As larger overdensities tend to be less dense than smaller ones, this density dependence of the free-fall time explains why smaller objects formed before larger ones. For an $M \sim 10^{14} M_\odot$ cluster with $R_f \sim 1$ Mpc, the associated free-fall time is $t_{\text{ff}} \sim 5$ Gyr. As such, these objects only formed recently ($t_{\text{coll}} \sim 10$ Gyr), while smaller—galaxy-scale—haloes formed at a much earlier epoch. Once formed, dark matter haloes provide potential wells into which baryonic matter falls. As the primordial hydrogen and helium gas cools to $T \sim 10$ – 20 K, clouds of molecular gas are able to coagulate and eventually collapse down to form the first stars and galaxies ([Couchman and Rees, 1986](#)).

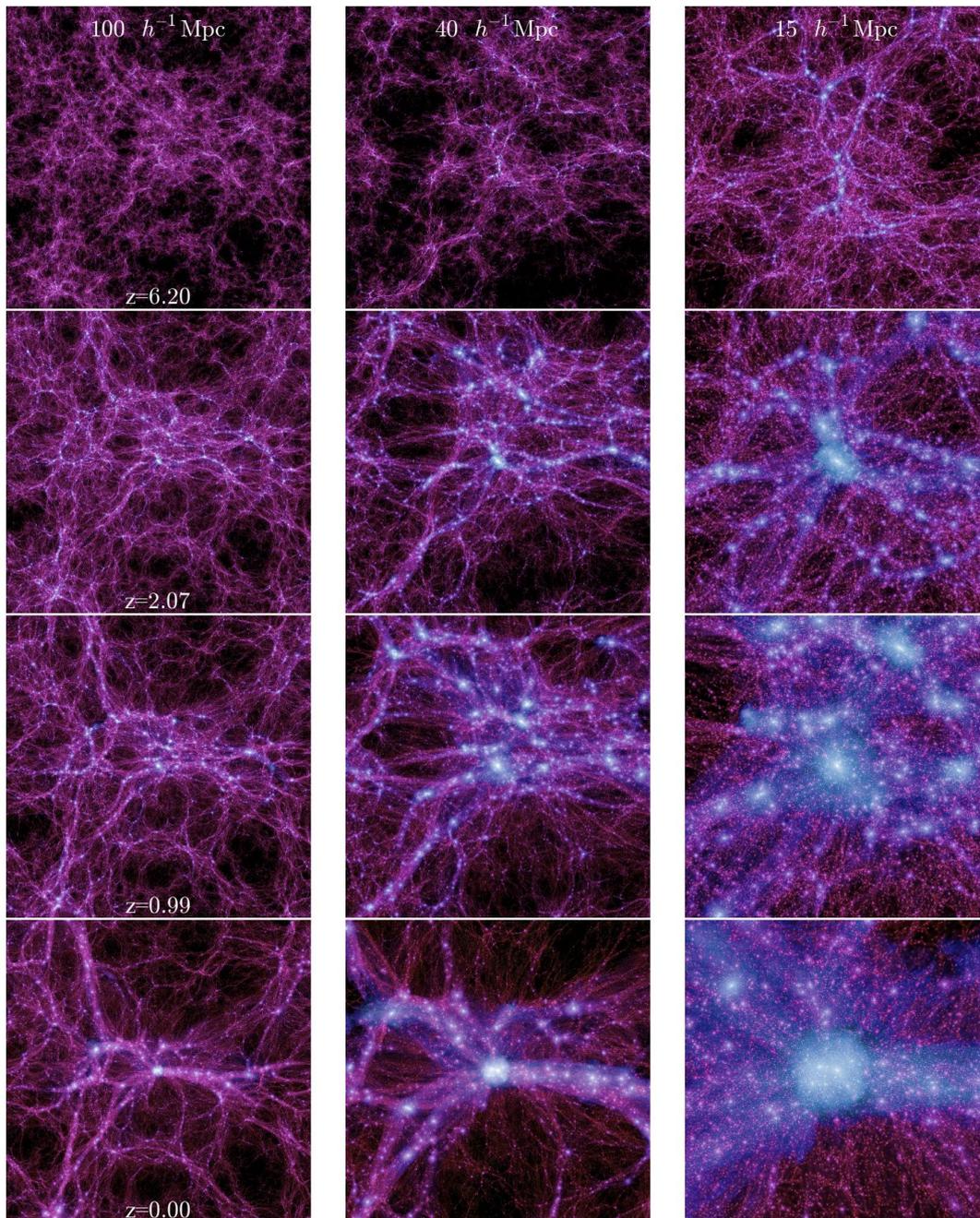


Figure 1.3: The formation of an $M \sim 10^{15} M_{\odot}$ cluster from the Millennium-II simulation. Shown at four different epochs (top to bottom: $z = 6.2, 2.07, 0.99,$ and 0) on three different scales (left to right: $100, 40$ and $15 h^{-1} \text{cMpc}$). Figure taken from [Boylan-Kolchin et al. \(2009\)](#).

As the first stars, galaxies and quasars are formed, the intense ultraviolet (UV) radiation emitted by these sources began the process of reionisation (Loeb and Barkana, 2001) at redshift $z \sim 10\text{--}30$. The process of reionisation transformed the intergalactic medium (IGM) from atomic hydrogen into a highly ionised state that is optically thin to Lyman continuum photons (Loeb and Barkana, 2001; Barkana and Loeb, 2001; Morales et al., 2010). The epoch of reionisation also marks a period in the history of the Universe when baryonic matter became a significant component in the formation and evolution of structure, due to the feedback from luminous sources producing of order one photon per baryon (Morales et al., 2010).

As cosmic time progressed, the numerous dark matter haloes distributed throughout the Universe were drawn together via gravitational forces, merging in turn to form more and more massive haloes (e.g. Somerville et al., 2012). A visual representation of this process, as observed in the Millennium-II dark matter simulation, is shown in Fig. 1.3. This shows the formation of a $M \sim 10^{15} M_{\odot}$ cluster from redshift $z = 6.2$ to $z = 0$. The galaxies within these dark matter haloes also undergo mergers themselves, over time evolving from star-forming spiral (late-type) galaxies into quiescent elliptical (early-type) galaxies (see Conselice, 2014, for a review). The present day Universe is found to have $\Omega_{\Lambda} = 0.685$, $\Omega_{\text{m}} = 0.315$ and $\Omega_{\text{b}} = 0.049$ (Planck Collaboration et al., 2020)—where Ω_{Λ} , Ω_{m} and Ω_{b} represent the fraction of the total matter-energy density of the Universe that is attributed to dark energy, matter and baryonic matter respectively—expanding at a rate quantified by the Hubble constant $H_0 = 67.36 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Finally, at the right-hand edge of Fig. 1.1, the merging of dark matter haloes culminates in the massive galaxy clusters observed in the present day Universe.

1.2 Lyman- α Forest

1.2.1 Definition and Properties

The Lyman- α (Ly- α) forest is the name given to a collection of absorption lines, each corresponding to the $n = 1$ to $n = 2$ transition of neutral hydrogen, that are seen in the spectra of quasi-stellar objects (QSOs) (Rauch, 1998). This phenomenon was first predicted by Gunn and Peterson (1965); Scheuer (1965); Shklovskii (1965) independently of one another, with these predictions promptly being followed by observations of discrete Ly- α lines by Lynds and Stockton (1966); Burbidge et al. (1966); Stockton and Lynds (1966); Kinman (1966). The

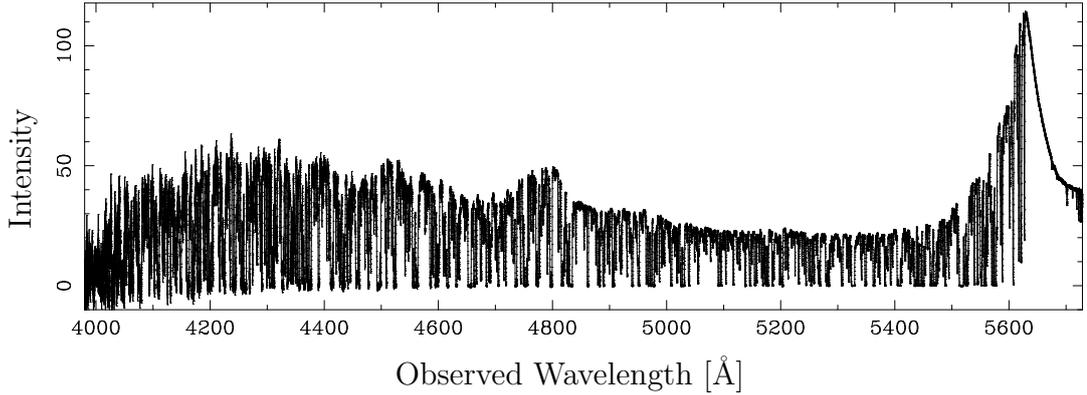


Figure 1.4: An example of the Ly- α forest, as seen in the high resolution spectrum of QSO1422+23 which lies at redshift $z = 3.62$. The rest-frame Ly- α emission peak of the QSO is at $\lambda \sim 5650 \text{ \AA}$, with the Ly- α absorption lines corresponding to intervening neutral gas clouds seen at shorter wavelengths. Figure taken from Rauch (1998).

suggestion that the bulk of the absorption seen in QSO spectra blueward of the rest frame Ly- α line was caused by intervening H I absorption (Lynds and Wills, 1970) was eventually confirmed by the detection of corresponding higher order Lyman series lines (e.g. Baldwin et al., 1974). The Ly- α absorption line occurs at a wavelength of 1216 \AA , corresponding to the energy gap of 10.2 eV . As light from a background QSO passes through multiple clouds of neutral hydrogen gas at slightly different redshifts, absorption lines are created at progressively shorter rest frame wavelengths. The result seen in the observed spectra is hundreds of distinct absorption lines blueward of 1216 \AA in the rest frame of the QSO. An example of such as spectra is shown in Fig. 1.4. As such, this phenomenon was eventually given the name “the Ly- α forest” by Weymann et al. (1981).

Since its discovery, the Ly- α forest has been an invaluable tool to astronomers and cosmologists. The Ly- α forest provides direct observational evidence of the IGM which accounts for the majority of the baryonic matter in the high redshift Universe according to the current cosmological paradigm. In a Λ CDM Universe undergoing hierarchical structure formation, as one moves further back in time the matter content of the Universe will tend to move from being concentrated in galaxies to being part of the diffuse IGM. If the IGM was composed of neutral H I, a ‘Gunn-Peterson trough’ (a spectral region of zero transmission) would be observed blueward of the Ly- α emission of the QSO. However, given that a Ly- α forest is observed instead, this suggests that the IGM instead exists as a highly ionised gas in photoionisation equilibrium with the cosmic UV background,

$$n_e n_p \alpha(T) = n_{\text{HI}} \Gamma. \quad (1.1)$$

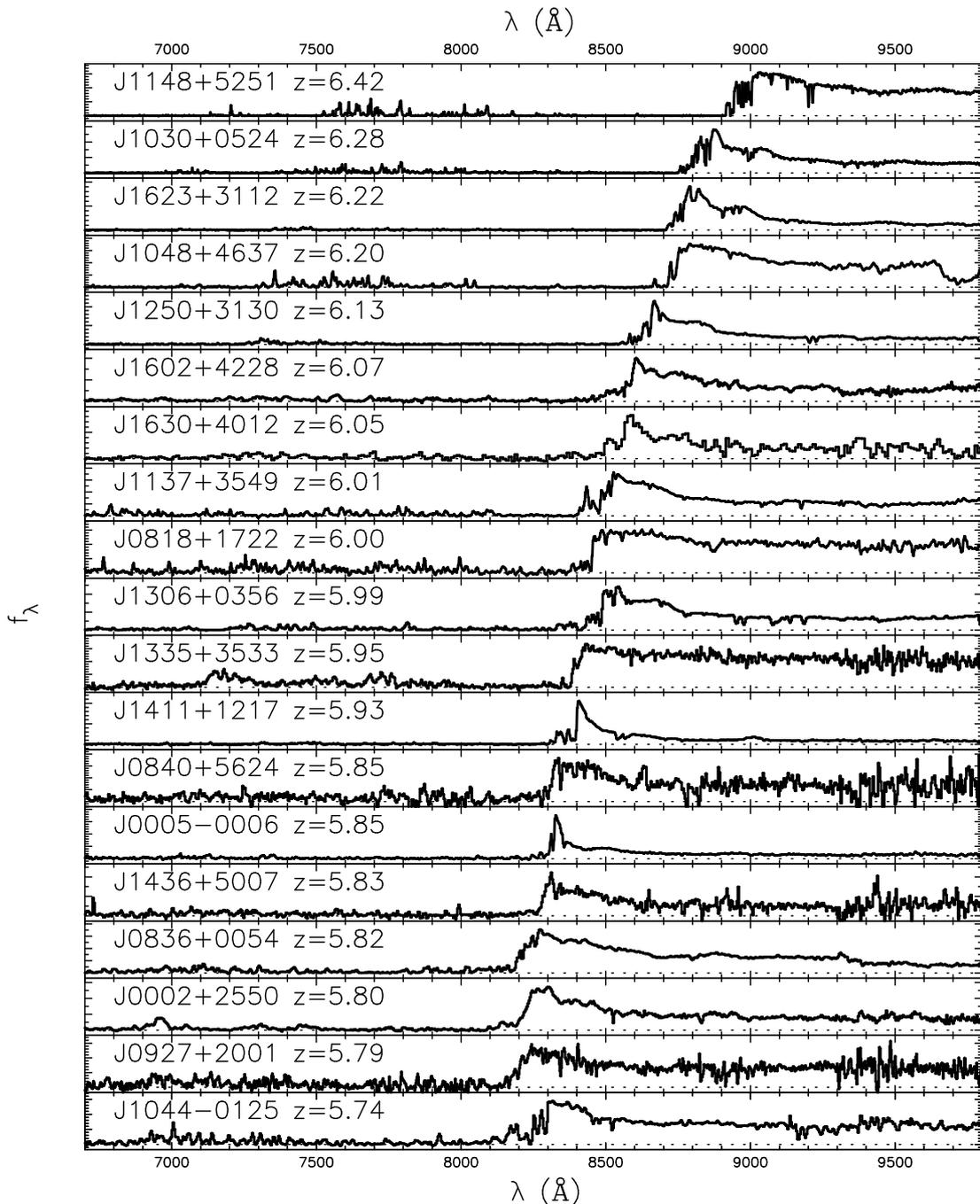


Figure 1.5: A sample of 19 QSO spectra from $z = 5.74$ to $z = 6.42$ showing the onset of Gunn-Peterson troughs. Figure taken from [Fan et al. \(2006\)](#).

Here, Γ is the photoionisation rate and α is the temperature dependent recombination coefficient, whilst n_e , n_p and n_{HI} refer to the number densities of electrons, protons and neutral hydrogen atoms respectively. This photoionisation equilibrium will apply at times greater than the recombination time, which is of order $t_{\text{rec}} \sim 10^4$ years for a highly ionised IGM.

The IGM has only existed in this highly ionised state since the epoch of reionisation, when the first luminous sources in the Universe were formed. The Ly- α

forest provides astronomers with one of the main observables of cosmic reionisation in the form of Gunn-Peterson troughs—extended saturated regions blueward of the QSO Ly- α emission line, caused by the intervening IGM having a neutral fraction significantly greater than $\sim 10^{-4}$. When observing QSO spectra at redshifts $z > 5$, one finds a dramatic increase in the opacity of the Ly- α forest with redshift (Becker et al., 2001; Fan et al., 2006; Mortlock et al., 2011; Becker et al., 2015). By measuring the onset of these Gunn-Peterson troughs (Fig. 1.5) astronomers have been able to determine that cosmic reionisation had completed around redshift $z \sim 5.5\text{--}6$ (Fan et al., 2001; Kulkarni et al., 2019).

As the amount of transmission observed in the Ly- α forest is sensitive to the amount of neutral hydrogen—and therefore the amount of baryonic matter—in the IGM, it is possible to make predictions on the values of Ω_b and Γ , by comparing measurements of the mean flux decrement present in Ly- α forest spectra extracted from cosmological simulations to that which is found in observations (Rauch et al., 1997).

1.2.2 Observational Properties

The simplest observable quantity of the Ly- α forest is the mean normalised flux $\langle F \rangle = \langle I_{\text{observed}}/I_{\text{emitted}} \rangle$ over a given rest-frame wavelength interval, commonly expressed in terms of an effective optical depth,

$$\tau_{\text{eff}} = -\ln \langle F \rangle \equiv \ln \langle e^{-\tau} \rangle, \quad (1.2)$$

where τ is the Ly- α optical depth in each pixel of an observed spectrum. The Gunn-Peterson optical depth is,

$$\tau_{\text{GP}} = \frac{\sigma_{\alpha} c}{H(z)} n_{\text{HI}}(z) \quad (1.3)$$

where σ_{α} is the Ly- α cross section. Hence, at a fixed redshift, the optical depth is directly proportional to the number density of H I in the scattering region, n_{HI} .

Ly- α absorbers can be grouped into three main classes based upon their column density. These are Ly- α forest systems, Lyman-limit systems, and damped Ly- α systems. Ly- α forest systems have column densities of $N_{\text{HI}} < 10^{17.2} \text{ cm}^{-2}$ and exist in a highly ionised state (neutral fraction $\lesssim 10^{-3}$). Gas clouds with column densities with $10^{17.2} \text{ cm}^{-2} \leq N_{\text{HI}} < 10^{19} \text{ cm}^{-2}$ are instead referred to as Lyman limit systems (LLSs) and, unlike Ly- α forest systems, they are optically thick at the Lyman continuum (Prochaska, 1999).

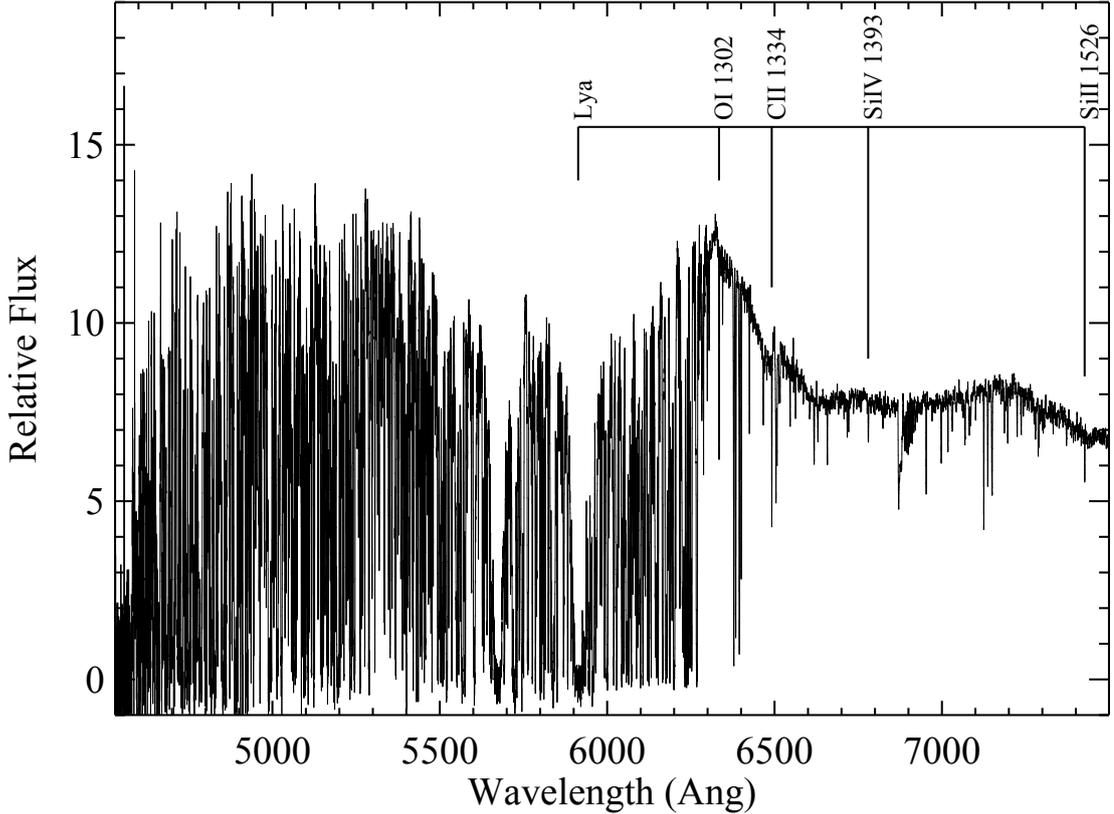


Figure 1.6: Example QSO spectrum (QSO PSS0209+0517) showing the Ly- α forest with a pair of DLAs at $\lambda = 5674 \text{ \AA}$ and $\lambda \simeq 5900 \text{ \AA}$. Figure taken from [Wolfe et al. \(2005\)](#)

Systems with even higher column densities ($N_{\text{HI}} > 10^{19} \text{ cm}^{-2}$) start to become almost entirely neutral due to self shielding. Self shielding of dense gas occurs when the recombination rate of ionised hydrogen within the cloud is sufficient to allow the outer layers of gas that are photoionised by UV radiation to recombine into neutral hydrogen. This prevents the ionising radiation from penetrating further into the gas cloud.

Systems with $10^{19} \text{ cm}^{-2} \leq N_{\text{HI}} < 10^{20.3} \text{ cm}^{-2}$ are known as super Lyman limit systems (SLLSs), or sub-DLAs, and those with $N_{\text{HI}} > 10^{20.3} \text{ cm}^{-2}$ as damped Ly- α systems (DLAs). Both of these classes of absorber correspond to physically small, highly dense, regions of hydrogen gas with ionised fractions approaching zero due to a high degree of self shielding ([Katz et al., 1996b](#)).

Fig. 1.6 shows an example QSO spectra containing two such DLAs. These DLAs represent high-redshift galaxies with a typical cross section of πr^2 with $r \sim 10 h^{-1} \text{ kpc}$ ([Schaye, 2001](#)). Despite their small physical extent, these systems produce absorption features with large damping wings that span $> 100 \text{ \AA}$. In the context of this thesis I refer to SLLSs and DLAs collectively as damped absorbers/systems.

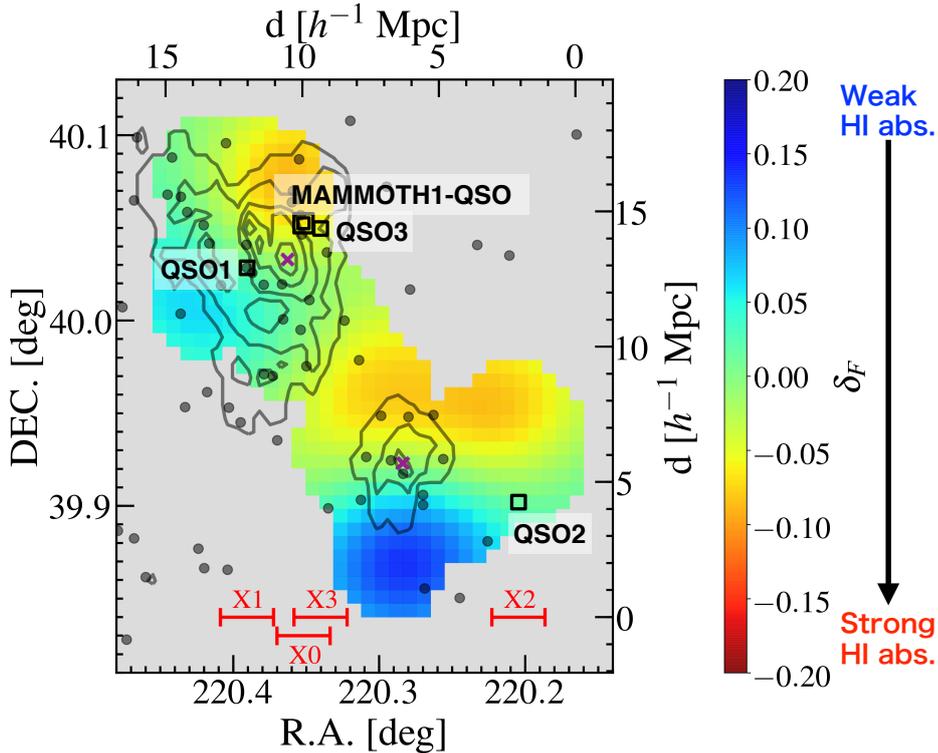


Figure 1.7: Figure 8 from Mukae et al. (2020a), showing a 2-D projection of a tomographic IGM map from redshift $z = 2.30$ to $z = 2.33$. The coloured contours show the reconstructed Ly- α flux overdensity, $\delta_F = (F/\langle F \rangle) - 1$ (where $\langle F \rangle$ is the mean Ly- α forest transmission), projected over this redshift range. Grey contours show the significance of overdensities of LAEs found by Cai et al. (2017a) (grey points) from 2σ to 6σ . Local LAE density maxima are marked as purple crosses. Black squares show the locations of coeval QSOs. The red bars labelled “X0”–“X3” show the R.A. ranges of slices shown in figure 7 from Mukae et al. (2020a).

1.2.3 The Proximity Effect

The present day IGM exists in a highly ionised state, predominantly due to photoionisation from the UV background. It has long been thought that sources of this photoionising UV background radiation include highly luminous QSOs (Arons and McCray, 1969; Rees et al., 1970) as well as star forming galaxies (Bolton et al., 2005). In regions nearby to the QSOs, the contribution to the photoionisation rate is,

$$\Gamma_{\text{HI}}^Q = \int \frac{d\nu}{h_p \nu} \frac{L_\nu^Q \sigma_{\text{HI}}}{4\pi r_L^2}, \quad (1.4)$$

where, L_ν^Q , is the luminosity of the QSOs at a frequency, ν ; r_L is the luminosity distance from said QSO, and σ_{HI} is the H I cross section. This dominates over the UV background value from fainter sources and therefore a region of enhanced photoionisation is created in the vicinity of the QSO.

This results in the so-called “proximity effect” (Bajtlik et al., 1988), where an increase in transmitted flux is observed in the Ly- α forest as one approaches the redshift of a QSO. This effect was first discussed under the name “inverse effect” by Weymann et al. (1981), who first proposed that this was due to increased ionisation nearby to a QSO, with measurements of the Ly- α forest absorption line density dN/dz by Carswell et al. (1982) supporting this hypothesis. Further observations made by Murdoch et al. (1986) and Hunstead et al. (1988) cemented the idea that the proximity effect is indeed caused by QSOs enhancing the ionisation rate in the regions surrounding them.

The proximity effect has provided yet another useful tool to astronomers as a measure of the UV background intensity (e.g. Bajtlik et al., 1988; Dall’Aglio et al., 2008; Calverley et al., 2011). More recently the proximity effect has also been proposed as a potential cause of offsets seen in Ly- α absorption tomographic IGM maps between the location of maximum galaxy density and the location of minimum reconstructed Ly- α flux (Mukae et al., 2020a, see also Chapter 2). This is shown in Fig. 1.7—revealing a $\sim 3 - 5 h^{-1}$ cMpc offset between the location of LAE and H I density peaks.

1.2.4 The Fluctuating Gunn-Peterson Approximation

A useful analytical model for the Ly- α forest is the Fluctuating Gunn-Peterson Approximation (FGPA; Weinberg et al., 1997; Croft et al., 1999) which I discuss again in Chapters 2 and 3. The FGPA allows one to directly link the Ly- α forest optical depth to a matter density Δ in the presence of a few reasonable approximations. First one must assume that the IGM is in photoionisation equilibrium with a spatially uniform UV background such that equation (1.1) applies, with the recombination rate being given by (Verner and Ferland, 1996),

$$\alpha(T) = 4.063 \times 10^{-13} \left(\frac{T}{10^4 \text{K}} \right)^{-0.72} \text{cm}^3 \text{s}^{-1}. \quad (1.5)$$

One then assumes the IGM to follow a strict temperature-density power law,

$$T = T_0 (\Delta)^{\gamma-1}, \quad (1.6)$$

where $\Delta \equiv \rho/\bar{\rho}$ is the normalised matter density (Hui and Gnedin, 1997). This is fairly robust for the low density IGM ($\Delta < 10$), where the assumption of photoionisation equilibrium holds and $T \sim 10^4$ K also means that the effects of collisional ionisation can be ignored. Under these conditions it is possible to

relate the neutral hydrogen number density, and therefore the optical depth, to the matter overdensity Δ . This gives the FGPA,

$$\tau_{\text{FGPA}}^{\alpha} \simeq 0.39 \Delta^{\beta} \frac{(1 + \mathcal{X}_{\text{He}})}{\Gamma_{-12}} \left(\frac{T_0}{10^4 \text{ K}} \right)^{-0.72} \left(\frac{\Omega_b h^2}{0.022} \right)^2 \times \left(\frac{\Omega_m h^2}{0.141} \right)^{-1/2} \left(\frac{1 - Y}{0.76} \right)^2 \left(\frac{1 + z}{3.4} \right)^{9/2}, \quad (1.7)$$

where $\beta = 2 - 0.72(\gamma - 1)$, $\Gamma_{-12} \equiv \Gamma_{\text{HI}}/10^{-12} \text{ s}^{-1}$, $\mathcal{X}_{\text{He}} = \eta Y/4(1 - Y)$ accounts for the contribution to the electron number density from helium (with $\eta = 1$ and $\eta = 2$ corresponding to singularly and doubly ionised helium respectively) and Y is the primordial helium fraction.

The FGPA has proved to be a valuable tool for modelling the Ly- α forest. Here Ω_m , Ω_b and h are set by the chosen cosmology (e.g. [Hinshaw et al., 2013](#); [Planck Collaboration et al., 2014, 2015, 2020](#)), T_0 and γ can be determined from observational measurements of the thermal history of the IGM (e.g. [Schaye et al., 2000](#); [Ricotti et al., 1999](#); [Garzilli et al., 2012](#); [Boera et al., 2014](#)), Γ_{-12} ([Haardt and Madau, 2012](#); [Faucher-Giguère et al., 2009](#); [Faucher-Giguère, 2020](#)) can be determined from observations of the luminosity function and IGM opacity, and Y is given from big bang nucleosynthesis ([Pospelov and Pradler, 2010](#)). It is then possible for one to calculate a Ly- α forest optical depth at a redshift z as a function of the matter overdensity.

1.3 Protoclusters

1.3.1 Definition and Properties

In the present day Universe, the largest virialised structures are galaxy clusters. A galaxy “cluster” is defined as a virialised collection of galaxies residing in a dark matter halo with a total mass greater than $10^{14} M_{\odot}$, whilst lower mass galaxy collections are referred to as “groups” ([Bower and Balogh, 2004](#)). Hierarchical structure formation in a Λ CDM Universe dictates that the progenitors of these galaxy clusters will be larger, more diffuse, structures composed of multiple lower mass dark matter haloes (e.g. [White and Frenk, 1991](#)) (see also Fig. 1.3). These “protoclusters” exist in the high redshift Universe as large, gravitationally bound but unvirialised, collections of dark matter haloes, spanning over $20 h^{-1} \text{ cMpc}$ at $z \sim 2$ ([Chiang et al., 2013](#); [Muldrew et al., 2015](#)) and representing the highest density regions on these scales at redshift $z \sim 2-3$. However, the term protocluster

is also used by some authors to refer only to the most massive progenitor (or main) halo (Muldrew et al., 2015). Many authors define a protocluster as a structure which has a density that is sufficiently large, and over a sufficient volume, for the structure to collapse to form a $M \geq 10^{14} M_{\odot}$ cluster by redshift $z = 0$ (Overzier, 2016). Chiang et al. (2013) used the Millennium Simulations to show that 80% of $(13.4 h^{-1} \text{cMpc})^3$ regions with densities greater than three times the mean density at redshift $z = 2$, will collapse to form a cluster with $M_{z=0} \geq 10^{14} M_{\odot}$. However, in this thesis I define protoclusters as the totality of matter at some redshift $z > 0$ that will collapse to form a cluster, where a cluster is defined as a massive halo with $M > 10^{14} M_{\odot}$ found using a friends-of-friends (FoF) algorithm at $z = 0$. The work within this thesis focuses on protoclusters at $z \sim 2.4$, which is chosen in order to be comparable to previous simulation work (Cai et al., 2016) and observational results (Cai et al., 2017b; Lee et al., 2018; Mukae et al., 2020a; Newman et al., 2020) at this epoch.

Protoclusters are objects of great interest to astronomers and cosmologists alike for a number of reasons. In the local Universe, it has long been known that dense environments have a strong influence on many galaxy properties. These include: morphology (e.g. Dressler, 1980; Goto et al., 2003), with cluster environments being dominated by elliptical (early-type) galaxies in contrast to the field which is dominated by spiral (late-type) galaxies; star-formation rate (e.g. Lewis et al., 2002; Tanaka et al., 2004), with the star-formation rate in high-density environments being significantly suppressed with respect to low-density environments, and active galactic nuclei (AGN) activity (Dressler et al., 1985; Kauffmann et al., 2004), where AGN activity is found to be enhanced in galaxies in low-density environments in comparison to those in clusters.

By studying the progenitors of cluster regions it is possible to learn more about the mechanisms that cause these prominent differences today. Indeed, studies of galaxies residing in protoclusters have shown that galaxies in these regions have different properties to those in the field (e.g. Steidel et al., 2005; Hatch et al., 2016; Shi et al., 2020), implying that the environments in which galaxies are situated affected their formation even at redshift $z > 2$. Studies of AGN within protocluster regions have suggested that AGN activity is enhanced in these high-density regions (e.g. Pentericci et al., 2002; Lehmer et al., 2009; Krishnan et al., 2017). This is contrary to the AGN–environment relation observed at low redshifts. Similarly, studies of galaxies at high redshift have also found the relationship between density and star-formation rate to be inverted relative the the redshift $z = 0$ relation (e.g. Elbaz et al., 2007; Hatch et al., 2011b; Koyama et al., 2013).

Furthermore, environmental quenching processes—such as ram pressure stripping (e.g. Gunn and Gott, 1972; Abadi et al., 1999), whereby the gas that comprises the interstellar medium is removed from the galaxy due to pressure from the intra-cluster medium, and ‘strangulation’ (e.g. Larson et al., 1980; Balogh and Morris, 2000), which refers to the situation where a galaxy is no longer able to accrete gas from the IGM, are expected to be more efficient in high-density regions (Contini et al., 2020).

Protoclusters provide astronomers with a laboratory for studying the formation and evolution of galaxies in dense environments during the epoch of maximal star formation rate density, also called “cosmic noon”. During this time quasar activity was also at its peak and today’s massive galaxies assembled the majority of their mass (Madau and Dickinson, 2014). In recent years there has also been a renewed interest in galaxy protoclusters for their populations of so called dusty star forming galaxies (DSFGs) (Greenslade et al., 2018). These galaxies are undergoing star bursts which are thought to be instrumental in the quenching process that transforms galaxies from blue star forming spirals, to red quenched ellipticals.

To cosmologists, protoclusters give insight into the process of large scale structure formation and the growth rates of protoclusters can be used to help constrain cosmological models (Kravtsov and Borgani, 2012). Protoclusters are also thought to be drivers of the thermal energy history of the Universe (Chiang et al., 2020) and play a significant role in cosmic reionisation (Chiang et al., 2017), the process by which the Universe changed from an opaque neutral environment to the highly ionised state observed today.

1.3.2 Identification of Protoclusters

Discovering protoclusters presents a significant challenge to observational astronomers. A small number of protoclusters have been detected using X-ray observations and the Sunyaev-Zel’dovich effect (Sunyaev et al., 1972), which identifies the $T \sim 10^7$ K intracluster medium within massive collapsed clusters and groups (e.g. Bleem et al., 2015), up to redshifts of $z = 1.7$ (Strazzullo et al., 2019). The majority of protoclusters lack a hot X-ray emitting intra-cluster medium (Overzier, 2016), as the majority of the gas within these objects has yet to be shock heated due to its relatively low density. As a result, protoclusters are primarily discovered by the detection of galaxy overdensities.

Protoclusters are large ($> 10 h^{-1}$ cMpc; [Muldrew et al., 2015](#)) and diffuse objects, with only a modest galaxy density enhancement over the field ($\delta_m \sim 1\text{--}6$ at $z = 2$; [Chiang et al., 2013](#)) when compared to mature clusters. They often comprise of galaxies with faint apparent magnitudes which do not yet lie on a well defined red sequence ([Gladders and Yee, 2000](#)).

“Blind Surveys”

Until recently, there have been two main methods of discovering protoclusters in observations. The first of these is through serendipitous discoveries in large galaxy redshift surveys, in so-called “blind surveys”. Indeed multiple protoclusters have been discovered in photometric ([Daddi et al., 2009](#); [Chiang et al., 2014](#)), narrow band ([Hu et al., 2021](#)) and spectroscopic surveys ([Steidel et al., 2005](#); [Cucciati et al., 2014](#); [Lemaux et al., 2014](#); [Chiang et al., 2015](#); [Lemaux et al., 2017](#)). However, these blind surveys are limited in their efficacy by their small areas. Protoclusters are large objects with number densities of $< 10^{-6}$ cMpc $^{-3}$ and as such large areas of the sky must be surveyed at sufficient depth to detect galaxies at $z \sim 2$. Present surveys, such as the COSMOS survey ([Scoville et al., 2007](#)), only cover areas of $\lesssim 1$ deg 2 and as a result have only been successful in discovering a few tens of protoclusters over the last decade. Whilst these limitations have held back blind surveys as a protocluster discovery tool in the past, ongoing and future surveys are set to greatly improve on this. The Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP)² is set to survey 1400 deg 2 to a depth of $r = 26$ and the Large Synoptic Survey Telescope (LSST)³ is expected to cover 18,000 deg 2 down to $r = 27.8$ in the fully stacked data. These surveys show great promise, with [Toshikawa et al. \(2018\)](#) having already discovered over 179 protocluster candidates at $z \sim 4$ using only a small subset of the HSC-SSP data.

Biased Tracers

In an effort to overcome the shortcomings of blind surveys, one can instead direct a focused survey around luminous tracer objects. These objects consist of various forms of massive galaxies (detailed further below). Given that these massive galaxies ($M_* \sim 10^{11.5} M_\odot$) will reside within massive dark matter haloes ($M_h \sim 10^{13} M_\odot$)—which in turn are preferentially located in regions that are overdense on larger scales—they are expected to be correlated with high density regions at redshift $z \sim 2$, or thought to be the progenitors of present day brightest cluster galaxies.

²<https://hsc.mtk.nao.ac.jp/ssp/>

³<https://www.lsst.org>

The first of these objects are high redshift radio galaxies, which have long been known to have properties which suggest that they are the progenitors of central cluster galaxies (e.g. [Best et al., 1998](#); [Collet et al., 2015](#)). By first identifying candidate emission line galaxies (e.g. Ly α or H α emitters) at a similar redshift to the radio galaxy using narrow-band imaging, then confirming candidates with spectroscopy, many protoclusters have been found in searches around radio galaxies ([Le Fèvre et al., 1996](#); [Venemans et al., 2007](#); [Hatch et al., 2011a](#); [Cooke et al., 2014](#)).

Similarly, QSOs at high redshifts are also often associated with cluster formation. Studying their immediate environments at $z \simeq 2 - 6$ has led to the discovery of several protoclusters ([Wold et al., 2003](#); [Zheng et al., 2006](#); [Capak et al., 2011](#); [Wylezalek et al., 2013](#); [Morselli et al., 2014](#); [Adams et al., 2015](#)). In addition to this, simulations have suggested that quasars at $z \gtrsim 6$ very strongly correlate with highly overdense regions ([Springel et al., 2005](#)), motivating searches around these objects which have had mixed results (e.g. [Morselli et al., 2014](#); [Simpson et al., 2014](#)). However, there is currently some disagreement as to whether QSOs do indeed reside in over dense locations, which larger data sets will be required to resolve. For example, [Husband et al. \(2013\)](#) and [Morselli et al. \(2014\)](#) both find that QSOs at $z \sim 5-6$ reside in high density regions, whilst [Bañados et al. \(2013\)](#) and [Mazzucchelli et al. \(2017\)](#) find that $z \sim 5.7$ QSOs are found in environments consistent with mean density. Observations also suggest that QSOs live in host haloes of mass $\sim 10^{12} M_{\odot}$ (e.g. [White et al., 2012](#)). Haloes in this mass range populate the average environment at $z \sim 2-3$, but are very rare objects at $z \gtrsim 5$. This redshift dependence on the expected environment of QSOs may be responsible for some of the conflicting results. Indeed, [Fanidakis et al. \(2013\)](#) used a semi-analytic model to conclude that QSOs live in $\sim 10^{12} M_{\odot}$ mass haloes and average environments at $z \sim 3$. They also find that QSOs at $z \sim 5$ are preferentially located in galaxy overdensities. However, these overdensities are not as large as those expected at the peaks of the dark matter distribution at $z \sim 5$, suggesting that QSOs at this redshift do not probe the progenitors of $z = 0$ superclusters.

More recently, it has been suggested that Ly- α blobs (LABs) should be good tracers of overdense regions, as they are believed to be large reservoirs of cool gas which are being ionised by either starbursts or AGN ([Ao et al., 2015](#)). Indeed, studies of known overdensities have revealed the presence of these Ly- α blobs (e.g. [Matsuda et al., 2004](#)) and protoclusters have been discovered in searches around previously known LABs ([Hennawi et al., 2015](#); [Bădescu et al., 2017](#); [Cai et al., 2017b](#)).

Finally, abundances of the starbursting, dusty star forming galaxies have also been found to correlate with protocluster environments (Casey et al., 2014; Greenslade et al., 2018; Harikane et al., 2019).

Biased tracers are a very useful tool for the discovery of protoclusters as their high luminosity allows them to be detected in significantly shallower surveys than the fainter constituent galaxies, greatly reducing the observational cost. It is via surveys around these biased tracers that the majority of presently spectroscopically confirmed protoclusters at $z \gtrsim 1.3$ have been discovered. However, this biased tracer method is not without its own drawbacks. By using these objects as beacons for a follow up survey, one necessarily introduces a strong selection bias into any resulting protocluster sample. Additionally there is some uncertainty as to whether these objects probe a significant fraction of the true protocluster populations (Orsi et al., 2016), thus potentially leading to a highly incomplete sample.

Absorption From The IGM

The shortcomings of these two established methods has motivated research towards finding a new method that can be used to discover protoclusters. In recent years there has been an emergence of new research into using the Ly- α forest (see Section 1.2) as a potential probe of protocluster regions. As protoclusters are traced not only by galaxy overdensities, but also by hydrogen gas in the IGM (Adelberger et al., 2003; Mukae et al., 2017), the residual neutral hydrogen (HI) in the highly ionised IGM will leave an imprint in the spectra of background quasars within the Ly- α forest. Whilst the bulk of literature on this approach is fairly recent, the idea of using the Ly- α forest to locate protoclusters was first discussed by Francis and Hewett (1993). Shortly afterwards, the first protocluster to be discovered using this approach was presented in Francis et al. (1996). However, at the time this was not a feasible method for discovering protoclusters in large numbers due to the limited data available. The resurgence of interest into this method has coincided with the recent completion of large QSO surveys, such as the Baryon Oscillation Spectroscopic Survey (BOSS) (Dawson et al., 2013), the Extended Baryon Oscillation Spectroscopic Survey (eBOSS), as well as forthcoming surveys such as the Dark Energy Spectroscopic Instrument (DESI) (Vargas-Magana et al., 2019) and WEAVE-QSO (Pieri et al., 2016). Furthermore, surveys such as LATIS (Newman et al., 2020) and CLAMATO (Lee et al., 2018) have been able to incorporate Ly- α forest spectra from bright star forming galaxies in order to obtain an even greater transverse sightline density.

The use of the Ly- α forest as a protocluster discovery technique has been approached from two different directions. The first of these is by using spectra from an ensemble of background sources to reconstruct maps of the IGM in 3-D (Pichon et al., 2001; Caucci et al., 2008; Lee et al., 2014; Stark et al., 2015; Lee et al., 2018; Horowitz et al., 2019; Newman et al., 2020; Porqueres et al., 2020), a technique known as Ly- α forest tomography, and then locating regions of significantly enhanced absorption which should correspond to the locations of protoclusters. Ideally, this method requires one to have a mean transverse sightline separation on the order of a few h^{-1} cMpc (Lee et al., 2014; Stark et al., 2015). Whilst lower density maps can be constructed using a large sample of QSO sight-lines alone (Ravoux et al., 2020), recent observational advances have allowed for higher resolution maps to be made by incorporating spectra from background star-forming galaxies (Lee et al., 2016, 2014; Mukae et al., 2020a; Newman et al., 2020). The Ly- α forest transmission from these spectra over a given area of the sky are mapped onto a 3-D grid using a Wiener filter. Fluctuations in the Ly- α forest transmission are defined as,

$$\delta_F = \frac{f}{C \langle F \rangle(z)} - 1 \quad (1.8)$$

where f is the observed spectral flux density, C is the estimated continuum and $\langle F \rangle(z)$ is the mean Ly- α forest transmitted flux. An example rendering of the 3-D reconstruction of the IGM in the LATIS (Newman et al., 2020) fields is shown in Fig. 1.8. Lee et al. (2016) have used a theoretical framework obtained from N-body simulations (Lee et al., 2014; Stark et al., 2015) to detect a protocluster at redshift $z = 2.44$. Additionally these tomographic IGM maps have been used by Krolewski et al. (2018) to detect large cosmic voids.

Finally, these Ly- α transmission maps can be combined with existing galaxy surveys covering the same volume to allow astronomers to gain a greater understanding of the connection between the IGM and galaxies at redshift $z > 2$. This in turn can be used to improve the accuracy of the tomographic maps themselves (Mukae et al., 2017; Momose et al., 2021; Mukae et al., 2020b; Liang et al., 2021; Horowitz et al., 2021).

The second of these approaches is to instead identify strong coherent large scale Ly- α absorption within individual spectra of background QSOs (Cai et al., 2016). Whilst this approach does not give the same utility as the tomographic approach, it has the advantage of requiring a far lower sightline density and is therefore much better suited to the detection of protoclusters at redshifts where the density of suitable background sources is too low for tomography to be used.

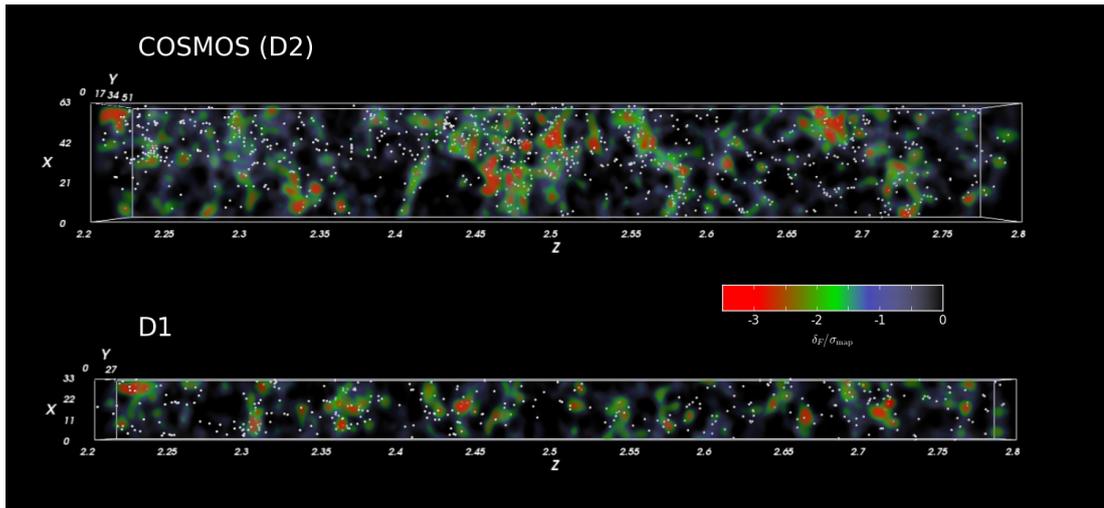


Figure 1.8: Figure 25 from Newman et al. (2020), showing renderings of the IGM opacity in the LATIS fields. The maps for both the COSMOS (top) and D1 (bottom) fields are smoothed on a $2 h^{-1}$ cMpc scale and viewed using a side-on projection. The x and y axes are plotted in h^{-1} cMpc while the z axes is shown in units of redshift. Red colours represent regions of most negative δ_F (or strongest absorption), blue colours show regions of more modest absorption and regions with $\delta_F > 0$ are completely transparent. White points denote the location of coeval galaxies.

Cai et al. (2016) used mock spectra from a $1 h^{-3}$ cGpc³ collisionless dark matter simulation combined with LyMAS (an approximate scheme for modelling Ly- α absorption, Peirani et al., 2014) to identify a close association between proto-clusters and so called Coherently Strong intergalactic Ly- α absorption systems (CoSLAs). Here CoSLAs are defined as $15 h^{-1}$ cMpc regions of Ly- α forest spectra with an effective optical depth $\tau_{\text{eff}} > 4.5 \langle \tau \rangle$, where $\langle \tau \rangle = -\log \langle F \rangle$. This approach was used successfully in the discovery of a massive overdensity at $z = 2.32$ (Cai et al., 2017a) in the vicinity of six such CoSLAs, which is shown in Fig. 1.9, with a further two proto-clusters having recently been discovered at $z = 2.24$ in a similar fashion (Shi et al., 2021). However, this method has yet to have been applied on a large scale.

Whilst this IGM approach to protocluster location has proved promising, much of the theoretical framework is based upon studies using N-body simulations in conjunction with the fluctuating Gunn-Peterson approximation (see Section 1.2.4) or LyMAS (Peirani et al., 2014). Until now these methods had not been investigated using full hydrodynamical simulations which provide a self consistent forward modelling of high column density absorbers as well as the effects of shock heating and AGN and stellar feedback. Furthermore the existing studies have not connected to $z = 0$ data, instead focusing on the—more observationally motivated—approach of relating optical depth to mass overdensities at the same

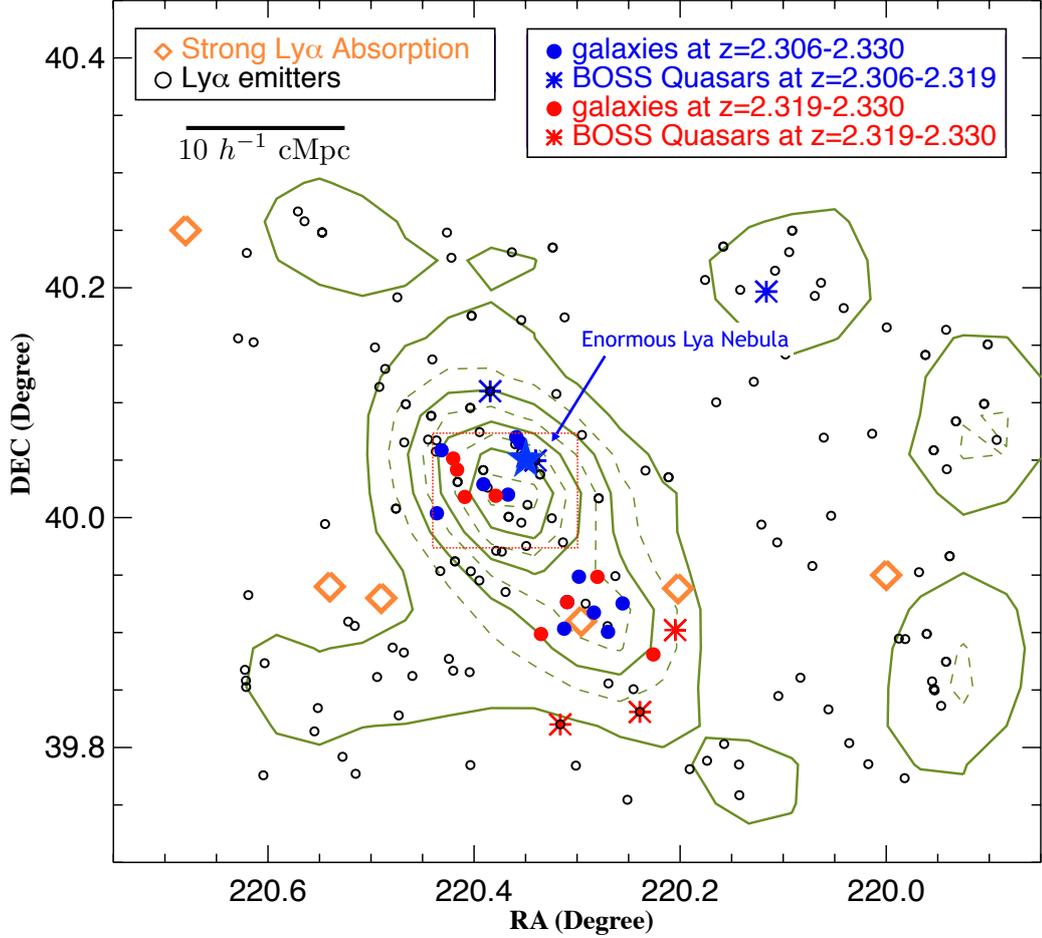


Figure 1.9: Figure 6 from Cai et al. (2017b), showing the BOSS1441 galaxy overdensity at redshift $z = 2.32 \pm 0.02$. Orange diamonds show the locations of the six CoSLAs and the blue/red asterisks denote the locations of BOSS QSOs over $30 h^{-1} \text{ cMpc}$ at redshift $z \sim 2.3$. The black hollow circles show the location of candidate LAEs while the spectroscopically confirmed galaxies are represented by the blue/red filled circles. The green contours show the LAE density maps, with each contour representing a density increase of $0.1 \text{ galaxies per arcmin}^2$ up to a maximum of $1 \text{ galaxy per arcmin}^2$ for the central contour.

redshift. Finally, previous theoretical studies have assumed a spatially uniform UV background in their models. As a result of this, the impact of the proximity effect on tomographically reconstructed IGM maps has not yet been studied. It is possible that the increased transmission expected around luminous QSOs could induce an ionisation bias to the results. This will be examined further in Chapter 2.

1.4 Cosmological Simulations

A fundamental limitation of observational extragalactic astronomy is that one is restricted to only viewing objects at a single instance in time, or more precisely, over a time period that is insignificant in comparison to the timescale over which galaxies and large scale structure evolves. A natural consequence of this is that one is not able to directly link progenitor objects to their descendants, instead having to defer to statistical relations between progenitor and descendant populations. By contrast, cosmological simulations allow astronomers to follow individual objects through cosmic time for a given evolutionary model. This allows simulations to be used to test hypotheses informed by observations and provide predictions for further observations, which in turn can be used to better constrain theoretical models.

The simplest type of cosmological simulations are N-body simulations. These simulations consider particles which are only acted upon by the gravitational forces of every other particle in the simulation volume. One could argue that the first example of such a simulation was performed by [Holmberg \(1941\)](#), who used 74 light bulbs to investigate the tidal interactions of two galaxies in close proximity. Today cosmological N-body simulations are run using large supercomputers such as Hazel Hen at the High Performance Computing Center Stuttgart (HLRS)⁴ and Curie supercomputer at the Très Grand Centre de calcul (TGCC)⁵. This vast and rapid increase in computing power has allowed the number of mass elements to increase dramatically since the work of [Holmberg \(1941\)](#), with the largest N-body simulation today consisting of $\simeq 4.4 \times 10^{12}$ particles simulated over a timescale of $\simeq 13.7 \times 10^9$ years ([Cheng et al., 2020](#)).

Despite this relatively simple modelling process, N-body simulations have been invaluable to astronomers and cosmologists. Using N-body simulations has improved understanding of the halo mass function, the dark matter distribution of the Universe and the structure of dark matter haloes (see [Vogelsberger et al., 2020](#), and references therein). Furthermore, when used in combination with the FGPA, N-body simulations are able to recreate the low density Ly- α forest with a good degree of accuracy (e.g. [Wang et al., 2013](#); [Davies et al., 2018](#); [Choudhury et al., 2021](#)). However, in order to have a fully self-consistent model of the IGM it is necessary to also model the baryonic physics at play on small scales.

⁴<https://www.hlrs.de/systems/cray-xc40-hazel-hen/>

⁵<http://www-hpc.cea.fr/en/complexe/tgcc-curie.htm>

These so-called hydrodynamical simulations model the baryon content of the Universe as a fluid, with the dark matter contribution following a simpler N-body prescription. In the context of the Ly- α forest, hydrodynamical simulations allow astronomers to consider the effects that shock heating, collisional ionisation and—with a prescription for modelling the self-shielding of dense gas—obtain the correct incidence of high column density and damped Ly- α systems.

Due to the increased computational cost of hydrodynamical simulations, the trade off is that they must span smaller volumes than their N-body counterparts. The largest dark matter only simulations such as Millennium-XXL (Angulo et al., 2012), EuclidFlagship (Potter et al., 2017) and DEUS Full Observable Λ CDM Universe Simulation (Reverdy et al., 2015), cover volumes of $4,110^3$ Mpc 3 , $20,000^3$ Mpc 3 and $29,167^3$ Mpc 3 respectively; on the other hand, the largest hydrodynamical simulations are several orders of magnitude smaller, with IllustrisTNG300 (Springel et al., 2018), BAHAMAS (McCarthy et al., 2017) and BlueTides (Feng et al., 2016) covering only 302.6^3 Mpc 3 , 571^3 Mpc 3 and 574^3 Mpc 3 respectively.

For the work contained within this thesis, two types of hydrodynamical simulation code are used. The first is smoothed particle hydrodynamics (SPH), which was first developed by Gingold and Monaghan (1977). In an SPH simulation the fluid is modelled as an ensemble of particles which each have a given smoothing length, h , which represents the characteristic radius of a smoothing kernel, $W(|\mathbf{r}_i - \mathbf{r}_j|, h)$. In order to retrieve the fluid properties of any one particle in the simulation, one must sum the contributions to the given property from all the particles within the smoothing length of the target particle. The ability of particles to have independent values of h , that may also vary in time, allows SPH simulations to have an adaptive resolution whereby the smoothing length is small (and thus the resolution high) in regions of high density, whilst in low density regions (where resolution is less important) the smoothing length can be increased.

The second approach is a moving mesh method. Here, rather than being modelled as particles, the simulation volume is divided into a series of Voronoi cells. Each of these Voronoi cells encloses the same amount of mass, such that the physical size of the cells decreases with increased matter density (see Springel, 2010). This allows the moving mesh to dynamically adjust its spatial resolution to be greater in high density regions (and *visa versa*), without any additional input. Unlike a simple or adaptive grid method, in a moving mesh the mesh-generating points are not constrained to a Cartesian grid, but instead move unrestricted along with the local flow of the baryonic matter. This geometric flexibility allows the mesh to have no preferred directions and be Galilean invariant.

Mock Ly- α spectra can be extracted from these hydrodynamical simulations following the SPH interpolation scheme detailed in appendix A4 of [Theuns et al. \(1998\)](#). A sightline is divided into a number of evenly spaced bins (or pixels) in velocity space and the SPH kernel is then used to calculate the density and density-weighted temperature and velocity for each bin. These values are then used to calculate the optical depth of the bin k adjacent to bin j as follows,

$$\tau(k) = \frac{\sigma_\alpha}{\sqrt{\pi}} \frac{c}{\sqrt{2k_B T(j)/m_p}} \rho_{\text{HI}}(j) \Delta_x \exp \left[- \left(\frac{v(k) - v(j)}{\sqrt{2k_B T(j)/m_p}} \right)^2 \right], \quad (1.9)$$

where σ_α is the Ly- α cross-section, Δ_x is the proper distance interval of the bins and $\sqrt{2k_B T(j)/m_p} = b$ is the Doppler parameter of bin j . This same methodology can be applied to moving mesh simulations by treating each cell as an SPH particle with a smoothing length,

$$h_i = \left(\frac{3N_{\text{sph}} V_i}{4\pi} \right)^{1/3}, \quad (1.10)$$

where V_i is the volume of the cell and N_{sph} is the chosen number of smoothing neighbours.

As discussed in Section 1.3.1, the ability of simulations to follow structure formation throughout cosmic time means that I can use a “true” definition of a protocluster. A protocluster can then be defined in simulations to be the ensemble of mass at a given redshift that will collapse down to form a protocluster at $z = 0$. This allows us to both comment on the redshift $z = 0$ descendants of these objects, but also to determine what fraction of the true protocluster population is recovered using a given selection technique.

1.5 Summary

Galaxy protoclusters are important objects in the study of galaxy formation and evolution, large scale structure formation and cosmology. Traditional methods of discovering these objects, however, are restricted to observationally expensive and highly incomplete “blind surveys” or focused surveys around biased tracer objects, which in turn may return incomplete and necessarily selection biased protocluster samples. With the advent of recent spectroscopic surveys, there has been renewed interest into the field of utilising IGM absorption as a method for discovering protocluster regions in the early Universe.

In this thesis hydrodynamical simulations will be used to gain new insights into these IGM absorption methods of protocluster discovery. Much of the theoretical work underpinning these methods has been undertaken with N-body simulations in combination with either the FGPA or LyMAS. As such, it is currently unclear what the impact of self consistently modelled high column density absorbers, shock heating, stellar and AGN feedback, as well as non-uniform UV backgrounds will be on these methods.

Over the course of this thesis I aim to address three main questions:

Q1: What role does ionisation bias play in tomographically reconstructed maps of the IGM?

In Chapter 2 attention is focused on the IGM tomography method (Stark et al., 2015; Lee et al., 2018; Newman et al., 2020). I use hydrodynamical simulations to investigate the impact that an ionisation bias caused by the proximity effect and shocked gas has on the selection of protoclusters in these tomographic IGM maps. I also examine the relationship between Ly- α absorption and coeval galaxies in the tomographic maps.

Q2: How well do Coherently Strong Ly- α Absorption systems trace protocluster regions at redshift $z \sim 2.4$?

In Chapter 3 focus is moved to the use of Coherently Strong Ly- α Absorption systems as a probe of protocluster regions at redshift $z \sim 2.4$, following Cai et al. (2016). In this chapter I investigate how the relationship between mass overdensity and effective optical depth on $15 h^{-1}$ cMpc scales is affected by the use of hydrodynamical simulations. I then determine how effectively one can select protoclusters using coherent Ly- α absorption in individual QSO sightlines.

Q3: To what extent does simulation box size explain the discrepancies between Coherently Strong Ly- α Absorption system populations in large N-body simulations and comparatively small hydrodynamical simulations?

In Chapter 4 I explore the main caveat of Chapter 3, determining the extent to which simulation box size—and the larger mass protoclusters present in larger simulation volumes—impact on the results of Chapter 3.

Finally the main conclusions of the above work is discussed in Chapter 5, along with an outlook of where this area of research will be heading over the coming years.

Chapter 2

The effect of local ionisation on the Lyman- α absorption signatures of protoclusters at redshift $z \sim 2.4$

The work contained within this chapter has been submitted to Monthly Notices of the Royal Astronomical Society as a paper and is currently in review. The contents have been modified slightly by the insertion of the Appendix of the paper into the main body of this chapter after Section 2.3, as well as a shortened introduction and other minor formatting adjustments.

2.1 Introduction

A key component of recent intergalactic medium (IGM) tomography studies are numerical simulations of the Lyman- α (Ly- α) transmitted flux; these are used to translate the 3-D 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 and Peterson \(1965\)](#) approximation (FGPA; see Section 1.2.4 in Chapter 1) to the density field from large collisionless dark matter simulations ([Stark et al., 2015](#); [Lee et al., 2016](#); [Newman et al., 2020](#)). However, 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) dense gas that is also hot and predominantly collisionally ionised, 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 ultraviolet (UV) background due to the presence of bright, rare sources (i.e. the proximity effect; see Section 1.2.3 in Chapter 1). Indeed, the presence of hot, highly ionised 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 $\sim 3\text{--}5 h^{-1} \text{cMpc}$ 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 ionising background may explain this offset, by changing the distribution of neutral hydrogen in the surrounding IGM (see also [Momose et al., 2021](#), for a similar result obtained from the cross-correlation of galaxies and Ly- α tomographic maps).

In this chapter we investigate this issue of “ionisation bias” further using state-of-the-art hydrodynamical simulations from the IllustrisTNG project. We explore how local ionisation variations in the IGM—either due to the presence of hot, collisionally ionised gas or the proximity effect—impact on the detectability of protoclusters with Ly- α forest tomography. Furthermore, we also assess how these variations may affect the relationship between the Ly- α transmission in tomographic maps and the distribution of coeval galaxies, and how the assumption of the FGPA may bias constraints on protocluster mass. In Section 2.2 we introduce the hydrodynamical simulations and local ionisation models used throughout this chapter, and then examine the expected Ly- α transmission profiles around dark matter haloes in Section 2.3. In Section 2.4, we then briefly consider the impact that mass resolution and simulation box size has on these transmission profiles. We discuss Ly- α transmission tomographic maps of protoclusters and their relationship with coeval Ly- α emitting galaxies in Section 2.5, and assess the role that local ionisation variations may play in Ly- α tomography measurements. Finally, we conclude in Section 2.6. Throughout this chapter, we refer to comoving distance units using the prefix “c” and to proper distance units using the prefix “p”.

Table 2.1: Hydrodynamical simulations used in this chapter. 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]	$N_{\text{gas+DM}}$	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 Simulating $Ly-\alpha$ absorption from protoclusters

2.2.1 Cosmological hydrodynamical simulations

In this chapter 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 ([Pillepich et al., 2018](#); [Springel et al., 2018](#); [Naiman et al., 2018](#); [Marinacci et al., 2018](#); [Nelson 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 Section 2.4 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 sub-grid physics model on our results. The key differences between IllustrisTNG and Illustris-1 are summarised in table 2 of [Nelson et al. \(2019\)](#). These differences include changes to the stellar and active galactic nuclei (AGN) feedback implementations, and the addition of ideal magneto-hydrodynamics in IllustrisTNG ([Pakmor et al., 2011](#)). There are also small differences in the Λ CDM 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 chapter are summarised in Table 2.1. For each simulation we use the snapshots and halo catalogues at $z = 2.44$ and $z = 0$.

2.2.2 Protocluster identification

Throughout this chapter we take advantage of the ability of simulations to connect physical structures at different instances in time. 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% 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. 2.1. A 2-D projection of the normalised gas density, $\Delta = \rho/\langle\rho\rangle$, and gas temperature, T , are shown within $\pm 1 h^{-1} \text{cMpc}$ 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 Chapter 3, a wide range of protocluster morphologies are expected using our chosen definition, where typically $R_{95} = 5\text{--}10 h^{-1} \text{cMpc}$. On average, the gas in protoclusters will exhibit slightly higher densities, temperatures and neutral hydrogen fractions compared to the surrounding IGM.

2.2.3 Local ionisation models

The primary focus of this chapter is assessing the impact that local variations in the IGM ionisation state may have on the identification of protoclusters using $L\gamma\text{-}\alpha$ absorption. We now turn to describing the three different ionisation models we use for this purpose.

In our fiducial ionisation model we assume a spatially uniform UV background using the Faucher-Giguère (2020) synthesis model. For reference, the Faucher-Giguère (2020) model has an H I photo-ionisation rate $\Gamma_{\text{HI}} = 9.76 \times 10^{-13} \text{s}^{-1}$ at $z = 2.44$, which is consistent with independent constraints on Γ_{HI} from the

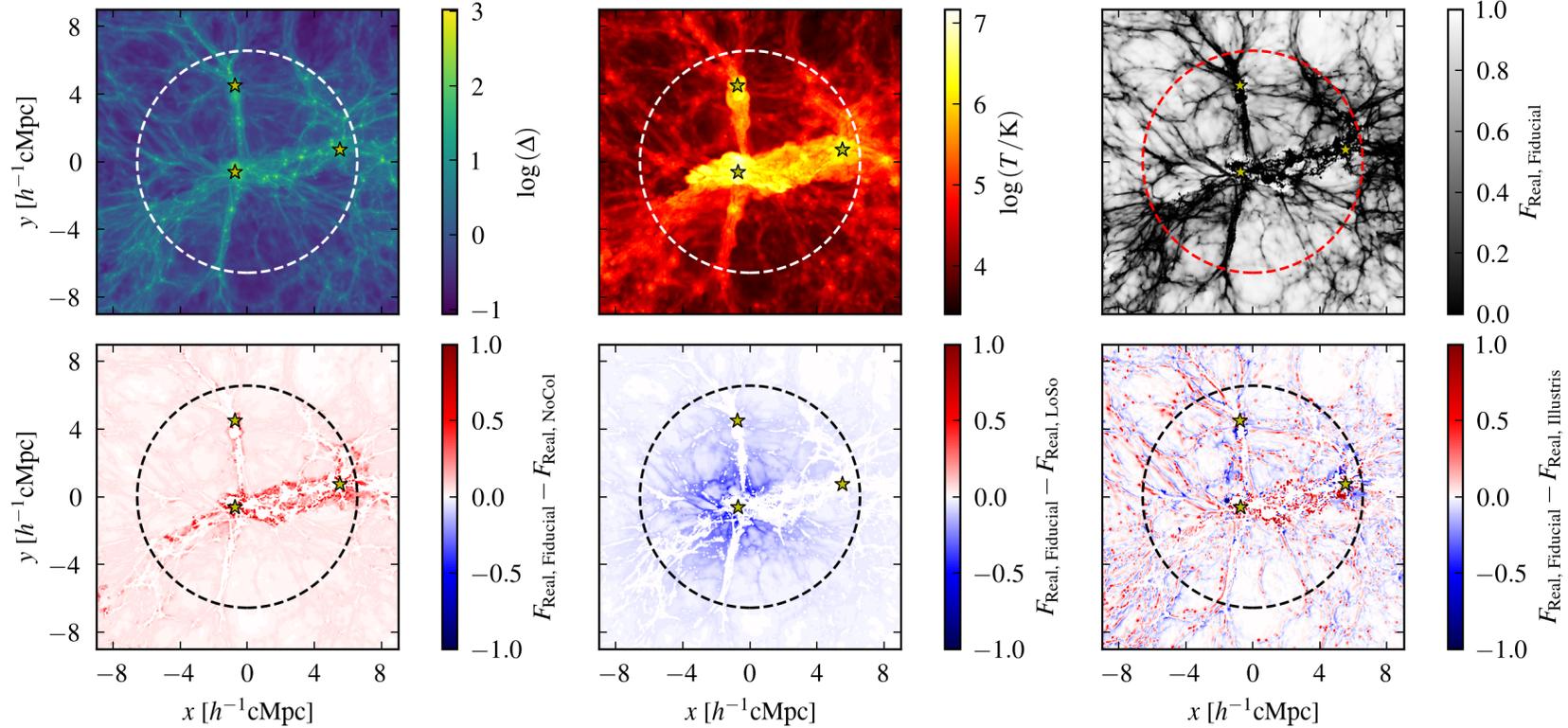


Figure 2.1: Top: A series of 2-D projections of the gas overdensity (left), gas temperature (centre) and the real space Ly- α forest transmission F_{real} (right, see text for details) for a protocluster with redshift $z = 0$ mass $M_{z=0} = 10^{14.46} M_{\odot}$ in the TNG100-1 simulation at $z = 2.44$. The slices are projected over a distance of $2 h^{-1} \text{cMpc}$ 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 *LoSo* model (see Section 2.2.3 for details). Bottom: Slices showing the difference in F_{real} between our fiducial model and our *NoCol* model (left), *LoSo* model (centre), and the Illustris-1 simulation which uses a different sub-grid physics implementation (right). Here red represents a larger Ly- α transmission (less absorption) than the fiducial model, while blue represents a smaller transmission (more absorption).

$Ly-\alpha$ forest opacity (Becker and Bolton, 2013). We calculate neutral hydrogen fractions in each cell of the simulation under the assumption of ionisation equilibrium by using the coupled equations given by Katz et al. (1996a), 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 $N_{\text{HI}} \geq 10^{17.2} \text{ cm}^{-2}$). This post-processing method for modelling self-shielding leads to an inconsistency in the underlying gas temperature, due to the photoheating applied from the UV background assuming an optically thin limit. However, it has been shown that this effect does not have a significant impact on the simulated H I CDDF (Pontzen et al., 2008; McQuinn and Switzer, 2010; Altay et al., 2011). The procedure described above reproduces the shape of the observed H I column density distribution over the range $10^{12} \text{ cm}^{-2} \leq N_{\text{HI}} \leq 10^{22} \text{ cm}^{-2}$ very well (see Fig. 4.1 in Chapter 4).

In addition to our fiducial model, we investigate two further, alternative ionisation models. In the first we assume a spatially uniform UV background, but now ignore the effects of collisional ionisation 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 TNG100-1 volume. Collisional ionisation will be particularly important for the ionisation state of gas around haloes, where gas is heated to $T > 10^6 \text{ K}$ by gravitational infall and AGN or supernovae feedback (see e.g. the protocluster in the upper central panel of Fig. 2.1). Neglecting collisional ionisation in these hot, dense regions will result in an overestimate of the H I fraction, and hence an overestimate of the $Ly-\alpha$ 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 $N_{\text{HI}} \geq 10^{17.2} \text{ cm}^{-2}$ that arise from cool, dense gas. We refer to this model as “No Collisional”—shortened to *NoCol*—throughout this chapter. 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 $Ly-\alpha$ optical depth to the underlying gas or dark matter density (e.g. Stark et al., 2015; Newman et al., 2020). The FGPA assumes photo-ionisation equilibrium in an IGM which follows a power-law temperature density relation, $T = T_0 \Delta^{\gamma-1}$, which is good approximation only for gas with $\Delta \leq 10$ (see e.g. Rauch, 1998).

Our second alternative ionisation model includes the effect of local enhancements in the IGM ionisation state due to quasi-stellar objects (QSOs) and AGN (i.e. the

proximity effect, [Murdoch et al., 1986](#); [Bajtlik et al., 1988](#), see also Section 1.2.3 in Chapter 1). At the redshift we consider in this chapter, $z = 2.44$, 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 AGN in close proximity to protoclusters could mean the background photo-ionisation 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 ionisation level of neutral hydrogen following the simple model described in [Bolton and 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

$$N_{\text{AGN}} = V \int_{L_{\text{min}}}^{\infty} \phi(L_{1450}) dL_{1450}, \quad (2.1)$$

where $\phi(L_{1450})$ is the AGN luminosity function from [Kulkarni et al. \(2019\)](#) at $z = 2.44$. We assume a minimum luminosity of $L_{\text{min}} = 10^{43.2} \text{ erg s}^{-1}$, 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). An AGN will see an effective decrease in its luminosity as a result of neutral hydrogen in the interstellar medium and CGM preventing some photons from escaping the AGN host galaxy. Whilst we do not model this effect explicitly, assigning luminosity values based upon the observed luminosity function means that the escape fraction is included in a statistical sense. This process yields 281 AGN within the TNG100-1 volume, with a median $M_{1450} = -18.9$ and a minimum of $M_{1450} = -24.7$.

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

$$L(\nu) \propto \begin{cases} \nu^{-0.61} & (912 \text{ \AA} < \lambda \leq 1450 \text{ \AA}), \\ \nu^{-1.70} & (\lambda \leq 912 \text{ \AA}). \end{cases} \quad (2.2)$$

We then compute the specific intensity, $J(\mathbf{r}, \nu)$, of the ionising 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,

$$J(\mathbf{r}, \nu) = \frac{1}{4\pi} \sum_{i=1}^{N_{\text{AGN}}} \frac{L_i(\mathbf{r}, \nu)}{4\pi |\mathbf{r}_i - \mathbf{r}|^2}, \quad (2.3)$$

where $|\mathbf{r}_i - \mathbf{r}|$ is the distance of the i^{th} AGN from \mathbf{r} . Finally we compute the spatially varying photo-ionisation rate from the AGN by evaluating

$$\Gamma_{\text{HI}}(\mathbf{r}) = \int_{\nu_{\text{HI}}}^{4\nu_{\text{HI}}} \frac{4\pi J(\mathbf{r}, \nu)}{h_{\text{P}}\nu} \sigma_{\text{HI}}(\nu) d\nu, \quad (2.4)$$

where $\sigma_{\text{HI}}(\nu)$ is the photo-ionisation cross-section from [Verner et al. \(1996\)](#) and ν_{HI} is the frequency at the hydrogen Lyman limit. The photo-ionisation 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 photo-ionisation rate from Eq. (2.4) 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 ionisation fraction. Throughout this chapter we shall refer to this as our “local sources”—shortened to *LoSo*—model.

In Fig. 2.1, we perform an initial assessment of the effect of these ionisation models on the average $Ly\text{-}\alpha$ forest transmission. We consider a region of width $\Delta R = 2 h^{-1} \text{cMpc}$ centred around the protocluster, and obtain an estimate of the real space transmitted flux, F_{real} , from the column density, N_{HI} , in each pixel following a similar approach to [Kulkarni et al. \(2015\)](#), where

$$F_{\text{real}} = \exp\left(-\frac{3\lambda_{\text{Ly}\alpha}^3 \gamma_{\text{Ly}\alpha} N_{\text{HI}}}{8\pi H(z) \Delta R}\right). \quad (2.5)$$

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

In the top right panel of Fig. 2.1 we show F_{real} for the example TNG100-1 protocluster at $z = 2.44$ in the fiducial ionisation model, whilst in the lower panels we show the difference in F_{real} between the fiducial model and the *NoCol* model (left), *LoSo* model (centre), and the same region in the Illustris-1 simulation under the assumption of a spatially uniform UV background (right). In the *NoCol* model we observe a decrease in the transmission from the filamentary structure consisting of overdense gas ($\Delta \sim 10\text{--}100$) at high temperatures ($T \sim 10^5\text{--}10^7 \text{K}$). This corresponds to gas that has been heated by shocks and outflows and is therefore collisionally ionised in the fiducial model. Hence, we expect that ignoring hot, collisionally ionised gas will underestimate the $Ly\text{-}\alpha$ transmission from the gas

in protoclusters. By contrast, in the *LoSo* model we observe an increase in the $Ly-\alpha$ transmission in the protocluster, with a magnitude that decreases radially and is generally more pronounced in cooler, less dense regions where photoionisation dominates. Finally, comparing the different sub-grid physics implementation used in Illustris-1 to the fiducial TNG100-1 model, 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 ionisation state of the hydrogen gas.¹ The exact nature of these changes will, however, depend on the specific feedback model chosen. Using a simulation with a jet mode (e.g. SIMBA, Davé et al., 2019) or thermal (e.g. EAGLE, Schaye et al., 2015) feedback model instead, may result in further small changes to the underlying gas temperature. Note the transmission from the low density IGM with $\Delta \lesssim 1$ remains unchanged, however, as the gas in voids is largely unaffected by AGN or supernovae driven winds at $z = 2.44$ (e.g. Theuns et al., 2002; Viel et al., 2013).

2.2.4 Mock $Ly-\alpha$ absorption spectra

In the remainder of this chapter we will analyse the $Ly-\alpha$ absorption associated with protoclusters using simulated $Ly-\alpha$ forest spectra. The procedure used is described in detail in Chapter 3, but we also give a brief summary here. Mock $Ly-\alpha$ 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

$$h_i = \left(\frac{3N_{\text{sph}}V_i}{4\pi} \right)^{1/3}. \quad (2.6)$$

We assume $N_{\text{sph}} = 64$ for the number of smoothing neighbours. The interpolation scheme described by Theuns et al. (1998) is then used to extract $Ly-\alpha$ 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 $Ly-\alpha$ forest effective optical depth, $\tau_{\text{eff}} = -\ln\langle F \rangle = 0.20$ 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

¹This comparison is not exact, however, due to the slightly different cosmological parameters used in Illustris-1 and TNG-1 (see Table 2.1)

velocity window of width Δv —around the IGM mean value

$$\delta_{\text{F}} = \frac{\langle F \rangle_{\Delta v}}{\langle F \rangle} - 1. \quad (2.7)$$

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.

2.3 Ly- α absorption profiles around haloes

We perform a consistency test of our mock Ly- α absorption spectra in Fig. 2.2, where we show the transmitted flux contrast for our different ionisation models around haloes in three mass bins: $M \geq 10^{12.8} M_{\odot}$ (left), $10^{12.4} M_{\odot} \leq M < 10^{12.8} M_{\odot}$ (centre) and $10^{12.1} M_{\odot} \leq M < 10^{12.4} M_{\odot}$ (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 $1.96 h^{-1} \text{ cMpc}$. We then calculate the mean transmission within 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-Ly- α 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 et al., 2017; Sorini et al., 2020; Nagamine et al., 2021). In general, differences in stellar and AGN feedback implementations, halo mass and numerical resolution all play an important role. For the radial profiles around haloes with $M \geq 10^{12.8} M_{\odot}$, our fiducial (black), *LoSo* (green) and Illustris-1 (blue) models are all within a 1σ deviation from the data points at $0.3 h^{-1} \text{ cMpc} \lesssim b \lesssim 1 h^{-1} \text{ cMpc}$ (whilst our *NoCol* model deviates by $\sim 1\text{--}2\sigma$ over these impact parameters)². This relative agreement changes at large scales where all of our models begin to deviate from the data, reaching a $\sim 20\sigma$ difference³ at $b = 25 h^{-1} \text{ cMpc}$. The differences we find here are consistent with

²The uncertainties for the fiducial and *LoSo* models on these scales are comparable to those for the Prochaska et al. (2013) data at $\pm \sim 0.05\text{--}0.1$

³This does not account for any systematic errors

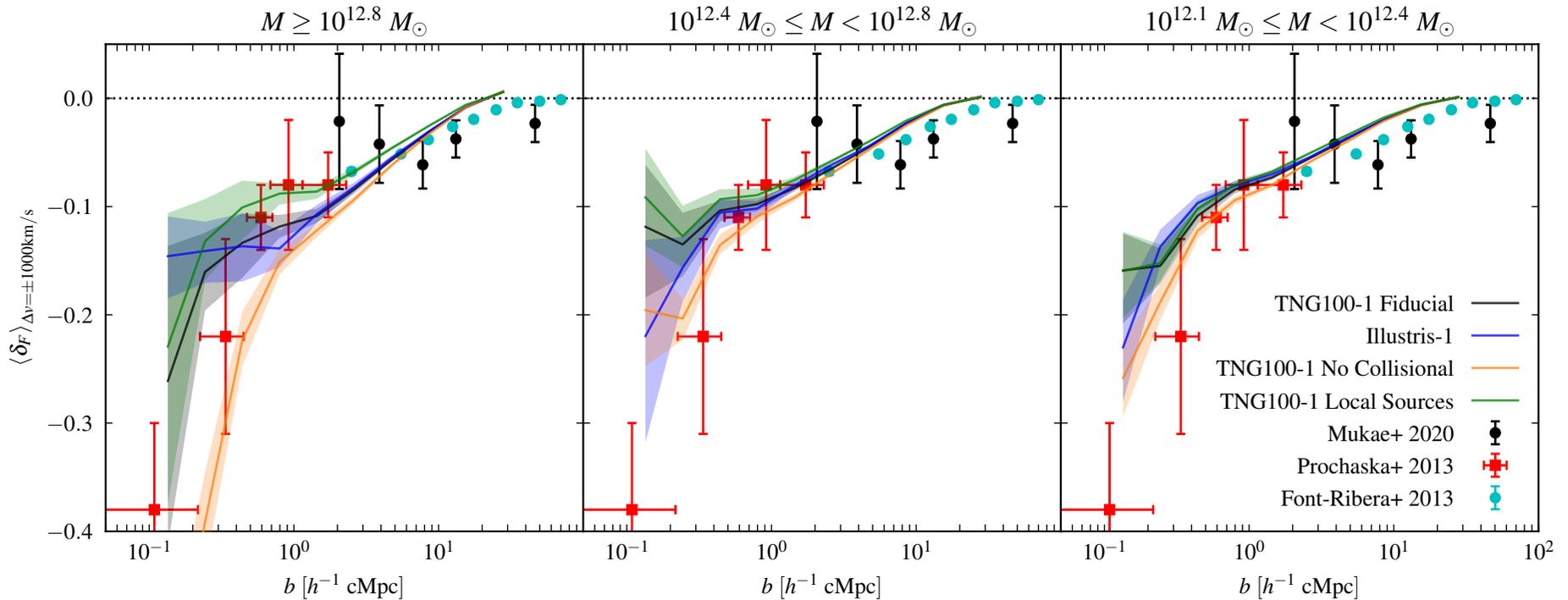


Figure 2.2: The median of the relative Ly- α transmission, δ_F , averaged over a velocity window of width $\Delta v = 2000 \text{ km s}^{-1}$, as a function of impact parameter, b , around haloes in three mass bins at redshift $z = 2.44$: $M \geq 10^{12.8} M_{\odot}$ (left), $10^{12.4} M_{\odot} \leq M < 10^{12.8} M_{\odot}$ (centre) and $10^{12.1} M_{\odot} \leq M < 10^{12.4} M_{\odot}$ (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 fiducial model from the TNG100-1 simulation (black curves) is compared to the *NoCol* model (orange curves) and *LoSo* model (green curves). The corresponding haloes in the Illustris-1 simulation are shown by the blue curves. The shaded regions bound 68% 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).

earlier work, where the relative transmission at small scales, $b < 0.5 h^{-1}$ cMpc, in Illustris-1 (blue curves) and TNG100-1 (black curves) underpredicts 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 ionised 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 ($b > 1 h^{-1}$ 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 ($50 h^{-1}$ cMpc) of the SIMBA simulation. We test this hypothesis by analysing the TNG-300-1 simulation in Section 2.4, and conclude that TNG100-1 predicts transmission that is systematically higher than the Font-Ribera et al. (2013) data for the same reason.

The main focus of this study is the differences caused by the various ionisation models. 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 $\text{Ly-}\alpha$ transmission relative to the fiducial TNG100-1 model, respectively. In the *NoCol* model (orange curves) this is due to neglecting collisional ionisation from hot circumgalactic gas within $b \leq 1 h^{-1}$ cMpc, 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 ionisation (and hence gas temperature) plays an important role in setting the $\text{Ly-}\alpha$ transmission for $b < 1 h^{-1}$ cMpc (see also Sorini et al., 2018).

The increased transmission in the *LoSo* model (green curves) due to enhanced ionisation 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.⁴ 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

⁴If 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 photo-ionisation and light travel time effects due to flickering AGN emission may also result in gas that is less highly ionised (e.g. Oppenheimer and Schaye, 2013; Schmidt et al., 2019).

at $1 h^{-1} \text{cMpc} \leq b \leq 5 h^{-1} \text{cMpc}$. This is because the enhanced photo-ionisation rate only begins to dominate over collisional ionisation at $b \gtrsim 1 h^{-1} \text{cMpc}$. By contrast, the haloes in the lower two mass bins host either fainter AGN, or are unoccupied. As a result, we find the local source model does not have a significant effect on the $Ly-\alpha$ 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 $f_{\text{duty}} < 1$ (e.g. Shankar et al., 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 $b < 5 h^{-1} \text{cMpc}$, although due to the large error bars the significance is not high. This is consistent with the interpretation advanced by Mukae et al. (2020a) that the MAMMOTH1-QSO tomographic map exhibits a QSO proximity zone.

Since the *LoSo* and *NoCol* models effectively bracket the plausible range in the $Ly-\alpha$ transmission profiles, we proceed to investigate the effect these models have on the expected $Ly-\alpha$ transmission associated with protoclusters in TNG100-1. We expect the different sub-grid physics implementation in Illustris-1 will sit between the extremes explored by these models, and so we do not investigate it further.

2.4 The effect of box size and mass resolution on $Ly-\alpha$ absorption around haloes

Following on from Fig. 2.2 and the associated discussion in Section 2.3, the effect of simulation mass resolution and box size on the transmission profiles around haloes are shown in Fig. 2.3 and Fig. 2.4. Here, in addition to the fiducial 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 2.1. There is generally very good agreement between the different simulations in Fig. 2.3, suggesting that our results should be sufficiently converged with respect to mass resolution. However, for the simulations with varying box sizes in Fig. 2.4 we observe larger differences. In the case of the highest mass bin (left panel), we observe a divergence between the two simulations at impact parameters $b < 1 h^{-1} \text{cMpc}$. 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 $T > 10^6 \text{K}$ gas with correspondingly low HI 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.

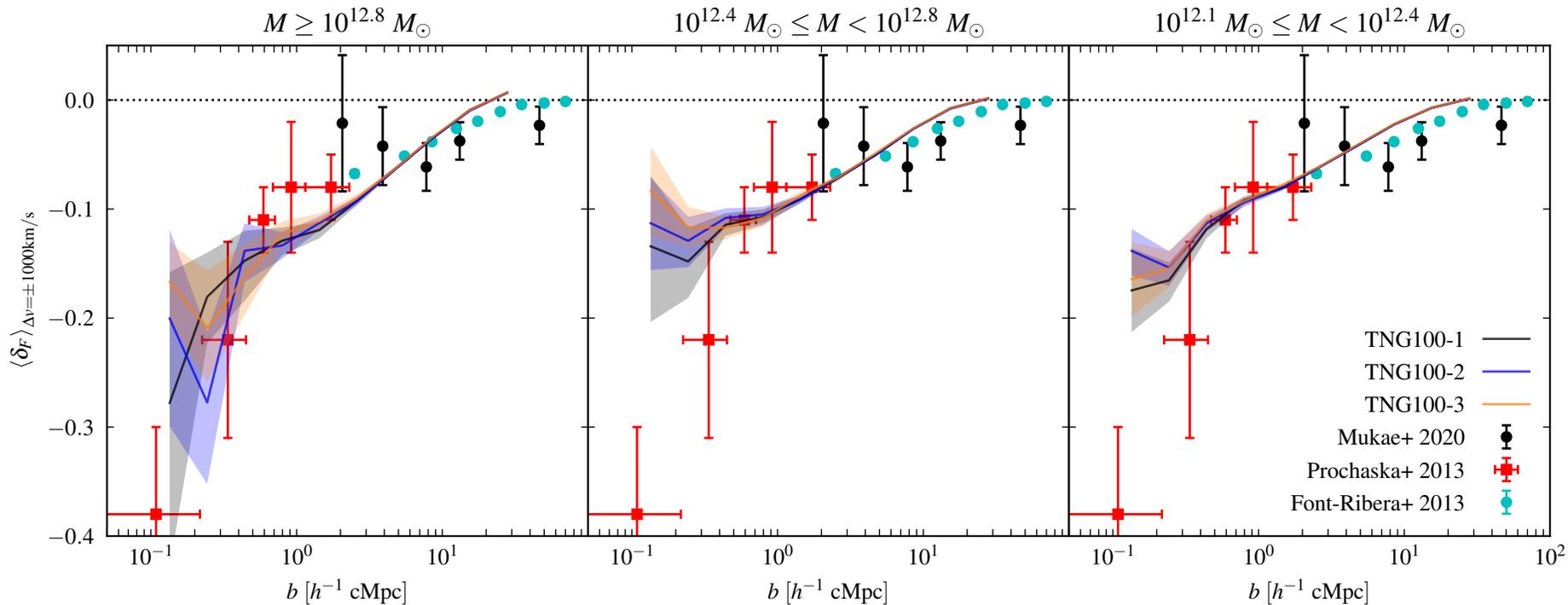


Figure 2.3: As for Fig. 2.2, but comparing simulations with fixed box size and different mass resolutions. The fiducial TNG100-1 simulation (black curves) is compared to the TNG100-2 (blue curves) and TNG100-3 (orange curves) simulations. These have dark matter particle masses a factor of 8 and 64 times larger than the TNG100-1 simulation, respectively. The transmission profiles are consistent within the 68% scatter around the median, shown by the shaded regions.

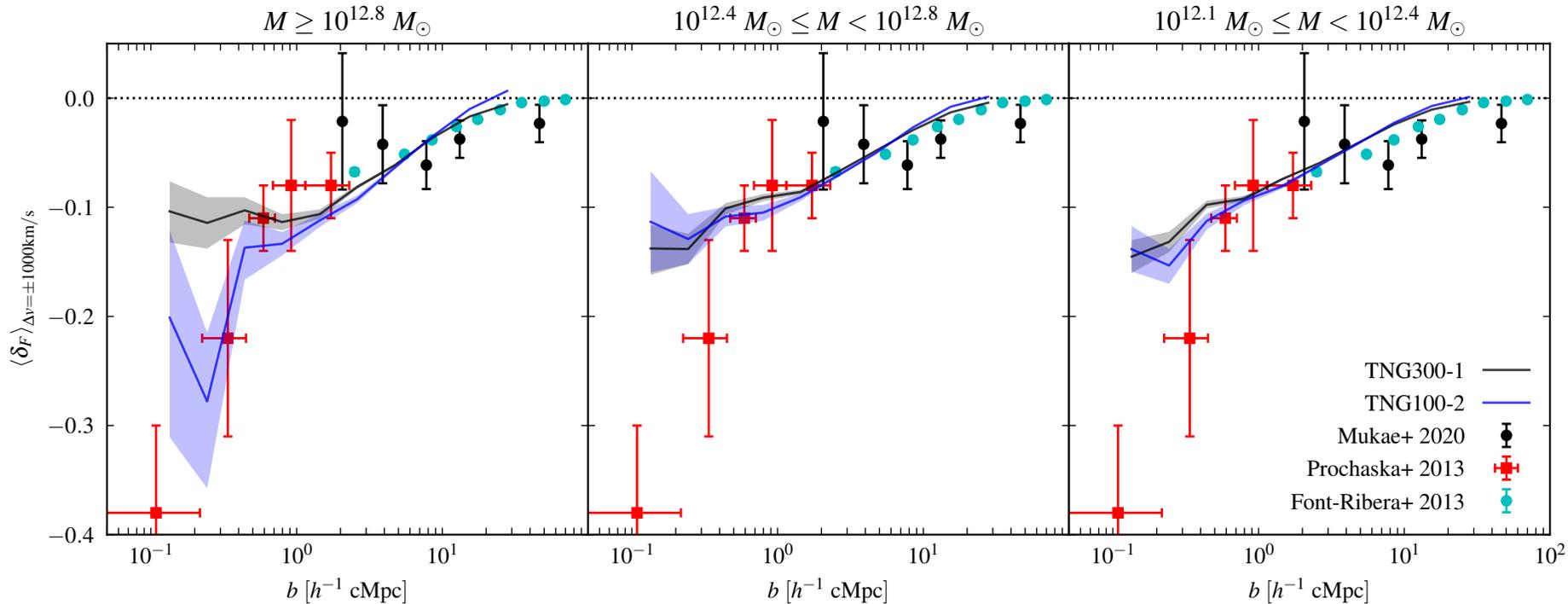


Figure 2.4: As for Fig. 2.2, but comparing simulations with different box sizes and very similar mass resolutions. The TNG100-2 simulation (black curves) uses the same box size as the fiducial TNG100-1 model, whereas the TNG300-1 simulation (blue curves) has a volume 27 times larger. The TNG300-1 simulation 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 panel. The TNG300-1 model also exhibits more transmission at impact parameters $b > 1 h^{-1} \text{ cMpc}$ around the most massive haloes.

On larger scales ($b > 1 h^{-1} \text{cMpc}$), 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 where the difference between the two is reduced to 4σ (down from 20σ for TNG100-1). 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 higher than the [Font-Ribera et al. \(2013\)](#) measurements.

2.5 The effect of local ionisation on the Ly- α transmission around protoclusters

2.5.1 Smoothed Ly- α forest tomographic maps

We now turn to investigate how the Ly- α transmission around protoclusters is altered by changes in the local ionisation state of the IGM. For simplicity, we do not create the Ly- α transmission maps by forward modelling the observational data in detail (e.g. [Stark et al., 2015](#)). Instead, we use the noiseless spectra drawn from the simulations to directly create idealised mock Ly- α tomographic maps of the relative transmission, δ_{F} , around each of the 22 protoclusters in the TNG100-1 volume.

We first extract spectra in a 90×90 grid in a $15 h^{-2} \text{cMpc}^2$ area centred on each protocluster's centre of mass, following the procedure described in Section 2.2.4. This yields an average line of sight transverse separation of $0.17 h^{-1} \text{cMpc}$. We then construct Ly- α tomographic maps by obtaining the average Ly- α transmission over velocity windows, $\Delta v = 1000 \text{km s}^{-1}$ and then smoothing the relative transmission, δ_{F} , in the transverse direction using a Gaussian with standard deviation $4 h^{-1} \text{cMpc}$. 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. The velocity window, Δv , is chosen to match the full width at half maximum of the Gaussian filter at $z = 2.44$. Finally, we normalise each tomographic map by the dispersion of δ_{F} obtained from the full simulation volume. We obtain a dispersion of $\sigma = 0.076$, $\sigma = 0.081$ and $\sigma = 0.070$ for the fiducial, *NoCol* and *LoSo* models, respectively. The dispersion is slightly increased in the *NoCol* model with respect to fiducial, because ignoring collisional ionisation decreases the neutral hydrogen fraction in dense, hot gas, thus increasing the

amount of strong Ly- α absorption. Conversely, the dispersion is reduced relative to fiducial in our *LoSo* model, as the ionising sources (AGN) are placed into high density regions, thus reducing the incidence of strong Ly- α absorption.

The resulting Ly- α tomographic maps for three different protoclusters are shown in each row of Fig. 2.5. The different local ionisation models for the protoclusters are displayed in each column, with the yellow stars in the right column showing the location of coeval AGN within 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, $M_{z=0} = 10^{14.18} M_{\odot}$), a region where the average Ly- α transmission within R_{95} is relatively high (middle row, $M_{z=0} = 10^{14.46} M_{\odot}$) and a region that is more representative of the average protocluster in TNG100-1 (lower row, $M_{z=0} = 10^{14.43} M_{\odot}$). Note the protocluster in the middle row is also displayed in Fig. 2.1. The position of the maps is selected using the velocity window within R_{95} where δ_F/σ is minimised, similar to how these structures are identified within observed tomographic maps.

In each map we also mark the locations of individual sight lines that contain coherently strong Ly- α absorption systems (CoSLAs) using red crosses. Following Cai et al. (2017a), CoSLAs are defined as sight lines that exhibit a fluctuation in the Ly- α forest effective optical depth, $\delta_{\tau_{\text{eff}}} > 3.5$, over a scale of $15 h^{-1} \text{ cMpc}$, after excluding any Ly- α absorbers with damping wings, $N_{\text{HI}} \geq 10^{19} \text{ cm}^{-2}$ (see also Chapter 3 for further details). This allows us to assess how CoSLAs are distributed relative to the Ly- α tomographic 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 sub-halo in TNG100-1. We convert the instantaneous SFR into a luminosity at 1216 \AA using the relation from Dijkstra (2017) for a Salpeter (1955) initial mass function,

$$L_{\text{Ly}\alpha} = 1.0 \times 10^{42} \text{ erg s}^{-1} f_{\text{esc}}^{\text{Ly}\alpha} \left(\frac{\text{SFR}}{M_{\odot} \text{ yr}^{-1}} \right). \quad (2.8)$$

We have implicitly assumed a Lyman continuum escape fraction $f_{\text{esc}}^{\text{LyC}} \simeq 0$ in Eq. (2.8), and $f_{\text{esc}}^{\text{Ly}\alpha}$ is the volume averaged effective Ly- α escape fraction inferred by Hayes et al. (2011),

$$f_{\text{esc}}^{\text{Ly}\alpha} = C_{\text{Ly}\alpha} 10^{-0.4 A_{\text{Ly}\alpha}}, \quad (2.9)$$

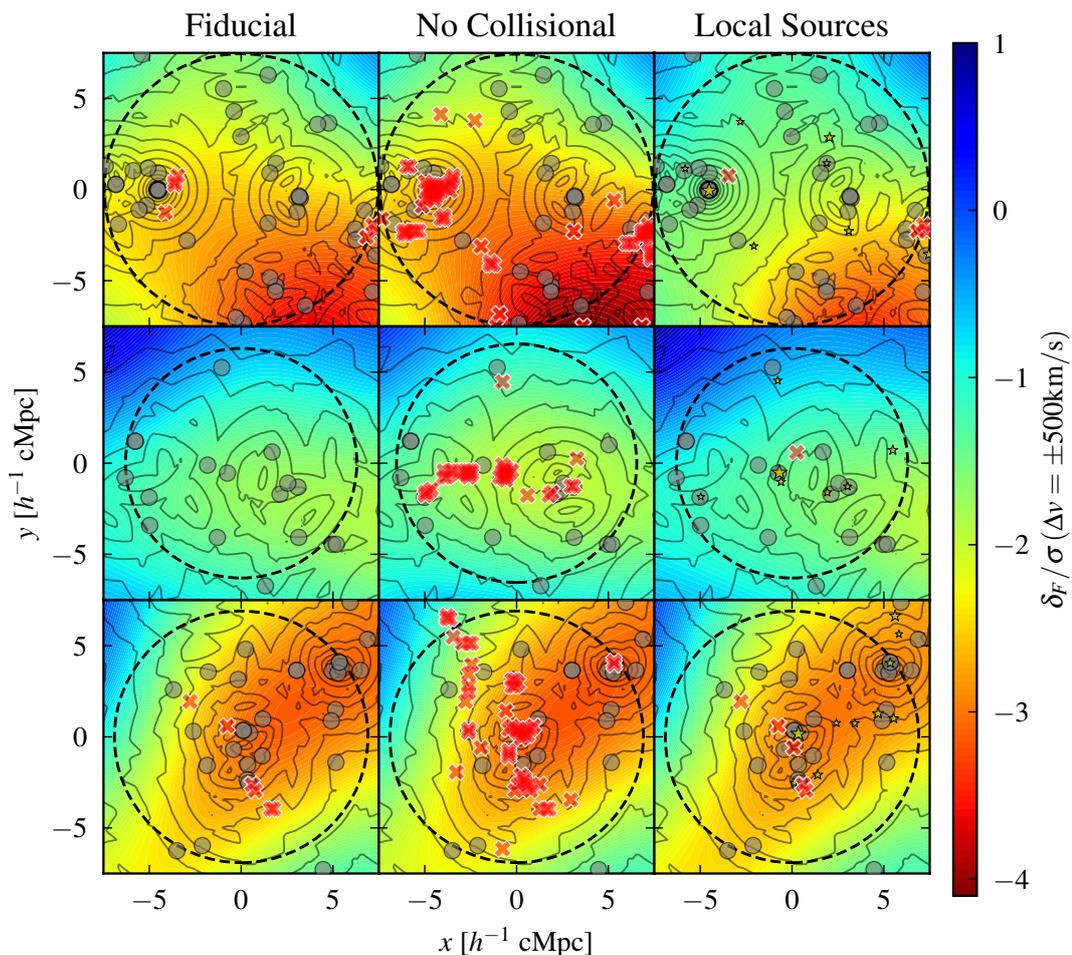


Figure 2.5: Smoothed tomographic maps showing δ_F/σ for three protoclusters in the TNG100-1 simulation at redshift $z = 2.44$. Positive (negative) values of δ_F/σ correspond to Ly- α transmission that is larger (smaller) relative to the average for the IGM. The maps are obtained by averaging the Ly- α forest over a $\Delta v = 1000 \text{ km s}^{-1}$ window, smoothing in the transverse direction with a Gaussian filter with standard deviation $4 h^{-1} \text{ cMpc}$, and then centring on velocity windows $\Delta v = \pm 500 \text{ km s}^{-1}$ where δ_F/σ is minimised within R_{95} for each protocluster. Each row shows a different protocluster for the fiducial (left), no collisional (centre) and local sources (right) ionisation models. On top of each map we display the locations of coeval Ly- α emitting galaxies that satisfy the criteria $L_{\text{Ly}\alpha} > 10^{41.5} \text{ erg s}^{-1}$ and $\text{EW}_{\text{Ly}\alpha} > 15 \text{ \AA}$ (grey filled circles) using a simple empirical model (see text for details). The grey contours show the logarithm of LAE overdensity, $\log(1 + \delta_{\text{LAE}}) = \log(\rho_{\text{LAE}}/\langle\rho_{\text{LAE}}\rangle)$, obtained from the distance to the 5th 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 2-D cross-section of the sphere of radius R_{95} , the radius that contains 95% of the $z = 0$ mass, $M_{z=0}$, that intersects the velocity window for each protocluster. From top to bottom, the selected protoclusters have $M_{z=0} = 10^{14.18} M_{\odot}$, $M_{z=0} = 10^{14.46} M_{\odot}$ and $M_{z=0} = 10^{14.43} M_{\odot}$. Note the protocluster in the middle row is also shown in Fig. 2.1.

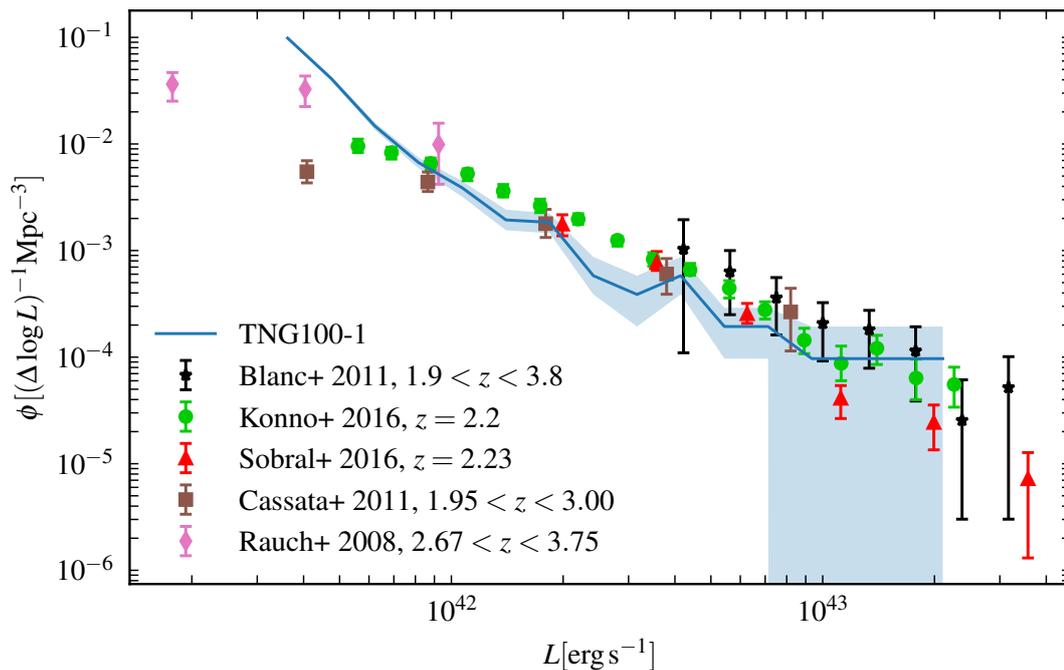


Figure 2.6: The Ly- α luminosity function for all galaxies in the TNG100-1 simulation at $z = 2.44$ following the application of our simple Ly- α luminosity model. The solid blue line shows the median result from 1000 bootstrapping iterations, while the shaded region shows the resulting 16th–84th percentile spread. We also plot observational data from [Blanc et al. \(2011\)](#) (black stars, $1.9 < z < 3.8$), [Konno et al. \(2016\)](#) (green circles, $z = 2.2$), [Sobral et al. \(2016\)](#) (red triangles, $z = 2.23$), [Cassata, P. et al. \(2011\)](#) (brown squares, $1.95 < z < 3.00$) and [Rauch et al. \(2008\)](#) (pink diamonds, $2.67 < z < 3.75$).

where $C_{\text{Ly}\alpha} = 0.445$ ([Hayes et al., 2011](#)). The quantity $A_{\text{Ly}\alpha}$ is derived using the relation between extinction at H α wavelengths and stellar mass derived by [Garn and Best \(2010\)](#), and then converting to $A_{\text{Ly}\alpha}$ using the [Calzetti et al. \(2000\)](#) dust law. We also estimate the rest frame equivalent width in Angstroms, $\text{EW}_{\text{Ly}\alpha}$, for each Ly- α emitter (LAE) using the relation $\text{EW}_{\text{Ly}\alpha} = f_{\text{esc}}^{\text{Ly}\alpha} / 0.0048$ from [Sobral and Matthee \(2019\)](#). The LAEs displayed in the maps are selected by requiring $L_{\text{Ly}\alpha} > 10^{41.5} \text{ erg s}^{-1}$ and $\text{EW}_{\text{Ly}\alpha} > 15 \text{ \AA}$ (e.g. [Shimakawa et al., 2017](#)). The grey dashed contours correspond to the logarithm of LAE overdensity, $\log(1 + \delta_{\text{LAE}}) = \log(\rho_{\text{LAE}} / \langle \rho_{\text{LAE}} \rangle)$, 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, and furthermore does not follow the complex Ly- α radiative transfer within the interstellar medium of the galaxies (e.g. [Laursen et al., 2011](#); [Gurung-López et al., 2019](#)). As such, while it is consistent with average LAE properties by design, it will still likely underestimate the variation in $L_{\text{Ly}\alpha}$ for a given stellar mass.

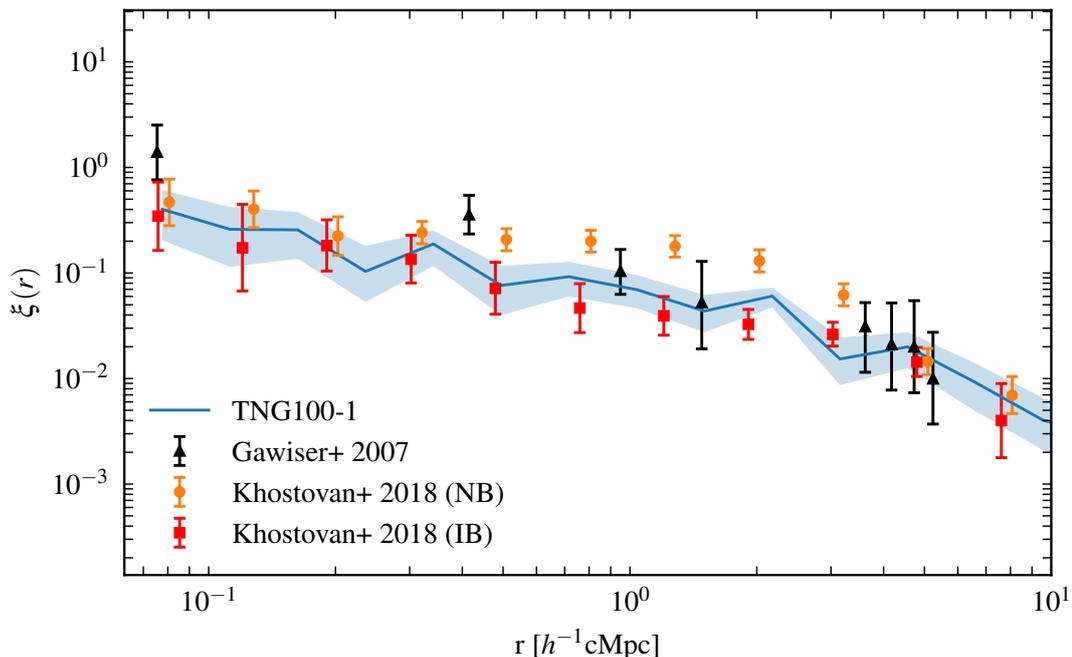


Figure 2.7: The two-point correlation function for LAEs satisfying $L_{\text{Ly}\alpha} > 10^{41.5} \text{ erg s}^{-1}$ and $\text{EW}_{\text{Ly}\alpha} > 15 \text{ \AA}$, as determined by applying our Ly- α luminosity model to the TNG100-1 simulation at $z = 2.44$. The solid blue line shows the median result from 1000 bootstrapping iterations, with the 16th–84th percentile spread of this process shown by the blue shaded region. For comparison, observational results from [Gawiser et al. \(2007\)](#) (black triangles) as well as narrow band (orange circles) and intermediate band (red squares) results from [Khostovan et al. \(2019\)](#) are shown.

Following the application of our Ly- α luminosity model to the TNG100-1 simulation volume, we plot the resulting Ly- α luminosity function in Fig. 2.6 as a sanity check. When comparing to observational data from [Rauch et al. \(2008\)](#); [Blanc et al. \(2011\)](#); [Cassata, P. et al. \(2011\)](#); [Konno et al. \(2016\)](#); [Sobral et al. \(2016\)](#) we find our model to mostly be in reasonably good agreement. Similarly we show the two-point correlation function for our model LAEs in Fig. 2.7. We again find that our model is in reasonably good agreement with the observational data, suggesting that (in addition to having a sensible distribution of luminosities) our LAEs are clustered in a manner that is similar to what is observed in the real Universe. These results suggest that the simple model we employ should be adequate to provide an indicative distribution of LAEs within our idealised IGM maps.

We first consider the fiducial ionisation model, displayed in the left column of Fig. 2.5. We observe an anti-correlation between the LAE density and δ_{F}/σ for all three protoclusters, and any CoSLAs are typically situated where the galaxy clustering is strongest. All maps generally exhibit a smaller δ_{F}/σ (less

Ly- α transmission) where the LAE density is largest. However, for the proto-cluster displayed in the top row of Fig. 2.5 there is an offset between where the LAEs are most strongly clustered around a massive halo with $M = 10^{13.5} M_{\odot}$ at $(x, y) \simeq (-4, 0) h^{-1} \text{cMpc}$, and the largest Ly- α transmission decrement at $(x, y) \simeq (3, -6) h^{-1} \text{cMpc}$. 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 $(x, y) \simeq (-4, 0) h^{-1} \text{cMpc}$ has been heated to $T > 10^6 \text{K}$ and is therefore highly ionised, whereas the IGM associated with the Ly- α transmission decrement in the lower right of the map is overdense but significantly cooler, with $T < 10^5 \text{K}$. The protocluster displayed in the middle row (see also the same object in Fig. 2.1) 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 regions of $T > 10^6 \text{K}$ gas around the protocluster centre of mass that further increases the Ly- α transmission.

In the central column of Fig. 2.5 we show the Ly- α tomographic maps for the same three protoclusters, but now using the *NoCol* ionisation model. As expected, all three protoclusters exhibit smaller values of δ_{F}/σ where the LAE density is largest (see $(x, y) \simeq (-4, 0) h^{-1} \text{cMpc}$ in the top row, $(x, y) \simeq (2, 0) h^{-1} \text{cMpc}$ in the middle row, and $(x, y) \simeq (0, 0) h^{-1} \text{cMpc}$ in the bottom row). By contrast, there is no significant change in δ_{F}/σ where the LAE density is lower. This is because hot, collisionally ionised 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* ionisation 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 ionisation 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 $N_{\text{HI}} > 10^{19} \text{cm}^{-2}$ in the fiducial model (i.e. damped systems) are over-ionised and have much *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 Chapter 3, 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. 2.5 we show the Ly- α forest tomographic maps for the *LoSo* ionisation 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. 2.2, where the brightest AGN in the model can ionise their surroundings up to $\sim 5 h^{-1}$ cMpc from the centre of their host halo. Whilst more subtle reductions in opacity can be seen around the QSOs at $(x, y) \simeq (0, 0) h^{-1}$ cMpc in both the middle and bottom rows of Fig. 2.5, the largest proximity zone in the simulation volume is shown in the upper right panel of Fig. 2.5. 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 $40 h^{-1}$ cMpc, has been reported by Mukae et al. (2020a), who find an HI underdensity in the CLAMATO tomographic maps associated with a LAE overdensity. These authors suggest this is due to the enhanced ionisation of the IGM by multiple nearby QSO proximity regions.

Due to the increased level of ionisation 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. 2.5. This is the result of absorption systems that are classified as damped ($N_{\text{HI}} > 10^{19} \text{ cm}^{-2}$) 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 absorbers in the fiducial model are reduced due to the proximity effect, and these regions are then classified as CoSLAs.

2.5.2 Protocluster masses and the correlation between LAEs and Ly- α transmission in smoothed tomographic maps

It is apparent from our qualitative discussion of the IGM tomographic maps in Section 2.5.1 that local ionisation plays an important role in the correlation between the tomographically reconstructed Ly- α transmission, coeval galaxies and

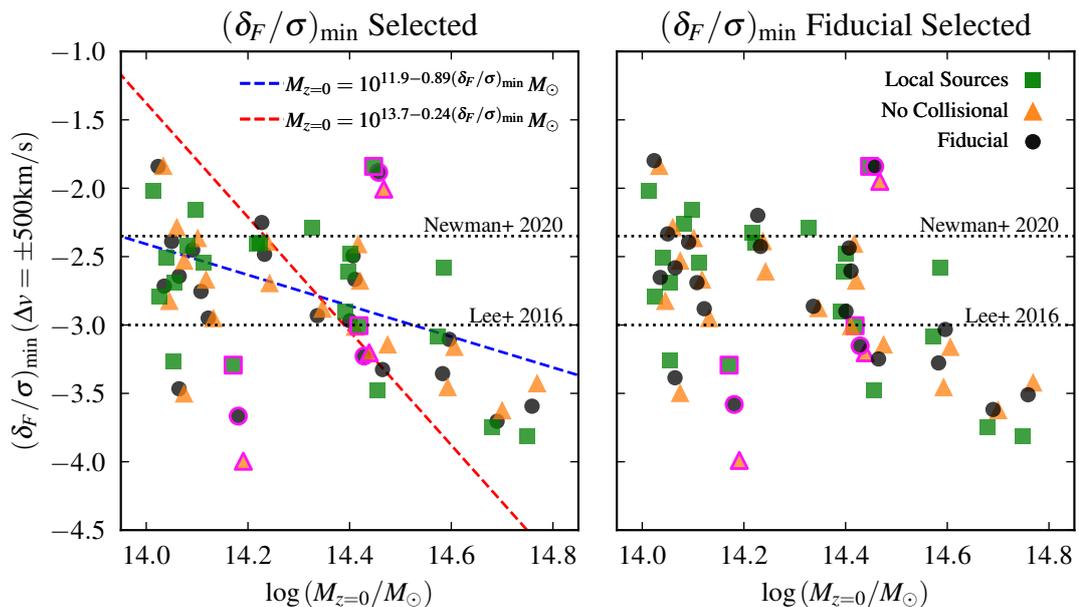


Figure 2.8: Scatter plot showing the relationship between the minimum relative Ly- α transmission, $(\delta_F/\sigma)_{\min}$, and the $z = 0$ cluster mass, $M_{z=0}$, in the smoothed tomographic maps for the 22 protoclusters in the TNG100-1 volume at redshift $z = 2.44$. We show the results for our fiducial (black circles), *NoCol* (orange triangles) and *LoSo* (green squares) ionisation 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. 2.5. The left panel shows the $(\delta_F/\sigma)_{\min}$ identified within R_{95} for each model. The right panel instead shows $(\delta_F/\sigma)_{\min}$ obtained when centring at the same physical location as $(\delta_F/\sigma)_{\min}$ in the fiducial model, where we have also fixed $\sigma = \sigma_{\text{fid}}$ for all the ionisation 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.

the distribution of coherent Ly- α absorption systems within *individual* protoclusters. However, we now instead consider how local ionisation variations impact on two different 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. 2.8. Candidate protoclusters are identified in both the CLAMATO and LATIS tomographic surveys by applying (δ_F/σ) thresholds to the smoothed Ly- α

2.5. THE EFFECT OF LOCAL IONISATION ON THE Ly- α
TRANSMISSION AROUND PROTOCLUSTERS

Table 2.2: Table showing the fraction of protoclusters from the TNG100-1 simulation that would be selected according to the thresholds defined by Lee et al. (2016) and Newman et al. (2020), for each of our three ionisation models. These data correspond to the results shown in the left panel of Fig. 2.8.

Selection threshold	Fiducial	NoCol	LoSo
$(\delta_F/\sigma) < -3.00$ (Lee+16)	8/22	9/22	7/22
$(\delta_F/\sigma) < -2.35$ (Newman+20)	19/22	19/22	18/22

tomographic maps. Newman et al. (2020) define their matter overdensity/proto-cluster candidates as regions with $\delta_F/\sigma < -2.35$ in the LATIS survey, whilst the CLAMATO survey Lee et al. (2016) use a more conservative value of $\delta_F/\sigma < -3$.

The left panel of Fig. 2.8 shows $(\delta_F/\sigma)_{\min}$ obtained from our smoothed tomographic 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 Chapter 3) means the minimum transmission can be located anywhere up to $\sim 10 h^{-1}$ cMpc from the true protocluster centre of mass. For comparison, the right panel shows $(\delta_F/\sigma)_{\min}$ when the maps for all models are instead all centred at the location of $(\delta_F/\sigma)_{\min}$ in the fiducial model. Additionally, in this case we use $\sigma = \sigma_{\text{fid}}$ for all three models. This allows us to focus on how the ionisation 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 ionisation models are small in the left hand panel of Fig. 2.8, with the largest differences occurring for the protocluster with $M_{z=0} = 10^{14.18} M_{\odot}$ that harbours the brightest AGN/most massive halo in the TNG100-1 volume (see the upper panels of Fig. 2.5). This suggests that smoothing the Ly- α tomographic maps on $\sim 4 h^{-1}$ cMpc scales mitigates 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 ionisation models occur on scales $< 1 h^{-1}$ cMpc (see Fig. 2.2). For comparison, we find that when selecting the same physical locations in the tomographic maps and fixing $\sigma = \sigma_{\text{fid}}$ (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 ionisation models are again fairly modest for most protoclusters.

From the left hand panel of Fig. 2.8, the protocluster completeness for the selection thresholds $(\delta_F/\sigma) < -2.35$ (< -3) in the fiducial model is 86% (36%), and this remains similar for both the *NoCol* and *LoSo* models (see Table 2.2).

However, we find the correlation 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. 2.8). Our best fit 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 panel of Fig. 2.8. This implies that, for a given (δ_F/σ) , our 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 following gas with $T > 10^6$ K from shocks and AGN feedback (see also figure 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 $256 h^{-1}$ cMpc.

In the smoothed tomographic maps in Fig. 2.5 we also observed an anti-correlation between the LAE overdensity, δ_{LAE} , and the relative transmission δ_F/σ . In Fig. 2.9 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 ionisation models are shown in the individual panels. The filled diamonds correspond to the LAEs in the tomographic 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 $\delta_{\text{LAE}} \lesssim 1.5$. The Spearman's rank correlation coefficient for the LAEs with $\delta_{\text{LAE}} < 1.5$ is -0.4 in all three models, consistent with a weak anti-correlation. This is followed by a flattening at $\delta_{\text{LAE}} \gtrsim 1.5$ due to the $4 h^{-1}$ cMpc Gaussian smoothing we apply to the tomographic maps; adopting a smaller smoothing scale reduces this apparent flattening. As was the case in Fig. 2.8, 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 δ_F/σ - δ_{LAE} plane, and are unimportant for the shape of the median trend. The colours of each point in Fig. 2.9 show the Ly- α luminosity of the LAEs, which are selected using the criteria $L_{\text{Ly}\alpha} > 10^{41.5}$ erg s $^{-1}$ and $\text{EW}_{\text{Ly}\alpha} > 15$ Å. There is no correlation (Spearman's rank coefficient -0.04 in all three models) apparent between δ_F/σ and the LAE luminosity, $L_{\text{Ly}\alpha}$ in our maps.

We also perform a comparison 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). 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

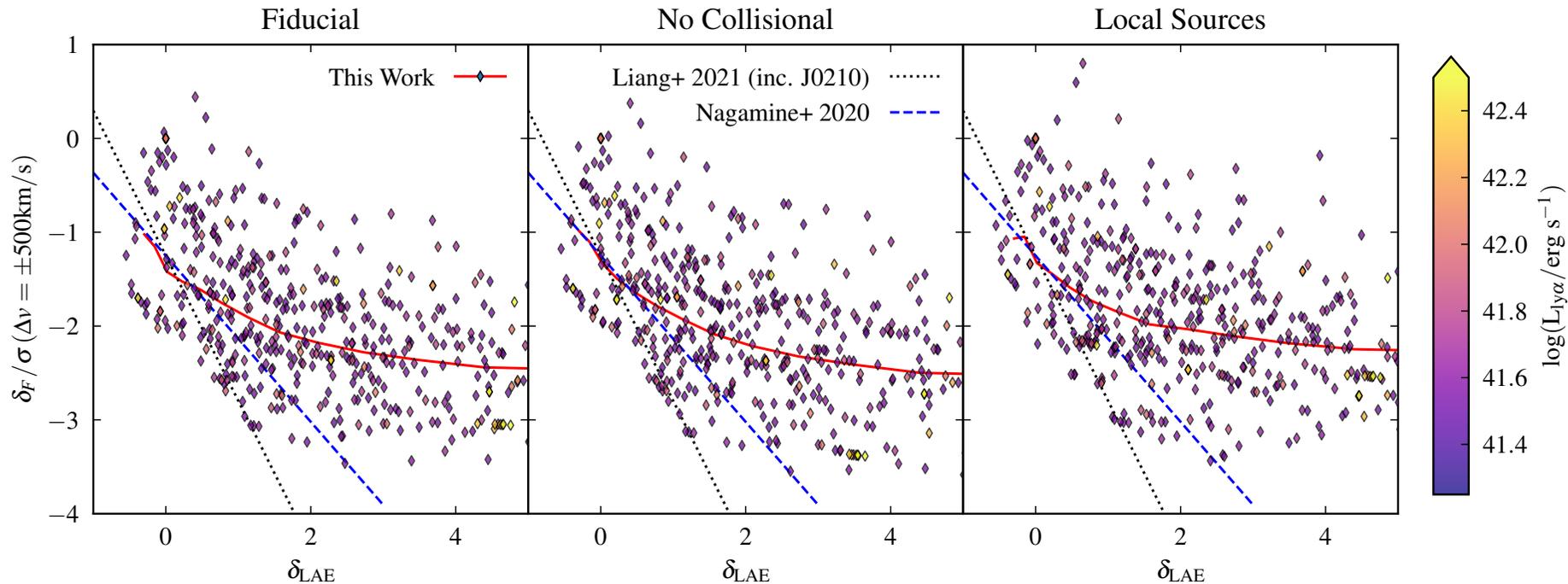


Figure 2.9: Scatter plot showing the relationship between the relative transmission, δ_F/σ , and the LAE overdensity, δ_{LAE} , in our smoothed tomographic maps at redshift $z = 2.44$. From left to right, we show the results for our fiducial, no collisional and local source ionisation models. The coloured diamonds correspond to the positions of each LAE with $L_{\text{Ly}\alpha} > 10^{41.5} \text{ erg s}^{-1}$ and $\text{EW}_{\text{Ly}\alpha} > 15 \text{ \AA}$ within R_{95} of all 22 protoclusters in the TNG100-1 volume. The colour scale shows the $\text{Ly-}\alpha$ luminosity of the LAEs, and the red curve shows the median trend obtained by randomly sampling within R_{95} of the smoothed protocluster maps. Our simulated data are compared to best fit linear relations to observational data from [Liang et al. \(2021\)](#) (black dotted line), and the results from GADGET-3/Osaka cosmological simulations ([Nagamine et al., 2021](#)). The linear fit from [Liang et al. \(2021\)](#) has been renormalised to coincide with the [Nagamine et al. \(2021\)](#) result at $\delta_{\text{LAE}} = 0$ (see text for details).

(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} using different selection criteria within cylindrical apertures. Our results are therefore not directly comparable, and in particular we expect the normalisation of the linear fits to their data to differ from this chapter (see figure 10 and the related discussion in Liang et al. (2021)). Following these authors, we therefore focus instead on comparing the gradient of the anti-correlation. We use their fit obtained when including the field BOSS J0210+0052, which has 11 proximate QSOs associated with it. In addition, we also compare to the linear fit in the $\delta_{\text{F}}-\delta_{\text{gal}}$ plane obtained by Nagamine et al. (2021) using the GADGET3-Osaka simulations, obtained using a cylindrical aperture matched to the Liang et al. (2021) measurement. To compare the gradients, we renormalise the Liang et al. (2021) fit to match the Nagamine et al. (2021) relation at $\delta_{\text{LAE}} = 0$, which is very similar to the median relation (red curve in Fig. 2.9) obtained in this chapter.

The slope of the linear fits are broadly consistent with the weak anti-correlation we observe at $\delta_{\text{LAE}} \lesssim 1.5$, consistent with the suggestion by Liang et al. (2021) that LAEs are preferentially located in regions with increased Ly- α absorption and hence larger HI densities at $z \simeq 2-3$. The scatter we observe in Fig. 2.9 is furthermore similar to the theoretical result from the GADGET-3-Osaka simulations (not shown in Fig. 2.9), but smaller than the observational data presented by Liang et al. (2021). In this chapter, we find the distribution of δ_{F} is not significantly altered in our different ionisation models, suggesting that local ionisation variations from (isotropic) AGN emission may not be the full explanation for the large observed scatter. We note, however, 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, $f_{\text{esc}}^{\text{Ly}\alpha}$, 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 address this question further (e.g. Barnes et al., 2011; Laursen et al., 2011; Gurung-López et al., 2019)

2.6 Summary

In this chapter we have investigated the effect that local ionisation variations in the intergalactic medium (IGM), due to the proximity effect from active galactic nuclei (AGN) and hot, $T > 10^6$ K gas from shocks and AGN feedback, have on

the Lyman- α (Ly- α) absorption signature of protoclusters in the IllustrisTNG simulations at redshift $z \simeq 2.4$. We consider three different local ionisation models in our analysis: a fiducial model with a spatially uniform ultraviolet (UV) background model, a second model that ignores the effect of collisional ionisation and self-shielding on the H I fraction in the IGM, and finally, a model where we incorporate spatial variations in the UV background due to the proximity effect from AGN. The impact of this potential “ionisation bias” in the Ly- α transmission profiles around massive haloes and idealised, smoothed Ly- α tomographic maps is then investigated. We quantify this by computing the relative Ly- α transmission averaged over a velocity window Δv , $\delta_{\text{F}} = (\langle F \rangle_{\Delta v} / \langle F \rangle) - 1$, 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 tomographic 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:

- We find local ionisation effects have a significant impact on Ly- α absorption in the vicinity of massive dark matter haloes with $M \geq 10^{12.8} M_{\odot}$ for impact parameters $b \lesssim 1 h^{-1} \text{cMpc}$ (see also Sorini et al., 2018). In particular, the presence of hot ($T > 10^6 \text{K}$) collisionally ionised gas will strongly increase δ_{F} within $\sim 1 h^{-1} \text{cMpc}$ 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 $1 h^{-1} \text{cMpc} \leq b \leq 5 h^{-1} \text{cMpc}$, corresponding to distances where photo-ionisation of the IGM begins to dominate over collisional ionisation. However, both of these effects become less important on larger scales ($b \gtrsim 5 h^{-1} \text{cMpc}$) and around less massive haloes with $M < 10^{12.8} M_{\odot}$ in our model.
- We construct idealised mock Ly- α tomographic maps around the 22 protoclusters with $M_{z=0} \geq 10^{14} M_{\odot}$ in the TNG100-1 volume (cf. Lee et al., 2016; Newman et al., 2020). We find that local ionisation effects play an important role in the correlation between the Ly- α transmission, coeval galaxies and CoSLAs within individual protoclusters. In particular, we find a spatial offset of $\sim 9 h^{-1} \text{cMpc}$ between a LAE overdensity around a massive halo, and the largest Ly- α flux decrement in a protocluster with $M_{z=0} = 10^{14.18} M_{\odot}$. This offset is due to collisionally ionised gas with temperature $T > 10^6 \text{K}$ 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 ionisation models, largely as a result of changes in the number of self-shielded, damped Ly- α absorbers with $N_{\text{HI}} \geq 10^{19} \text{ cm}^{-2}$ (see also Chapter 3).

- After smoothing the simulated Ly- α tomographic maps with a Gaussian of standard deviation $4 h^{-1} \text{ cMpc}$ (i.e. the standard approach for observational surveys) we find that local ionisation effects have a rather limited impact on the completeness of protocluster identification if using a fixed identification threshold of $(\delta_{\text{F}}/\sigma)_{\text{min}} \leq -2.35$ (Newman et al., 2020) or $(\delta_{\text{F}}/\sigma)_{\text{min}} \leq -3.00$ (Lee et al., 2016). For an ensemble of 22 protoclusters drawn from the TNG100-1 volume, we obtain a completeness of 82–86% and 32–41%, respectively for these thresholds if applied across all three of our ionisation models. These results suggest that, in addition to optimising protocluster detection (Stark et al., 2015), smoothing the Ly- α tomographic maps on $4 h^{-1} \text{ cMpc}$ scales also mitigates for a possible ‘ionisation bias’ in the inferred cluster masses. Within our model, this is because the largest differences in the Ly- α forest transmission typically occur on smaller scales 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 tomographic Ly- α maps using the fluctuating Gunn-Peterson approximation. We find $M_{z=0} = 10^{11.9-0.89(\delta_{\text{F}}/\sigma)_{\text{min}}} M_{\odot}$ for the 22 protoclusters in our fiducial model.
- 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 and Matthee, 2019) is used to populate the IGM tomographic maps with Ly- α emitting galaxies. In agreement with recent results from observations (Mukae et al., 2017; Liang et al., 2021) and hydrodynamical simulations (Nagamine et al., 2021), we observe a modest anti-correlation (Spearman’s rank correlation coefficient of ~ -0.4) between δ_{F} and the LAE overdensity, δ_{LAE} , for all three of our ionisation models at $\delta_{\text{LAE}} \lesssim 1.5$. This is consistent with recent suggestions that these galaxies

are 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 is also scatter around the median $\delta_{\text{F}}/\sigma-\delta_{\text{LAE}}$ relation that—along with the anti-correlation—is not significantly affected by changes in our ionisation model. The scatter is, however, smaller than observed by [Liang et al. \(2021\)](#). We find that local variations in δ_{F} due to (isotropic) AGN emission will not resolve this discrepancy. Instead, we speculate the origin of the larger scatter in the [Liang et al. \(2021\)](#) data may be associated with complex Ly- α radiative transfer effects that are not captured in our model.

In summary, models that incorporate local ionisation 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 reconstructions of the IGM density field on large scales ($\gtrsim 4 h^{-1}$ cMpc) using Ly- α forest tomography should already be relatively robust to variations in local ionisation when identifying protoclusters at $z \simeq 2.4$.

Chapter 3

Characterising Protoclusters With Line of Sight Lyman- α Absorption

The work within this chapter is adapted from the paper [Miller et al. \(2019\)](#) in Monthly Notices of the Royal Astronomical Society Volume 489 Issue 4 Pages 5381–5397. Here the introduction is shortened where material has already been covered in Chapter 1, as well as other minor formatting changes.

3.1 Introduction

Guided by mock spectra extracted from a $1 h^{-3} \text{cGpc}^3$ collisionless dark matter simulation combined with an approximate scheme for modelling Lyman- α (Ly- α) absorption ([Peirani et al., 2014](#)), [Cai et al. \(2016\)](#) (hereafter [C16](#)) identified protoclusters as being closely associated with what they call Coherently Strong intergalactic Ly- α absorption systems (CoSLAs). Whilst this approach has been successful at discovering a massive overdensity at $z = 2.32$ using a group of CoSLAs ([Cai et al., 2017a](#)), fully hydrodynamical simulations that directly connect the gas distribution at $z > 2$ to present-day clusters are required to establish whether coherent Ly- α absorption accurately tracks the progenitors of the largest collapsed clusters at $z = 0$, or if a more complex relationship between mass and Ly- α absorption may disrupt this picture. Furthermore, for protocluster searches that employ individual sight lines to detect CoSLAs, high column density Ly- α

absorbers that possess large damping wings can be a significant contaminant (cf. C16).

In this chapter we shall address these points by investigating the properties of Ly- α absorption systems within protoclusters at $z > 2.4$ using fully hydrodynamical simulations performed by the Sherwood (Bolton et al., 2017), EAGLE (Schaye et al., 2015) and Illustris (Vogelsberger et al., 2014) projects. With a typical volume of order 10^6 cMpc^3 , these simulations only contain of order ~ 10 clusters with masses $M_{z=0} \geq 10^{14} M_{\odot}$. On the other hand, unlike larger collisionless dark matter simulations, they include a full treatment of the gas physics using a variety of different (sub-grid) feedback models. More importantly, however, these simulations also have sufficient resolution to correctly incorporate Ly- α absorption features over a wide range of H I column densities, including absorption systems with damping wings. Our approach therefore provides a complementary perspective on the relationship between Ly- α absorption on small scales and the distribution of mass in protoclusters at $z \simeq 2.4$. Furthermore, as we explicitly track the formation of structure to $z = 0$ in the simulations, we are able to assess the completeness and contamination of a sample of protocluster candidates selected using coherent line of sight Ly- α absorption.

This chapter is structured as follows. We first introduce the hydrodynamical simulations we use in Section 3.2, and compare the H I column density distribution in each model to observational constraints in Section 3.3. In Section 3.4 we describe the basic properties of protoclusters in the simulations. We characterise mass overdensities on large scales by their Ly- α absorption in Section 3.5, and examine their connection to protoclusters in Section 3.6. Finally, the effectiveness of using line of sight Ly- α absorption to detect candidate protoclusters at $z \simeq 2.4$ is discussed in Section 3.7, before we summarise and conclude in Section 3.8. Throughout the chapter, we refer to comoving distance units using the prefix “c”.

3.2 Hydrodynamical simulations

The hydrodynamical simulations used in this chapter are summarised in Table 3.1. Outputs at two different redshifts were used for each model: the output closest¹ to $z = 2.4$, along with the corresponding output at $z = 0$. Below, we briefly describe the relevant properties of each simulation in turn, although further, extensive descriptions of these simulations may be found elsewhere.

¹This corresponds to $z = 2.4$ for Sherwood, $z = 2.478$ for EAGLE and $z = 2.44$ for Illustris.

Table 3.1: Hydrodynamical simulations used in this chapter. The columns list, from left to right: the simulation name, the box size in h^{-1} cMpc, the total number of resolution elements, the dark matter and gas particle masses (or equivalently for Illustris, the typical hydrodynamical cell mass), and the number of objects that have a total mass at $z = 0$ (as defined by a friends-of-friends halo finder) in the range $M_{z=0} \geq 10^{14} M_{\odot}$ (clusters), $10^{13.75} \leq M_{z=0}/M_{\odot} < 10^{14}$ (large groups) and $10^{13.5} \leq M_{z=0}/M_{\odot} < 10^{13.75}$ (small groups).

Name	Box size [h^{-1} cMpc]	N_{tot}	M_{dm} [M_{\odot}]	M_{gas} [M_{\odot}]	$M_{z=0} \geq 10^{14} M_{\odot}$ Clusters	$10^{13.75} \leq \frac{M_{z=0}}{M_{\odot}} < 10^{14}$ Large groups	$10^{13.5} \leq \frac{M_{z=0}}{M_{\odot}} < 10^{13.75}$ Small groups
Sherwood	80	2×1024^3	5.07×10^7	9.41×10^6	29	23	46
EAGLE	67.77	2×1504^3	9.70×10^6	1.81×10^6	10	15	33
Illustris	75	3×1820^3	6.26×10^6	1.26×10^6	14	18	29

3.2.1 Sherwood

The Sherwood project (Bolton et al., 2017) consists of a set of large, high resolution simulations of the Ly- α forest performed using a modified version of the smoothed particle hydrodynamics (SPH) code P-GADGET-3, last described in Springel (2005). In this chapter we predominantly use the 80-1024-ps13 simulation from Bolton et al. (2017), which we shall refer to as “Sherwood”. This simulation was performed in a $80^3 h^{-3} \text{cMpc}^3$ volume using the star formation and galactic outflow model developed by Puchwein and Springel (2013). This assumes a Chabrier (2003) initial mass function (IMF) and a wind velocity, v_w , that is proportional to the galaxy escape velocity, such that the wind mass-loading scales as v_w^{-2} .

The ultraviolet (UV) background follows the spatially uniform Haardt and Madau (2012) model, which quickly reionises the intergalactic medium (IGM) at $z = 15$. This model assumes the hydrogen is optically thin and in photo-ionisation equilibrium, and has an H I photo-ionisation rate of $\Gamma_{\text{HI}} = 9.6 \times 10^{-13} \text{ s}^{-1}$ at $z = 2.4$. Additionally, a small boost to the He II photo-heating rate, $\epsilon_{\text{HeII}} = 1.7 \epsilon_{\text{HeII}}^{\text{HM12}}$, has been applied at $2.2 < z < 3.4$ to better match observational constraints on the IGM temperature during and after He II reionisation (Becker et al., 2011). The Sherwood models adopt a Planck 2013 consistent cosmology (Planck Collaboration et al., 2014) with $\Omega_m = 0.308$, $\Omega_\Lambda = 0.692$, $\Omega_b = 0.0481$, $\sigma_8 = 0.826$, $n_s = 0.963$ and $h = 0.678$.

In addition to the fiducial simulation described above, we also use the 80-1024 simulation described in Bolton et al. (2017). This does not follow a physically motivated star formation model, but is identical to our fiducial run in all other respects. Instead, gas particles with density $\Delta = \rho/\langle\rho\rangle > 1000$ and temperature $T < 10^5 \text{ K}$ are converted directly into collisionless star particles. This “quick Ly- α ” approach increases computational speed while having a minimal effect on the low column density absorbers in the Ly- α forest (Viel et al., 2004). As this removes all the cold, dense gas in the simulation, the incidence of H I absorbers with column densities $N_{\text{HI}} \gtrsim 10^{17} \text{ cm}^{-2}$ will be underpredicted. We include this approach here, however, as it will more closely approximate results from earlier work using post-processed dark matter simulations that have insufficient resolution for modelling dense gas. We refer to this model as Sherwood “QLy- α ”.

3.2.2 EAGLE

The EAGLE (Evolution and Assembly of GaLaxies and their Environments) simulation (Schaye et al., 2015; Crain et al., 2015; McAlpine et al., 2016) was performed with a customised version of P-GADGET-3, where the standard SPH approach has been modified following the ANARCHY scheme described by Schaye et al. (2015). The simulation we use in this chapter is the Ref-L0100N1504 model. This has a smaller box size than Sherwood, corresponding to $67.77^3 h^{-3} \text{cMpc}^3$, but with a factor of $\simeq 5$ better mass resolution (see Table 3.1). Star formation is modelled using the approach of Schaye and Dalla Vecchia (2008) assuming a Chabrier IMF, while stellar feedback follows the stochastic, thermal scheme described in Dalla Vecchia and Schaye (2012). In addition, EAGLE follows gas accretion onto black holes; feedback from active galactic nuclei (AGN) is included using the methodology of Booth and Schaye (2009).

The spatially uniform, optically thin UV background used in EAGLE follows Haardt and Madau (2001); this quickly reionises the hydrogen at $z = 9$. An additional 2 eV per proton of energy from He II reionisation is also added at $z \simeq 3.5$, resulting in gas temperatures at mean density that are $\sim 4000 \text{ K}$ larger compared to Sherwood by $z = 2.4$. The H I photo-ionisation rate at $z = 2.478$ in the Haardt and Madau (2001) model is $\Gamma_{\text{HI}} = 1.4 \times 10^{-12} \text{ s}^{-1}$, a factor of 1.5 larger than the more recent Haardt and Madau (2012) UV background used in Sherwood. EAGLE assumes a ΛCDM cosmology consistent with Planck Collaboration et al. (2014), where $\Omega_m = 0.307$, $\Omega_\Lambda = 0.693$, $\Omega_b = 0.04825$, $\sigma_8 = 0.8288$, $n_s = 0.9611$ and $h = 0.677$.

3.2.3 Illustris

Finally, we also use the Illustris-1 simulation (referred to as ‘‘Illustris’’ for the remainder of this chapter) in our analysis (Vogelsberger et al., 2014; Nelson et al., 2015). Unlike Sherwood and EAGLE, Illustris is performed with the moving-mesh hydrodynamics code AREPO (Springel, 2010). Illustris has a slightly smaller volume than Sherwood ($75^3 h^{-3} \text{cMpc}^3$), but it has the highest mass resolution of the three simulations we consider. The star formation, stellar feedback and AGN feedback models are described in detail by Vogelsberger et al. (2013). Star formation in Illustris also uses a Chabrier IMF and is based upon the Springel and Hernquist (2003) model. The stellar feedback uses a variable winds approach, where the wind velocity, v_w , is scaled to the local dark matter velocity dispersion.

The spatially uniform UV background follows [Faucher-Giguère et al. \(2009\)](#), which quickly reionises the hydrogen in the simulation at $z = 10.5$. This UV background model has a HI photo-ionisation rate $\Gamma_{\text{HI}} = 6.1 \times 10^{-13} \text{ s}^{-1}$ at $z = 2.44$ —a factor of 0.6 smaller than [Haardt and Madau \(2012\)](#)—and produces gas temperatures at mean density around 1500 K lower at $z = 2.4$ compared to Sherwood.

Furthermore, unlike Sherwood and EAGLE, Illustris uses an on-the-fly prescription for the self-shielding of hydrogen from Lyman continuum photons, following the approach of [Rahmati et al. \(2013\)](#). As we will discuss below, incorporating self-shielding is necessary for correctly capturing the incidence of absorption systems with column densities $N_{\text{HI}} \geq 10^{17.2} \text{ cm}^{-2}$ (i.e. absorption systems that are optically thick to Lyman continuum photons). The Illustris simulations assume a WMAP-9 consistent cosmology ([Hinshaw et al., 2013](#)), with $\Omega_m = 0.2726$, $\Omega_\Lambda = 0.7274$, $\Omega_b = 0.0456$, $\sigma_8 = 0.809$, $n_s = 0.963$ and $h = 0.704$.

3.2.4 Generation of mock Ly- α absorption sight-lines

Mock Ly- α absorption spectra were extracted along sight lines drawn from the Sherwood simulation using the SPH interpolation scheme described by [Theuns et al. \(1998\)](#), combined with the Voigt profile approximation from [Tepper-García \(2006\)](#). Each sight line consists of 1024 pixels, and is drawn in a direction parallel to the simulation boundaries, starting from a position selected at random on the projection axis. A total of 30,000 sight lines (10,000 along each projection axis) were extracted from Sherwood, corresponding to an average transverse separation of $0.8 h^{-1} \text{ cMpc}$. The transmitted flux in each pixel is given by $F = e^{-\tau}$, where τ is the Ly- α optical depth.

Sight lines were extracted with the same average transverse separation from EAGLE and Illustris, although there are some small differences in the methodology due to the different hydrodynamics schemes employed by these models. For EAGLE, the M_4 cubic spline kernel ([Monaghan and Lattanzio, 1985](#)) used in the standard version of P-GADGET-3 was replaced with the C_2 [Wendland \(1995\)](#) kernel when performing the SPH interpolation. In Illustris there are no smoothing lengths, h_i , associated with the hydrodynamic cells. Instead, we assign these based on the volume, V_i , of each Voronoi cell, where

$$h_i = \left(\frac{3N_{\text{sph}}V_i}{4\pi} \right)^{1/3}, \quad (3.1)$$

and we adopt $N_{\text{sph}} = 64$ for the number of smoothing neighbours.

We furthermore set all star-forming hydrogen gas with $n_{\text{H}} > 0.13 \text{ cm}^{-3}$ to be fully neutral in Illustris, correcting for the unphysical neutral hydrogen fractions produced by the sub-grid star formation model. In addition, since Illustris already incorporates self-shielded hydrogen on-the-fly, the Rahmati et al. (2013) prescription for self-shielding was applied in post-processing to both Sherwood and EAGLE.

Lastly, in order to correct for the approximately factor of two uncertainty in the UV background H I photo-ionisation rate Γ_{HI} (Bolton et al., 2005), we rescale the optical depths in each pixel of the mock spectra by a constant, such that the Ly- α forest effective optical depth obtained from all the mock sight lines, $\tau_{\text{eff}} = -\ln \langle F \rangle$, where $\langle F \rangle$ is the mean transmitted flux, matches observational constraints (Theuns et al., 1998; Lukić et al., 2014). We use the τ_{eff} measurements from Becker et al. (2013) for this purpose. At $z = 2.4$, these data correspond to $\tau_{\text{eff}} = 0.20$.

3.3 The H I column density distribution function

Before proceeding to analyse the properties of protoclusters in Ly- α absorption, we must first verify if the simulations reproduce the observed distribution of H I column densities² at $z \simeq 2.4$. The column density distribution function (CDDF) obtained from the three simulations are displayed Figure 3.1, along with observational measurements from Kim et al. (2013), Noterdaeme et al. (2012) and Prochaska and Wolfe (2009). The simulated CDDFs at $N_{\text{HI}} < 10^{17} \text{ cm}^{-2}$ are calculated by integrating the H I number density in each pixel in the mock sight lines over 50 km s^{-1} windows (Gurvich et al., 2017). However, as absorption systems with $N_{\text{HI}} > 10^{17} \text{ cm}^{-2}$ are comparatively rare, we instead compute the CDDF by projecting the H I density for the entire simulation box onto a 2-D grid consisting of 20000^2 pixels (Altay et al., 2011; Rahmati et al., 2013; Bird et al., 2014; Villaescusa-Navarro et al., 2018). The discontinuity seen at $N_{\text{HI}} = 10^{17} \text{ cm}^{-2}$ in Figure 3.1 represents the point we switch from the integration method to the projection method.

²We will refer to Ly- α absorbers in four groups based on their column densities: Ly- α forest ($N_{\text{HI}} < 10^{17.2} \text{ cm}^{-2}$), Lyman-limit systems (LLSs, $10^{17.2} \leq N_{\text{HI}}/\text{cm}^{-2} < 10^{19}$), super Lyman-limit systems (SLLSs, $10^{19} \leq N_{\text{HI}}/\text{cm}^{-2} < 10^{20.3}$) and damped Ly- α absorbers (DLAs, $N_{\text{HI}} \geq 10^{20.3} \text{ cm}^{-2}$). We will also refer to SLLSs and DLAs collectively as “damped systems”.

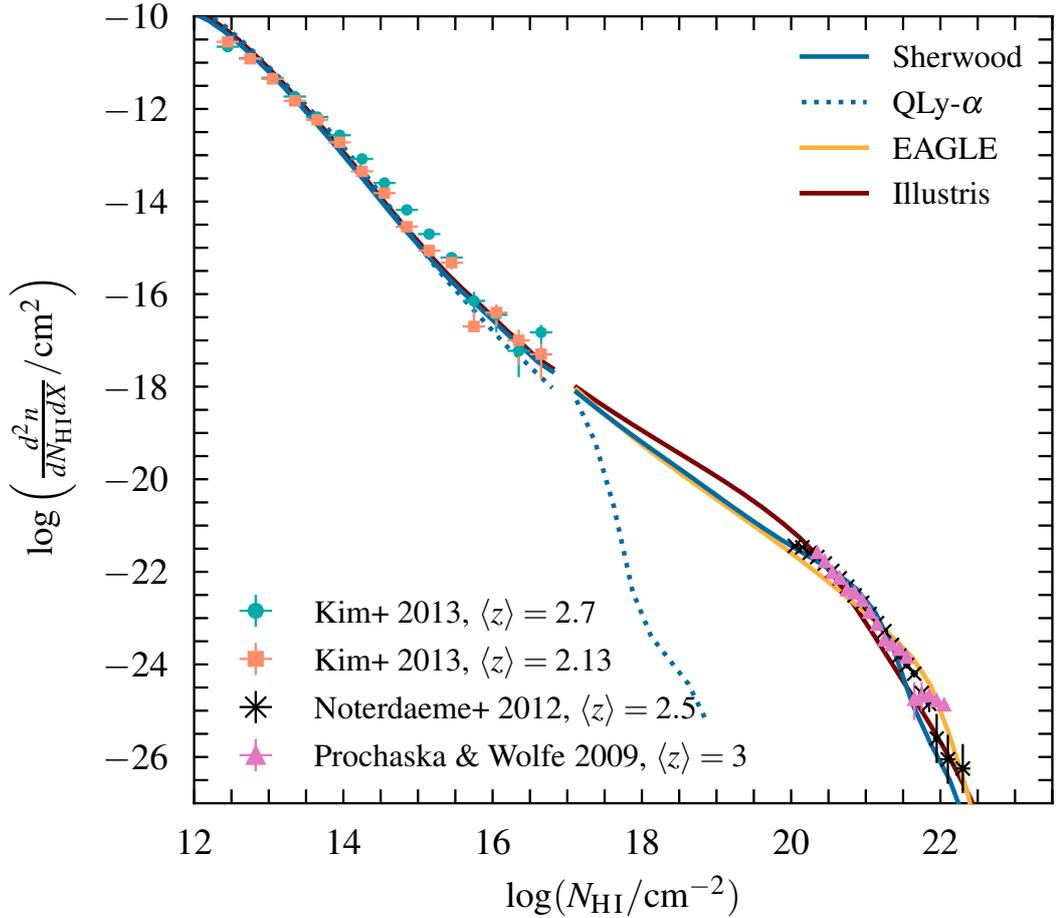


Figure 3.1: The HI column density distribution function (CDDF) at $z \simeq 2.4$ obtained from Sherwood (solid blue curve), EAGLE (orange curve) and Illustris (brown curve). For comparison, the dotted blue line represents the CDDF from the QLy- α simulation. Observational data from Kim et al. (2013) at $\langle z \rangle = 2.7$ and $\langle z \rangle = 2.13$ have been added at $N_{\text{HI}} < 10^{17} \text{ cm}^{-2}$, while data points from Noterdaeme et al. (2012) at $\langle z \rangle = 2.5$ and Prochaska and Wolfe (2009) $\langle z \rangle = 3$ are displayed at $N_{\text{HI}} > 10^{20} \text{ cm}^{-2}$.

The fiducial Sherwood simulation (solid blue curve), as well as the EAGLE (orange curve) and Illustris (brown curve) are in good agreement with observational data over a wide range of column densities (to within a factor of ~ 2), although the level of agreement at $N_{\text{HI}} > 10^{21.3} \text{ cm}^{-2}$ is possibly fortuitous given that a treatment of molecular hydrogen is not included in our analysis (Altay et al., 2013; Crain et al., 2017). The Illustris simulation also predicts a greater incidence of systems with $10^{17} \leq N_{\text{HI}}/\text{cm}^{-2} \leq 10^{20}$, relative to EAGLE and Sherwood. The reason for this difference is unclear, although we speculate this may be in part because the self-shielding correction is included on-the-fly in Illustris. The photo-heating (and hence pressure smoothing) experienced by the high column density, self-shielded gas will therefore differ from the post-processed EAGLE and Sherwood runs.

For comparison, we also show the CDDF from the QLy- α model in Figure 3.1 (blue dotted curve). The QLy- α model is only a good match to the observational data for the Ly- α forest at $N_{\text{HI}} \lesssim 10^{17} \text{ cm}^{-2}$. This is because most of the overdense gas that forms the strongest absorption systems has been converted into collisionless star particles (see Section 3.2.1). Overall, this comparison demonstrates that Sherwood, EAGLE and Illustris will adequately capture the incidence of Ly- α absorbers at $z \simeq 2.4$ over a wide range of HI column densities.

3.4 Simulated protoclusters

3.4.1 Protocluster definitions

We now turn to describe the protoclusters in the simulations at $z \sim 2.4$. We identify the protoclusters by tracking all particles in a friends-of-friends (FoF) group³ at $z = 0$. We will refer to the total mass of the $z = 0$ FoF group as the mass of the protocluster, $M_{z=0}$. Earlier studies (e.g. [Muldrew et al., 2015](#)) have instead identified simulated protoclusters by tracing the merger tree of $z = 0$ haloes back to the redshift of interest. However, as we are primarily interested in following large scale Ly- α absorption, in this chapter we also choose to follow the mass that is not bound in haloes at $z \simeq 2.4$.

We consider three mass bins in our analysis: $M_{z=0} \geq 10^{14} M_{\odot}$ (clusters), $10^{13.75} M_{\odot} \leq M_{z=0} < 10^{14} M_{\odot}$ (large groups) and $10^{13.5} M_{\odot} \leq M_{z=0} < 10^{13.75} M_{\odot}$ (small groups). We refer to the cluster progenitors as protoclusters, and the group progenitors as protogroups; we include the latter to provide a comparison to lower mass systems. The number of FoF groups in each bin is summarised in Table 3.1. We also define the size of each of the protoclusters/groups at $z \simeq 2.4$ using R_{95} , which corresponds to the radius of a sphere around the protocluster centre of mass that contains 95% of $M_{z=0}$.

Additionally, following an analysis of the Millennium simulation ([Springel et al., 2005](#)) by [Lovell et al. \(2018\)](#), for each protocluster we calculate two parameters derived from the principal semi-axes of a triaxial mass distribution, where $a \geq b \geq c$. The first is the axis ratio $s = c/a$ which provides a measure of sphericity, with $s = 1$ corresponding to a spherical distribution and $s \sim 0$ corresponding to

³Note this means that our quoted protocluster masses will be systematically larger than virial mass estimators such as M_{200} —the total mass enclosed within a sphere whose mean density is 200 times the critical density. For example, [White \(2000\)](#) demonstrate that FoF halo masses will be approximately 10% greater than M_{200} .

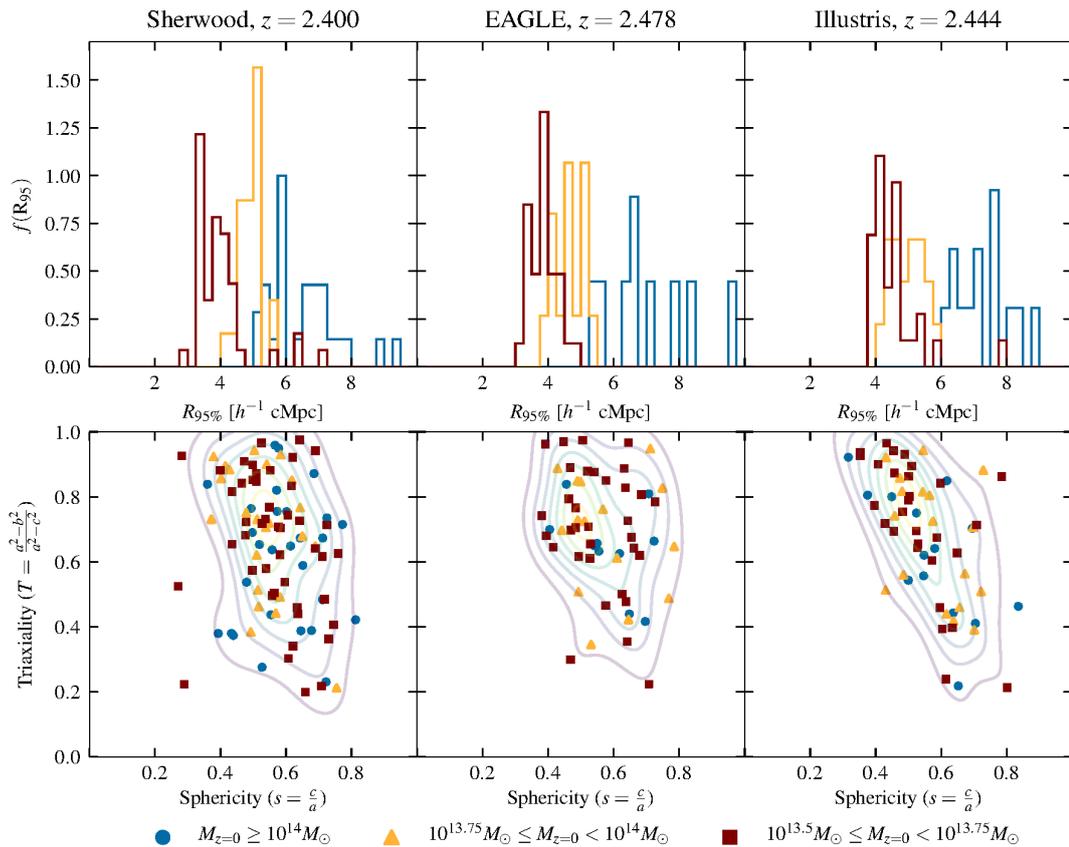


Figure 3.2: Top: Distribution of R_{95} for the protoclusters (blue) and large and small protogroups (orange and brown) in Sherwood (left), EAGLE (centre) and Illustris (right) at $z \sim 2.4$. Bottom: Scatter plots showing the triaxiality (where $T = 1$ and $T = 0$ are prolate and oblate spheroids, respectively) against the sphericity (where $s = 1$ and $s = 0$ are spherical and aspherical, respectively) of the protoclusters and protogroups, where a , b and c are the principal semi-axes and $a \geq b \geq c$. Contours show the kernel density estimate for all points.

a highly aspherical distribution. The second is the triaxiality parameter defined by Franx et al. (1991) as

$$T = \frac{a^2 - b^2}{a^2 - c^2}. \quad (3.2)$$

This quantifies whether the mass distribution resembles a prolate ($T \sim 1$) or oblate ($T \sim 0$) spheroid.

3.4.2 The size and shape of simulated protoclusters

The R_{95} distributions in the three mass bins are shown in the top row of Figure 3.2 for each simulation. The protoclusters tend to be larger than the protogroups, and typically have $R_{95} = 5\text{--}10 h^{-1} \text{ cMpc}$. These sizes are broadly consistent with the 90% stellar mass radii recovered from the Millennium simulation at $z = 2$ by

Muldrew et al. (2015), as well as the radial extent determined by Hatch et al. (2011a) for overdensities around radio galaxies at $z \sim 2.4$. The properties displayed in Figure 3.2 are also largely consistent between the different simulations, suggesting that (as expected on large scales) variations in the hydrodynamics schemes and sub-grid physics have little impact on the overall size and shape of the protoclusters. Furthermore, in all three cases, the smaller, lower mass protogroups do not differ significantly in shape from the protoclusters.

As noted by Lovell et al. (2018), however, a simple triaxial model does not fully capture the distribution of groups and filaments within protoclusters. In Fig. 3.3, we therefore display two-dimensional projections of protoclusters selected from each of the three simulations: these have masses $M_{z=0} = 10^{14.3} M_{\odot}$ (Sherwood), $M_{z=0} = 10^{14.4} M_{\odot}$ (EAGLE) and $M_{z=0} = 10^{14.3} M_{\odot}$ (Illustris). Several different protocluster morphologies are apparent, ranging from a structure dominated by a massive central halo (Sherwood) to a more diffuse structure with multiple, lower mass haloes (e.g. EAGLE). Fig. 3.3 also shows the location of the most massive FoF group (yellow stars) along with the next nine most massive FoF groups (white stars) in each protocluster. These are located in overdense regions where, in general, the gas temperatures are higher as a result of shock heating and feedback from stellar winds and/or black hole accretion. It is also apparent that regions with high H I fractions trace the overdense gas; photo-ionisation equilibrium with the UV background means the H I number density, n_{HI} , scales with the square of the gas density.

Finally, in Figure 3.4 we show the volume weighted distribution of the gas density, temperature and H I fraction for the protoclusters compared to the “field” (i.e. regions outwith R_{95}) in each of the simulations. The differences in gas properties between the protoclusters and the field are small, but as might be anticipated from an inspection of Fig. 3.3, the protoclusters are slightly more dense, hotter and have a larger H I fraction. Figure 3.4 also highlights the differences in gas properties between the three different simulations. Although the gas density distribution is similar in all three cases, there are differences in the temperature and H I fraction distribution that arise from the different UV background models (see Section 3.2). Illustris (EAGLE) has a slightly larger (smaller) average H I fraction compared to Sherwood. This is due to the smaller (larger) H I photo-ionisation rate used in the UV background model and—to a lesser extent—the dependence of the H I fraction on the lower (higher) gas temperature through the H II recombination rate.

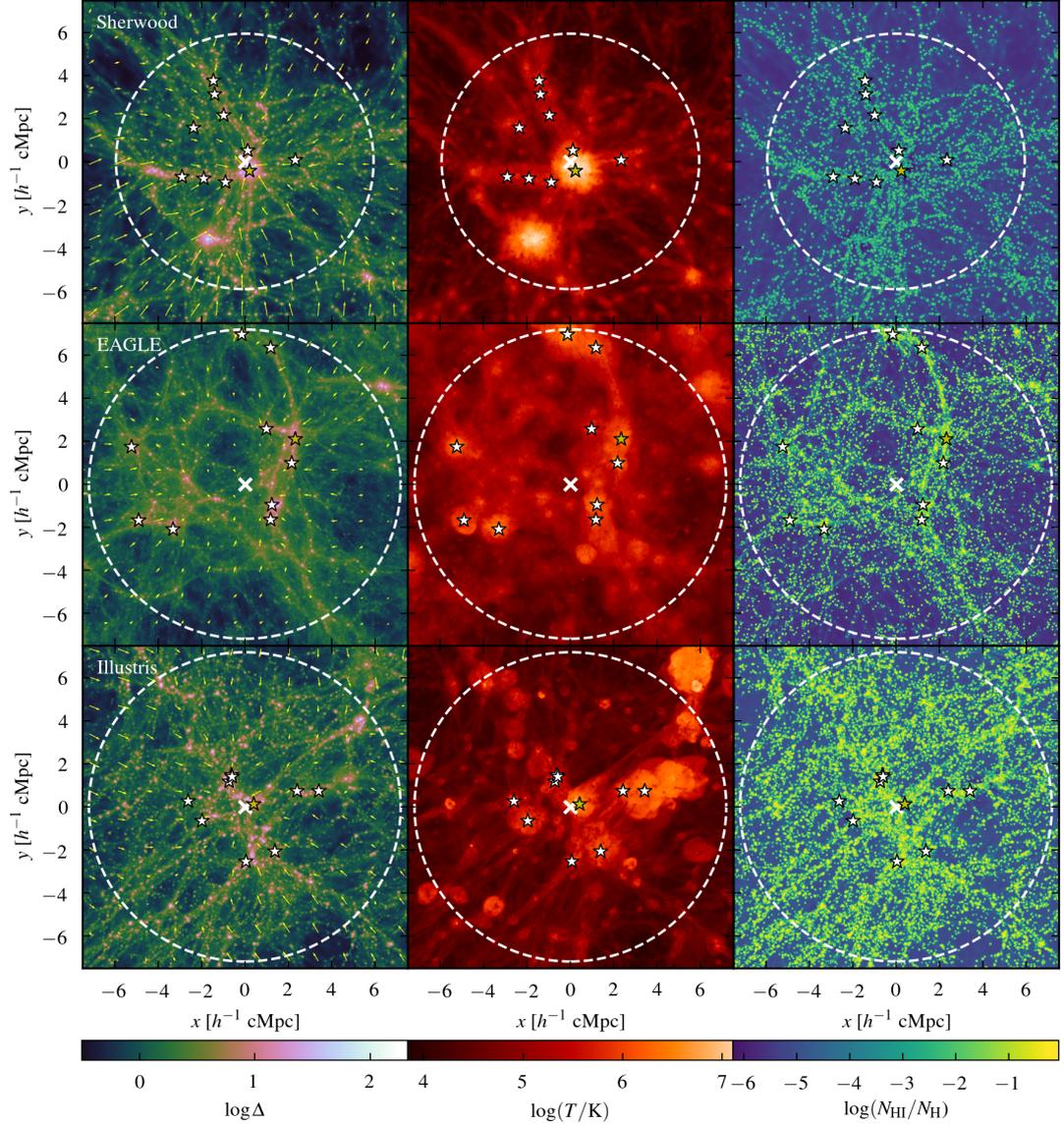


Figure 3.3: Projected maps of the normalised gas density, temperature and HI fraction for protoclusters at $z \sim 2.4$ with masses $M_{z=0} = 10^{14.3} M_{\odot}$, $M_{z=0} = 10^{14.4} M_{\odot}$ and $M_{z=0} = 10^{14.3} M_{\odot}$ in Sherwood (top row), EAGLE (middle row) and Illustris (bottom row). The projection depth along the z -axis is $15 h^{-1}$ cMpc, and is centred on the centre of mass of each protocluster. The dashed circles show R_{95} , and the star symbols correspond to the locations of the 10 most massive FoF groups within R_{95} . The yellow stars correspond to the most massive FoF groups in each protocluster, which have mass $M = 10^{13.6} M_{\odot}$, $M = 10^{12.8} M_{\odot}$ and $M = 10^{13.1} M_{\odot}$ in Sherwood, EAGLE and Illustris, respectively. The yellow arrows overlaid on the left column show the peculiar velocity field in the protoclusters relative to the centre of mass.

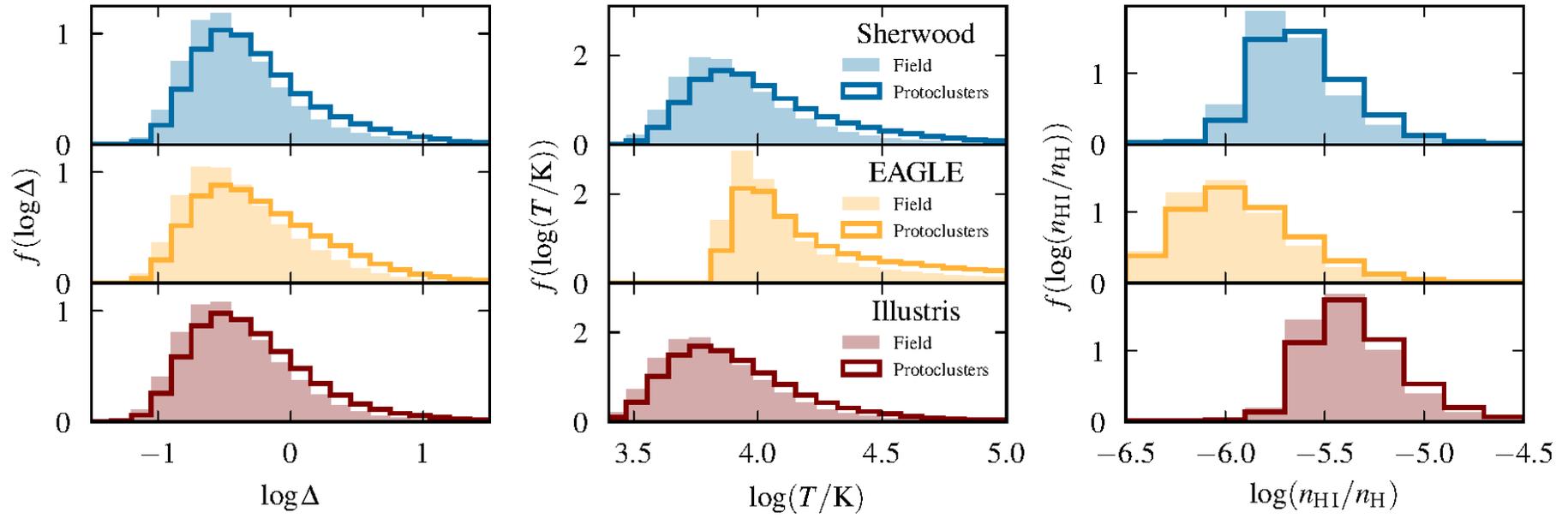


Figure 3.4: Volume weighted distributions of the normalised gas density, temperature and H I fraction in all protoclusters (line histograms) compared to the field (filled histograms) in Sherwood (blue), EAGLE (orange) and Illustris (brown).

3.5 Characterising mass overdensities in Ly- α absorption on $15 h^{-1}$ cMpc scales

Observationally, protoclusters are identified as large scale, high density regions. For this reason, we first investigate the Ly- α absorption properties of mass overdensities within the hydrodynamical simulations, with no consideration as to whether these eventually collapse to form a cluster by $z = 0$. From this, we determine the effectiveness of using Ly- α absorption to detect large scale overdense regions at $z = 2.4$. It is important to note, however, that this does not address how effectively Ly- α absorption probes the gas at $z \simeq 2.4$ that *actually* collapses to form clusters by $z = 0$ —we consider this further in Section 3.6.

3.5.1 Relationship between mass and effective optical depth on $15 h^{-1}$ cMpc scales

In order to assess how well Ly- α absorption traces large scale, high density regions, we consider the correlation between mass and Ly- α effective optical depth on $15 h^{-1}$ cMpc scales. We choose this scale as it corresponds to the characteristic size of protoclusters at $z \geq 2$ (Chiang et al., 2013; Muldrew et al., 2015), and also follows the scale adopted by C16.

Our procedure throughout this chapter is as follows. First, we sum the masses of all particles in $15^3 h^{-3}$ cMpc³ cubic volumes along every simulated Ly- α forest sight line (i.e. 30,000 in total for Sherwood, each with length $80 h^{-1}$ cMpc), and compute the mass overdensity, $\delta_m = \frac{m - \langle m \rangle}{\langle m \rangle}$ in each volume. Hence, δ_m is a 3D average, where m is the total mass within each $15^3 h^{-3}$ cMpc³ volume, and $\langle m \rangle$ is the mass contained in a $15^3 h^{-3}$ cMpc³ volume with density equal to the mean density, $\langle \rho \rangle = \rho_{\text{crit}} \Omega_m (1 + z)^3$, where

$$\langle m \rangle = 4.25 \times 10^{14} M_{\odot} \left(\frac{h}{0.678} \right)^{-1} \left(\frac{\Omega_m}{0.308} \right). \quad (3.3)$$

Masses above (below) this threshold represent overdense (underdense) volumes on $15 h^{-1}$ cMpc scales. Next, we associate every δ_m with the Ly- α forest effective optical depth, $\tau_{\text{eff}} = -\ln \langle F \rangle$, obtained from the $15 h^{-1}$ cMpc segments of simulated spectrum that pass directly through the centre of each $15^3 h^{-3}$ cMpc³

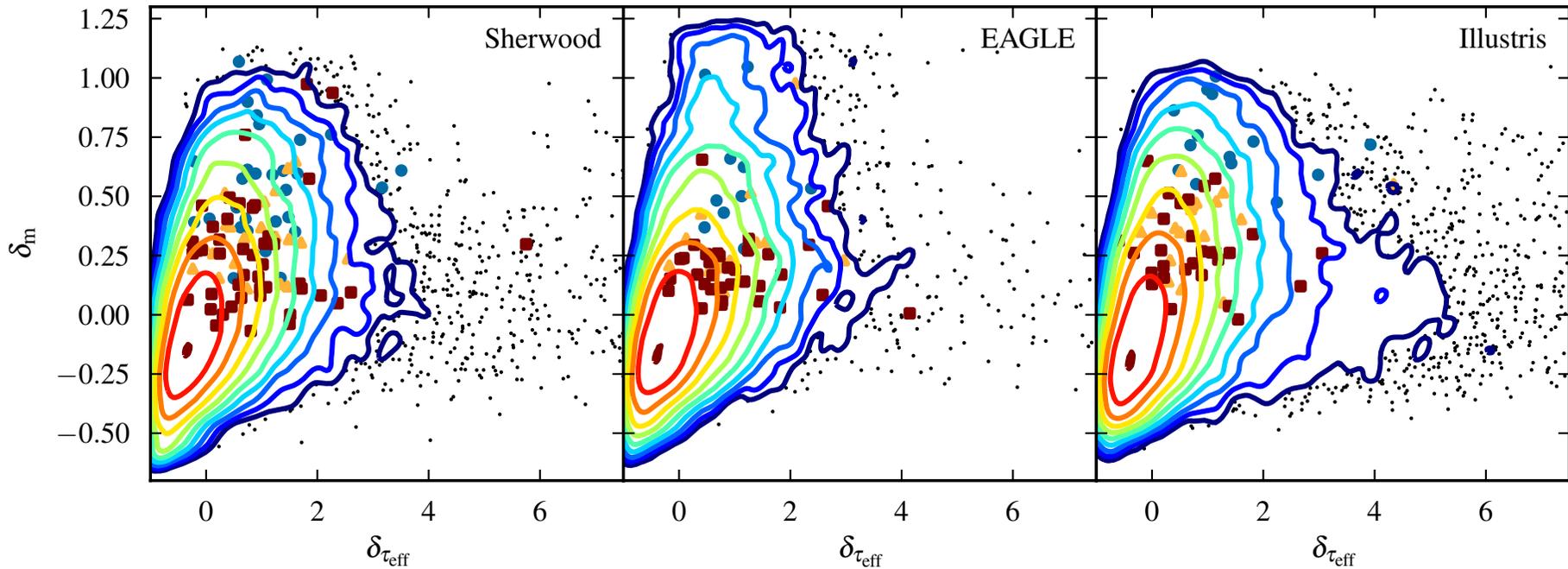


Figure 3.5: The relationship between δ_m and $\delta_{\tau_{\text{eff}}}$ on $15 h^{-1}$ cMpc scales (following Cai et al., 2016). Left to right, the panels show the results from Sherwood, EAGLE and Illustris. The contours represent the number density of data points relative to the central contour, with the number density decreasing by 1/3 dex with each contour. Black points represent every system which lies outside of these contours (i.e. with relative number density $< 10^{-3}$). Additional points corresponding to the centre of mass of the protoclusters (blue circles) large protogroups (orange triangles) and small protogroups (brown squares) are also displayed.

volume⁴. We then compute $\delta_{\tau_{\text{eff}}} = \frac{\tau_{\text{eff}} - \langle \tau_{\text{eff}} \rangle}{\langle \tau_{\text{eff}} \rangle}$, where τ_{eff} is the effective optical depth measured in each $15 h^{-1}$ cMpc spectral segment, and $\langle \tau_{\text{eff}} \rangle = 0.20$ is obtained from Becker et al. (2013) at $z = 2.4$. In this way, we associate $\delta_{\tau_{\text{eff}}}$ from every 1-D $15 h^{-1}$ cMpc segment with δ_m in the surrounding 3D $15^3 h^{-3}$ cMpc³ volume (see also C16).

The relationship between δ_m and $\delta_{\tau_{\text{eff}}}$, averaged over $15 h^{-1}$ cMpc scales, is displayed in Fig. 3.5. In Sherwood, EAGLE and Illustris the bulk of the simulation box corresponds to volumes close to the mean density, and there is a weak positive correlation between δ_m and $\delta_{\tau_{\text{eff}}}$. The centre of mass for all protoclusters and most protogroups (shown by the blue, orange and brown symbols) reside in overdense volumes, although note these are not usually associated with segments of high Ly- α opacity; the protocluster centre of mass will not always coincide with a halo or large value of τ_{eff} (cf. Fig. 3.3). The differences between the three simulations are minimal at $\delta_{\tau_{\text{eff}}} \lesssim 2$, where they each show a positive correlation with Spearman’s rank coefficients of 0.53–0.55. This suggests that variations in numerical methodology have little impact on the properties of the low density gas on large scales. By contrast, more significant differences between the simulations are apparent at $\delta_{\tau_{\text{eff}}} > 2$, corresponding to segments containing high column density, self-shielded absorption systems. Here, we find that the fraction of systems with $\delta_{\tau_{\text{eff}}} > 2$ is 80% (20%) greater in Illustris (Sherwood) than it is in EAGLE. Furthermore, the distribution of systems at $\delta_{\tau_{\text{eff}}} > 4$ also differs somewhat. When considering these systems only, we find a median $\delta_{\tau_{\text{eff}}}$ of 9.2, 9.9 and 6.7—with 16th–84th percentile spreads of 19.9, 29.5 and 9.9—for Sherwood, EAGLE and Illustris respectively.

These systems reside in both overdense and underdense volumes. The number of $\delta_{\tau_{\text{eff}}} > 2$ segments is greatest in Illustris, which is the simulation with the largest number of absorption systems at $10^{17} \leq N_{\text{HI}}/\text{cm}^{-2} \leq 10^{20}$ (see Fig. 3.1).

Haloes and galaxies are biased tracers of mass overdensity, therefore we expect a correlation between the presence of the high column density systems—responsible for the high $\delta_{\tau_{\text{eff}}}$ tail in Fig. 3.5—and δ_m . In Fig. 3.6 the fraction of $15 h^{-1}$ cMpc segments containing an absorption system, classed as either a DLA, SLLS, LLS or as Ly- α forest based on the largest constituent column density in each segment, is shown in bins of δ_m for each of the simulations. In all three simulations, the vast majority of the volume—irrespective of overdensity—is traced by Ly- α forest absorption, with 97.1%, 98.0% and 93.5% of $15 h^{-1}$ cMpc segments containing systems with a maximum column density of $N_{\text{HI}} < 10^{17.2} \text{ cm}^{-2}$ in Sherwood,

⁴That is, the average transmission for each 1-D spectral segment, $\langle F \rangle = e^{-\tau_{\text{eff}}}$, is obtained by averaging over all the pixels in a section of simulated spectrum with length $15 h^{-1}$ cMpc.

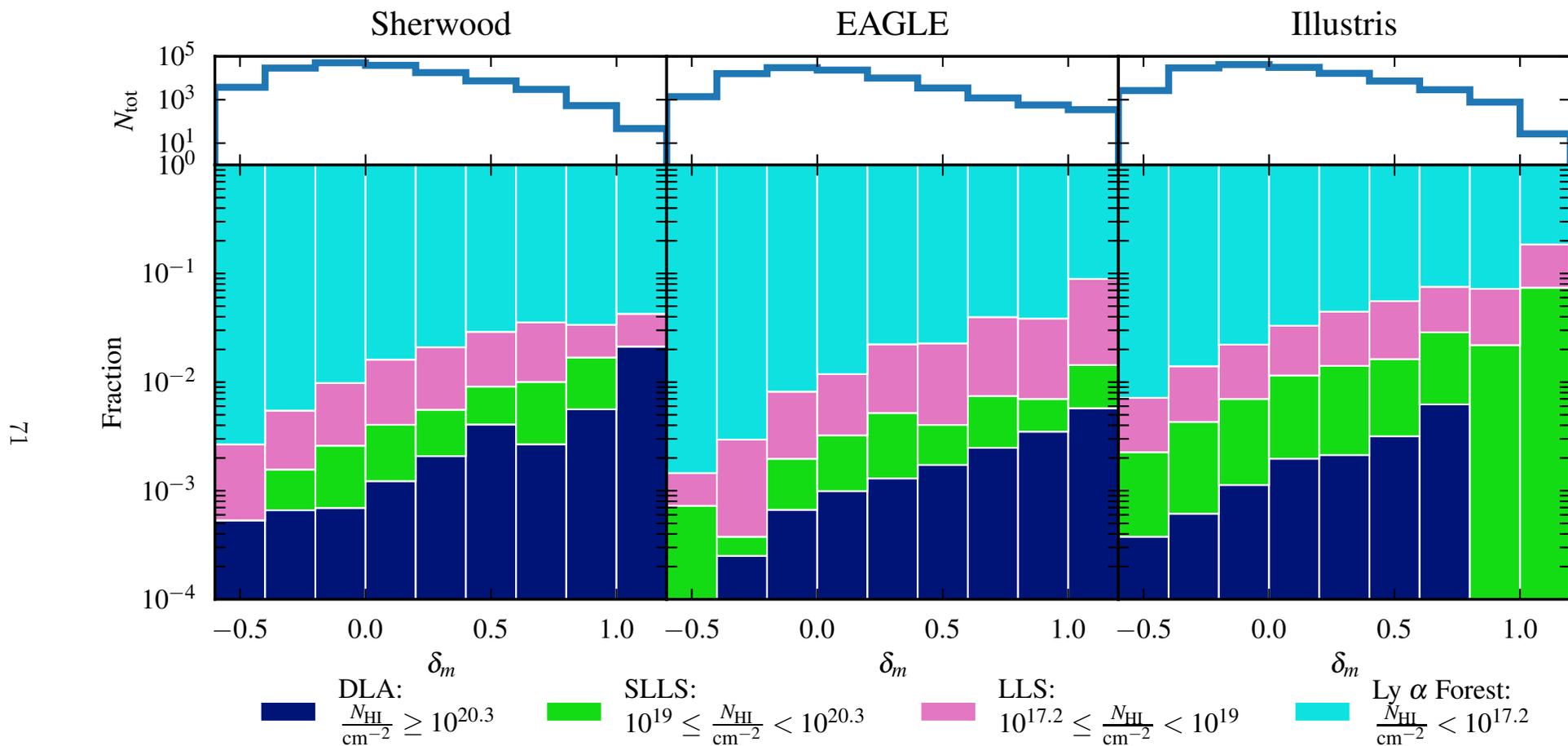


Figure 3.6: The fraction of $15 h^{-1}$ cMpc segments in bins of $\Delta\delta_m = 0.2$ whose largest constituent HI column density measured on 50 km s^{-1} scales corresponds to either Ly- α forest, LLSs, SLLSs or DLAs in Sherwood (left), EAGLE (centre) and Illustris (right). The upper panels display the total number of $15 h^{-1}$ cMpc segments in each bin.

EAGLE and Illustris, respectively. At all δ_m , Illustris has a greater fraction of segments containing SLLSs than the other simulations, with $\sim 20\%$ of all segments associated with volumes of $\delta_m > 1$ containing a LLS or SLLS. Note, however, that DLAs in Illustris are only present in segments associated with volumes of $\delta_m \leq 0.8$, implying that absorption systems in the most overdense volumes in Illustris have slightly lower H I fractions relative to EAGLE and Sherwood. This is broadly consistent with the H I column density distribution, where Illustris contains an excess of absorption systems with $10^{17} \leq N_{\text{HI}}/\text{cm}^{-2} \leq 10^{20}$ relative to Sherwood and EAGLE. Fig. 3.6 also demonstrates that damped systems with $N_{\text{HI}} > 10^{19} \text{ cm}^{-2}$ are present at all overdensities on $15 h^{-1} \text{ cMpc}$ scales, explaining the scatter with δ_m in the high $\delta_{\tau_{\text{eff}}}$ tail observed in Fig. 3.5.

This highlights the importance of self-consistently modelling the incidence of high column density H I absorption systems when identifying overdense volumes using $\text{Ly-}\alpha$ absorption. Damped systems produce large values of $\delta_{\tau_{\text{eff}}}$ regardless of large-scale environment, and so segments of high $\delta_{\tau_{\text{eff}}}$ will not uniquely probe overdense volumes. In addition to this, in all three simulations there is an increase in the fraction of segments containing systems with $N_{\text{HI}} \geq 10^{19} \text{ cm}^{-2}$ with increasing δ_m . This implies that highly overdense volumes, such as those that may collapse to form a cluster by $z = 0$, will contain a greater fraction of sight lines that pass through a damped system.

Throughout this work, we are comparing an averaged δ_m over a large volume to a much more localised $\delta_{\tau_{\text{eff}}}$ that is averaged along a single skewer. As the gas distribution is not smooth within this volume and $\delta_{\tau_{\text{eff}}}$ is sensitive to the gas density along the sight line, this will be a significant factor in the large scatter we see in the relation between the two quantities. If an average $\delta_{\tau_{\text{eff}}}$ from many sight lines spanning the full $15 h^{-3} \text{ cMpc}^3$ volume was used instead, one would expect to observe a tighter relation, with less dynamic range along the $\delta_{\tau_{\text{eff}}}$ axis. However, in regions where such a density of sight lines is available, the IGM tomography method would be better suited for identifying protoclusters than using CoSLAs as tracer objects (see section 1.3.2).

3.5.2 CoSLAs: are they mass overdensities?

From Fig. 3.6, it is evident that the majority of overdense volumes are traced by absorption from the $\text{Ly-}\alpha$ forest and LLSs. We now explore whether any of these overdense volumes would be classified as being associated with Coherently

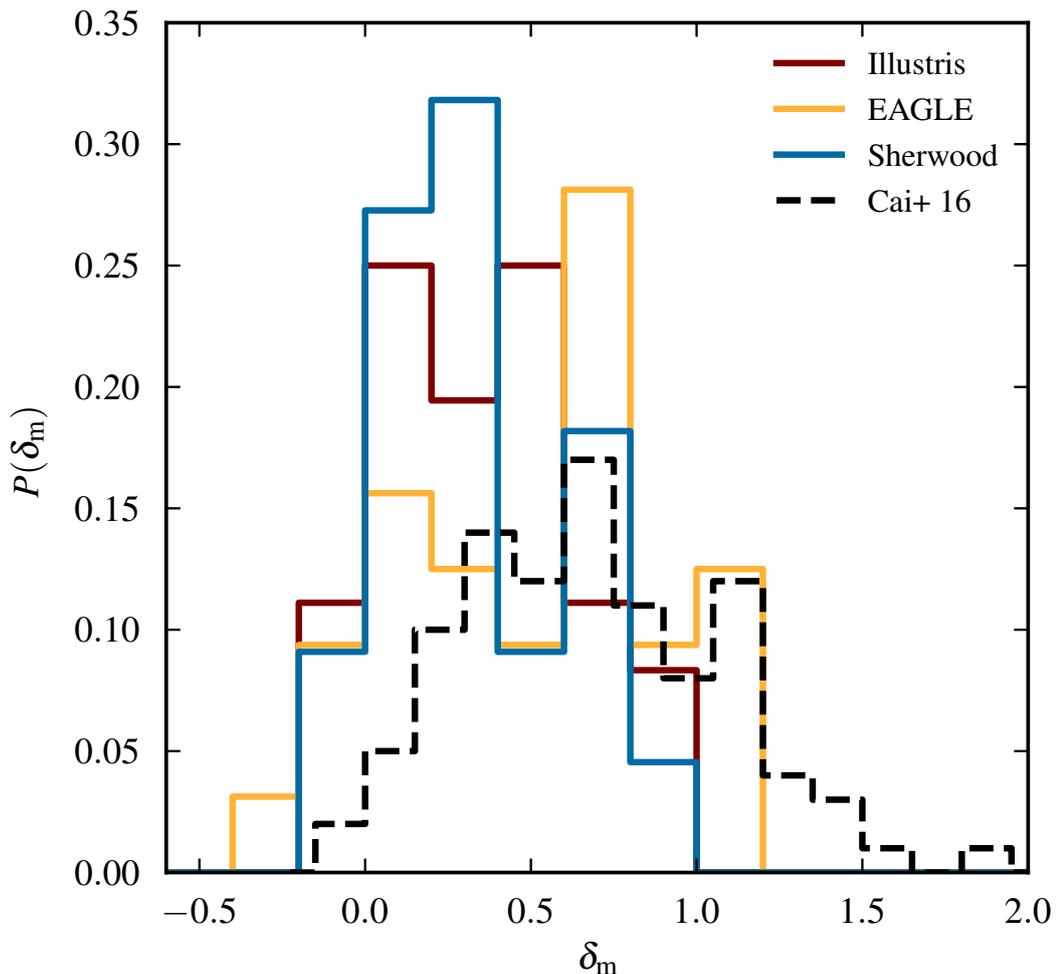


Figure 3.7: Probability distribution of δ_m for $15 h^{-1} \text{ cMpc}$ volumes with an associated $\delta_{\tau_{\text{eff}}} > 3.5$ (i.e. CoSLAs) after removing all sight lines containing damped systems in Sherwood (blue), EAGLE (orange) and Illustris (brown). The corresponding distribution from the analysis of a $1 h^{-3} \text{ cGpc}^3$ post-processed collisionless dark matter simulation by [C16](#) is displayed as a black dashed line. The total number of CoSLAs identified in each simulation is: Illustris (36), EAGLE (32), Sherwood (22).

Strong intergalactic Ly- α Absorption systems (CoSLAs) within the hydrodynamical simulations. [C16](#) associate protoclusters with CoSLAs, defining these as all segments of Ly- α absorption with $\delta_{\tau_{\text{eff}}} > 3.5$ on $15 h^{-1} \text{ cMpc}$ scales after excluding any high column density absorbers with damping wings. Therefore, before selecting CoSLAs, we first remove all sight lines that contain column densities $N_{\text{HI}} > 10^{19} \text{ cm}^{-2}$ in the simulated spectra. We take this approach to ensure that we not only remove the segments contain the damped system itself, but also any neighbouring segments where extended damping wings from these systems are still present. Due to the correlation between δ_m and the fraction of damped systems (see Fig. 3.6), this process preferentially removes segments corresponding to overdense volumes.

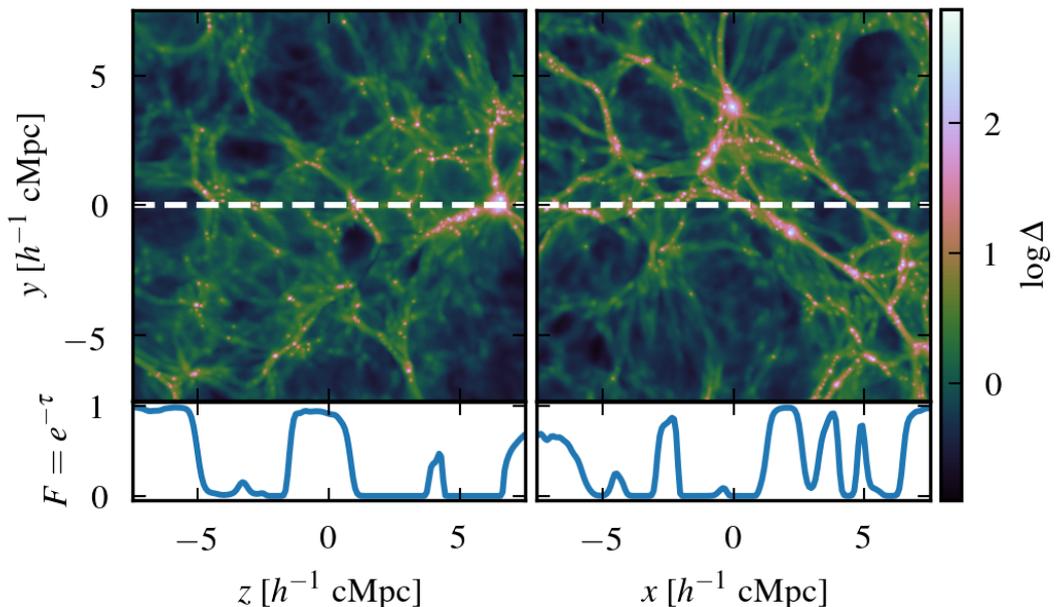


Figure 3.8: Examples of two $15 h^{-1}$ cMpc volumes with an associated $\delta_{\tau_{\text{eff}}} > 3.5$ (i.e. CoSLAs) in Sherwood, one of which is drawn from an underdense volume with $\delta_m = -0.09$ (left), with the other corresponding to an overdense volume with $\delta_m = 0.98$ (right). In each case, the upper panels display the normalised gas density projected over a distance of $1 h^{-1}$ cMpc centred on the sight line along which the Ly- α absorption spectrum is extracted (white dashed line), while the lower panels show the corresponding simulated Ly- α absorption. The underdense CoSLA has $\delta_{\tau_{\text{eff}}} = 3.76$ and a maximum $N_{\text{HI}} = 10^{18} \text{ cm}^{-2}$, while the overdense CoSLA has $\delta_{\tau_{\text{eff}}} = 3.68$ and a maximum $N_{\text{HI}} = 10^{15.3} \text{ cm}^{-2}$.

C16 used a collisionless dark matter simulation coupled with LyMAS (Peirani et al., 2014) to show that almost all CoSLAs correspond to mass overdensities in the range $0 < \delta_m < 2$, with $\sim 70\%$ having $\delta_m \geq 0.5$. Chiang et al. (2013) showed that, at similar redshifts and scales ($z = 2$ and $16.3 h^{-1}$ cMpc respectively), more than 80% of volumes with $\delta_m \geq 0.8$ will collapse to form clusters with $M_{z=0} > 10^{14} M_{\odot}$ by $z = 0$. The expectation is therefore that most CoSLAs will also probe structures that collapse to form clusters by $z = 0$.

In Fig. 3.7, we examine the δ_m associated with CoSLAs within each of the three hydrodynamical simulations. The distributions for Sherwood and Illustris are consistent with one another, with a median $\delta_m = 0.36 \pm 0.07$ and $\delta_m = 0.35 \pm 0.08$, respectively, where we estimate the uncertainty by bootstrapping with replacement. EAGLE has a higher median $\delta_m = 0.59 \pm 0.11$, formally differing from Illustris and Sherwood by 1.8σ . However, given the relatively small number of sight lines with coherent absorption on $15 h^{-1}$ cMpc scales in the simulations, this may reflect differences in the rare, massive structures within each simulation. Regardless of the median of each distribution, however, in all three simulations the

CoSLAs cover a broad δ_m range including $\delta_m \leq 0$, indicating that they do not exclusively probe large scale mass overdensities at $z = 2.4$. As a further illustration, Fig. 3.8 displays examples of CoSLAs from both underdense ($\delta_m = -0.09$) and overdense ($\delta_m = 0.98$) volumes in Sherwood. In both cases, we observe that a CoSLA arises from the alignment of structure along the line of sight. This can be due to LLSs associated with the alignment of several haloes punctuated by voids (left panel), or multiple lower column density absorbers associated with a more extended, filamentary structure (right panel).

These results differ from C16, who found CoSLAs to have a median $\delta_m = 0.75 \pm 0.03$ and almost exclusively correspond to overdense volumes. To explore the reason for this difference, we also consider the incidence of CoSLAs in the QLy- α simulation. The QLy- α model does not produce any damped systems and under produces LLSs (see Fig. 3.1) due to missing high density gas, but is otherwise identical to Sherwood. Consequently, the QLy- α model should better approximate the lower resolution simulation⁵ used in C16.

In Fig. 3.9, the δ_m distribution of CoSLAs in Sherwood is compared to the QLy- α simulation (this time in terms of number in each bin). The CoSLAs identified in QLy- α are coloured according to the maximum column density drawn from the matching locations within the Sherwood simulation. There are several points to note from Fig. 3.9. First, QLy- α contains two CoSLAs that are in segments containing SLLSs or DLAs in Sherwood (meaning these segments were discarded), while Sherwood contains eight segments classified as CoSLAs that are not present in QLy- α . These eight CoSLAs all contain LLSs or high column density Ly- α forest systems in Sherwood, but have lower column densities—and therefore lower values of $\delta_{\tau_{\text{eff}}}$ —in QLy- α . Furthermore, all these additional systems have $\delta_m < 0.4$, which acts to dilute the correlation between δ_m and $\delta_{\tau_{\text{eff}}}$. This emphasises the importance of correctly capturing LLSs in the models; not only will a fraction of the CoSLAs in the QLy- α model be contaminated by the presence of damped systems, but additional systems with $10^{16} \lesssim N_{\text{HI}}/\text{cm}^{-2} \leq 10^{19}$ are missed, resulting in an erroneously strong correlation between δ_m and $\delta_{\tau_{\text{eff}}}$ for CoSLAs. Second, the CoSLAs in QLy- α have a median $\delta_m = 0.40 \pm 0.02$, with $\sim 50\%$ of CoSLAs having $\delta_m \geq 0.5$. This median δ_m is consistent with the Sherwood simulation, although note it is still smaller when compared with C16.

⁵C16 use a collisionless dark matter simulation in a $1 h^{-3} \text{cGpc}^3$ volume with 1024^3 particles, yielding a particle mass of around $9.6 \times 10^{10} M_{\odot}$. For comparison, the typical dark matter particle mass needed to resolve the small scale structure of the Ly- α forest at $z \simeq 2$ is $\sim 10^7 M_{\odot}$ (Bolton and Becker, 2009).

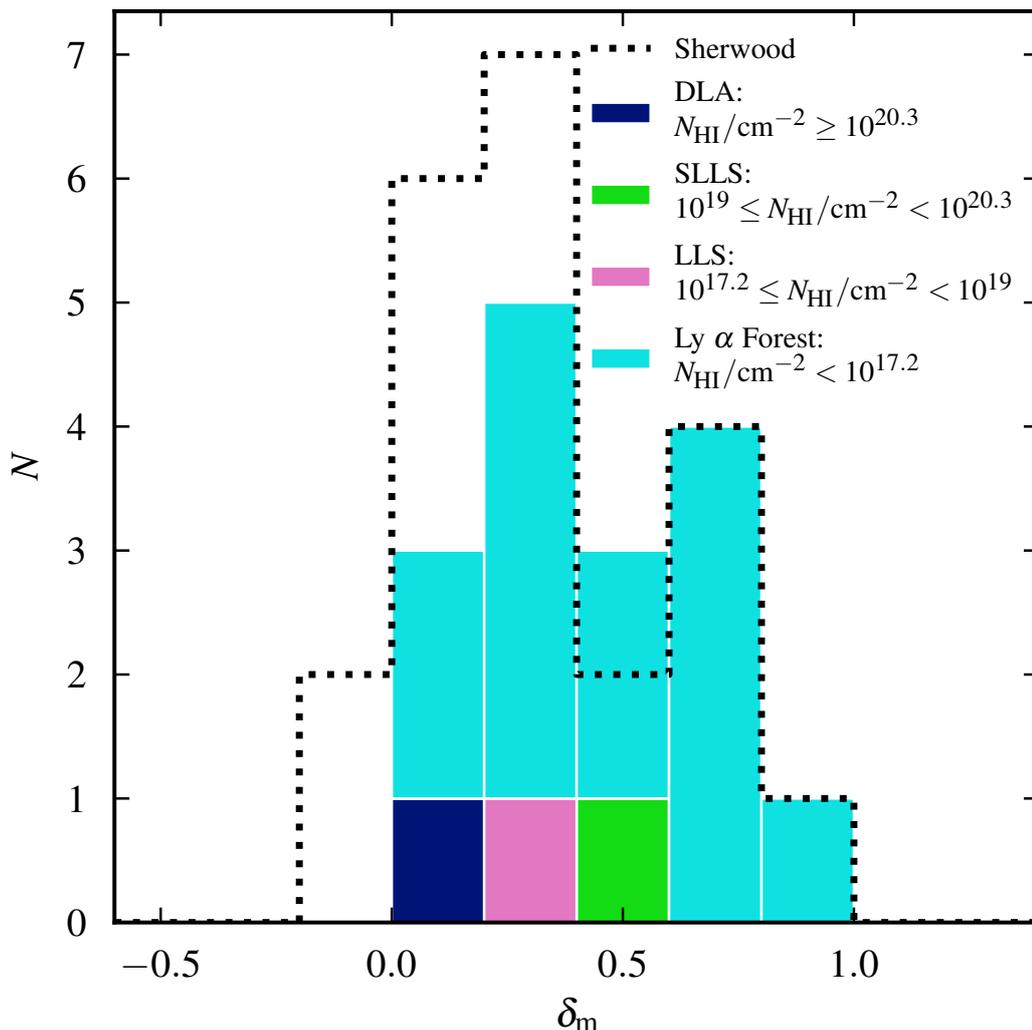


Figure 3.9: The filled histogram displays the number of $15 h^{-1} \text{ cMpc}$ volumes associated with segments of $\delta_{\tau_{\text{eff}}} > 3.5$ (i.e. CoSLAs) in bins of $\Delta\delta_m = 0.2$ for the QLY- α simulation. The colour of each stacked bar indicates the class of absorber each segment corresponds to in Sherwood, which is identical to QLY- α except for the lack of any cold ($T < 10^5 \text{ K}$), dense ($\Delta > 10^3$) gas. The corresponding distribution for Sherwood is shown as the black dotted line, where there is a greater frequency of CoSLAs with $\delta_m < 0.4$. The total number of $15 h^{-1} \text{ cMpc}$ segments with $\delta_{\tau_{\text{eff}}} > 3.5$ in each simulation is: Sherwood (22), QLY- α (16).

3.5.3 Effect of box size and mass resolution on CoSLAs

Another important factor to consider is the box size (and therefore maximum cluster mass) and mass resolution used in the simulations. We examine how mass resolution and box size impact on the relationship between mass overdensities and $\delta_{\tau_{\text{eff}}}$ on $15 h^{-1} \text{ cMpc}$ scales using four QLY- α runs drawn from the Sherwood simulation suite. We use QLY- α simulations for these tests as the 80-2048 and 160-2048 boxes have only been run using QLY- α . These simulations are summarised

3.5. CHARACTERISING MASS OVERDENSITIES IN $Ly-\alpha$ ABSORPTION ON $15 h^{-1} \text{ cMpc}$ SCALES

Table 3.2: Hydrodynamical simulations used to test the effect of box size and mass resolution on CoSLAs. All four simulations were performed using the quick- $Ly-\alpha$ approach that excludes dense, star forming gas (see Section 3.2). The columns list, from left to right: the simulation name, box size, number of resolution elements, dark matter and gas particle masses, and the number of CoSLAs per unit volume obtained using sight lines with an average transverse separation of $0.8 h^{-1} \text{ cMpc}$. All of the $80 h^{-1} \text{ cMpc}$ boxes were performed with initial conditions generated using the same random seed.

Name	Box size [$h^{-1} \text{ cMpc}$]	N_{tot}	M_{dm} [M_{\odot}]	M_{gas} [M_{\odot}]	N_{CoSLAs}/V [$h^3 \text{ cMpc}^{-3}$]
80-512	80	2×512^3	4.06×10^8	7.52×10^7	4.9×10^{-5}
80-1024	80	2×1024^3	5.07×10^7	9.41×10^6	3.1×10^{-5}
80-2048	80	2×2048^3	6.34×10^6	1.17×10^6	1.4×10^{-5}
160-2048	160	2×2048^3	5.07×10^7	9.41×10^6	5.5×10^{-5}

in Table 3.2. As previously, sight lines were extracted with the same average transverse separation of $0.8 h^{-1} \text{ cMpc}$ from all runs.

The effect of an increase in box size on the δ_m and $\delta_{\tau_{\text{eff}}}$ distributions for CoSLAs is displayed in Fig. 3.10. The δ_m distribution in the largest volume simulation (160-2048) extends toward higher values as a consequence of the rarer, more massive systems present in the larger volume. The number of CoSLAs per unit volume increases by 77% in the 160-2048 simulation (see Table 3.2) relative to the fiducial 80-1024 (QLy- α) model, although both the 80-1024 and 160-2048 models share similar median values of $\delta_m = 0.40 \pm 0.02$ and $\delta_m = 0.48 \pm 0.05$, respectively. The corresponding $\delta_{\tau_{\text{eff}}}$ distribution displays similar behaviour, with a tail that extends to larger values of $\delta_{\tau_{\text{eff}}}$ in the 160-2048 model. The effect of the ~ 1000 times larger volume used by C16 therefore likely explains some of the differences relative to this chapter. The CoSLAs selected in C16 (shown by the dashed black histogram in the upper panel of Fig. 3.10) probe rarer, more massive systems that are not present in our small, higher resolution models. We are unable to reliably assess how effectively CoSLAs probe overdense $15^3 h^{-3} \text{ cMpc}^3$ volumes with $\delta_m > 1$ using our fiducial $80^3 h^{-3} \text{ cMpc}^{-3}$ simulation box. However, this still does not fully explain the greater incidence of CoSLAs at overdensities of $\delta_m < 0.5$ within the hydrodynamical simulations compared to C16, even when using QLy- α models that underpredict the number of CoSLAs at $\delta_m \leq 0.4$ due to missing LLSs. A possible explanation is that the lack of high τ_{eff} systems at low δ_m is even more pronounced in much larger collisionless simulations, due to unresolved high column density absorption from small scale structure, but the exact reason for this difference remains unclear (this is addressed in more detail in Chapter 4).

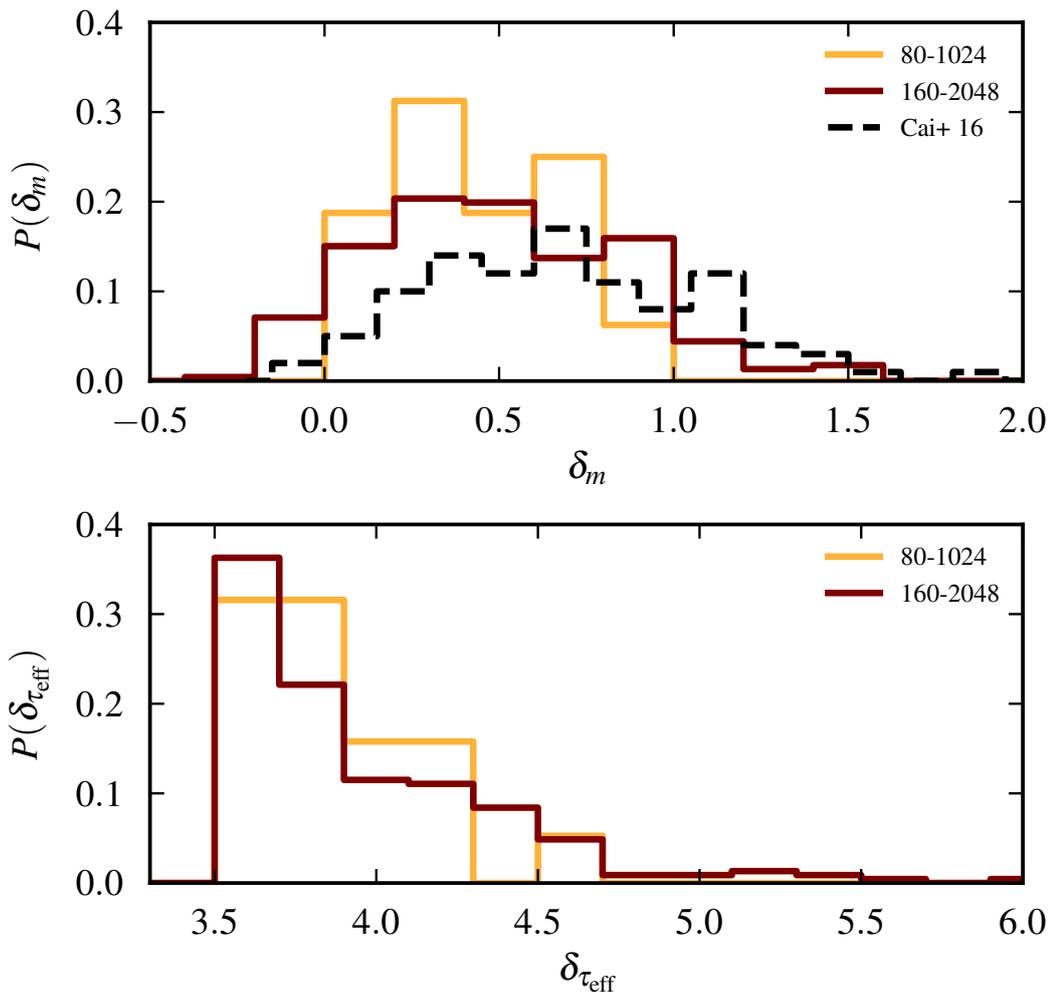


Figure 3.10: The probability distribution of CoSLAs as a function of δ_m (upper panel) and $\delta_{\tau_{\text{eff}}}$ (lower panel) for the QLy- α model (80-1024) compared to a simulation with the same mass resolution but a volume eight times larger (160-2048). The δ_m distribution from C16 is displayed as the dashed black histogram in the upper panel.

Mass resolution also has an important—but more subtle—effect on the typical $\delta_{\tau_{\text{eff}}}$ and δ_m associated with CoSLAs. The top panel of Fig. 3.11 highlights how the δ_m distribution of CoSLAs is shifted toward lower values for the 80-512 simulation when compared to the two higher resolution models. The median $\delta_m = 0.38 \pm 0.01$ for the 80-512 model, compared to median values of $\delta_m = 0.40 \pm 0.02$ and $\delta_m = 0.42 \pm 0.04$ for 80-1024 (QLy- α) and 80-2048, respectively. However, the average number of CoSLAs per unit volume also decreases by a factor ~ 2 when increasing the mass resolution by a factor of 8 (see Table 3.2), suggesting this quantity is not yet converged with mass resolution.

The lower panel of Fig. 3.11 elucidates the origin of this behaviour: at higher mass resolution the high $\delta_{\tau_{\text{eff}}}$ tail of the CoSLA distribution is truncated in comparison with the lower resolution runs. This is a consequence of the failure of the

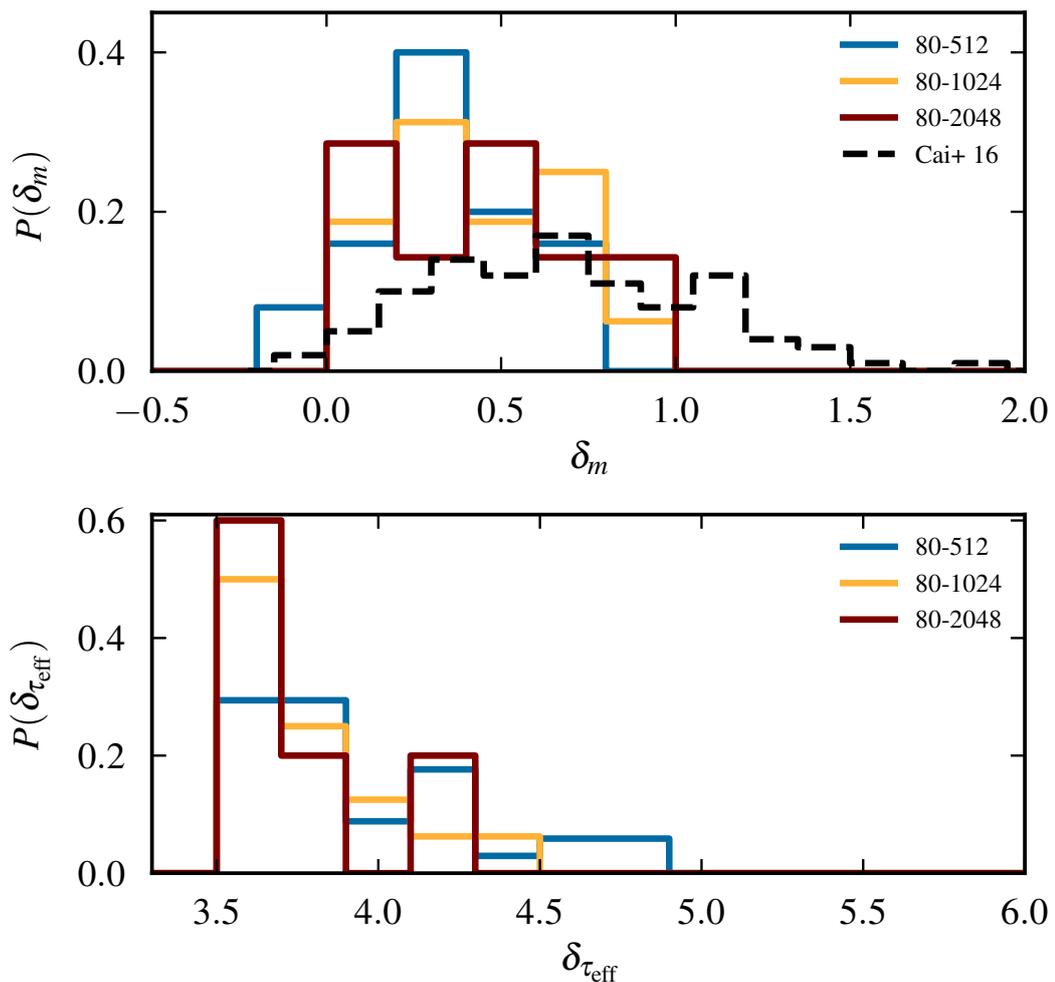


Figure 3.11: The probability distribution of CoSLAs as a function of δ_m (upper panel) and $\delta_{\tau_{\text{eff}}}$ (lower panel) for the QLy- α model (80-1024) compared to a simulations with the same box size but a particle mass that is eight times larger (80-512) or smaller (80-2048). The δ_m distribution from C16 is displayed as the dashed black histogram in the upper panel.

lower resolution runs to correctly resolve the structure of the Ly- α absorption, particularly in underdense regions, which are too opaque in low mass resolution models (Bolton and Becker, 2009). The effect of this poorly resolved underdense gas is to increase the number of volumes on $15 h^{-1} \text{cMpc}$ scales that are associated with segments which exceed the CoSLA identification threshold of $\delta_{\tau_{\text{eff}}} = 3.5$. An illustration of this effect is shown in Fig. 3.12, where Ly- α absorption from a randomly selected segment in the 80-2048 simulation (dark red) is compared directly to the 80-512 simulation (blue). This implies that—in addition to not capturing absorption from high column density, self-shielded absorbers—low mass resolution models that do not adequately resolve the structure of the Ly- α forest on small scales will overpredict the incidence of CoSLAs above a fixed $\delta_{\tau_{\text{eff}}}$ threshold.

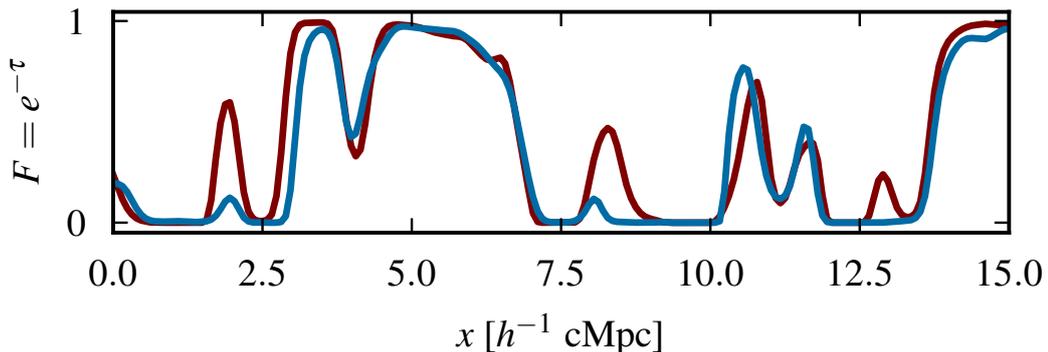


Figure 3.12: Simulated CoSLAs from the same segment drawn from the 80-512 (blue) and 80-2048 (dark red) models. This segment corresponds to $\delta_{\tau_{\text{eff}}} = 3.90$ in 80-512 but only $\delta_{\tau_{\text{eff}}} = 3.67$ in 80-2048, due to the increased transmission from the better resolved underdense gas in the higher resolution model.

In addition to this effect of lower resolution simulations producing higher column densities over $15 h^{-1} \text{ cMpc}$, [van deVoort et al. \(2018\)](#) showed a competing effect of greater numbers of dense clumps being resolved in very high resolution simulations. Using a mass resolution of $5.4 \times 10^4 M_{\odot}$ coupled with a 1 ckpc spatial refinement, the authors found an increase in the covering fraction of absorbers with $N_{\text{HI}} \lesssim 10^{19} \text{ cm}^{-2}$ over what was found without the additional spatial refinement (with the covering fraction of LLSs increasing from 18% to 30%). This results in the median neutral hydrogen column density increasing by 1.6 dex. Future study determining the resolution at which one effect begins to dominate over the other, as well as which effect is larger, is needed to fully understand the effect of mass resolution on the incidence of CoSLAs.

3.6 Characterising protoclusters in $\text{Ly-}\alpha$ absorption

We now address the question of how protoclusters, as opposed to overdensities, are characterised by $\text{Ly-}\alpha$ absorption. Whilst—observationally—protoclusters are identified as overdense regions at $z \sim 2$, when using simulations we have access to *a priori* knowledge of which structures will collapse to form a cluster along with the associated cluster mass. We now use this simulation based “true” definition of a protocluster to examine the opacity of protoclusters at $z = 2.4$ on $15 h^{-1} \text{ cMpc}$ scales. We do not expect variations in the numerical methodology used by Sherwood, EAGLE and Illustris to significantly change these results, so from this point onward we focus on analysing only the Sherwood simulation.

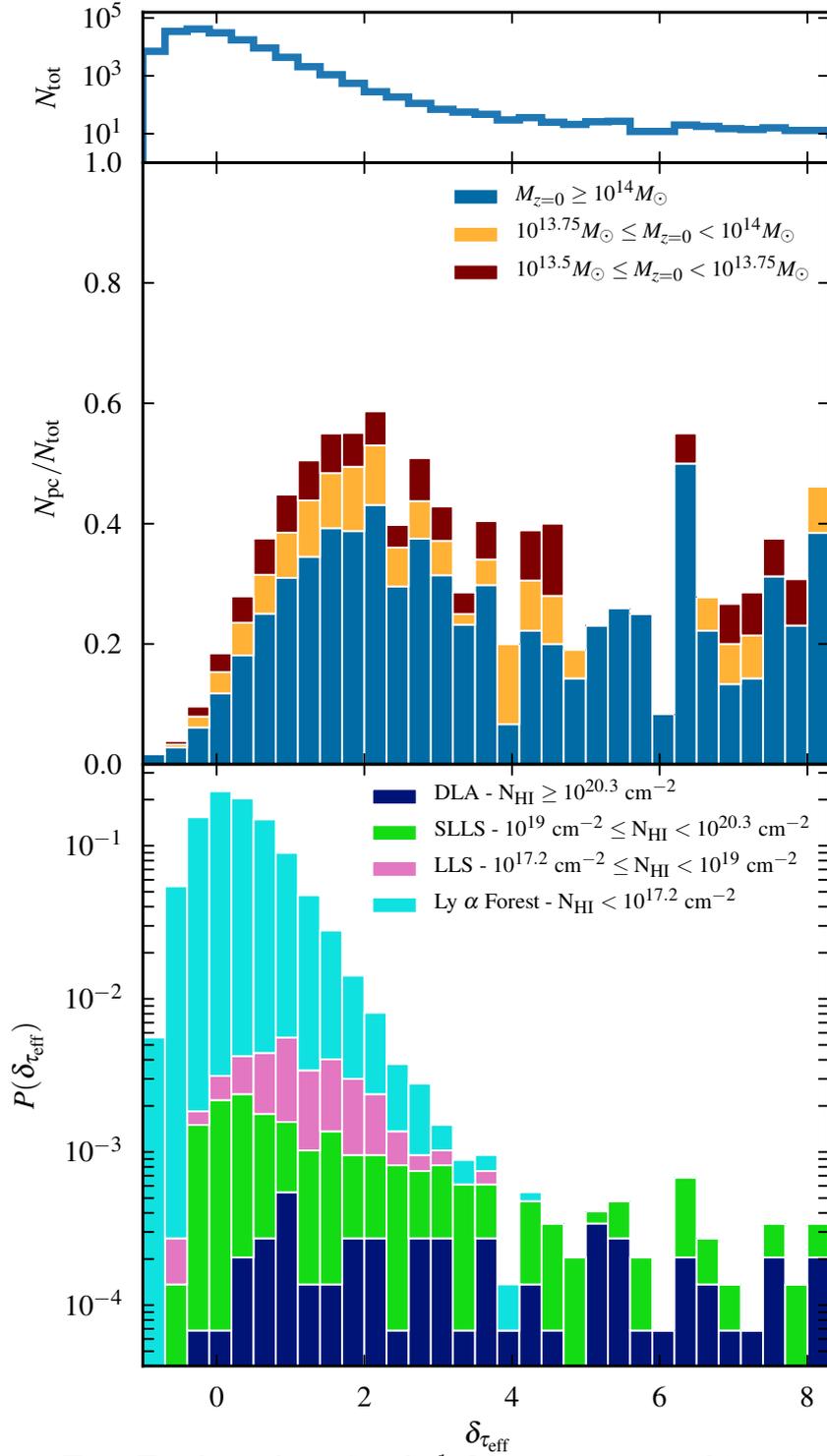


Figure 3.13: Top: Total number of $15 h^{-1} \text{cMpc}$ segments in bins of $\Delta\delta_{\tau_{\text{eff}}} = 0.3$. Middle: The fraction of $15 h^{-1} \text{cMpc}$ segments in bins of $\Delta\delta_{\tau_{\text{eff}}} = 0.3$ associated with protoclusters or protogroups. The colour of each stacked bar indicates the mass of the protocluster or protogroup associated with the absorption. Bottom: probability distribution for all $15 h^{-1} \text{cMpc}$ segments that are associated with protoclusters. The colour of each stacked bar indicates the maximum column density within each segment.

In Fig. 3.13, the fraction of all $15 h^{-1}$ cMpc segments in bins of $\Delta\delta_{\tau_{\text{eff}}} = 0.3$ that are associated⁶ with protoclusters (blue shading) or protogroups (orange and brown shading) are displayed in the central panel. In the event that a system is associated with more than one protocluster, it is associated with the most massive one only. The corresponding number of $15 h^{-1}$ cMpc segments in each bin is shown in the upper panel, with most segments close to the mean, $\delta_{\tau_{\text{eff}}} = 0$.

Note that in almost all bins, the majority of the “associated” segments are in the protoclusters with $M_{z=0} \geq 10^{14} M_{\odot}$. This in part reflects the fact that the protoclusters occupy a larger fraction of the simulation volume relative to the protogroups⁷. Typically fewer than 10% of segments with $\delta_{\tau_{\text{eff}}} \lesssim 0$ are associated with protoclusters, increasing to $\sim 40\%$ at $\delta_{\tau_{\text{eff}}} = 2$. At $\delta_{\tau_{\text{eff}}} \gtrsim 2$, however, this fraction declines, where on average only $\sim 25\%$ of segments are associated with protoclusters. Fig. 3.13 therefore demonstrates that for $\delta_{\tau_{\text{eff}}} < 2$ there is a positive, albeit weak, correlation between $\delta_{\tau_{\text{eff}}}$ and the likelihood of the associated volume collapsing to form a cluster by $z = 0$.

In the lower panel of Fig. 3.13, the probability distribution for $15 h^{-1}$ cMpc segments residing within the protoclusters is displayed as a function of $\delta_{\tau_{\text{eff}}}$, with the shading indicating the fraction of segments in each bin that contains a given maximum H I column density. The majority of segments passing through protoclusters only contain Ly- α forest absorption systems and have $\delta_{\tau_{\text{eff}}} \sim 0$: 28% of the segments which are associated with protoclusters are Ly- α forest with $\delta_{\tau_{\text{eff}}} < 0$, and 84% are entirely Ly- α forest with $\delta_{\tau_{\text{eff}}} < 1$. Thus, while the majority of segments with $\delta_{\tau_{\text{eff}}} < 1$ do not correspond to protoclusters, but rather to “field” regions, the majority of segments associated *with* protoclusters lie in the same $\delta_{\tau_{\text{eff}}}$ range and share identical characteristics. This reflects the fact that protoclusters at $z \simeq 2.4$ are extended over large scales (Muldrew et al., 2015); this means that it is very difficult—or impossible—to disentangle most protocluster sight lines from the field on $15 h^{-1}$ cMpc scales.

At larger values of $\delta_{\tau_{\text{eff}}}$ the segments associated with protoclusters become dominated by SLLSs and DLAs. At $\delta_{\tau_{\text{eff}}} > 2.9$, the majority of segments associated with protocluster contain damped systems, and 96% of the segments with $\delta_{\tau_{\text{eff}}} \gtrsim 3.5$ (i.e. the threshold for CoSLAs) that are associated with protoclusters arise from high density gas producing damped Ly- α absorption. At $\delta_{\tau_{\text{eff}}} > 3.5$

⁶A $15 h^{-1}$ cMpc segment is associated with a protocluster or protogroup if at least one third of the segment passes within R_{95} of any protocluster/group.

⁷Protoclusters occupy 7.7% of the Sherwood simulation volume at $z = 2.4$, whereas the large and small protogroups together only occupy 5.0%.)

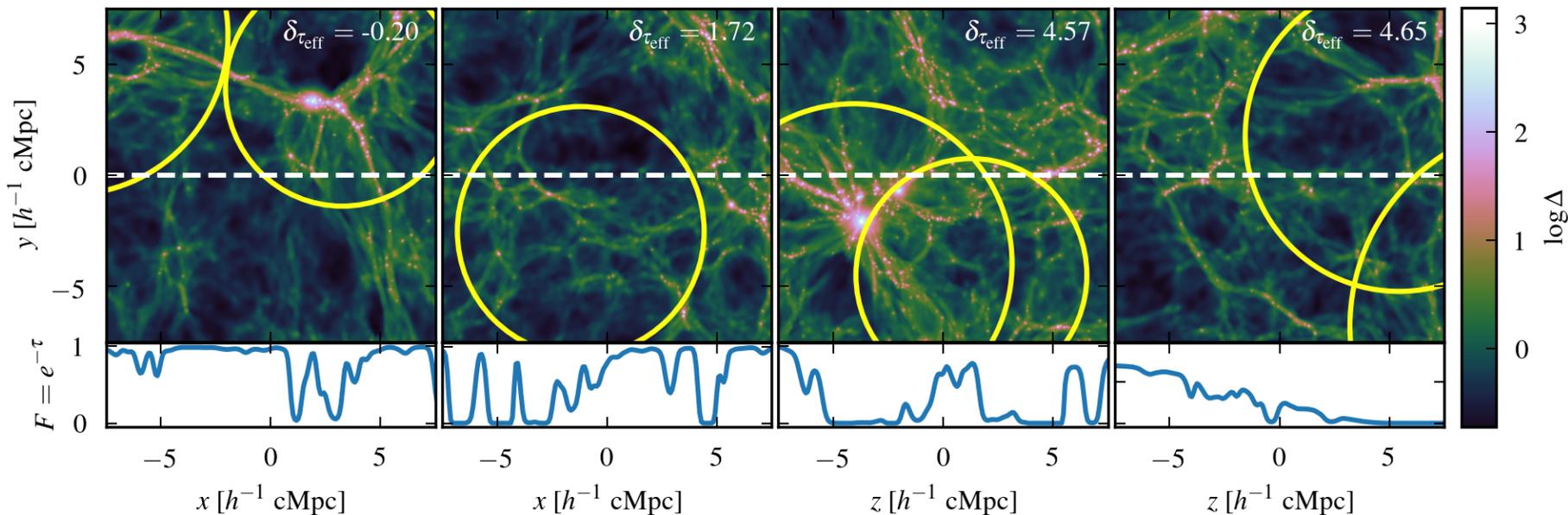


Figure 3.14: Example of four different $15 h^{-1} \text{ cMpc}$ segments that pass through protoclusters in Sherwood. The top panels show projected maps of the normalised gas density in a $15 h^{-2} \text{ cMpc}^2$ slice with a projection depth of $1 h^{-1} \text{ cMpc}$. The white dashed line shows the direction in which the Ly- α absorption spectrum was extracted, while the yellow circles indicate the cross section of all protoclusters that intersect with the slice. The bottom panels show the corresponding mock Ly- α absorption spectra. From left to right, each segment corresponds to: a low opacity segment with $\delta\tau_{\text{eff}} = -0.20$ containing a maximum H I column density $N_{\text{HI}} = 10^{13.8} \text{ cm}^{-2}$, a higher opacity segment with $\delta\tau_{\text{eff}} = 1.72$ and maximum $N_{\text{HI}} = 10^{14.8} \text{ cm}^{-2}$, a CoSLA with $\delta\tau_{\text{eff}} = 4.57$ and maximum $N_{\text{HI}} = 10^{16.9} \text{ cm}^{-2}$, and a DLA with $\delta\tau_{\text{eff}} = 4.65$ and maximum $N_{\text{HI}} = 10^{20.5} \text{ cm}^{-2}$.

SLLSs and DLAs are the dominant probe of gas that forms clusters by $z = 0$, but—critically for protocluster selection by Ly- α absorption—the middle panel of Figure 3.13 shows that these are not uniquely associated with protoclusters at $z \simeq 2.4$, but are more likely to be in the field.

Fig. 3.14 displays selected examples of segments that pass through protoclusters that cover the full range of $\delta_{\tau_{\text{eff}}}$ observed in the simulation. Most sight lines passing through protoclusters with $\delta_{\tau_{\text{eff}}} < 2$ are exemplified by the two left-hand panels in Fig. 3.14. In these cases the sight line passes through two filamentary structures perpendicular to the sight line (left panel, $\delta_{\tau_{\text{eff}}} = -0.20$) or through a small region of overdense gas within a protocluster (second left panel, $\delta_{\tau_{\text{eff}}} = 1.72$), but in both cases do not exhibit extended regions of saturated Ly- α absorption.

By contrast, in the two right-hand panels of Fig. 3.14, two examples of segments that pass through protoclusters and have $\delta_{\tau_{\text{eff}}} > 3.5$ (i.e. the CoSLA threshold) are displayed. The right-most panel ($\delta_{\tau_{\text{eff}}} = 4.65$) displays strong absorption due to the presence of an extended damping wing arising from a DLA; there is no large scale gas overdensity present along the sight line. The segment displayed on the second right ($\delta_{\tau_{\text{eff}}} = 4.57$) has no damped absorbers, but instead the absorption arises from haloes and a filament that are aligned with the sight line and are associated with two intersecting protoclusters. The high value of $\delta_{\tau_{\text{eff}}}$ in this case results from extended Ly- α forest absorption. This implies that it is the orientation of dense gas with respect to the sight line that has the greatest impact on the value of $\delta_{\tau_{\text{eff}}}$, rather than either the overdensity δ_{m} (see also Fig. 3.8) or the presence of a protocluster.

3.7 Selecting protoclusters with line of sight Ly- α absorption

Now that we have characterised the Ly- α absorption associated with protoclusters, we finally turn to examine the effectiveness with which one can select protocluster regions using line of sight Ly- α absorption. We have established that most of the sight lines that pass through protoclusters with $M_{z=0} \sim 10^{14} M_{\odot}$ exhibit low values of $\delta_{\tau_{\text{eff}}}$, and it is impossible to distinguish these from the field. On the other hand, for large values of $\delta_{\tau_{\text{eff}}}$ most of the segments associated with protoclusters are the result of damped absorption systems with $N_{\text{HI}} > 10^{19} \text{ cm}^{-2}$, but these systems do not uniquely trace protoclusters.

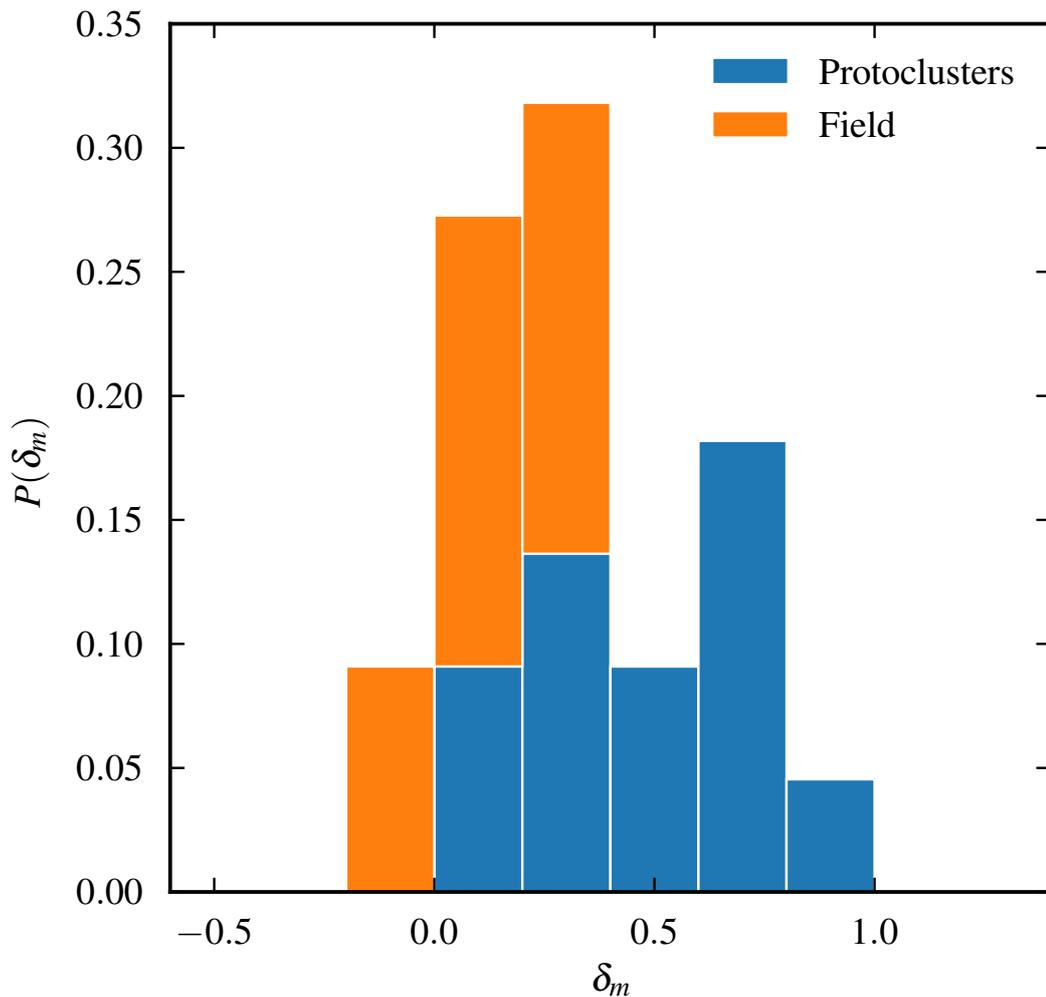


Figure 3.15: The probability distribution of CoSLAs identified in Sherwood as a function of δ_m , coloured by whether the CoSLA is associated with a protocluster (12; blue) or associated with the field (10; orange).

In addition, in the simulations there are a small subset of high opacity segments within protoclusters that do not exhibit damped absorption (e.g. the example CoSLA shown in the second from right panel in Fig. 3.14). This is demonstrated further in Fig. 3.15, which shows the distribution of CoSLAs in Sherwood split into two groups as a function of δ_m : those that are associated with a protocluster and those that are not. We find 55% of CoSLAs are associated with protoclusters with a median $\delta_m = 0.4 \pm 0.1$, in comparison to CoSLAs that probe the field, with median $\delta_m = 0.06 \pm 0.05$. This indicates that, while the two populations cannot be separated by their Ly- α absorption spectra alone, approximately half of CoSLAs—which in general arise from the alignment of overdense structure along the line of sight—are indeed associated with protoclusters. Thus we find that CoSLAs are not a unique tracer of protoclusters, even when assuming perfect removal of damped systems from the mock data.

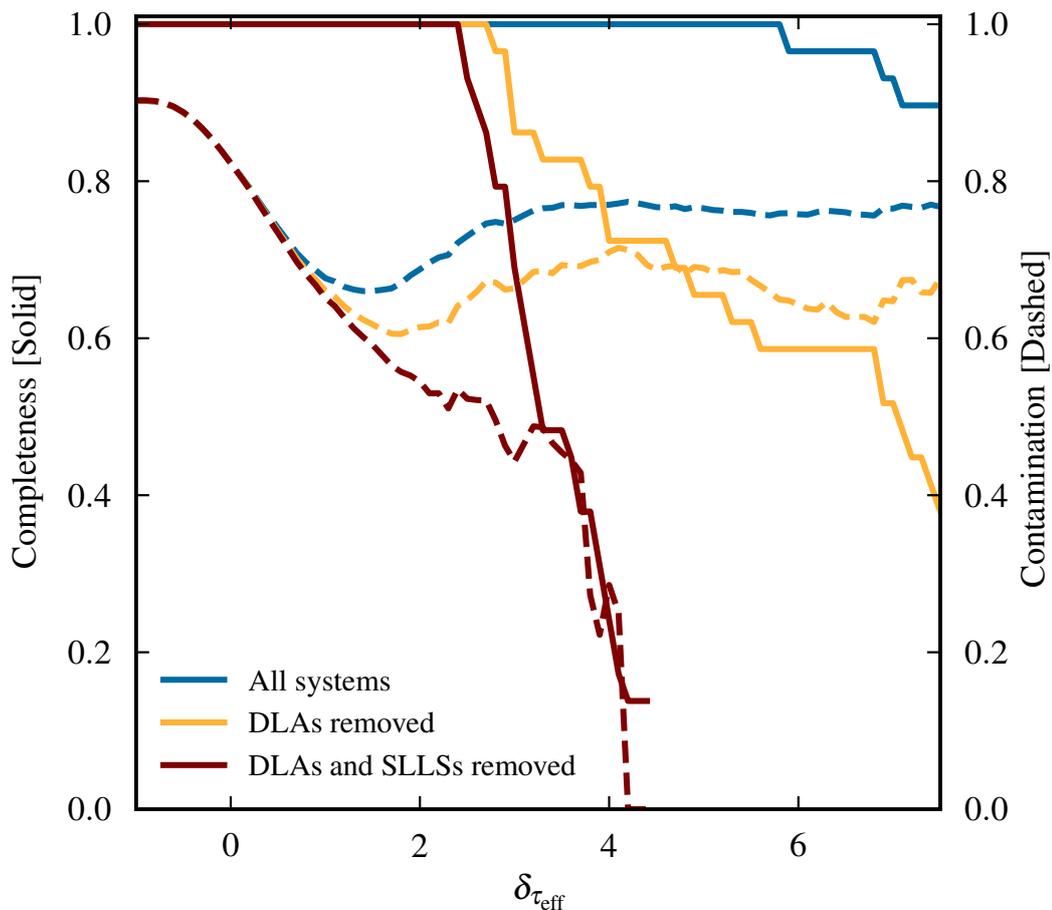


Figure 3.16: Contamination (dashed) and “completeness” (solid) for protoclusters (with $M_{z=0} \geq 10^{14} M_{\odot}$) when selecting all segments above a fixed threshold of $\delta\tau_{\text{eff}}$ on $15 h^{-1}$ cMpc scales (see text for details). The blue curves correspond to the case where all absorption is considered, orange curves to the case where sight lines containing DLAs are removed, and brown curves to the case where SLLSs and DLAs are removed. Results are displayed only where there is more than one $15 h^{-1}$ cMpc segment in each bin.

The effective optical depth threshold that defines a CoSLA as a segment exhibiting $\delta\tau_{\text{eff}} > 3.5$ on $15 h^{-1}$ cMpc scales is to some extent arbitrary. We may therefore consider whether a fixed threshold in $\delta\tau_{\text{eff}}$ can be chosen that minimises the contamination within a sample of protoclusters selected from the mock spectra. We define the contamination as $N(\geq \delta\tau_{\text{eff}})/N_{\text{tot}}(\geq \delta\tau_{\text{eff}})$, where $N(\geq \delta\tau_{\text{eff}})$ is the number of $15 h^{-1}$ cMpc segments above the $\delta\tau_{\text{eff}}$ threshold that do not pass through at least $5 h^{-1}$ cMpc of a protocluster volume, and $N_{\text{tot}}(\geq \delta\tau_{\text{eff}})$ is the total number of $15 h^{-1}$ cMpc segments above the $\delta\tau_{\text{eff}}$ threshold in the entire simulation volume. We also define the sample “completeness” for a given $\delta\tau_{\text{eff}}$ threshold as the fraction of protoclusters with at least one sight line that exhibits absorption on $15 h^{-1}$ cMpc scales above the $\delta\tau_{\text{eff}}$ threshold. This is defined as $N_{\text{pc}}(\geq \delta\tau_{\text{eff}})/N_{\text{pc,tot}}$, where $N_{\text{pc}}(\geq \delta\tau_{\text{eff}})$ is the number of unique protoclusters probed by $15 h^{-1}$ cMpc

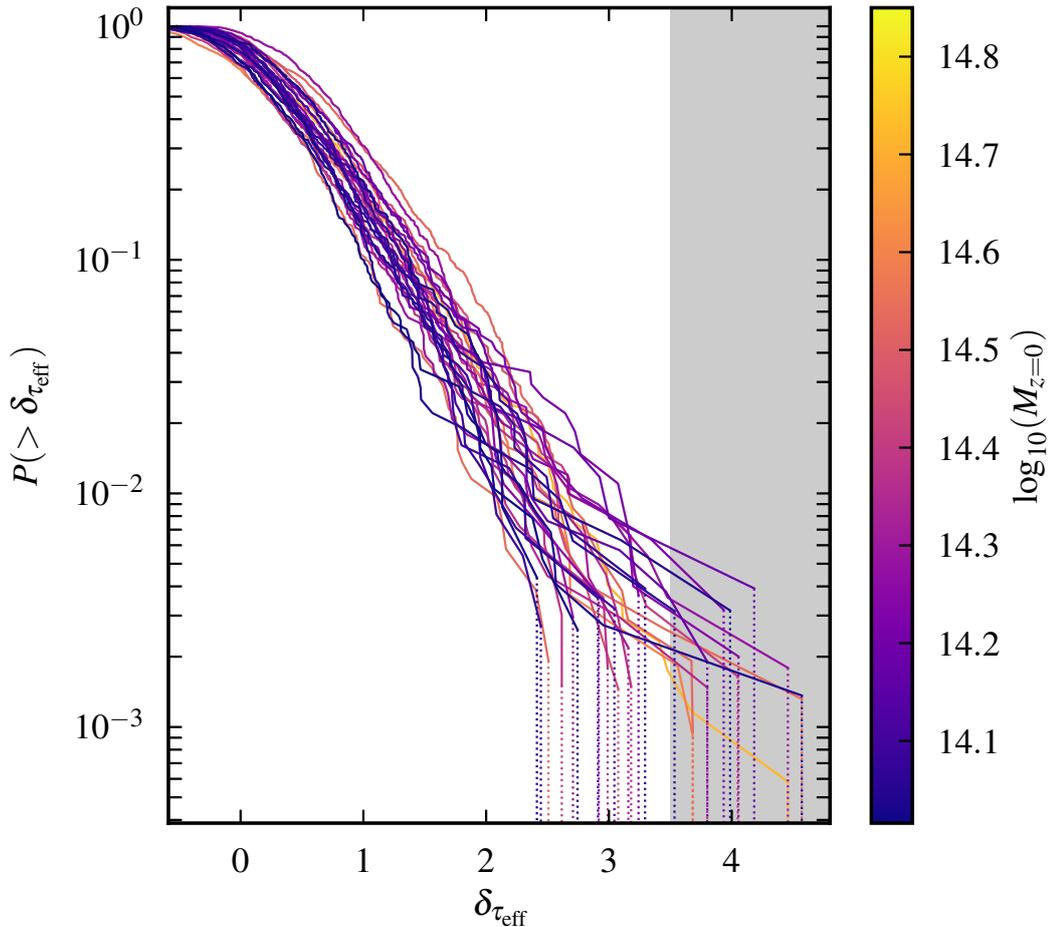


Figure 3.17: Reverse cumulative distribution functions, $P(> \delta_{\tau_{\text{eff}}})$, for all $15 h^{-1} \text{cMpc}$ segments that are associated with each of the 29 protoclusters in Sherwood. Each line is coloured according to the mass of the cluster at $z = 0$. The grey shaded region indicates the area above the CoSLA selection threshold of $\delta_{\tau_{\text{eff}}} > 3.5$.

segments above the $\delta_{\tau_{\text{eff}}}$ threshold, and $N_{\text{pc,tot}}$ is the total number of protoclusters in the simulation (i.e. 29 for Sherwood, see Table 3.1).

These values are plotted as a function of the $\delta_{\tau_{\text{eff}}}$ threshold in Fig. 3.16 for three different cases: where absorption in all the sight lines is considered (blue curves), where sight lines containing DLAs are removed from the sample (orange curves), and the case where both SLLSs and DLAs are removed (brown curves). When considering all sight lines, the protocluster sample remains 100% complete using a threshold up to $\delta_{\tau_{\text{eff}}} \sim 6$ because all protoclusters contain DLA systems that produce such high $\delta_{\tau_{\text{eff}}}$. The contamination also remains very high at $\sim 75\%$, which reflects the fact that the damped Ly- α absorbers do not uniquely probe protocluster environments. If we instead assume perfect removal of all sight lines containing DLAs, the contamination rate falls slightly to around 65% for the

segments with the strongest absorption. Finally, removing all the $15 h^{-1}$ cMpc segments containing SLLSs as well as DLAs causes the contamination to drop to zero at $\delta_{\tau_{\text{eff}}} = 4.1$, i.e. all the remaining segments with this $\delta_{\tau_{\text{eff}}}$ uniquely trace protocluster gas. Consequently, if we assume perfect removal of damped absorption systems, it is possible to obtain a perfectly clean sample of protoclusters by applying a sufficiently high threshold in $\delta_{\tau_{\text{eff}}}$.

This clean sample of protoclusters is, however, highly incomplete as not all protoclusters will exhibit strong coherent Ly- α absorption on $15 h^{-1}$ cMpc scales. This can be seen in Fig. 3.16, where the drop in contamination is accompanied by the completeness falling to only 17%. This means that less than a fifth⁸ of protoclusters have any sight lines with $\delta_{\tau_{\text{eff}}} > 4.1$. This is further exemplified in Fig. 3.17, which shows the reverse cumulative distribution function of $\delta_{\tau_{\text{eff}}}$ for each of the 29 protoclusters in Sherwood. The colour of each line corresponds to the mass of the resulting cluster at $z = 0$. This demonstrates that high $\delta_{\tau_{\text{eff}}}$ segments that pass through protoclusters are rare: less than 0.1% of segments that pass through $M_{z=0} \sim 10^{14} M_{\odot}$ protoclusters have $\delta_{\tau_{\text{eff}}} > 3.5$ —corresponding to the CoSLA threshold defined by C16, shown by the grey shading.

All the CoSLAs associated with protoclusters in Sherwood are listed in detail in Table 3.3: a total of 14 unique protoclusters over a mass range of $10^{14.0}$ – $10^{14.7} M_{\odot}$ are probed by 12 unique CoSLAs, and 42% of the CoSLAs associated with protoclusters pass through more than one protocluster. A sample of four of the CoSLAs associated with protoclusters in Sherwood are displayed in Fig. 3.18. We observe again that typically there is an alignment of structure along the sight line that causes the coherent Ly- α absorption. Orientation of structure to the line of sight, rather than association with a protocluster of a given mass or an overdensity, appears to be a critical factor that determines the extended nature of the Ly- α absorption.

In this context, we briefly note that Finley et al. (2014) argued for the detection of an intergalactic filament based on observations of multiple LLSs and SLLSs with $N_{\text{HI}} > 10^{18} \text{ cm}^{-2}$ along two closely separated quasar sight lines at $z = 2.69$. The seven strong H I absorption systems observed by Finley et al. (2014) span $\sim 1700 \text{ km s}^{-1}$, corresponding to $16.6 h^{-1}$ cMpc at $z = 2.69$. While Finley et al. (2014) could not definitively rule out association of the H I absorbers with a protocluster, their favoured interpretation is broadly consistent with our analysis.

⁸The completeness is likely to be strongly dependent on the protocluster mass. Since we do not probe massive protoclusters—due to our small box size—this completeness is likely to be a lower limit.

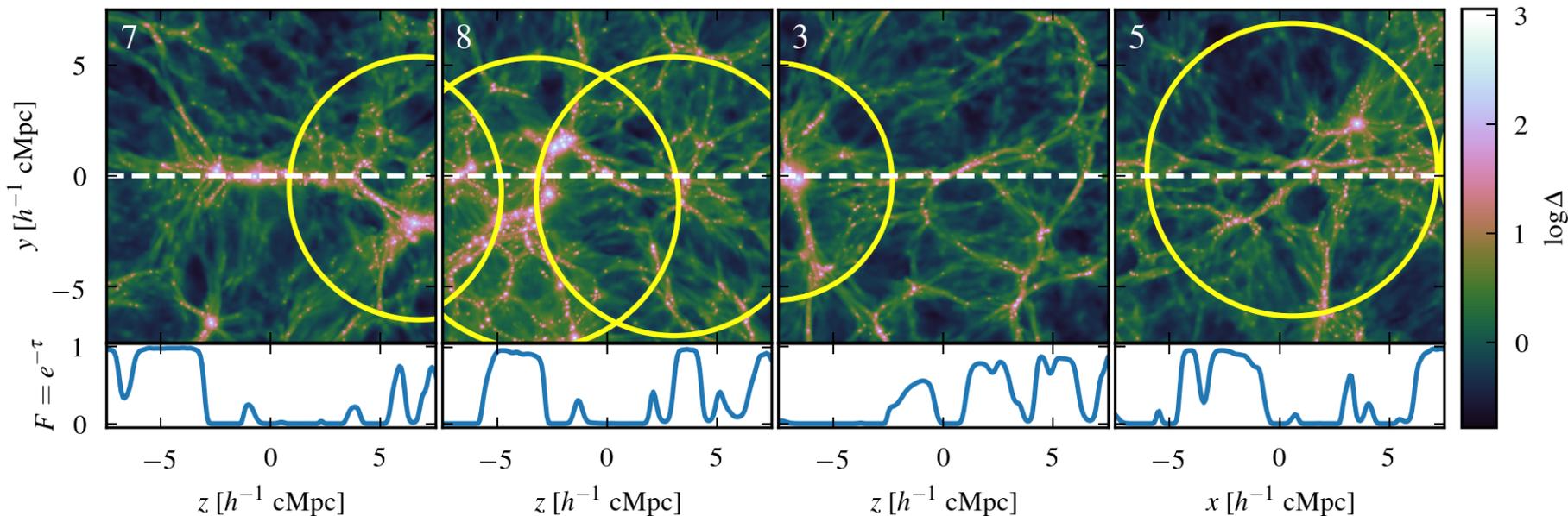


Figure 3.18: Example of four different CoSLAs that pass through protoclusters in Sherwood. The top panels show projected maps of the normalised gas density in a $15 h^{-2}$ cMpc 2 slice with a projection depth of $1 h^{-1}$ cMpc. The white dashed line shows the direction in which the Ly- α absorption spectrum was extracted. The bottom panels show the corresponding mock Ly- α absorption spectra. The numbers in the upper left of each panel correspond to the CoSLAs listed in Table 3.3. From left to right, the maximum H I column density associated with each CoSLA is $N_{\text{HI}} = 10^{16.5} \text{ cm}^{-2}$, $10^{14.9} \text{ cm}^{-2}$, $10^{19.0} \text{ cm}^{-2}$ and $10^{14.7} \text{ cm}^{-2}$.

Table 3.3: All CoSLAs (defined as $15 h^{-1} \text{cMpc}$ segments with $\delta_{\tau_{\text{eff}}} > 3.5$ that do not contain a damped Ly- α absorber) in Sherwood that are associated with protoclusters. A total of 14 protoclusters are probed by 12 unique CoSLAs. Note that each CoSLA may probe multiple protoclusters.

CoSLA	$\delta_{\tau_{\text{eff}}}$	δ_m	Protocluster $M_{z=0}$ [$\log_{10}(\text{M}_{\odot})$]
1	4.57	0.56	14.54, 14.04
2	4.46	0.80	14.72, 14.28
3	4.18	0.38	14.19
4	4.05	0.37	14.25
5	4.05	0.42	14.41
6	3.99	0.12	14.05
7	3.93	0.35	14.22
8	3.80	0.72	14.35, 14.18
9	3.68	0.98	14.59, 14.53
10	3.67	0.70	14.72, 14.53
11	3.56	0.62	14.59
12	3.53	0.12	14.06

Taken a step further, this suggests that CoSLAs may in fact be a tracer of extended filamentary structure in the early Universe. An intriguing possibility is that CoSLAs, and their association (if any) with galaxies and/or metal absorption lines, may therefore provide a route to identifying and studying filamentary environments at $z > 2$. To distinguish between a lone filament and a protocluster using Ly- α absorption, IGM tomography would need to be used. The ability to do this will only further increase as 30m class telescopes come online.

3.8 Summary

In this chapter we have used state of the art hydrodynamical simulations from the Sherwood (Bolton et al., 2017), EAGLE (Schaye et al., 2015; Crain et al., 2015; McAlpine et al., 2016) and Illustris (Vogelsberger et al., 2014; Nelson et al., 2015) projects to examine the signature of protoclusters observed in Lyman- α (Ly- α) absorption at $z \simeq 2.4$. Building upon earlier work using low resolution collisionless dark matter simulations (e.g. Stark et al., 2015; Cai et al., 2016), here we use models that resolve small scale structure in the intergalactic medium (IGM), correctly reproduce the incidence of H I absorption systems (including high column densities that are self-shielded to Lyman continuum photons), and track the formation of structure to $z = 0$.

We examine the impact of small scale gas structure on the signature of large scale overdensities on $15 h^{-1}$ cMpc scales, finding that the simulation mass resolution required for resolving Ly- α absorption on small scales is also a requirement for correctly modelling the average τ_{eff} on $15 h^{-1}$ cMpc scales. This is necessary for capturing the incidence of HI absorption systems over a wide range of column densities, including damped systems and Lyman-limit systems (LLSs). At the same time, however, adequate mass resolution is necessary for correctly capturing the opacity of underdense regions in the IGM. A mass resolution that is too low will overpredict the typical Ly- α effective optical depth on $15 h^{-1}$ cMpc scales.

We furthermore assess the prevalence of coherent Ly- α absorption within proto-clusters at high redshift. Our main conclusions may be summarised as follows:

- We confirm there is a weak correlation between the mass overdensity, δ_m , and the effective optical depth relative to the mean, $\delta_{\tau_{\text{eff}}}$, on $15 h^{-1}$ cMpc scales, in the simulations; however, there is a large amount of scatter that, particularly at large values of $\delta_{\tau_{\text{eff}}}$, means it is not possible to uniquely identify large scale overdensities with strong Ly- α absorption. This remains true even if first removing all damped Ly- α absorption systems that arise from dense, neutral gas on small scales.
- We examine the properties of coherently strong intergalactic Ly- α absorption systems (CoSLAs) in the simulations. CoSLAs—defined by C16 as regions on $15 h^{-1}$ cMpc scales with 4.5 times the average Ly- α effective optical depth after excluding damped absorption systems—are rare objects, only accounting for 0.1% of all $15 h^{-1}$ cMpc spectral segments drawn from the models. They probe a wide range in mass overdensity, δ_m , including underdense regions on $15 h^{-1}$ cMpc scales, and so do not uniquely trace significant mass overdensities.
- Protoclusters with $M_{z=0} \simeq 10^{14} M_{\odot}$ exhibit a broad range of signatures in Ly- α absorption, with $\delta_{\tau_{\text{eff}}}$ ranging from -0.5 to > 8 . However, the vast majority (84%) of sight lines passing through what we define as protoclusters in the simulations contain only low column density Ly- α forest absorption and have $\delta_{\tau_{\text{eff}}} < 1$. This signature is identical to the field, so the majority of sight lines through protoclusters do not bear a telltale signature in Ly- α absorption. Most sight lines with high $\delta_{\tau_{\text{eff}}}$ are a result of passing through a damped absorption system. A small subset of sight lines through protoclusters do, however, exhibit high $\delta_{\tau_{\text{eff}}}$ due to coherently strong intergalactic Ly- α absorption systems i.e. CoSLAs. LLSs and high column density Ly- α forest absorbers are typically responsible for this absorption.

- Assuming the perfect removal of damped Ly- α absorption systems, CoSLAs are a good but non-unique probe of protoclusters at $z \simeq 2.4$. In the Sherwood simulation, approximately half of CoSLAs with $\delta_{\tau_{\text{eff}}} > 3.5$ trace protoclusters with $10^{14} \leq M_{z=0}/M_{\odot} \leq 10^{14.7}$. The other 46% of CoSLAs arise from LLSs that are aligned along the line of sight.
- We find that threshold of $\delta_{\tau_{\text{eff}}} > 4.1$ —corresponding to regions on $15 h^{-1}$ cMpc scales with 5.1 times the average Ly- α effective optical depth after excluding damped absorption systems—enables us to select a completely pure sample of protoclusters from simulated spectra. However, we sample protoclusters with a mean transverse separation of $0.8 h^{-1}$ cMpc and find that CoSLAs are rare within the volumes that these protoclusters occupy: less than 0.1% of sight lines that pass through at least $5 h^{-1}$ cMpc of a protocluster volume exhibit $\delta_{\tau_{\text{eff}}} > 4.1$, excluding absorption caused by damped systems. This means that any sample of protoclusters selected with the CoSLA technique will be incomplete. We stress, however, that throughout this chapter we are limited by the box size of the hydrodynamical simulations. In particular, there are no $10^{15} M_{\odot}$ cluster progenitors in any of the models considered here; it is possible that more massive protoclusters have a stronger association with CoSLAs.

Finally, we note that visual inspection of CoSLAs suggests that coherent Ly- α absorption typically selects structure orientated along the line of sight to the observer, regardless of whether or not this is associated with a protocluster. With the advent of large spectroscopic QSO surveys such DESI (Vargas-Magana et al., 2019) and WEAVE-QSO (Pieri et al., 2016) in the next few years, further investigation of the potential of CoSLAs for identifying intergalactic filaments in the high redshift Universe may be a worthwhile endeavour.

Chapter 4

Characterising Protoclusters With Line of Sight Lyman- α Absorption Redux - The Impact of Box Size and the Most Massive Protoclusters

4.1 Introduction

One of the main caveats of Chapter 3 is the limited box size of the Sherwood simulations. With a box side length of only $80 h^{-1} \text{ cMpc}$ the analysis was restricted to a sample of only 29 $M_{z=0} \geq 10^{14} M_{\odot}$ protoclusters. Furthermore, due to these box size restrictions it was not possible to assess the impact of the most massive ($M_{z=0} \geq 10^{15} M_{\odot}$) protoclusters—with the most massive protocluster in the Sherwood simulation having a $z = 0$ mass of only $M_{z=0} = 10^{14.72} M_{\odot}$. In particular, this limits the ability to make comparisons to the earlier work of [Cai et al. \(2016\)](#) (referred to throughout this chapter as [C16](#)) as the simulation they used had a box side length of $1 h^{-1} \text{ cGpc}$ and therefore contained many $M_{z=0} \geq 10^{15} M_{\odot}$ protoclusters. This may be responsible for some of the differences seen between their work and the results in Chapter 3.

In this chapter I will work to address this caveat, using the larger TNG300-1 simulation volume that became available after the publication of Chapter 3. This

provides significantly increased numbers of protoclusters and proto-groups, as well as a sample of seven $M_{z=0} \geq 10^{15} M_{\odot}$ protoclusters, with a mass resolution on the same order of magnitude as the previously used simulations from the Sherwood project.

This chapter is structured as follows: in Section 4.2 I re-introduce the IllustrisTNG simulations that are used in this chapter. In Section 4.3 I outline the protocluster definition method and compare the basic properties of the protocluster sample from TNG300-1 to those from Chapter 3. In Section 4.4 I investigate how the relationship between the Ly- α forest optical depth and 3-D mass overdensity differs between the TNG100 and TNG300 simulations, whilst also making comparisons to the work of C16. Finally, in Section 4.5, I investigate how the relationship between CoSLAs and the volumes enclosing the progenitors of $M > 10^{14} M_{\odot}$ clusters varies between the TNG100 and TNG300 simulations. I then summarise the findings in Section 4.6. Throughout this chapter I refer to units of co-moving distance using the prefix “c”.

4.2 Hydrodynamical Simulations

In this chapter I use three publicly available simulations from the IllustrisTNG project (Nelson et al., 2018; Pillepich et al., 2018; Marinacci et al., 2018; Springel et al., 2018; Naiman et al., 2018): TNG100-1, TNG100-2 and TNG300-1. These models were first introduced in Chapter 2 (see Section 2.2). The key differences between these simulations are summarised in Table 4.1. All three simulations are run using the AREPO moving-mesh hydrodynamics code (Springel, 2010) along with a prescription for magneto-hydrodynamics (Pakmor et al., 2011), assuming a cosmology consistent with Planck Collaboration et al. (2015) where $\Omega_{\text{m}} = 0.3089$, $\Omega_{\Lambda} = 0.6911$, $\Omega_{\text{b}} = 0.0486$, $\sigma_8 = 0.8159$, $n_{\text{s}} = 0.9667$ and $h = 0.6774$.

The primary motivation for choosing these simulations is to assess how the results of Chapter 3 are affected by a larger simulation volume. However, the mass resolution of TNG300-1 is slightly poorer than the simulations used in Chapter 3. To address this the TNG100-2 and TNG100-1 simulations are also used to determine the effect of mass resolution on CoSLAs. TNG100-1 provides a simulation with similar box size and improved mass resolution to that of the Sherwood simulations used in Chapter 3, allowing an investigation of how the results are affected by the different active galactic nuclei and stellar feedback models—in addition to the inclusion of magneto-hydrodynamics—used in TNG100-1. However, given

Table 4.1: Hydrodynamical simulations used in this chapter. The columns list, from left to right: the simulation name, the box size in h^{-1} cMpc, the total number of resolution elements, the dark matter and the typical hydrodynamical cell mass, and the number of objects that have a total mass at $z = 0$ (as defined by a friends-of-friends halo finder) in the range $M_{z=0} \geq 10^{14} M_{\odot}$ (clusters), $10^{13.75} \leq M_{z=0}/M_{\odot} < 10^{14}$ (large groups) and $10^{13.5} \leq M_{z=0}/M_{\odot} < 10^{13.75}$ (small groups). For comparison, the Sherwood simulation from Chapter 3 is also listed (M_{gas} in this row refers instead to particle mass).

Name	Box size [h^{-1} cMpc]	$N_{\text{gas+DM}}$	M_{dm} [M_{\odot}]	M_{gas} [M_{\odot}]	$M_{z=0} \geq 10^{14} M_{\odot}$ Clusters	$10^{13.75} \leq \frac{M_{z=0}}{M_{\odot}} < 10^{14}$ Large groups	$10^{13.5} \leq \frac{M_{z=0}}{M_{\odot}} < 10^{13.75}$ Small groups
TNG100-1	75	2×1820^3	7.50×10^6	1.40×10^6	22	30	26
TNG100-2	75	2×910^3	5.97×10^7	1.12×10^7	22	30	26
TNG300-1	205	2×2500^3	5.960×10^7	1.10×10^7	426	467	776
Sherwood	80	2×1024^3	5.07×10^7	9.41×10^6	29	23	46

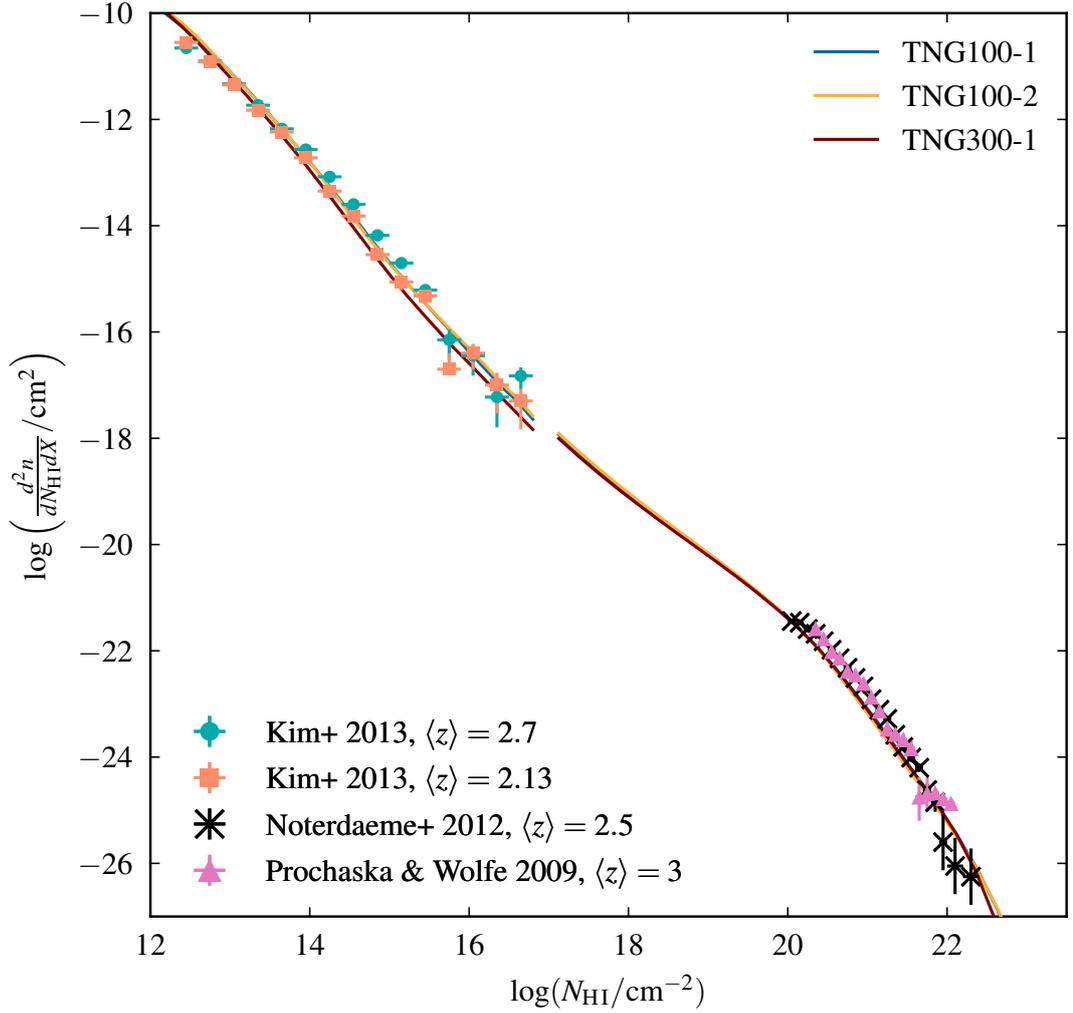


Figure 4.1: The H I column density distribution function (CDDF) at $z = 2.444$ obtained from the TNG100-1 (blue curve), TNG100-2 (orange curve) and TNG300-1 (dark red curve) simulations. At $N_{\text{HI}} < 10^{17} \text{ cm}^{-2}$, observational data from Kim et al. (2013) at $\langle z \rangle = 2.7$ and $\langle z \rangle = 2.13$ have been added. Similarly, data points from Noterdaeme et al. (2012) at $\langle z \rangle = 2.5$ and Prochaska and Wolfe (2009) at $\langle z \rangle = 3$ are plotted for $N_{\text{HI}} > 10^{20} \text{ cm}^{-2}$. This figure can be directly compared to Fig. 3.1 in Chapter 3.

the comparison between Sherwood, EAGLE and the original Illustris-1 simulations in Chapter 3, I expect the results from TNG100-1 to be consistent with these previous findings. TNG100-2 then provides a simulation that is identical to TNG100-1 in every way except mass resolution, where it matches TNG300-1 instead, this allows the effect of simulation box size to be isolated from differences in mass resolution.

4.2.1 Generation of mock Ly- α absorption sightlines

I follow the same methodology used for the Illustris-1 simulation in Chapter 3 to extract mock Ly- α absorption spectra from the IllustrisTNG simulations, using the combination of the SPH interpolation scheme described by Theuns et al. (1998), the Voigt profile approximation from Tepper-García (2006) and the prescription for self-shielding from Rahmati et al. (2013). I then scale the optical depths of the spectra by a constant to match the observational measurements of the mean effective optical depth from Becker et al. (2013), which at $z = 2.444$ corresponds to $\tau_{\text{eff}} = 0.21$. In all three simulations a total of 30,000 sightlines (10,000 along each projection axis) comprised of 1024 pixels are drawn at random positions in a direction parallel to the simulation boundaries. This corresponds to a mean transverse separation of $0.75 h^{-1}$ cMpc for TNG100-1 and TNG100-2, and $2.05 h^{-1}$ cMpc for TNG300-1.

In order to assess whether these simulations reproduce both low and high column density absorption systems in the correct abundances, I calculate the column density distribution function (CDDF) for each of the simulations used in this section alongside observational measurements from Kim et al. (2013), Noterdaeme et al. (2012) and Prochaska and Wolfe (2009) in Fig. 4.1. Here I use the same method as in Chapter 3. The CDDFs for all three simulations are in good agreement with both the observational data and each other, confirming the findings from Chapter 3 that the CDDF is not overly sensitive to changes in mass resolution or simulation box size.

4.3 Protocluster Properties

As was the case in Chapter 3, I make use of the ability to connect to $z = 0$ data to obtain a protocluster (and proto-group) sample in each simulation based upon the knowledge that the structure will collapse to form a cluster/group by redshift $z = 0$. I again select friends-of-friends haloes in three mass bins: $M_{z=0} \geq 10^{14} M_{\odot}$, $10^{13.75} M_{\odot} \leq M_{z=0} \leq 10^{14} M_{\odot}$ and $10^{13.5} M_{\odot} \leq M_{z=0} \leq 10^{13.75} M_{\odot}$ and I refer to these as ‘clusters’, ‘large groups’ and ‘small groups’ respectively. Every particle in each of these haloes is then traced back to the snapshot at $z = 2.444$ and these collections of particles are defined as protoclusters, large proto-groups and small proto-groups accordingly. The total number of protoclusters and proto-groups in each simulation is tabulated in table 4.1.

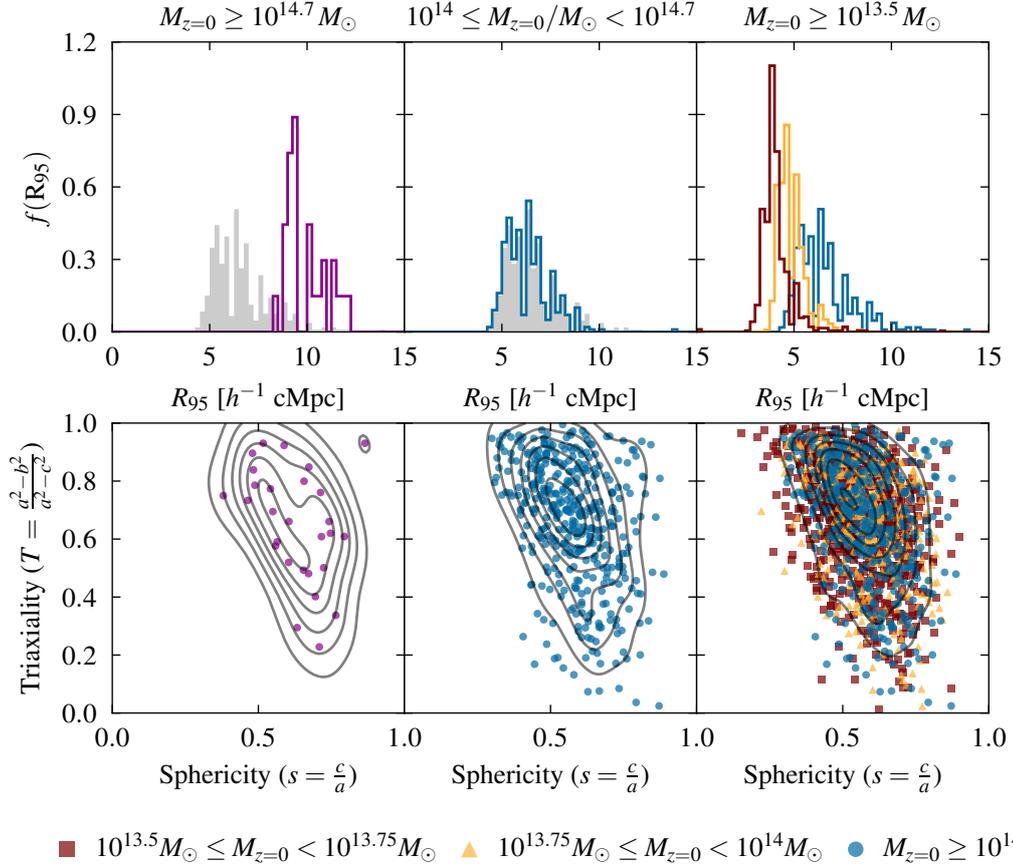


Figure 4.2: Top: R_{95} distributions for protoclusters with $M_{z=0} \geq 10^{14.7} M_{\odot}$ (left), protoclusters with $M_{z=0} < 10^{14.7} M_{\odot}$ (centre) and all protoclusters (blue) and large and small proto-groups (yellow and dark red) in the TNG300-1 simulations (right) at $z = 2.444$. In the left and centre panels, the grey silhouettes show the R_{95} distribution of all $M_{z=0} \geq 10^{14} M_{\odot}$ protoclusters as a fixed reference point. Bottom: The triaxiality-sphericity relation for protoclusters with $M_{z=0} \geq 10^{14.7} M_{\odot}$ (left), protoclusters with $M_{z=0} < 10^{14.7} M_{\odot}$ (centre), and (right) all protoclusters (blue circles), large proto-groups (orange triangles) and small proto-groups (dark red squares). $T = 1$ corresponds to prolate spheroids and $T = 0$ to oblate spheroids, while $s = 1$ denotes spherical objects and $s = 0$ highly aspherical ones. Both values are calculated from the principle semi-axes $a \geq b \geq c$. Contours show the kernel density estimate for all points in each panel. The right-hand panel of this figure can be directly compared to Fig. 3.2 in Chapter 3.

As in Chapter 3 (see Section 3.4), I define the region of space that these objects occupy at $z = 2.444$ using their CoM and R_{95} . In the top row of Fig. 4.2, the distributions of R_{95} for protoclusters and proto-groups in the TNG300-1 simulation is shown in the right-hand column. Unlike in Chapter 3, I have also subdivided the protocluster sample into two mass bins above and below $M_{z=0} = 10^{14.7} M_{\odot}$ in order to observe the properties of the most massive protoclusters. In the central column the distribution for protoclusters with $10^{14} \leq M_{z=0}/M_{\odot} < 10^{14.7}$ (corresponding to the range of cluster masses found in both the TNG100 and Sherwood simulations) is shown, containing 398 objects. The left-hand column then shows

the distribution for the $M_{z=0} \geq 10^{14.7} M_{\odot}$ protoclusters in the TNG300-1 simulation, of which there are 28. The results are consistent with those in Fig. 3.2 from Chapter 3; the more massive objects tend to have larger values of R_{95} with a significant overlap between the mass bins (the largest ‘small proto-groups’ have $R_{95} \sim 6 h^{-1} \text{cMpc}$ and the smallest protoclusters have $R_{95} \sim 4 h^{-1} \text{cMpc}$ in the TNG300-1 simulation). Comparing Fig. 3.2 from Chapter 3 to Fig. 4.2 shows the improved statistics provided by the larger simulation volume, and the high R_{95} tail is now populated with several protoclusters having $R_{95} \gtrsim 10 h^{-1} \text{cMpc}$ —the majority of these having $M_{z=0} \geq 10^{14.7} M_{\odot}$.

In the bottom row of Fig. 4.2 the shape of protoclusters and proto-groups in each simulation are shown, taking the same approach as in Chapter 3. In Fig. 4.2, for all mass ranges, the protoclusters and proto-groups tend to be prolate objects with a moderate sphericity on average; a weak anti-correlation exists between T and s , such that more oblate protoclusters are also slightly more spherical than prolate ones. These results are consistent with those shown in Fig. 3.2 from Chapter 3, confirming that—even at the highest masses—protocluster shape is not strongly influenced by its mass or physical extent.

4.4 Characterising mass overdensities in Ly- α absorption on $15 h^{-1} \text{cMpc}$ scales

As discussed in Chapter 3, whilst the protocluster definition described in Section 4.3 is useful from the perspective of simulations, it is not applicable to observations given that they are limited to only a single snapshot in time. In this section I take an observationally motivated approach and investigate the relationship between mass overdensities and effective optical depth on $15 h^{-1} \text{cMpc}$ scales, regardless of whether these regions will actually collapse to form a cluster by $z = 0$.

Here I again use the procedure used to obtain Fig. 3.5 in Chapter 3 to show the relationship between δ_m from the 3-D $15^3 (h^{-1} \text{cMpc})^3$ volumes and associated $\delta_{\tau_{\text{eff}}}$ in Fig. 4.3. I find the results for all three simulations to be consistent with those from Chapter 3, with a significant majority of the volumes in each of the three simulations residing close to the mean density with their associated effective optical depths close to the mean transmission. A weak positive correlation between δ_m and $\delta_{\tau_{\text{eff}}}$ is again found up to $\delta_{\tau_{\text{eff}}} \sim 2$ in all three simulations. For comparison, the same relation from C16 is shown in Fig. 4.4. The correlation

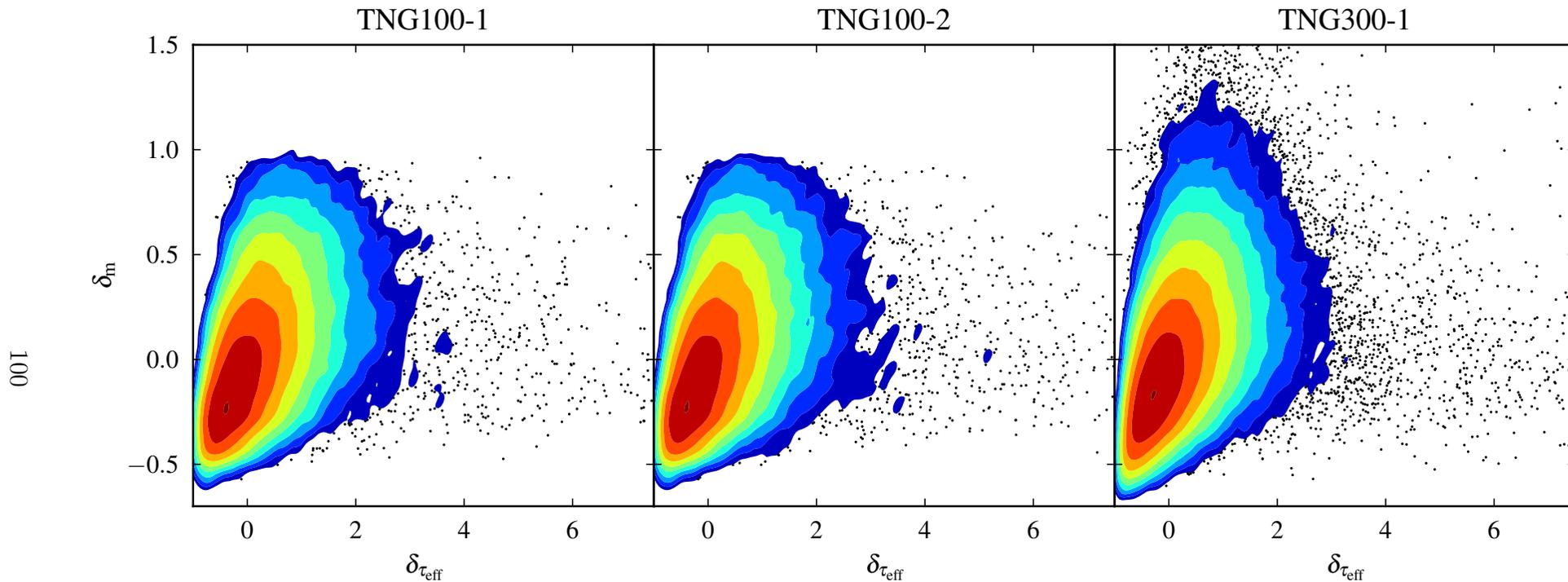


Figure 4.3: The δ_m - $\delta_{\tau_{\text{eff}}}$ relation for all $15 h^{-1}$ cMpc volumes and their corresponding sightline segments in the TNG100-1 (left), TNG100-2 (centre) and TNG300-1 (right) simulations. Contours show the number density of point relative to the central contour in steps of 1/3 dex. Black points show the locations of each system outside of these contours (where the relative number density is below 10^{-3}). This figure can be directly compared to Fig. 3.5 in Chapter 3.

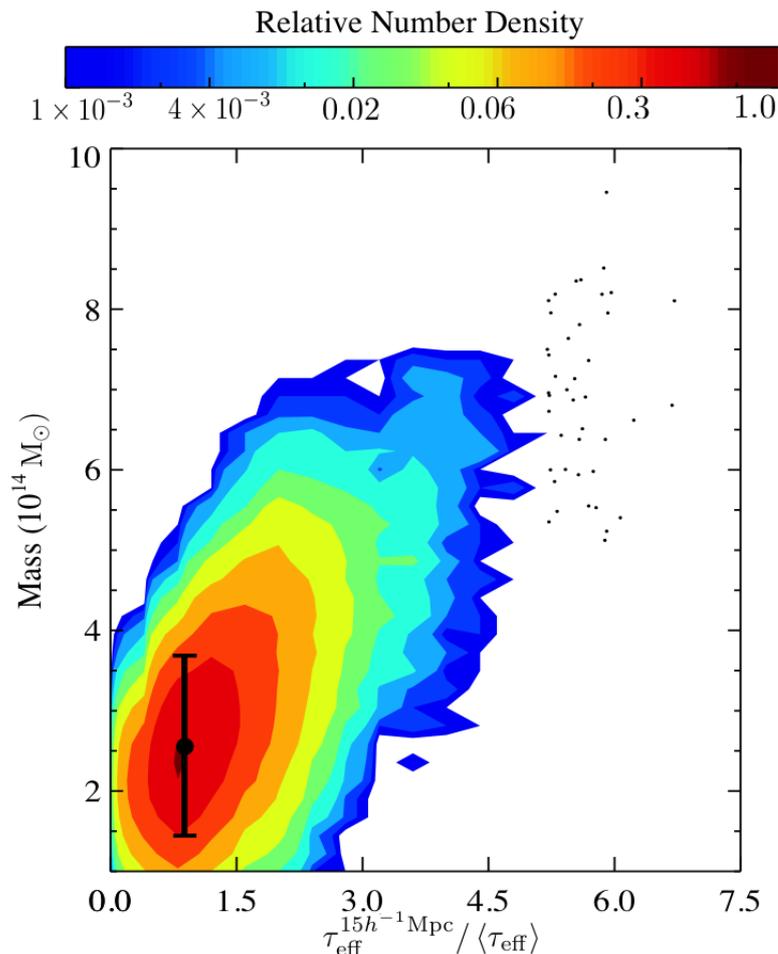


Figure 4.4: Figure 2 from C16, showing the relationship between 3-D mass and 1-D effective optical depth on $15h^{-1}cMpc$ scales using their LyMAS simulation. Contours represent the number density of point relative to the central contour (brown) which is normalised to unity. The filled black circle and error bar denotes the median mass within $15^3 h^{-3} cMpc^3$ ($2.6 \times 10^{14} M_{\odot}$) along with the 1σ spread. Black dots represent the systems with the highest effective optical depths. N.B. In this plot a mass of $5.2 \times 10^{14} M_{\odot}$ corresponds to $\delta_m = 1$ in Fig. 4.3, whilst the x-axis can be converted to that of Fig. 4.3 by subtracting 1.

between mass and effective optical depth is significantly enhanced in Fig. 4.4 when compared to Fig. 4.3, with all of the systems at $\delta_{\tau_{\text{eff}}} \gtrsim 4$ having a mass corresponding to $\delta_m > 1$.

Whilst there appears to be very little difference between TNG100-1 and TNG100-2 in the $\delta_{\tau_{\text{eff}}} \lesssim 2$ regime (consistent with the findings of Chapter 3; see Fig. 3.5), in TNG300-1 there are more points at $\delta_m > 1$. Contrary to what would be expected from the results of C16—who found these regions to almost exclusively have $\delta_{\tau_{\text{eff}}} > 0$ (see Fig. 4.4)—I find these higher mass volumes follow the same relation seen at lower masses, with a wide range of $\delta_{\tau_{\text{eff}}}$ values being represented. These results suggest that a full hydrodynamical modelling of the intergalactic medium is more important than simulation box size for the purposes of accurately

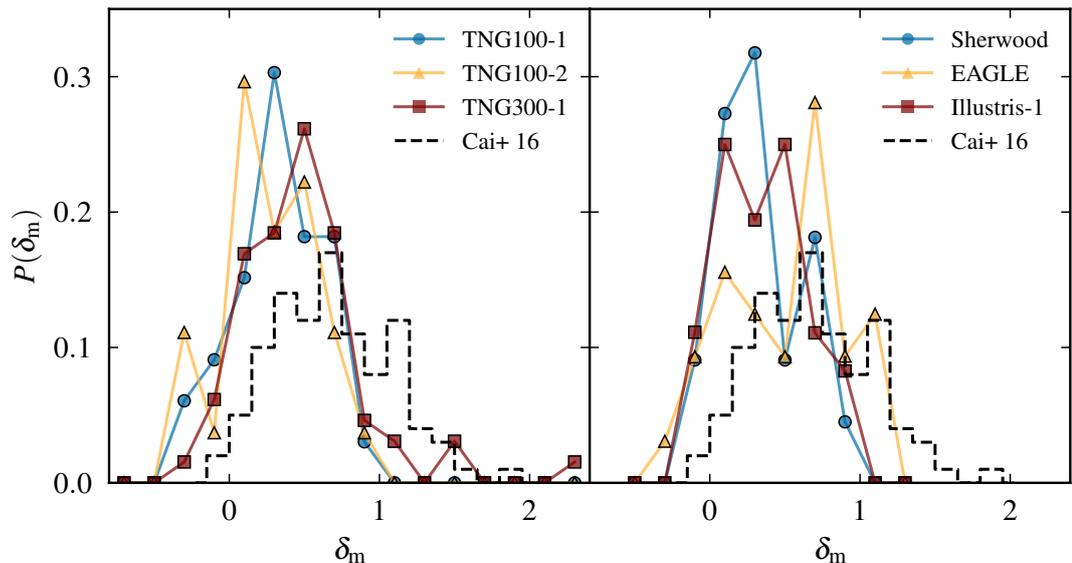


Figure 4.5: Left: Probability distributions of the mass overdensity (δ_m) associated with CoSLAs in the TNG100-1 (blue circles), TNG100-2 (yellow triangle) and TNG300-1 (dark red squares) simulations. Right: Fig. 3.7 remade in the same style as the left-hand panel. In both panels the dashed black histogram shows the corresponding PDF from C16.

predicting the δ_m - $\delta_{\tau_{\text{eff}}}$ relation.

At values of $\delta_{\tau_{\text{eff}}} \gtrsim 2$, where the spectra increasingly contain self-shielded absorption systems with large damping wings, all three simulations again show similar results to each other and to the results from Fig. 3.5 in Chapter 3. The sightline segments in this regime are associated with both over and under dense volumes, with there being no correlation between δ_m and $\delta_{\tau_{\text{eff}}}$. This is again a significantly different distribution to the corresponding points in Fig. 4.4, where the systems with the highest effective optical depths instead exclusively occupy to $\delta_m \gtrsim 1$. The difference between Fig. 4.3 and Fig. 4.4 at $\delta_{\tau_{\text{eff}}} > 4$ will be primarily due to the bulk of $\delta_{\tau_{\text{eff}}} > 4$ systems with $\delta_m \sim 0$ in Fig. 4.3 corresponding to damped systems. The simulations used throughout this thesis self-consistently model the incidence of damped absorption systems, however, C16 used LyMAS to extract $Ly-\alpha$ forest optical depths from an N-body simulation. As the simulations used to calibrate LyMAS do not contain a prescription for self-shielding, they only reproduce the low density (optically thin) $Ly-\alpha$ forest. Because of this, there are no damped systems present in the data used by C16 to produce Fig. 4.4.

Next I turn to look at the range of δ_m values that are probed by CoSLAs, defined by C16 as $15 h^{-1} cMpc$ segments of $Ly-\alpha$ forest spectra with $\delta_{\tau_{\text{eff}}} > 3.5$. There are 33 (27) CoSLAs in the TNG100-1 (TNG100-2) simulations compared to 65 in TNG300-1. This difference between the two box sizes is consistent with the

fact that the TNG300-1 sightlines are over twice as long as those from the smaller simulation volumes, resulting in a number of CoSLAs per unit redshift of $dn/dz = 9.34 \times 10^{-3}$ for TNG300-1 versus $dn/dz = 1.00 \times 10^{-2}$ for TNG100-2 at redshift $z = 2.444$.

In Fig. 4.5 I show the δ_m probability distribution function (PDF) for CoSLAs in the TNG100-1 (blue circles), TNG100-2 (yellow triangles) and TNG300-1 (dark red squares) with the results from C16 (black dashed histogram), alongside a reproduction of Fig. 3.7 from Chapter 3. For each PDF I compute the median δ_m along with the 68th percentile spread via bootstrapping. The results for both TNG100 simulations are consistent with the results from Chapter 3.7. The median δ_m of CoSLAs is 0.36 ± 0.05 (0.33 ± 0.13) in the TNG100-1 (TNG100-2) simulation. When increasing the box size from $75 h^{-1} \text{cMpc}$ in TNG100-2 to $205 h^{-1} \text{cMpc}$ in TNG300-1, the median δ_m of CoSLAs also increases to 0.45 ± 0.05 , a change of 0.86σ . This is in part a result of the additional high δ_m tail at $\delta_m > 1$. I also note that, in absolute terms, this increase is 1.5 times that of the increase seen between Sherwood 80-1024 and 160-2048. However, due to the relatively large uncertainty in δ_m for TNG100-2, this is actually a less significant increase than the 1.5σ found in Chapter 3. Finally, the significantly larger change in simulation volume between TNG100 and TNG300 ($V_{\text{TNG300}}/V_{\text{TNG100}} = 20.4$) compared to Sherwood 80-1024 and 160-2048 ($V_{160-2048}/V_{80-1024} = 8$), suggests that further box size increases may yield diminishing returns with regard to increases in the median CoSLA mass overdensity.

The median δ_m of CoSLAs in TNG300-1 is still lower than that of C16, who find a median of $(7.0 \pm 1.6) \times 10^{14} M_{\odot} \equiv \delta_m = 0.641 \pm 0.375$ using N-body simulations. There are also significantly more CoSLAs with $\delta_m < 0$ in TNG300-1. Hence, this further bolsters the suggestion that the differences between the work in Chapter 3 and C16 are not primarily due to the differences in simulation volume and a lack of $\delta_m > 1$ regions.

Instead, it is possible that not self-consistently modelling the effect of shock heated, collisionally ionised, gas will result in an overestimate of the $\text{Ly-}\alpha$ absorption associated with highly overdense regions. Self-consistent modelling of shock heated gas is present in the hydrodynamical simulations used throughout this thesis but not in the N-body simulation used by C16. The impact that correctly modelling the effect of collisional ionisation in such regions has on the incidence of CoSLAs was demonstrated in Fig. 2.5 from Chapter 2. In the central column of Fig. 2.5, where the effects of collisional ionisation and self-shielding are ignored, there is roughly an order of magnitude increase in the number of

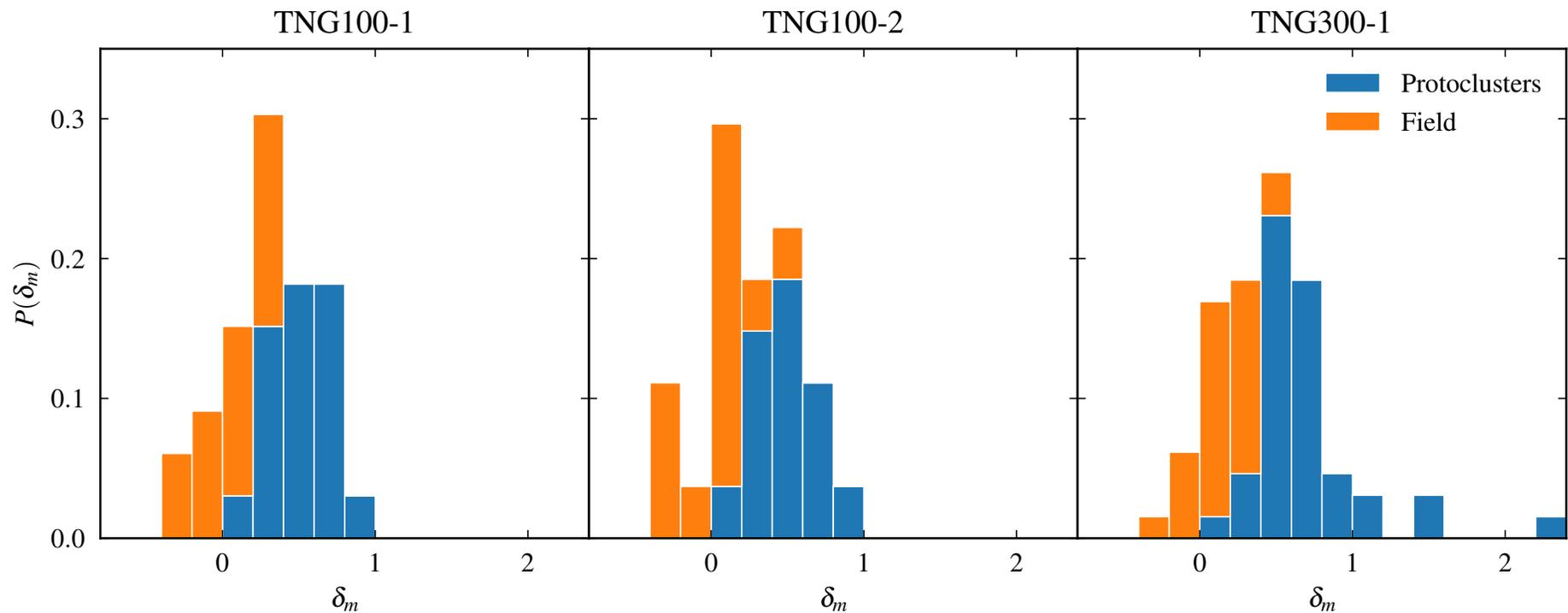


Figure 4.6: PDFs for CoSLAs identified in TNG100-1 (left), TNG100-2 (centre) and TNG300-1 (right) as a function of δ_m , coloured according to whether they are associated with a protocluster (blue) or the field (orange). This figure can be directly compared to Fig. 3.15 in Chapter 3.

CoSLAs present in comparison to the left-hand column, where these effects are self-consistently modelled. In particular, when moving from the *NoCol* model to the fiducial model used in Chapter 2 (see Section 2.2.3 for details of these models) CoSLAs are removed from the centre of haloes and filaments, where the dense gas has been shock heated to $T \sim 10^5$ K.

4.5 The Fraction of CoSLAs in Protoclusters

Having taken an observationally motivated approach, looking at the relationship between CoSLAs and mass overdensities at $z = 2.444$, I now take a simulation based approach and consider the objects that will collapse to form clusters by $z = 0$. Here I follow the same approach used in Chapter 3, defining a sightline as associated with a protocluster/proto-group if over one third of the pixels that comprise it fall within R_{95} of a protocluster/proto-group CoM.

In Fig. 4.6 I show the same mass PDF for CoSLAs as Fig. 4.5, however, this time I colour the fractions of each bin according to whether the volumes/segments are associated with a protocluster or not (reproducing Fig. 3.15 from Chapter 3). In all three cases I find results that are once again consistent with those found in Chapter 3. No CoSLAs with $\delta_m < 0$ are associated with protoclusters. The fraction of CoSLAs that are associated with protoclusters then increases with δ_m , such that by $\delta_m > 0.6$ all CoSLAs are associated with protoclusters.

Furthermore, I find that 57.6% (51.9%) of CoSLAs in the TNG100-1 (TNG100-2) simulation are associated with protoclusters, similar to the 54.5% of CoSLAs which were found to be associated with protoclusters from the Sherwood simulation in Chapter 3. This fraction increases slightly in the TNG300-1 simulation, where 60.0% of CoSLAs are associated with protoclusters. These results confirm the findings from Chapter 3 that the majority of CoSLAs are associated with protocluster regions at redshift $z = 2.444$, making them a good but non-unique probe of protocluster regions.

4.6 Summary

In this chapter I have used three of the IllustrisTNG simulations (TNG100-1, TNG100-2 and TNG300-1) to investigate the impact of box size on the results of Chapter 3. Using TNG300-1 provides a $\sim 20\times$ increase in the number of

protoclusters within the simulation volume and also allows the investigation of the progenitors of the most massive clusters. I have compared the populations of coherently strong Lyman- α absorption systems (CoSLAs) from this larger ($205^3 h^{-3} \text{cMpc}^3$) simulation volume, to those found in simulations with smaller ($75^3 h^{-3} \text{cMpc}^3$) box sizes, in order to determine to what extent simulation box size can explain the discrepancies observed between the results of Chapter 3 (where an $80^3 h^{-3} \text{cMpc}^3$ simulation volume was used) and those of [Cai et al. \(2016\)](#) (who used a $1^3 h^{-3} \text{cGpc}^3$). The main conclusions are as follows:

- The results from TNG100-1 and TNG100-2 are in good agreement with the findings in Chapter 3. There is a weak correlation between δ_m and $\delta_{\tau_{\text{eff}}}$ in the TNG100 simulations, consistent with that seen for the Sherwood, EAGLE and Illustris-1 simulations in Chapter 3. This is significantly less pronounced than the relation found by [C16](#).
- I again find that only CoSLAs with $\delta_m > 0$ are associated with protocluster regions, with over half of all CoSLAs being associated with a protocluster at redshift $z = 2.444$. This is again consistent with the fraction found for the Sherwood simulation in Chapter 3. These results further support the suggestion from Chapter 3 that the differences in feedback modelling between different hydrodynamical simulations does not have a large impact on the incidence of CoSLAs.
- I confirm that CoSLAs are rare objects, finding $dn/dz \sim 1 \times 10^{-2}$ at redshift $z = 2.444$ for all three of the IllustrisTNG simulations used in this Chapter.
- With the better statistics afforded by TNG300-1 I find that the relationship between mass overdensity, δ_m , and the effective optical depth relative to the mean, $\delta_{\tau_{\text{eff}}}$, is consistent with that found using a smaller box size of $75 h^{-1} \text{cMpc}$. In particular, I observe the $\delta_m > 1$ systems present in the larger volume to continue to exist at a wide range of $\delta_{\tau_{\text{eff}}}$ values. This is in contrast to the findings of [C16](#), who find such regions to preferentially be associated with $\delta_{\tau_{\text{eff}}} > 1$.
- The median mass overdensity, δ_m , associated with a CoSLA in TNG300-1 is 0.45 ± 0.05 . Whilst this represents a modest increase over the median $\delta_m = 0.33 \pm 0.13$ found in the smaller TNG100-2 simulation, it is still lower than the value of $\delta_m = 0.64 \pm 0.38$ found by [C16](#). This suggests that box size and the presence of larger mass overdensities does not fully explain the differences between the median CoSLA mass found in Chapter 3 and

C16. I suggest that the main cause of these differences is instead the self-consistent modelling of hot gas surrounding the massive haloes that reside within protocluster regions at redshift $z \sim 2.4$.

Chapter 5

Conclusions and Outlook

“You ever feel like every time we get close to getting the answers, somebody changes the question?”

—Peter Bishop, Fringe S3E10, *The Firefly*

5.1 Key Conclusions

In this thesis, state-of-the-art hydrodynamical simulations from the Sherwood (Bolton et al., 2017), EAGLE (Schaye et al., 2015), Illustris (Vogelsberger et al., 2014; Nelson et al., 2015) and IllustrisTNG (Springel et al., 2018; Nelson et al., 2018; Marinacci et al., 2018; Naiman et al., 2018; Pillepich et al., 2018) projects have been used to study the Lyman- α (Ly- α) absorption characteristics of proto-cluster regions at redshift $z \simeq 2.4$, corresponding broadly to the redshift at which Ly- α forest based protocluster detection methods have been focused to date. The study of protocluster regions using the Ly- α forest is an area of research that has been rapidly growing in recent years, proving to be a promising new angle from which to approach the discovery of these objects. Quasi-stellar object (QSO) surveys such as the Baryon Oscillation Spectroscopic Survey (BOSS) (Dawson et al., 2013; Pâris et al., 2014) have provided data for the identification of coherently strong Ly- α absorption (Cai et al., 2017b), and larger surveys that include Ly- α spectra from star forming galaxies have been used to reconstruct 3-D maps of the intergalactic medium (IGM) (Lee et al., 2018; Newman et al., 2020). The aim of this thesis was to address some of the main shortcomings in the theoretical models used to calibrate these methods.

The key questions I address along with the main conclusions are as follows:

Q1: What role does ionisation bias play in tomographically reconstructed maps of the IGM?

Using the TNG100-1 simulation from the IllustrisTNG project, the potential role of ionisation bias in tomographically reconstructed maps of the IGM was investigated. Collisional ionisation of hot gas surrounding haloes that host groups of Ly- α emitters (LAEs) was found to impact the correlation between LAE density and reconstructed Ly- α flux within individual protoclusters. In a protocluster with $M_{z=0} = 10^{14.18} M_{\odot}$ this effect was observed to result in a $\sim 9 h^{-1}$ cMpc offset between the LAE density peak and the location of the largest Ly- α flux decrement, qualitatively similar to protocluster found in the CLAMATO tomographic maps (Lee et al., 2018; Mukae et al., 2020a). By employing a simple—but observationally motivated—fluctuating ultraviolet background model, I accounted for the increased photoionisation rate in the vicinity of luminous QSOs. Using this model, I found the spatial offset to be further enhanced by the proximity effect of a $M_{1450} = -24.7$ active galactic nuclei (AGN) hosted within the halo corresponding to the peak in LAE density. However, both collisional ionisation and the proximity effect from QSOs were found to only have a significant impact on scales smaller than the typical smoothing scale employed by observational IGM tomography surveys ($4 h^{-1}$ cMpc). As a result of this, ionisation biases should have little impact on protocluster selections made from these maps using fixed $(\delta_{\text{F}}/\sigma)_{\text{min}}$ identification thresholds, with a completeness of 82–86% and 32–41% found across all three models for thresholds of $(\delta_{\text{F}}/\sigma)_{\text{min}} \leq -2.35$ and $(\delta_{\text{F}}/\sigma)_{\text{min}} \leq -3.00$ respectively. However, I note that calibrating tomographic IGM maps using the fluctuating Gunn-Peterson approximation (FGPA) may lead to systematically low estimates of $M_{z=0}$ for the most massive protoclusters due to the presence of hot gas around haloes. Using the 22 protoclusters within the fiducial TNG100-1 model, the relation $M_{z=0} = 10^{11.9-0.89(\delta_{\text{F}}/\sigma)_{\text{min}}} M_{\odot}$ was found. Additionally, for LAE overdensities $\delta_{\text{LAE}} \lesssim 1.5$, a weak anti-correlation between δ_{LAE} and δ_{F} was found with Spearman’s rank correlation coefficient of ~ -0.4 , supporting recent suggestions that these galaxies preferentially exist in regions with enhanced Ly- α absorption.

Q2: How well do Coherently Strong Ly- α Absorption systems trace protocluster regions at redshift $z \sim 2.4$?

Using the Sherwood simulation, I discovered that CoSLAs ($15 h^{-1}$ cMpc spectra sections with $\delta_{\tau_{\text{eff}}} > 3.5$ and $N_{\text{HI,max}} < 10^{19} \text{ cm}^{-2}$) were a non-unique but effective probe of protoclusters at redshift $z \simeq 2.4$, with 56% of CoSLAs found to be associated with protocluster regions. Furthermore, by choosing a strict, model

dependant $\delta_{\tau_{\text{eff}}}$ threshold, it was possible to eliminate the contamination from CoSLAs that arise from filamentary structure and Lyman-limit systems (LLSs) in the diffuse IGM that are preferentially aligned along the line of sight. However, in doing so one is left with a highly incomplete protocluster sample. Additionally, CoSLAs were found to be very rare objects; with an average transverse sightline spacing of $0.75 h^{-1} \text{ cMpc}$, only ~ 0.1 percent of $80 h^{-1} \text{ cMpc}$ sightlines at redshift $z \simeq 2.4$ contain a CoSLA within them, corresponding to a $dn/dz = 7.69 \times 10^{-3}$.

Q3: To what extent does simulation box size explain the discrepancies between Coherently Strong Ly- α Absorption system populations in large N-body simulations and comparatively small hydrodynamical simulations?

Finally, in Chapter 4, I used simulations from the IllustrisTNG project to further investigate the impact of box size on the relation between protocluster regions and CoSLAs. Using the TNG300-1 simulation, I confirmed that CoSLAs were rare objects (with $dn/dz \sim 1 \times 10^{-2}$ at redshift $z = 2.444$) as well as a good, but non-unique, probe of protocluster regions. 60% of the CoSLAs in the TNG300-1 volume were found to be associated with at least one protocluster. With the significantly increased range of overdensities probed within the TNG300-1 simulation volume, the $\delta_{\text{m}}-\delta_{\tau_{\text{eff}}}$ relation is found to be consistent with that seen in the smaller simulation volumes (TNG100, Sherwood, EAGLE and Illustris-1). In contrast to the findings of Cai et al. (2016), overdensities that exist at δ_{m} are found to cover a wide range of $\delta_{\tau_{\text{eff}}}$ rather than strongly favouring $\delta_{\tau_{\text{eff}}} > 1$. Similarly, CoSLAs in TNG300-1 are again found to correspond to both underdense and overdense regions with a median $\delta_{\text{m}} = 0.45 \pm 0.05$. This CoSLA mass distribution is consistent with that found in Chapter 3. I found a median δ_{m} that is still some way below the $\delta_{\text{m}} = 0.64$ found by Cai et al. (2016) with many more CoSLAs found at $\delta_{\text{m}} < 0$ in TNG300-1. Finally, I found the results from the TNG100 simulations to be consistent with those in Chapter 3, both in terms of the relation between δ_{m} and $\delta_{\tau_{\text{eff}}}$, as well as the characteristics of the CoSLA populations. This further supports the statement made in Chapter 3 that the differences in subgrid physics between different hydrodynamical simulations do not significantly affect the $\delta_{\text{m}}-\delta_{\tau_{\text{eff}}}$ relation on these scales. This work suggests that the differences found between hydrodynamical simulations in Chapter 3 and Cai et al. (2016) are not fully explained by the differences in simulation volume. Instead, results from Chapter 2 suggest that the differences between the hydrodynamical and N-body simulations are primarily caused by the modelling of the hot, collisionally ionised, gas that surrounds massive haloes found within proto-

clusters. The effect that collisional ionisation has on the incidence of CoSLAs was shown in Chapter 2, Fig. 2.5. The LyMAS scheme employed by [Cai et al. \(2016\)](#) is only able to account for this shock heated gas in statistical sense, whilst hydrodynamical simulations are able to accurately model the effect in a self-consistent fashion.

5.2 Limitations of this Thesis

There remains some scope for improving upon the numerical modelling in this thesis. In particular, the dynamic range of the hydrodynamical simulations should ideally be larger. A mass resolution of $M_{\text{gas}} \sim 10^6 M_{\odot}$ is required to resolve Ly- α absorption from the IGM at $z \simeq 2$ ([Bolton and Becker, 2009](#), Chapter 3). While the Ly- α forest in the TNG100-1 simulation in Chapter 2 and Sherwood simulation in Chapter 3 is therefore well resolved, the statistics are somewhat limited with only 22 and 29 protoclusters with $M_{z=0} \geq 10^{14} M_{\odot}$ respectively. Additionally, the lack of any $M_{z=0} \geq 10^{15} M_{\odot}$ clusters in either of these simulations means that it is not possible to determine the impact of the most massive structures. Though the results from Chapter 4 suggest that simulation box size does not significantly affect the outcomes of the research described in Chapter 3, the IllustrisTNG300-1 simulation still only contains a handful of the most massive protoclusters and therefore limits the conclusions that can be drawn regarding such objects.

Furthermore, in the local sources model from Chapter 2, isotropic AGN emission is assumed with non-equilibrium ionisation and light travel time effects being ignored. By instead modelling emission from AGN in the form of diametric cones with some beaming angle, as well as a temporally varying ionisation calculation, a more accurate local sources model could be constructed. Finally, in Chapter 2 the model used to obtain the distribution of LAEs in the tomographic maps does not capture the effect of the local IGM opacity on Ly- α emitter visibility. In reality, it should be expected that the observed luminosities of LAEs residing in overdense parts of the IGM will be preferentially diminished as a result of the IGM scattering Ly- α photons. This effect may lead to the correlation between LAE density and IGM opacity being decreased. As such, 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., 2019](#)). It should also be noted that, whilst the model for the Ly- α luminosity of galaxies produces luminosity and two-point correlation

functions that are in reasonably good agreement with observational data, the lack of a perfect LAE model may introduce a bias in the LAE distribution compared with the real Universe.

5.3 Outlook

In Chapters 3 and 4 individual CoSLAs were found to be a good but non-unique probe of protocluster regions. Despite Chapters 3 and 4 finding that CoSLAs are very rare objects with $dn/dz \sim 10^{-2}$ at redshift $z \sim 2.4$, current generation surveys such as WEAVE-QSO (Pieri et al., 2016) and DESI (DESI Collaboration et al., 2016) are expected to obtain spectra for 400,000 and 2.4 million QSOs respectively; this suggests that many CoSLAs should be discovered between these two surveys alone, and the findings from Chapters 3 and 4 also suggest that over half of these CoSLAs should be located within a $M_{z=0} \geq 10^{14} M_{\odot}$ protocluster. These numbers are only set to increase with next generation instruments capable of obtaining spectra for many objects over a large area simultaneously—such as 4MOST ($\sim 2,400$ objects over 4 square degrees; de Jong et al., 2019), the Subaru Prime Focus Spectrograph (2,400 objects over 1.3 degrees; Takada et al., 2014) and the Maunakea Spectroscopic Explorer ($\sim 1,000$ objects over ~ 90 arcseconds; McConnachie et al., 2016)—coming online over the few years, allowing for even greater numbers of QSO spectra to be obtained.

However, I only studied the relation between individual CoSLAs and protocluster regions. The $M_{z=0} \gtrsim 10^{15} M_{\odot}$ protocluster found by Cai et al. (2017a) was located by a group of 6 CoSLAs over $30 h^{-1} \text{cMpc}$. Whilst this is a particularly rare structure, further study into the relation between multiple CoSLAs and protocluster regions may find groups of CoSLAs to be a more reliable probe of protoclusters than individual CoSLAs, but at the expense of completeness. This thesis suggests that a theoretical study should ideally be carried out using large scale hydrodynamical simulations. However, given the rarity of CoSLA groups, current generation hydrodynamical simulations with sufficient resolution to correctly model the small scale components of the Ly- α forest ($M_{\text{gas}} \sim 10^6 M_{\odot}$; see Chapter 3 Section 3.5.3 and Chapter 4) may not follow volumes large enough to obtain robust number statistics. At present the largest simulation volume that approaches a sufficient mass resolution is TNG300-1 from the IllustrisTNG project. At $205^3 h^{-3} \text{cMpc}^3$, TNG300-1 only contains a handful of the most massive ($M_{z=0} \gtrsim 10^{15} M_{\odot}$) protoclusters; an order of magnitude increase in these numbers can be expected with a box size of $\sim 440^3 h^{-3} \text{cMpc}^3$ which currently

represents a significant increase over the largest currently existing hydrodynamical simulations such as BAHAMAS (McCarthy et al., 2017) and BlueTides (Feng et al., 2016), which have already had to sacrifice mass resolution (with BlueTides additionally stopping at high redshift) in order to achieve these volumes.

Furthermore, Fig. 3.18 in Chapter 3 shows tantalising hints that CoSLAs may also be associated with filaments which are preferentially aligned with the QSO sightline. This supports the suggestion from Finley et al. (2014) who claimed a detection of an intergalactic filament at $z = 2.69$, based on observations of seven strong ($N_{\text{HI}} > 10^{18} \text{ cm}^{-2}$) HI absorption systems spanning $16.6h^{-1} \text{ cMpc}$. A potential area of future study could be to quantify the relationship between CoSLAs and filamentary structure, by applying a filament finding algorithm such as DisPerSE (Sousbie, 2011) to a hydrodynamical simulation volume such as IllustrisTNG. Should CoSLAs be confirmed as good probes of filaments within these simulations, then they may provide a novel approach for detecting and studying filamentary environments in the early ($z > 2$) Universe.

The field of Ly- α forest tomography is set to be one of increasing interest in the near future, with new 30 meter class telescopes (The Giant Magellan Telescope¹, first light 2029; The Extremely Large Telescope², first light 2025; The Thirty Meter Telescope³, first light 2027) coming online this decade. Using this next generation of ground-based telescopes, astronomers will be able to obtain spectra from galaxies down to magnitudes of $r < 25.5$. This 1 mag increase in depth in turn results in an order of magnitude increase in the number of background sources for IGM tomography, allowing the resolution of these maps to increase dramatically to $\sim 0.3 \text{ pMpc}$ (Newman et al., 2019). Whilst the current generation of tomography surveys are only capable of revealing protocluster structures and large cosmic voids, this next generation will be able to resolve the IGM on scales only slightly larger than individual galaxy host haloes. Such tomographic surveys will allow astronomers to begin to construct 3-D profiles of the circumgalactic medium around individual galaxies (Rudie et al., 2019). The results from Chapter 2 suggest that the impact of local ionisation enhancements will be significant at these scales, meaning that both the proximity effect and the impact of collisional ionisation will need to be considered during the interpretation of these maps.

In addition to the improvement set to come in observational equipment, work is

¹<https://www.gmto.org/>

²<https://elt.eso.org/>

³<https://www.tmt.org/>

being done to improve the reconstruction techniques used to produce the current generation of tomographic IGM maps. For example, [Li et al. \(2021\)](#) find their ORCA method (Optimized Reconstruction with Constraints on Absorption) provides an improvement in the accuracy of reconstructed flux maps equivalent to an increase of 30-40% in the number of background sources when compared to the traditional Wiener filter method. The results from Chapter 2 suggest that, in order to make accurate predictions of a protocluster’s redshift $z = 0$ mass from its absorption signature in tomographic maps, one must account for the effect of hot gas surrounding dense haloes. Using large volume, high resolution, hydrodynamical simulations such as TNG300-1 may allow astronomers to calibrate this relation—building on the work of [Horowitz et al. \(2021\)](#) using an N-body simulation—and gain a better understanding of how well a protocluster’s $z = 0$ mass can be predicted in this way. Finally, extracting the signature of light echoes from tomographic maps around highly luminous QSOs with $M_{1450} < -27.5$ may prove to be a promising avenue for placing constraints on their optically bright lifetime ([Schmidt et al., 2019](#)). This could be further investigated by improving the local sources model employed in Chapter 2 to account for the effects of light travel time.

Advances in machine learning may also lead to improvements in future simulation work. Recently NVIDIA released version 2 of their Deep Learning Super Sampling technology⁴, which uses a neural network (that has been trained using ultra-high resolution and low resolution renders from the same computer game) to produce upscaled images with a similar fidelity to what would be expected from a native render at the higher resolution. It may be possible to apply a similar technique to hydrodynamical simulations, training a neural network on a small high resolution simulation and then using that to artificially increase the resolution of a larger but lower resolution simulation. Similarly, machine learning could potentially be used to train a neural network that operates in the same way as LyMAS ([Peirani et al., 2014](#)), using a small, high resolution, hydrodynamical simulation as training data for a neural network that could then be used to ‘paint on’ the baryonic physics to a much larger N-body simulation. By exploiting these machine learning methods, it may be possible to exceed the current limitations on hydrodynamical simulations in terms of box size and mass resolution.

In summary, this thesis adds to a rapidly growing volume of research into the Ly- α absorption properties of protocluster regions. Ly- α absorption in protoclusters is a field of research that is maturing from its developmental phase and providing

⁴<https://www.nvidia.com/en-gb/geforce/news/nvidia-dlss-2-0-a-big-leap-in-ai-rendering/>

us with exciting new insights into large scale structure in the early Universe. In the coming years Ly- α forest tomography is only set to further increase in its application and utility, with improvements to computational techniques and the construction of next generation 30 meter class telescopes allowing astronomers to probe the anatomy of the IGM in unprecedented detail.

Bibliography

- Abadi, M. G., Moore, B., and Bower, R. G. (1999). Ram pressure stripping of spiral galaxies in clusters. *Monthly Notices of the Royal Astronomical Society*, 308(4):947–954.
- Adams, S. M., Martini, P., Croxall, K. V., Overzier, R. A., and Silverman, J. D. (2015). Discovery of an overdensity of Lyman alpha emitters around a $z \sim 4$ QSO with the Large Binocular Telescope. *Monthly Notices of the Royal Astronomical Society*, 448(2):1335–1344.
- Adelberger, K. L., Steidel, C. C., Shapley, A. E., and Pettini, M. (2003). Galaxies and Intergalactic Matter at Redshift $z \sim 3$: Overview. *The Astrophysical Journal*, 584(1):45–75.
- Altay, G., Theuns, T., Schaye, J., Booth, C. M., and Vecchia, C. D. (2013). The impact of different physical processes on the statistics of lyman-limit and damped lyman α absorbers. *Monthly Notices of the Royal Astronomical Society*, 436(3):2689–2707.
- Altay, G., Theuns, T., Schaye, J., Crighton, N. H., and Dalla Vecchia, C. (2011). Through thick and thin H i absorption in cosmological simulations. *The Astrophysical Journal Letters*, 737(2):37.
- Angulo, R. E., Springel, V., White, S. D., Jenkins, A., Baugh, C. M., and Frenk, C. S. (2012). Scaling relations for galaxy clusters in the millennium-XXL simulation. *Monthly Notices of the Royal Astronomical Society*, 426(3):2046–2062.
- Ao, Y., Matsuda, Y., Beelen, A., Henkel, C., Cen, R., De Breuck, C., Francis, P. J., Kovács, A., Lagache, G., Lehnert, M., Mao, M. Y., Menten, K. M., Norris, R. P., Omont, A., Tatemastu, K., Weiß, A., and Zheng, Z. (2015). What powers Ly α blobs? *Astronomy & Astrophysics*, 581:A132.

- Arons, J. and McCray, R. (1969). Interaction of Cosmic Gamma Rays with Intergalactic Matter. *The Astrophysical Journal*, 158(8):L91.
- Bădescu, T., Yang, Y., Bertoldi, F., Zabludoff, A., Karim, A., and Magnelli, B. (2017). Discovery of a Protocluster Associated with a Ly α Blob Pair at $z = 2.3$. *The Astrophysical Journal*, 845(2):172.
- Bajtlik, S., Duncan, R. C., and Ostriker, J. P. (1988). Quasar ionization of Lyman-alpha clouds - The proximity effect, a probe of the ultraviolet background at high redshift. *The Astrophysical Journal*, 327:570.
- Baldwin, J. A., Robinson, L. B., Wampler, E. J., Burbidge, E. M., Burbidge, G. R., and Hazard, C. (1974). An analysis of the spectrum of the large-redshift quasi-stellar object OQ 172. *The Astrophysical Journal*, 193:513.
- Balogh, M. L. and Morris, S. L. (2000). H α photometry of Abell 2390. *Monthly Notices of the Royal Astronomical Society*, 318(3):703–714.
- Bañados, E., Venemans, B., Walter, F., Kurk, J., Overzier, R., and Ouchi, M. (2013). The Galaxy Environment of a QSO at $z \sim 5.7$. *The Astrophysical Journal*, 773(2):178.
- Barkana, R. and Loeb, A. (2001). In the beginning: The first sources of light and the reionization of the universe. *Physics Report*, 349(2):125–238.
- Barnes, L. A., Haehnelt, M. G., Tescari, E., and Viel, M. (2011). Galactic winds and extended Ly α emission from the host galaxies of high column density quasi-stellar object absorption systems. *Monthly Notices of the Royal Astronomical Society*, 416(3):1723–1738.
- Becker, G. D. and Bolton, J. S. (2013). New measurements of the ionizing ultraviolet background over $2 < z < 5$ and implications for hydrogen reionization. *Monthly Notices of the Royal Astronomical Society*, 436(2):1023–1039.
- Becker, G. D., Bolton, J. S., Haehnelt, M. G., and Sargent, W. L. (2011). Detection of extended He ii reionization in the temperature evolution of the intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 410(2):1096–1112.
- Becker, G. D., Bolton, J. S., and Lidz, A. (2015). Reionisation and High-Redshift Galaxies: The View from Quasar Absorption Lines. *Publications of the Astronomical Society of Australia*, 32:29.

- Becker, G. D., Hewett, P. C., Worseck, G., and Prochaska, J. X. (2013). A refined measurement of the mean transmitted flux in the Ly α forest over $2 < z < 5$ using composite quasar spectra. *Monthly Notices of the Royal Astronomical Society*, 430:2067–2081.
- Becker, R. H., Fan, X., White, R. L., Strauss, M. A., Narayanan, V. K., Lupton, R. H., Gunn, J. E., Annis, J., Bahcall, N. A., Brinkmann, J., Connolly, A. J., Csabai, I., Czarapata, P. C., Doi, M., Heckman, T. M., Hennessy, G. S., Ivezić, Z., Knapp, G. R., Lamb, D. Q., McKay, T. A., Munn, J. A., Nash, T., Nichol, R., Pier, J. R., Richards, G. T., Schneider, D. P., Stoughton, C., Szalay, A. S., Thakar, A. R., and York, D. G. (2001). Evidence for Reionization at $z \sim 6$: Detection of a Gunn-Peterson Trough in a $z = 6.28$ Quasar. *The Astronomical Journal*, 122(6):2850–2857.
- Best, P. N., Longair, M. S., and Röttgering, H. J. A. (1998). HST, radio and infrared observations of 28 3CR galaxies at $z \sim 1.1$: old stellar populations in central cluster galaxies. *Monthly Notices of the Royal Astronomical Society*, 295(3):549–567.
- Bird, S., Vogelsberger, M., Haehnelt, M., Sijacki, D., Genel, S., Torrey, P., Springel, V., and Hernquist, L. (2014). Damped Lyman α absorbers as a probe of stellar feedback. *Monthly Notices of the Royal Astronomical Society*, 445:2313–2324.
- Blanc, G. A., Adams, J. J., Gebhardt, K., Hill, G. J., Drory, N., Hao, L., Bender, R., Ciardullo, R., Finkelstein, S. L., Fry, A. B., Gawiser, E., Gronwall, C., Hopp, U., Jeong, D., Kelzenberg, R., Komatsu, E., MacQueen, P., Murphy, J. D., Roth, M. M., Schneider, D. P., and Tufts, J. (2011). The HETDEX Pilot Survey. II. The Evolution of the Ly α Escape Fraction from the Ultraviolet Slope and Luminosity Function of $1.9 < z < 3.8$ LAEs. *The Astrophysical Journal*, 736(1):31.
- Bleem, L. E., Stalder, B., De Haan, T., Aird, K. A., Allen, S. W., Applegate, D. E., Ashby, M. L., Bautz, M., Bayliss, M., Benson, B. A., Bocquet, S., Brodwin, M., Carlstrom, J. E., Chang, C. L., Chiu, I., Cho, H. M., Clocchiatti, A., Crawford, T. M., Crites, A. T., Desai, S., Dietrich, J. P., Dobbs, M. A., Foley, R. J., Forman, W. R., George, E. M., Gladders, M. D., Gonzalez, A. H., Halverson, N. W., Hennig, C., Hoekstra, H., Holder, G. P., Holzappel, W. L., Hrubes, J. D., Jones, C., Keisler, R., Knox, L., Lee, A. T., Leitch, E. M., Liu, J., Lueker, M., Luong-Van, D., Mantz, A., Marrone, D. P., McDonald, M., McMahan, J. J., Meyer, S. S., Mocuano, L., Mohr, J. J., Murray, S. S., Padin, S., Pryke, C., Reichardt, C. L., Rest, A., Ruel, J., Ruhl,

- J. E., Saliwanchik, B. R., Saro, A., Sayre, J. T., Schaffer, K. K., Schrabback, T., Shirokoff, E., Song, J., Spieler, H. G., Stanford, S. A., Staniszewski, Z., Stark, A. A., Story, K. T., Stubbs, C. W., Vanderlinde, K., Vieira, J. D., Vikhlinin, A., Williamson, R., Zahn, O., and Zenteno, A. (2015). Galaxy clusters discovered via the sunyaev-zel'dovich effect in the 2500-square-degree spt-sz survey. *The Astrophysical Journal, Supplement Series*, 216(2):27.
- Boera, E., Murphy, M. T., Becker, G. D., and Bolton, J. S. (2014). The thermal history of the intergalactic medium down to redshift $z = 1.5$: A new curvature measurement. *Monthly Notices of the Royal Astronomical Society*, 441(3):1916–1933.
- Bolton, J. S. and Becker, G. D. (2009). Resolving the high redshift Ly α forest in smoothed particle hydrodynamics simulations. *Monthly Notices of the Royal Astronomical Society: Letters*, 398(1):26–30.
- Bolton, J. S., Haehnelt, M. G., Viel, M., and Springel, V. (2005). The Lyman α forest opacity and the metagalactic hydrogen ionization rate at $z \sim 2 - 4$. *Monthly Notices of the Royal Astronomical Society*, 357(4):1178–1188.
- Bolton, J. S., Puchwein, E., Sijacki, D., Haehnelt, M. G., Kim, T. S., Meiksin, A., Regan, J. A., and Viel, M. (2017). The Sherwood simulation suite: Overview and data comparisons with the Lyman α forest at redshifts $2 \leq z \leq 5$. *Monthly Notices of the Royal Astronomical Society*, 464(1):897–914.
- Bolton, J. S. and Viel, M. (2011). The impact of spatial fluctuations in the ultraviolet background on intergalactic carbon and silicon. *Monthly Notices of the Royal Astronomical Society*, 414(1):241–252.
- Booth, C. M. and Schaye, J. (2009). Cosmological simulations of the growth of supermassive black holes and feedback from active galactic nuclei: method and tests. *Monthly Notices of the Royal Astronomical Society*, 398(1):53–74.
- Bower, R. G. and Balogh, M. L. (2004). The Difference Between Clusters and Groups: A Journey from Cluster Cores to Their Outskirts and Beyond. *Carnegie Observatories Astrophysics Series*, 3.
- Boylan-Kolchin, M., Springel, V., White, S. D., Jenkins, A., and Lemson, G. (2009). Resolving cosmic structure formation with the Millennium-II Simulation. *Monthly Notices of the Royal Astronomical Society*, 398(3):1150–1164.
- Burbidge, E. M., Lynds, C. R., and Burbidge, G. R. (1966). ON the Measurement and Interpretation of Absorption Features in the Spectrum of the Quasi-Stellar Object 3c 191. *The Astrophysical Journal*, 144:447.

- Cai, Z., Fan, X., Bian, F., Zabludoff, A., Yang, Y., Prochaska, J. X., McGreer, I., Zheng, Z.-Y., Kashikawa, N., Wang, R., Frye, B., Green, R., and Jiang, L. (2017a). MApping the Most Massive Overdensities (MAMMOTH) II – Discovery of an Extremely Massive Overdensity BOSS1441 at $z = 2.32$. *The Astrophysical Journal*, 839(2):131.
- Cai, Z., Fan, X., Peirani, S., Bian, F., Frye, B., McGreer, I., Prochaska, J. X., Lau, M. W., Tejos, N., Ho, S., and Schneider, D. P. (2016). MApping the Most Massive Overdensity Through Hydrogen (MAMMOTH). I. Methodology. *The Astrophysical Journal*, 833(2):135.
- Cai, Z., Fan, X., Yang, Y., Bian, F., Prochaska, J. X., Zabludoff, A., McGreer, I., Zheng, Z.-Y., Green, R., Cantalupo, S., Frye, B., Hamden, E., Jiang, L., Kashikawa, N., and Wang, R. (2017b). Discovery of an Enormous Ly α Nebula in a Massive Galaxy Overdensity at $z = 2.3$. *The Astrophysical Journal*, 837(1):71.
- Calverley, A. P., Becker, G. D., Haehnelt, M. G., and Bolton, J. S. (2011). Measurements of the ultraviolet background at $4.6 < z < 6.4$ using the quasar proximity effect. *Monthly Notices of the Royal Astronomical Society*, 412(4):2543–2562.
- Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., and Storchi-Bergmann, T. (2000). The Dust Content and Opacity of Actively Star-forming Galaxies. *The Astrophysical Journal*, 533(2):682–695.
- Capak, P. L., Riechers, D., Scoville, N. Z., Carilli, C., Cox, P., Neri, R., Robertson, B., Salvato, M., Schinnerer, E., Yan, L., Wilson, G. W., Yun, M., Civano, F., Elvis, M., Karim, A., Mobasher, B., and Staguhn, J. G. (2011). A massive protocluster of galaxies at a redshift of $z \approx 5.3$. *Nature*, 470(7333):233–235.
- Carswell, R. F., Whelan, J. A. J., Smith, M. G., Boksenberg, A., and Tytler, D. (1982). Observations of the spectra of Q0122 - 380 and Q1101 - 264. *Monthly Notices of the Royal Astronomical Society*, 198(1):91–110.
- Casey, C. M., Narayanan, D., and Cooray, A. (2014). Dusty star-forming galaxies at high redshift. *Physics Reports*, 541(2):45–161.
- Cassata, P., Le Fèvre, O., Garilli, B., Maccagni, D., Le Brun, V., Scodreggio, M., Tresse, L., Ilbert, O., Zamorani, G., Cucciati, O., Contini, T., Bielby, R., Mellier, Y., McCracken, H. J., Pollo, A., Zanichelli, A., Bardelli, S., Cappi, A., Pozzetti, L., Vergani, D., and Zucca, E. (2011). The vimos vlt

- deep survey: star formation rate density of $\text{Ly}\alpha$ emitters from a sample of 217 galaxies with spectroscopic redshifts $2 \leq z \leq 6.6$. *Astronomy & Astrophysics*, 525:A143.
- Caucci, S., Colombi, S., Pichon, C., Rollinde, E., Petitjean, P., and Sousbie, T. (2008). Recovering the topology of the intergalactic medium at $z \sim 2$. *Monthly Notices of the Royal Astronomical Society*, 386(1):211–229.
- Chabrier, G. (2003). Galactic Stellar and Substellar Initial Mass Function. *Publications of the Astronomical Society of the Pacific*, 115(809):763–795.
- Cheng, S., Yu, H. R., Inman, D., Liao, Q., Wu, Q., and Lin, J. (2020). CUBE-Towards an Optimal Scaling of Cosmological N-body Simulations. In *Proceedings - 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing, CCGRID 2020*, pages 685–690. Institute of Electrical and Electronics Engineers Inc.
- Chiang, Y.-K., Makiya, R., Ménard, B., and Komatsu, E. (2020). The cosmic thermal history probed by sunyaev–zeldovich effect tomography. *The Astrophysical Journal*, 902(1):56.
- Chiang, Y. K., Overzier, R., and Gebhardt, K. (2013). Ancient light from young cosmic cities: Physical and observational signatures of galaxy proto-clusters. *The Astrophysical Journal*, 779(2):127.
- Chiang, Y. K., Overzier, R., and Gebhardt, K. (2014). Discovery of a large number of candidate protoclusters traced by ~ 15 Mpc-scale galaxy overdensities in cosmos. *The Astrophysical Journal Letters*, 782(1).
- Chiang, Y. K., Overzier, R. A., Gebhardt, K., Finkelstein, S. L., Chiang, C. T., Hill, G. J., Blanc, G. A., Drory, N., Chonis, T. S., Zeimann, G. R., Hagen, A., Schneider, D. P., Jogee, S., Ciardullo, R., and Gronwall, C. (2015). Surveying galaxy proto-clusters in emission: A large-scale structure at $z = 2.44$ and the outlook for HETDEX. *The Astrophysical Journal*, 808(1).
- Chiang, Y.-K., Overzier, R. A., Gebhardt, K., and Henriques, B. (2017). Galaxy protoclusters as drivers of cosmic star formation history in the first 2 gyr. *The Astrophysical Journal*, 844(2):L23.
- Choudhury, T. R., Paranjape, A., and Bosman, S. E. I. (2021). Studying the Lyman α optical depth fluctuations at $z \sim 5.5$ using fast semi-numerical methods. *Monthly Notices of the Royal Astronomical Society*, 501:5782–5796.

- Collet, C., Nesvadba, N. P. H., De Breuck, C., Lehnert, M. D., Best, P., Bryant, J. J., Dicken, D., Johnston, H., Hunstead, R., and Wylezalek, D. (2015). Defying jet-gas alignment in two radio galaxies at $z \sim 2$ with extended light profiles: Similarities to brightest cluster galaxies. *Astronomy & Astrophysics*, 579:A89.
- Conselice, C. J. (2014). The evolution of galaxy structure over cosmic time. *Annual Review of Astronomy and Astrophysics*, 52(1):291–337.
- Contini, E., Gu, Q., Ge, X., Rhee, J., Yi, S. K., and Kang, X. (2020). The Roles of Mass and Environment In The Quenching of Galaxies. Ii. *The Astrophysical Journal*, 889(2):156.
- Cooke, E. A., Hatch, N. A., Muldrew, S. I., Rigby, E. E., and Kurk, J. D. (2014). A $z = 2.5$ protocluster associated with the radio galaxy MRC 2104-242: Star formation and differing mass functions in dense environments. *Monthly Notices of the Royal Astronomical Society*, 440(4):3262–3274.
- Couchman, H. M. P. and Rees, M. J. (1986). Pregalactic evolution in cosmologies with cold dark matter. *Monthly Notices of the Royal Astronomical Society*, 221(1):53–62.
- Crain, R. A., Bahé, Y. M., Lagos, C. d. P., Rahmati, A., Schaye, J., McCarthy, I. G., Marasco, A., Bower, R. G., Schaller, M., Theuns, T., and van der Hulst, T. (2017). The EAGLE simulations: Atomic hydrogen associated with galaxies. *Monthly Notices of the Royal Astronomical Society*, 464(4):4204–4226.
- Crain, R. A., Schaye, J., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Vecchia, C. D., Frenk, C. S., McCarthy, I. G., Helly, J. C., Jenkins, A., Rosas-Guevara, Y. M., White, S. D., and Trayford, J. W. (2015). The EAGLE simulations of galaxy formation: Calibration of subgrid physics and model variations. *Monthly Notices of the Royal Astronomical Society*, 450(2):1937–1961.
- Croft, R. A. C., Weinberg, D. H., Pettini, M., Hernquist, L., and Katz, N. (1999). The Power Spectrum of Mass Fluctuations Measured from the Ly-alpha Forest at Redshift $z = 2.5$. *The Astrophysical Journal*, 520:1–23.
- Cucciati, O., Zamorani, G., Lemaux, B. C., Bardelli, S., Cimatti, A., Le Fèvre, O., Cassata, P., Garilli, B., Le Brun, V., Maccagni, D., Pentericci, L., Tasca, L. A. M., Thomas, R., Vanzella, E., Zucca, E., Amorin, R., Capak, P., Cassarà, L. P., Castellano, M., Cuby, J. G., de la Torre, S., Durkalec, A., Fontana, A.,

- Giavalisco, M., Grazian, A., Hathi, N. P., Ilbert, O., Moreau, C., Paltani, S., Ribeiro, B., Salvato, M., Schaerer, D., Scodreggio, M., Sommariva, V., Talia, M., Taniguchi, Y., Tresse, L., Vergani, D., Wang, P. W., Charlot, S., Contini, T., Fotopoulou, S., López-Sanjuan, C., Mellier, Y., and Scoville, N. (2014). Discovery of a rich proto-cluster at $z = 2.9$ and associated diffuse cold gas in the VIMOS Ultra-Deep Survey (VUDS). *Astronomy & Astrophysics*, 570:A16.
- Daddi, E., Dannerbauer, H., Stern, D., Dickinson, M., Morrison, G., Elbaz, D., Giavalisco, M., Mancini, C., Pope, A., and Spinrad, H. (2009). Two bright submillimeter galaxies in A $z = 4.05$ protocluster in goods-north, and accurate radio-infrared photometric redshifts. *The Astrophysical Journal*, 694(2):1517–1538.
- Dalla Vecchia, C. and Schaye, J. (2012). Simulating galactic outflows with thermal supernova feedback. *Monthly Notices of the Royal Astronomical Society*, 426(1):140–158.
- Dall’Aglia, A., Wisotzki, L., and Worseck, G. (2008). An unbiased measurement of the UV background and its evolution via the proximity effect in quasar spectra. *Astronomy & Astrophysics*, 491(2):465–481.
- Davé, R., Anglés-Alcázar, D., Narayanan, D., Li, Q., Rafieferantsoa, M. H., and Appleby, S. (2019). SIMBA: Cosmological simulations with black hole growth and feedback. *Monthly Notices of the Royal Astronomical Society*, 486(2):2827–2849.
- Davies, F. B., Becker, G. D., and Furlanetto, S. R. (2018). Determining the nature of late Gunn-Peterson troughs with galaxy surveys. *The Astrophysical Journal*, 860(2):155.
- Dawson, K. S., Kneib, J.-P., Percival, W. J., Alam, S., Albareti, F. D., Anderson, S. F., Armengaud, E., Aubourg, É., Bailey, S., Bautista, J. E., Berlind, A. A., Bershad, M. A., Beutler, F., Bizyaev, D., Blanton, M. R., Blomqvist, M., Bolton, A. S., Bovy, J., Brandt, W. N., Brinkmann, J., Brownstein, J. R., Burtin, E., Busca, N. G., Cai, Z., Chuang, C.-H., Clerc, N., Comparat, J., Cope, F., Croft, R. A. C., Cruz-Gonzalez, I., Costa, L. N. d., Cousinou, M.-C., Darling, J., Macorra, A. d. l., Torre, S. d. l., Delubac, T., Bourboux, H. d. M. d., Dwelly, T., Ealet, A., Eisenstein, D. J., Eracleous, M., Escoffier, S., Fan, X., Finoguenov, A., Font-Ribera, A., Frinchaboy, P., Gaulme, P., Georgakakis, A., Green, P., Guo, H., Guy, J., Ho, S., Holder,

D., Huehnerhoff, J., Hutchinson, T., Jing, Y., Jullo, E., Kamble, V., Kinemuchi, K., Kirkby, D., Kitaura, F.-S., Klaene, M. A., Laher, R. R., Lang, D., Laurent, P., Goff, J.-M. L., Li, C., Liang, Y., Lima, M., Lin, Q., Lin, W., Lin, Y.-T., Long, D. C., Lundgren, B., MacDonald, N., Maia, M. A. G., Malanushenko, E., Malanushenko, V., Mariappan, V., McBride, C. K., McGreer, I. D., Ménard, B., Merloni, A., Meza, A., Montero-Dorta, A. D., Muna, D., Myers, A. D., Nandra, K., Naugle, T., Newman, J. A., Noterdaeme, P., Nugent, P., Ogando, R., Olmstead, M. D., Oravetz, A., Oravetz, D. J., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Parejko, J. K., Pâris, I., Peacock, J. A., Petitjean, P., Pieri, M. M., Pisani, A., Prada, F., Prakash, A., Raichoor, A., Reid, B., Rich, J., Ridl, J., Rodriguez-Torres, S., Rosell, A. C., Ross, A. J., Rossi, G., Ruan, J., Salvato, M., Sayres, C., Schneider, D. P., Schlegel, D. J., Seljak, U., Seo, H.-J., Sesar, B., Shandera, S., Shu, Y., Slosar, A., Sobreira, F., Streblyanska, A., Suzuki, N., Taylor, D., Tao, C., Tinker, J. L., Tojeiro, R., Vargas-Magaña, M., Wang, Y., Weaver, B. A., Weinberg, D. H., White, M., Wood-Vasey, W. M., Yeche, C., Zhai, Z., Zhao, C., Zhao, G.-b., Zheng, Z., Zhu, G. B., and Zou, H. (2016). The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data. *The Astronomical Journal*, 151(2):44.

Dawson, K. S., Schlegel, D. J., Ahn, C. P., Anderson, S. F., Aubourg, É., Bailey, S., Barkhouser, R. H., Bautista, J. E., Beifiori, A., Berlind, A. A., Bhardwaj, V., Bizyaev, D., Blake, C. H., Blanton, M. R., Blomqvist, M., Bolton, A. S., Borde, A., Bovy, J., Brandt, W. N., Brewington, H., Brinkmann, J., Brown, P. J., Brownstein, J. R., Bundy, K., Busca, N. G., Carithers, W., Carnero, A. R., Carr, M. A., Chen, Y., Comparat, J., Connolly, N., Cope, F., Croft, R. A., Cuesta, A., Dacosta, L., Davenport, J. R., Delubac, T., De Putter, R., Dhital, S., Ealet, A., Ebelke, G. L., Eisenstein, D. J., Escoffier, S., Fan, X., Filiz Ak, N., Finley, H., Font-Ribera, A., Génova-Santos, R., Gunn, J. E., Guo, H., Haggard, D., Hall, P. B., Hamilton, J. C., Harris, B., Harris, D. W., Ho, S., Hogg, D. W., Holder, D., Honscheid, K., Huehnerhoff, J., Jordan, B., Jordan, W. P., Kauffmann, G., Kazin, E., Kirkby, D., Klaene, M. A., Kneib, J. P., Legoff, J. M., Lee, K. G., Long, D. C., Loomis, C. P., Lundgren, B., Lupton, R., Maia, M. A., Makler, M., Malanushenko, E., Malanushenko, V., Mandelbaum, R., Manera, M., Maraston, C., Margala, D., Masters, K. L., McBride, C. K., McDonald, P., McGreer, I., McMahan, R., Mena, O., Miralda-Escudé, J., Montero-Dorta, A. D., Montesano, F., Muna, D., Myers, A. D., Naugle, T., Nichol, R. C., Noterdaeme, P., Nuza, S. E., Olmstead, M. D., Oravetz, A., Oravetz, D., Owen, R., Padmanabhan,

N., Palanque-Delabrouille, N., Pan, K., Parejko, J., Pâris, I., Percival, W. J., Pérez-Fournon, I., Pérez-Ràfols, I., Petitjean, P., Pfaffenberger, R., Pforr, J., Pieri, M. M., Prada, F., Price-Whelan, A., Raddick, M. J., Rebolo, R., Rich, J., Richards, G. T., Rockosi, C., Roe, N., Ross, A. J., Ross, N. P., Rossi, G., Rubiño-Martin, J. A., Samushia, L., Sánchez, A., Sayres, C., Schmidt, S. J., Schneider, D. P., Scóccola, C. G., Seo, H. J., Shelden, A., Sheldon, E., Shen, Y., Shu, Y., Slosar, A., Smee, S., Snedden, S. A., Stauffer, F., Steele, O., Strauss, M., Streblyanska, A., Suzuki, N., Swanson, M. C., Tal, T., Tanaka, M., Thomas, D., Tinker, J. L., Tojeiro, R., Tremonti, C., Vargas Magaña, M., Verde, L., Viel, M., Wake, D. A., Watson, M., Weaver, B., Weinberg, D. H., Weiner, B. J., West, A. A., White, M., Wood-Vasey, W. M., Yèche, C., Zehavi, I., Zhao, G. B., and Zheng, Z. (2013). The baryon oscillation spectroscopic survey of SDSS-III. *The Astronomical Journal*, 145(1):10.

de Jong, R. S., Agertz, O., Berbel, A. A., Aird, J., Alexander, D. A., Amarsi, A., Anders, F., Andrae, R., Ansarinejad, B., Ansorge, W., Antilogus, P., Anwand-Heerwart, H., Arentsen, A., Arnadottir, A., Asplund, M., Auger, M., Azais, N., Baade, D., Baker, G., Baker, S., Balbinot, E., Baldry, I. K., Banerji, M., Barden, S., Barklem, P., Barthélémy-Mazot, E., Battistini, C., Bauer, S., Bell, C. P., Bellido-Tirado, O., Bellstedt, S., Belokurov, V., Bensby, T., Bergemann, M., Bestenlehner, J. M., Bielby, R., Bilicki, M., Blake, C., Bland-Hawthorn, J., Boeche, C., Boland, W., Boller, T., Bongard, S., Bongiorno, A., Bonifacio, P., Boudon, D., Brooks, D., Brown, M. J., Brown, R., Brügger, M., Brynnel, J., Brzeski, J., Buchert, T., Buschkamp, P., Caffau, E., Caillier, P., Carrick, J., Casagrande, L., Case, S., Casey, A., Cesarini, I., Cescutti, G., Chapuis, D., Chiappini, C., Childress, M., Christlieb, N., Church, R., Cioni, M. R. L., Cluver, M., Colless, M., Collett, T., Comparat, J., Cooper, A., Couch, W., Courbin, F., Croom, S., Croton, D., Daguisé, E., Dalton, G., Davies, L. J., Davis, T., de Laverny, P., Deason, A., Dionies, F., Disseau, K., Doel, P., Döscher, D., Driver, S. P., Dwelly, T., Eckert, D., Edge, A., Edvardsson, B., Youssoufi, D. E., Elhaddad, A., Enke, H., Erfanianfar, G., Farrell, T., Fechner, T., Feiz, C., Feltzing, S., Ferreras, I., Feuerstein, D., Feuillet, D., Finoguenov, A., Ford, D., Fotopoulou, S., Fouesneau, M., Frenk, C., Frey, S., Gaessler, W., Geier, S., Fusillo, N. G., Gerhard, O., Giannantonio, T., Giannone, D., Gibson, B., Gillingham, P., González-Fernández, C., Gonzalez-Solares, E., Gottloeber, S., Gould, A., Grebel, E. K., Gueguen, A., Guiglion, G., Haehnelt, M., Hahn, T., Hansen, C. J., Hartman, H., Hauptner, K., Hawkins, K., Haynes, D., Haynes, R., Heiter, U., Helmi, A., Aguayo, C. H., Hewett, P., Hinton, S., Hobbs, D.,

Hoenig, S., Hofman, D., Hook, I., Hopgood, J., Hopkins, A., Hourihane, A., Howes, L., Howlett, C., Huet, T., Irwin, M., Iwert, O., Jablonka, P., Jahn, T., Jahnke, K., Jarno, A., Jin, S., Jofre, P., Johl, D., Jones, D., Jönsson, H., Jordan, C., Karovicova, I., Khalatyan, A., Kelz, A., Kennicutt, R., King, D., Kitaura, F., Klar, J., Klauser, U., Kneib, J. P., Koch, A., Kuposov, S., Kordopatis, G., Korn, A., Kosmalski, J., Kotak, R., Kovalev, M., Kreckel, K., Kripak, Y., Krumpe, M., Kuijken, K., Kunder, A., Kushniruk, I., Lam, M. I., Lamer, G., Laurent, F., Lawrence, J., Lehmitz, M., Lemasle, B., Lewis, J., Li, B., Lidman, C., Lind, K., Liske, J., Lizon, J. L., Loveday, J., Ludwig, H. G., McDermid, R. M., Maguire, K., Mainieri, V., Mali, S., Mandel, H., Mandel, K., Mannering, L., Martell, S., Delgado, D. M., Matijevic, G., McGregor, H., McMahan, R., McMillan, P., Mena, O., Merloni, A., Meyer, M. J., Michel, C., Micheva, G., Migniau, J. E., Minchev, I., Monari, G., Muller, R., Murphy, D., Muthukrishna, D., Nandra, K., Navarro, R., Ness, M., Nichani, V., Nichol, R., Nicklas, H., Niederhofer, F., Norberg, P., Obreschkow, D., Oliver, S., Owers, M., Pai, N., Pankratow, S., Parkinson, D., Paschke, J., Paterson, R., Pecontal, A., Parry, I., Phillips, D., Pillepich, A., Pinard, L., Pirard, J., Piskunov, N., Plank, V., Plüschke, D., Pons, E., Popesso, P., Power, C., Pragt, J., Pramskiy, A., Pryer, D., Quattri, M., de Andrade Queiroz, A. B., Quirrenbach, A., Rahurkar, S., Raichoor, A., Ramstedt, S., Rau, A., Recio-Blanco, A., Reiss, R., Renaud, F., Revaz, Y., Rhode, P., Richard, J., Richter, A. D., Rix, H. W., Robotham, A. S., Roelfsema, R., Romaniello, M., Rosario, D., Rothmaier, F., Roukema, B., Ruchti, G., Rupprecht, G., Rybizki, J., Ryde, N., Saar, A., Sadler, E., Sahlén, M., Salvato, M., Sassolas, B., Saunders, W., Saviauk, A., Sbordone, L., Schmidt, T., Schnurr, O., Scholz, R. D., Schwöpe, A., Seifert, W., Shanks, T., Sheinis, A., Sivov, T., Skúladóttir, Á., Smartt, S., Smedley, S., Smith, G., Smith, R., Sorce, J., Spitler, L., Starkenburg, E., Steinmetz, M., Stiliz, I., Storm, J., Sullivan, M., Sutherland, W., Swann, E., Tamone, A., Taylor, E. N., Teillon, J., Tempel, E., ter Horst, R., Thi, W. F., Tolstoy, E., Trager, S., Traven, G., Tremblay, P. E., Tresse, L., Valentini, M., van de Weygaert, R., van den Ancker, M., Veljanoski, J., Venkatesan, S., Wagner, L., Wagner, K., Walcher, C. J., Waller, L., Walton, N., Wang, L., Winkler, R., Wisotzki, L., Worley, C. C., Worseck, G., Xiang, M., Xu, W., Yong, D., Zhao, C., Zheng, J., Zscheyge, F., and Zucker, D. (2019). 4MOST: Project overview and information for the first call for proposals. *The Messenger*, 175:3.

DESI Collaboration, Aghamousa, A., Aguilar, J., Ahlen, S., Alam, S., Allen, L. E., Prieto, C. A., Annis, J., Bailey, S., Balland, C., Ballester, O., Baltay,

C., Beaufore, L., Bebek, C., Beers, T. C., Bell, E. F., Bernal, J. L., Besuner, R., Beutler, F., Blake, C., Bleuler, H., Blomqvist, M., Blum, R., Bolton, A. S., Briceno, C., Brooks, D., Brownstein, J. R., Buckley-Geer, E., Burden, A., Burtin, E., Busca, N. G., Cahn, R. N., Cai, Y.-C., Cardiel-Sas, L., Carlberg, R. G., Carton, P.-H., Casas, R., Castander, F. J., Cervantes-Cota, J. L., Claybaugh, T. M., Close, M., Coker, C. T., Cole, S., Comparat, J., Cooper, A. P., Cousinou, M. C., Croce, M., Cuby, J.-G., Cunningham, D. P., Davis, T. M., Dawson, K. S., de la Macorra, A., De Vicente, J., Delubac, T., Derwent, M., Dey, A., Dhungana, G., Ding, Z., Doel, P., Duan, Y. T., Ealet, A., Edelstein, J., Eftekhazadeh, S., Eisenstein, D. J., Elliott, A., Escoffier, S., Evatt, M., Fagrelus, P., Fan, X., Fanning, K., Farahi, A., Farihi, J., Favole, G., Feng, Y., Fernandez, E., Findlay, J. R., Finkbeiner, D. P., Fitzpatrick, M. J., Flaugher, B., Flender, S., Font-Ribera, A., Forero-Romero, J. E., Fosalba, P., Frenk, C. S., Fumagalli, M., Gaensicke, B. T., Gallo, G., Garcia-Bellido, J., Gaztanaga, E., Fusillo, N. P. G., Gerard, T., Gershkovich, I., Giannantonio, T., Gillet, D., Gonzalez-de Rivera, G., Gonzalez-Perez, V., Gott, S., Graur, O., Gutierrez, G., Guy, J., Habib, S., Heetderks, H., Heetderks, I., Heitmann, K., Hellwing, W. A., Herrera, D. A., Ho, S., Holland, S., Honscheid, K., Huff, E., Hutchinson, T. A., Huterer, D., Hwang, H. S., Laguna, J. M. I., Ishikawa, Y., Jacobs, D., Jeffrey, N., Jelinsky, P., Jennings, E., Jiang, L., Jimenez, J., Johnson, J., Joyce, R., Jullo, E., Juneau, S., Kama, S., Karcher, A., Karkar, S., Kehoe, R., Kennamer, N., Kent, S., Kilbinger, M., Kim, A. G., Kirkby, D., Kisner, T., Kitanidis, E., Kneib, J.-P., Kuposov, S., Kovacs, E., Koyama, K., Kremin, A., Kron, R., Kronig, L., Kueter-Young, A., Lacey, C. G., Lafever, R., Lahav, O., Lambert, A., Lampton, M., Landriau, M., Lang, D., Lauer, T. R., Goff, J.-M. L., Guillo, L. L., Van Suu, A. L., Lee, J. H., Lee, S.-J., Leitner, D., Lesser, M., Levi, M. E., L'Huillier, B., Li, B., Liang, M., Lin, H., Linder, E., Loebman, S. R., Lukić, Z., Ma, J., MacCrann, N., Magneville, C., Makarem, L., Manera, M., Manser, C. J., Marshall, R., Martini, P., Massey, R., Matheson, T., McCauley, J., McDonald, P., McGreer, I. D., Meisner, A., Metcalfe, N., Miller, T. N., Miquel, R., Moustakas, J., Myers, A., Naik, M., Newman, J. A., Nichol, R. C., Nicola, A., da Costa, L. N., Nie, J., Niz, G., Norberg, P., Nord, B., Norman, D., Nugent, P., O'Brien, T., Oh, M., Olsen, K. A. G., Padilla, C., Padmanabhan, H., Padmanabhan, N., Palanque-Delabrouille, N., Palmese, A., Pappalardo, D., Pâris, I., Park, C., Patej, A., Peacock, J. A., Peiris, H. V., Peng, X., Percival, W. J., Perruchot, S., Pieri, M. M., Pogge, R., Pollack, J. E., Poppett, C., Prada, F., Prakash, A., Probst, R. G., Rabinowitz, D., Raichoor, A., Ree, C. H., Refregier, A., Regal, X., Reid, B.,

Reil, K., Rezaie, M., Rockosi, C. M., Roe, N., Ronayette, S., Roodman, A., Ross, A. J., Ross, N. P., Rossi, G., Rozo, E., Ruhlmann-Kleider, V., Rykoff, E. S., Sabiu, C., Samushia, L., Sanchez, E., Sanchez, J., Schlegel, D. J., Schneider, M., Schubnell, M., Secroun, A., Seljak, U., Seo, H.-J., Serrano, S., Shafieloo, A., Shan, H., Sharples, R., Sholl, M. J., Shourt, W. V., Silber, J. H., Silva, D. R., Sirk, M. M., Slosar, A., Smith, A., Smoot, G. F., Som, D., Song, Y.-S., Sprayberry, D., Staten, R., Stefanik, A., Tarle, G., Tie, S. S., Tinker, J. L., Tojeiro, R., Valdes, F., Valenzuela, O., Valluri, M., Vargas-Magana, M., Verde, L., Walker, A. R., Wang, J., Wang, Y., Weaver, B. A., Weaverdyck, C., Wechsler, R. H., Weinberg, D. H., White, M., Yang, Q., Yeche, C., Zhang, T., Zhao, G.-B., Zheng, Y., Zhou, X., Zhou, Z., Zhu, Y., Zou, H., and Zu, Y. (2016). The DESI Experiment Part I: Science, Targeting, and Survey Design. *arXiv:1611.00036*.

Dijkstra, M. (2017). Lecture notes for the 46th Saas-Fee winterschool.

Dressler, A. (1980). Galaxy morphology in rich clusters - Implications for the formation and evolution of galaxies. *The Astrophysical Journal*, 236:351.

Dressler, A., Thompson, I. B., and Shectman Mount Wilson, S. A. (1985). Statistics of Emission-line Galaxies in Rich Clusters. *The Astrophysical Journal*, 288:481–486.

Elbaz, D., Daddi, E., Le Borgne, D., Dickinson, M., Alexander, D. M., Chary, R. R., Starck, J. L., Brandt, W. N., Kitzbichler, M., MacDonald, E., Nonino, M., Popesso, P., Stern, D., and Vanzella, E. (2007). The reversal of the star formation-density relation in the distant universe. *Astronomy & Astrophysics*, 468(1):33–48.

Fan, X., Narayanan, V. K., Strauss, M. A., White, R. L., Becker, R. H., Pentericci, L., and Rix, H.-W. (2001). Evolution of the Ionizing Background and the Epoch of Reionization from the Spectra of $z \sim 6$ Quasars. *The Astronomical Journal*, 123(3):1247–1257.

Fan, X., Strauss, M. A., Becker, R. H., White, R. L., Gunn, J. E., Knapp, G. R., Richards, G. T., Schneider, D. P., Brinkmann, J., and Fukugita, M. (2006). Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with $z \sim 6$ Quasars. II. A Sample of 19 Quasars. *The Astronomical Journal*, 132(1):117–136.

Fanidakis, N., Macciò, A. V., Baugh, C. M., Lacey, C. G., and Frenk, C. S. (2013). The most luminous quasars do not live in the most massive dark

- matter haloes at any redshift. *Monthly Notices of the Royal Astronomical Society*, 436(1):315–326.
- Faucher-Giguère, C. A., Feldmann, R., Quataert, E., Kereš, D., Hopkins, P. F., and Murray, N. (2016). A stellar feedback origin for neutral hydrogen in high-redshift quasar-mass haloes. *Monthly Notices of the Royal Astronomical Society: Letters*, 461(1):L32–L36.
- Faucher-Giguère, C. A., Lidz, A., Zaldarriaga, M., and Hernquist, L. (2009). A new calculation of the ionizing background spectrum and the effects of He II reionization. *The Astrophysical Journal*, 703(2):1416–1443.
- Faucher-Giguère, C.-A. (2020). A cosmic UV/X-ray background model update. *Monthly Notices of the Royal Astronomical Society*, 493(2):1614–1632.
- Feng, Y., Di-Matteo, T., Croft, R. A., Bird, S., Battaglia, N., and Wilkins, S. (2016). The BlueTides simulation: first galaxies and reionization. *Monthly Notices of the Royal Astronomical Society*, 455(3):2778–2791.
- Finley, H., Petitjean, P., Noterdaeme, P., and Pâris, I. (2014). Astrophysics A ~ 6 Mpc overdensity at $z \simeq 2.7$ detected along a pair of quasar sight lines: filament or protocluster? *Astronomy & Astrophysics*, 572:31.
- Fixsen, D. J. (2009). The Temperature of the Cosmic Microwave Background. *The Astrophysical Journal*, 707:916–920.
- Font-Ribera, A., Arnau, E., Miralda-Escudé, J., Rollinde, E., Brinkmann, J., Brownstein, J. R., Lee, K. G., Myers, A. D., Palanque-Delabrouille, N., Pâris, I., Petitjean, P., Rich, J., Ross, N. P., Schneider, D. P., and White, M. (2013). The large-scale quasar-Lyman α forest cross-correlation from BOSS. *Journal of Cosmology and Astroparticle Physics*, 2013(5):018.
- Francis, P., Woodgate, B., Warren, S., Moller, P., Mazzolini, M., Bunker, A., Lowenthal, J., Williams, T., Minezaki, T., Kobayashi, Y., and Yoshii, Y. (1996). A Group of Galaxies at Redshift 2.38. *The Astrophysical Journal*, 457:490.
- Francis, P. J. and Hewett, P. C. (1993). Superclustering of damped Lyman-alpha absorption systems at redshifts above two. *The Astronomical Journal*, 105:1633.
- Franx, M., Illingworth, G., and de Zeeuw, T. (1991). The ordered nature of elliptical galaxies - Implications for their intrinsic angular momenta and shapes. *The Astrophysical Journal*, 383:112–134.

- Fumagalli, M., Hennawi, J. F., Prochaska, J. X., Kasen, D., Dekel, A., Ceverino, D., and Primack, J. (2014). Confronting simulations of optically thick gas in massive halos with observations at $z = 2 - 3$. *The Astrophysical Journal*, 780(1):74.
- Garn, T. and Best, P. N. (2010). Predicting dust extinction from the stellar mass of a galaxy. *Monthly Notices of the Royal Astronomical Society*, 409(1):421–432.
- Garzilli, A., Bolton, J. S., Kim, T. S., Leach, S., and Viel, M. (2012). The intergalactic medium thermal history at redshift $z = 1.7 - 3.2$ from the Lyman alpha forest: a comparison of measurements using wavelets and the flux distribution. *Monthly Notices of the Royal Astronomical Society*, 424(3):1723–1736.
- Gawiser, E., Francke, H., Lai, K., Schawinski, K., Gronwall, C., Ciardullo, R., Quadri, R., Orsi, A., Barrientos, L. F., Blanc, G. A., Fazio, G., Feldmeier, J. J., sheng Huang, J., Infante, L., Lira, P., Padilla, N., Taylor, E. N., Treister, E., Urry, C. M., van Dokkum, P. G., and Virani, S. N. (2007). Ly α -emitting galaxies at $z = 3.1$: progenitors experiencing rapid star formation. *The Astrophysical Journal*, 671(1):278–284.
- Gingold, R. A. and Monaghan, J. J. (1977). Smoothed particle hydrodynamics: theory and application to non-spherical stars. *Monthly Notices of the Royal Astronomical Society*, 181(3):375–389.
- Gladders, M. D. and Yee, H. K. C. (2000). A New Method For Galaxy Cluster Detection I: The Algorithm. *The Astronomical Journal*, 120:2148–2162.
- Goto, T., Yamauchi, C., Fujita, Y., Okamura, S., Sekiguchi, M., Smail, I., Bernardi, M., and Gomez, P. L. (2003). The Morphology-Density Relation in the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 346(2):601–614.
- Greenslade, J., Clements, D. L., Cheng, T., De Zotti, G., Scott, D., Valiante, E., Eales, S., Bremer, M. N., Dannerbauer, H., Birkinshaw, M., Farrah, D., and Harrison, D. L. (2018). Candidate high- z protoclusters among the Planck compact sources, as revealed by Herschel-SPIRE. *Monthly Notices of the Royal Astronomical Society*, 476:14.
- Gunn, J. E. and Gott, J. Richard, I. (1972). On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *The Astrophysical Journal*, 176:1.

- Gunn, J. E. and Peterson, B. A. (1965). On the Density of Neutral Hydrogen in Intergalactic Space. *The Astrophysical Journal*, 142:1633.
- Gurung-López, S., Orsi, Á. A., Bonoli, S., Padilla, N., Lacey, C. G., and Baugh, C. M. (2019). Ly α emitters in a cosmological volume II: the impact of the intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 491(3):3266–3289.
- Gurvich, A., Burkhart, B., and Bird, S. (2017). The Effect of AGN Heating on the Low-redshift Ly α Forest. *The Astrophysical Journal*, 835(2):175.
- Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23(2):347–356.
- Haardt, F. and Madau, P. (2001). Modelling the uv/x-ray cosmic background with CUBA. *arXiv:astro-ph/0106018*.
- Haardt, F. and Madau, P. (2012). Radiative transfer in a clumpy universe. IV. New synthesis models of the cosmic UV/X-RAY background. *The Astrophysical Journal*, 746(2).
- Harikane, Y., Ouchi, M., Ono, Y., Fujimoto, S., Donevski, D., Shibuya, T., Faisst, A. L., Goto, T., Hatsukade, B., Kashikawa, N., Kohno, K., Hashimoto, T., Higuchi, R., Inoue, A. K., Lin, Y.-T., Martin, C. L., Overzier, R., Smail, I., Toshikawa, J., Umehata, H., Ao, Y., Chapman, S., Clements, D. L., Im, M., Jing, Y., Kawaguchi, T., Lee, C.-H., Lee, M. M., Lin, L., Matsuoka, Y., Marinello, M., Nagao, T., Onodera, M., Toft, S., and Wang, W.-H. (2019). SILVERRUSH. VIII. Spectroscopic Identifications of Early Large-scale Structures with Protoclusters over 200 Mpc at $z \sim 6 - 7$: Strong Associations of Dusty Star-forming Galaxies. *The Astrophysical Journal*, 883:142.
- Hatch, N. A., Cooke, E. A., Muldrew, S. I., Hartley, W. G., Almaini, O., Conselice, C. J., and Simpson, C. J. (2016). The impact of protocluster environments at $z = 1.6$. *Monthly Notices of the Royal Astronomical Society*, 464:876–884.
- Hatch, N. A., De Breuck, C., Galametz, A., Miley, G. K., Overzier, R. A., Röttgering, H. J., Doherty, M., Kodama, T., Kurk, J. D., Seymour, N., Venemans, B. P., Vernet, J., and Zirm, A. W. (2011a). Galaxy protocluster candidates around $z \sim 2.4$ radio galaxies. *Monthly Notices of the Royal Astronomical Society*, 410(3):1537–1549.

- Hatch, N. A., Kurk, J. D., Pentericci, L., Venemans, B. P., Kuiper, E., Miley, G. K., and Röttgering, H. J. (2011b). H α emitters in $z \sim 2$ protoclusters: Evidence for faster evolution in dense environments. *Monthly Notices of the Royal Astronomical Society*, 415(4):2993–3005.
- Hayes, M., Schaerer, D., Stlin, G., Mas-Hesse, J. M., Atek, H., and Kunth, D. (2011). On the redshift evolution of the Ly α escape fraction and the dust content of galaxies. *The Astrophysical Journal*, 730(1):8.
- Hennawi, J. F., Prochaska, J. X., Cantalupo, S., and Arrigoni-Battaia, F. (2015). Quasar quartet embedded in giant nebula reveals rare massive structure in distant universe. *Science*, 348(6236):779–783.
- Hinshaw, G., Larson, D., Komatsu, E., Spergel, D. N., Bennett, C. L., Dunkley, J., Nolta, M. R., Halpern, M., Hill, R. S., Odegard, N., Page, L., Smith, K. M., Weiland, J. L., Gold, B., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., Wollack, E., and Wright, E. L. (2013). Nine-year wilkinson microwave anisotropy probe (WMAP) observations: Cosmological parameter results. *The Astrophysical Journal, Supplement Series*, 208(2):19.
- Holmberg, E. (1941). On the Clustering Tendencies among the Nebulae. II. a Study of Encounters Between Laboratory Models of Stellar Systems by a New Integration Procedure. *The Astrophysical Journal*, 94:385.
- Horowitz, B., Lee, K. G., White, M., Krolewski, A., and Ata, M. (2019). TARDIS paper I: A constrained reconstruction approach to modelling the $z \sim 2.5$ cosmic web probed by lyman- α forest tomography. *The Astrophysical Journal*, 887(1):61.
- Horowitz, B., Zhang, B., Lee, K.-G., and Kooistra, R. (2021). TARDIS. II. Synergistic Density Reconstruction from Ly α Forest and Spectroscopic Galaxy Surveys with Applications to Protoclusters and the Cosmic Web. *The Astrophysical Journal*, 906:110.
- Hu, W., Wang, J., Infante, L., Rhoads, J. E., Zheng, Z. Y., Yang, H., Malhotra, S., Barrientos, L. F., Jiang, C., González-López, J., Prieto, G., Perez, L. A., Hibon, P., Galaz, G., Coughlin, A., Harish, S., Kong, X., Kang, W., Khostovan, A. A., Pharo, J., Valdes, F., Wold, I., Walker, A. R., and Zheng, X. Z. (2021). A Lyman- α protocluster at redshift 6.9. *Nature Astronomy*, pages 1–6.
- Hu, W. and White, M. (1997). Measuring the Curvature of the Universe. In *Proceedings of the XXXIth Moriond meeting*.

- Hui, L. and Gnedin, N. Y. (1997). Equation of state of the photoionized intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 292(1):27–42.
- Hunstead, R. W., Murdoch, H. S., Pettini, M., and Blades, J. C. (1988). The distribution of Lyman-alpha absorption lines in high-redshift QSOs - The case for evolution reexamined. *The Astrophysical Journal*, 329:527.
- Husband, K., Bremer, M. N., Stanway, E. R., Davies, L. J. M., Lehnert, M. D., and Douglas, L. S. (2013). Are $z \sim 5$ quasars found in the most massive high-redshift overdensities? *Monthly Notices of the Royal Astronomical Society*, 432(4):2869–2877.
- Katz, N., Weinberg, D. H., and Hernquist, L. (1996a). Cosmological Simulations with TreeSPH. *The Astrophysical Journal Supplement Series*, 105:19.
- Katz, N., Weinberg, D. H., Hernquist, L., and Miralda-Escudé, J. (1996b). Damped Lyman-Alpha and Lyman-Limit Absorbers in the Cold Dark Matter Model. *The Astrophysical Journal*, 457(2):L57.
- Kauffmann, G., White, S. D., Heckman, T. M., Ménard, B., Brinchmann, J., Charlot, S., Tremonti, C., and Brinkmann, J. (2004). The environmental dependence of the relations between stellar mass, structure, star formation and nuclear activity in galaxies. *Monthly Notices of the Royal Astronomical Society*, 353(3):713–731.
- Khostovan, A. A., Sobral, D., Mobasher, B., Matthee, J., Cochrane, R. K., Chartab, N., Jafariyazani, M., Paulino-Afonso, A., Santos, S., and Calhau, J. (2019). The clustering of typical Ly α emitters from $z \sim 2.5 - 6$: host halo masses depend on Ly α and UV luminosities. *Monthly Notices of the Royal Astronomical Society*, 489(1):555–573.
- Kim, T.-S., Partl, A. M., Carswell, R. F., and Müller, V. (2013). The evolution of HI and CIV quasar absorption line systems at $1.9 < z < 3.2$. *Astronomy & Astrophysics*, 552:77.
- Kinman, T. D. (1966). Object PHL 938 IN the Haro-Luyten Catalogue of Blue Stars. *The Astrophysical Journal*, 144:1232.
- Konno, A., Ouchi, M., Nakajima, K., Duval, F., Kusakabe, H., Ono, Y., and Shimasaku, K. (2016). Bright and Faint Ends of $ly\alpha$ Luminosity Functions at $z = 2$ Determined by the SUBARU Survey: Implications For AGNs, Magnification Bias, and ISM h i Evolution. *The Astrophysical Journal*, 823(1):20.

- Koyama, Y., Smail, I., Kurk, J., Geach, J. E., Sobral, D., Kodama, T., Nakata, F., Swinbank, A. M., Best, P. N., Hayashi, M., and Tadaki, K. i. (2013). On the evolution and environmental dependence of the star formation rate versus stellar mass relation since $z \sim 2$. *Monthly Notices of the Royal Astronomical Society*, 434(1):423–436.
- Kravtsov, A. and Borgani, S. (2012). Formation of Galaxy Clusters. *Annual Review of Astronomy and Astrophysics*, 50(1):353–409.
- Krishnan, C., Hatch, N. A., Almaini, O., Kocevski, D., Cooke, E. A., Hartley, W. G., Hasinger, G., Maltby, D. T., Muldrew, S. I., and Simpson, C. (2017). Enhancement of AGN in a protocluster at $z = 1.6$. *Monthly Notices of the Royal Astronomical Society*, 470:2170–2178.
- Krolewski, A., Lee, K.-G., White, M., Hennawi, J. F., Schlegel, D. J., Nugent, P. E., Lukić, Z., Stark, C. W., Koekemoer, A. M., Fèvre, O. L., Lemaux, B. C., Maier, C., Rich, R. M., Salvato, M., and Tasca, L. (2018). Detection of $z \sim 2.3$ Cosmic Voids from 3D Ly α Forest Tomography in the COSMOS Field. *The Astrophysical Journal*, 861(1):60.
- Kulkarni, G., Hennawi, J. F., Oñorbe, J., Rorai, A., and Springel, V. (2015). Characterizing the Pressure Smoothing Scale of the Intergalactic Medium. *The Astrophysical Journal*, 812(1):30.
- Kulkarni, G., Worseck, G., and Hennawi, J. F. (2019). Evolution of the AGN UV luminosity function from redshift 7.5. *Monthly Notices of the Royal Astronomical Society*, 488(1):1035–1065.
- Larson, R. B., Tinsley, B. M., and Caldwell, C. N. (1980). The evolution of disk galaxies and the origin of S0 galaxies. *The Astrophysical Journal*, 237:692.
- Laursen, P., Sommer-Larsen, J., and Razoumov, A. O. (2011). Intergalactic transmission and its impact on the Ly α line. *The Astrophysical Journal*, 728(1):52.
- Le Fèvre, O., Deltorn, J. M., Crampton, D., and Dickinson, M. (1996). Clustering Around the Radio Galaxy MRC 03162-257 at $z = 3.14$. *The Astrophysical Journal*, 471(1):L11–L14.
- Lee, K. G., Hennawi, J. F., White, M., Croft, R. A., and Ozbek, M. (2014). Observational requirements for Ly α forest tomographic mapping of large-scale structure at $z \sim 2$. *The Astrophysical Journal*, 788(1):49.

- Lee, K.-G., Hennawi, J. F., White, M., Prochaska, J. X., Font-Ribera, A., Schlegel, D. J., Rich, R. M., Suzuki, N., Stark, C. W., Fèvre, O. L., Nugent, P. E., Salvato, M., and Zamorani, G. (2016). Shadow of a Colossus: a $z = 2.44$ Galaxy Protocluster Detected in 3D Ly α Forest Tomographic Mapping of the COSMOS Field. *The Astrophysical Journal*, 817(2):160.
- Lee, K.-G., Krolewski, A., White, M., Schlegel, D., Nugent, P. E., Hennawi, J. F., Müller, T., Pan, R., Prochaska, J. X., Font-Ribera, A., Suzuki, N., Glazebrook, K., Kacprzak, G. G., Kartaltepe, J. S., Fèvre, O. L., Lemaux, B. C., Maier, C., Nanayakkara, T., Rich, R. M., Sanders, D. B., Salvato, M., Tasca, L., and Tran, K.-V. H. (2018). First Data Release of the COSMOS Lyman-Alpha Mapping And Tomography Observations: 3D Lyman- α Forest Tomography at $2.05 < z < 2.55$. *The Astronomical Journal Supplement Series*, 238(2).
- Lehmer, B. D., Alexander, D. M., Geach, J. E., Smail, I., Basu-Zych, A., Bauer, F. E., Chapman, S. C., Matsuda, Y., Scharf, C. A., Volonteri, M., and Yamada, T. (2009). The Chandra Deep Protocluster Survey: Evidence for an Enhancement in the AGN Activity in the SSA22 Protocluster at $z = 3.09$. *The Astrophysical Journal*, 691(1):687–695.
- Lemaux, B. C., Cucciati, O., Tasca, L. A. M., Le Fèvre, O., Zamorani, G., Cassata, P., Garilli, B., Le Brun, V., Maccagni, D., Pentericci, L., Thomas, R., Vanzella, E., Zucca, E., Amorín, R., Bardelli, S., Capak, P., Cassarà, L. P., Castellano, M., Cimatti, A., Cuby, J. G., de la Torre, S., Durkalec, A., Fontana, A., Giavalisco, M., Grazian, A., Hathi, N. P., Ilbert, O., Moreau, C., Paltani, S., Ribeiro, B., Salvato, M., Schaerer, D., Scodreggio, M., Sommariva, V., Talia, M., Taniguchi, Y., Tresse, L., Vergani, D., Wang, P. W., Charlot, S., Contini, T., Fotopoulou, S., Gal, R. R., Kocevski, D. D., López-Sanjuan, C., Lubin, L. M., Mellier, Y., Sadibekova, T., and Scoville, N. (2014). VIMOS Ultra-Deep Survey (VUDS): Witnessing the assembly of a massive cluster at $z \sim 3.3$. *Astronomy & Astrophysics*, 572:A41.
- Lemaux, B. C., Fèvre, O. L., Cucciati, O., Ribeiro, B., Tasca, L. A. M., Zamorani, G., Ilbert, O., Thomas, R., Bardelli, S., Cassata, P., Hathi, N. P., Pforr, J., Smolčić, V., Delvecchio, I., Novak, M., Berta, S., McCracken, H. J., Koekemoer, A., Amorín, R., Garilli, B., Maccagni, D., Schaerer, D., and Zucca, E. (2017). The VIMOS Ultra Deep Survey: Emerging from the Dark, a Massive Proto-Cluster at $z \sim 4.57$. *Astronomy & Astrophysics*, 615:77.
- Lewis, I., Balogh, M., De Propris, R., Couch, W., Bower, R., Offer, A., Bland-Hawthorn, J., Baldry, I. K., Baugh, C., Bridges, T., Cannon, R., Cole, S.,

- Colless, M., Collins, C., Cross, N., Dalton, G., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Hawkins, E., Jackson, C., Lahav, O., Lumsden, S., Maddox, S., Madgwick, D., Norberg, P., Peacock, J. A., Percival, W., Peterson, B. A., Sutherland, W., and Taylor, K. (2002). The 2dF galaxy redshift survey: The environmental dependence of galaxy star formation rates near clusters. *Monthly Notices of the Royal Astronomical Society*, 334(3):673–683.
- Li, Z., Horowitz, B., and Cai, Z. (2021). Improved $\text{Ly}\alpha$ tomography using optimized reconstruction with constraints on absorption (ORCA). *The Astrophysical Journal*, 916(1):20.
- Liang, Y., Kashikawa, N., Cai, Z., Fan, X., Prochaska, J. X., Shimasaku, K., Tanaka, M., Uchiyama, H., Ito, K., Shimakawa, R., Nagamine, K., Shimizu, I., Onoue, M., and Toshikawa, J. (2021). Statistical Correlation between the Distribution of $\text{Ly}\alpha$ Emitters and Intergalactic Medium HI at $z \sim 2.2$ Mapped by the Subaru/Hyper Suprime-Cam. *The Astrophysical Journal*, 907(1):3.
- Linde, A. D. (1982). A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108(6):389–393.
- Loeb, A. and Barkana, R. (2001). The Reionization of the Universe by the First Stars and Quasars. *Annu. Rev. Astron. Astrophys.*
- Lovell, C. C., Thomas, P. A., and Wilkins, S. M. (2018). Characterising and identifying galaxy protoclusters. *Monthly Notices of the Royal Astronomical Society*, 474:4612–4628.
- Lukić, Z., Stark, C. W., Nugent, P., White, M., Meiksin, A. A., and Almgren, A. (2014). The Lyman α forest in optically thin hydrodynamical simulations. *Monthly Notices of the Royal Astronomical Society*, 446(4):3697–3724.
- Lynds, C. R. and Stockton, A. N. (1966). The Large Redshift of the Quasi-Stellar Source 1116+12. *The Astrophysical Journal*, 144:446.
- Lynds, R. and Wills, D. (1970). The Unusually large redshift of 4C 05.34. *Nature*, 226(5245):532.
- Madau, P. and Dickinson, M. (2014). Cosmic Star Formation History. *Annual Review of Astronomy and Astrophysics*, 52:415–486.

- Marinacci, F., Vogelsberger, M., Pakmor, R., Torrey, P., Springel, V., Hernquist, L., Nelson, D., Weinberger, R., Pillepich, A., Naiman, J., and Genel, S. (2018). First results from the IllustrisTNG simulations: radio haloes and magnetic fields. *Monthly Notices of the Royal Astronomical Society*, 480(4):5113–5139.
- Matsuda, Y., Yamada, T., Hayashino, T., Tamura, H., Yamauchi, R., Ajiki, M., Fujita, S. S., Murayama, T., Nagao, T., Ohta, K., Okamura, S., Ouchi, M., Shimasaku, K., Shioya, Y., and Taniguchi, Y. (2004). A Subaru Search for Ly-alpha Blobs in and around the Proto-cluster Region at Redshift $z = 3.1$. *The Astronomical Journal*, 128(18):569–584.
- Mazzucchelli, C., Bañados, E., Decarli, R., Farina, E. P., Venemans, B. P., Walter, F., and Overzier, R. (2017). No Overdensity of Lyman-Alpha Emitting Galaxies Around a Quasar at $z \sim 5.7$. *The Astrophysical Journal*, 834(1):83.
- McAlpine, S., Helly, J. C., Schaller, M., Trayford, J. W., Qu, Y., Furlong, M., Bower, R. G., Crain, R. A., Schaye, J., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Jenkins, A., Rosas-Guevara, Y., White, S. D., Baes, M., Camps, P., and Lemson, G. (2016). The EAGLE simulations of galaxy formation: Public release of halo and galaxy catalogues. *Astronomy and Computing*, 15:72–89.
- McCarthy, I. G., Schaye, J., Bird, S., and Le Brun, A. M. (2017). The BAHAMAS project: Calibrated hydrodynamical simulations for large-scale structure cosmology. *Monthly Notices of the Royal Astronomical Society*, 465(3):2936–2965.
- McConnachie, A., Babusiaux, C., Balogh, M., Driver, S., Côté, P., Courtois, H., Davies, L., Ferrarese, L., Gallagher, S., Ibata, R., Martin, N., Robotham, A., Venn, K., Villaver, E., Bovy, J., Boselli, A., Colless, M., Comparat, J., Denny, K., Duc, P.-A., Ellison, S., de Grijs, R., Fernandez-Lorenzo, M., Freeman, K., Guhathakurta, R., Hall, P., Hopkins, A., Hudson, M., Johnson, A., Kaiser, N., Koda, J., Konstantopoulos, I., Koshy, G., Lee, K.-G., Nusser, A., Pancoast, A., Peng, E., Peroux, C., Petitjean, P., Pichon, C., Poggianti, B., Schmid, C., Shastri, P., Shen, Y., Willot, C., Croom, S., Lallement, R., Schimd, C., Smith, D., Walker, M., Willis, J., Colless, A. B. M., Goswami, A., Jarvis, M., Jullo, E., Kneib, J.-P., Konstantopoloulous, I., Newman, J., Richard, J., Sutaria, F., Taylor, E., van Waerbeke, L., Battaglia, G., Hall, P., Haywood, M., Sakari, C., Schmid, C., Seibert, A., Thirupathi, S., Wang, Y., Wang, Y., Babas, F., Bauman, S., Caffau, E., Laychak, M. B.,

Crampton, D., Devost, D., Flagey, N., Han, Z., Higgs, C., Hill, V., Ho, K., Isani, S., Mignot, S., Murowinski, R., Pandey, G., Salmon, D., Siebert, A., Simons, D., Starkenburg, E., Szeto, K., Tully, B., Vermeulen, T., Withington, K., Arimoto, N., Asplund, M., Aussel, H., Bannister, M., Bhatt, H., Bhargavi, S., Blakeslee, J., Bland-Hawthorn, J., Bullock, J., Burgarella, D., Chang, T.-C., Cole, A., Cooke, J., Cooper, A., Di Matteo, P., Favole, G., Flores, H., Gaensler, B., Garnavich, P., Gilbert, K., Gonzalez-Delgado, R., Guhathakurta, P., Hasinger, G., Herwig, F., Hwang, N., Jablonka, P., Jarvis, M., Kamath, U., Kewley, L., Borgne, D. L., Lewis, G., Lupton, R., Martell, S., Mateo, M., Mena, O., Nataf, D., Newman, J., Pérez, E., Prada, F., Puech, M., Recio-Blanco, A., Robin, A., Saunders, W., Smith, D., Stalin, C. S., Tao, C., Thanjuvur, K., Tresse, L., van Waerbeke, L., Wang, J.-M., Yong, D., Zhao, G., Boisse, P., Bolton, J., Bonifacio, P., Bouchy, F., Cowie, L., Cunha, K., Deleuil, M., de Mooij, E., Dufour, P., Foucaud, S., Glazebrook, K., Hutchings, J., Kobayashi, C., Kudritzki, R.-P., Li, Y.-S., Lin, L., Lin, Y.-T., Makler, M., Narita, N., Park, C., Ransom, R., Ravindranath, S., Reddy, B. E., Sawicki, M., Simard, L., Srianand, R., Storchi-Bergmann, T., Umetsu, K., Wang, T.-G., Woo, J.-H., and Wu, X.-B. (2016). The Detailed Science Case for the Maunakea Spectroscopic Explorer: the Composition and Dynamics of the Faint Universe. *arXiv:1606.00043*.

McQuinn, M. and Switzer, E. R. (2010). The He i 584 Å forest as a diagnostic of helium reionization. *Monthly Notices of the Royal Astronomical Society*, 408(3):1945–1955.

Meiksin, A., Bolton, J. S., and Puchwein, E. (2017). Gas around galaxy haloes - III: Hydrogen absorption signatures around galaxies and QSOs in the Sherwood simulation suite. *Monthly Notices of the Royal Astronomical Society*, 468(2):1893–1901.

Miller, J. S., Bolton, J. S., and Hatch, N. (2019). Searching for the shadows of giants: Characterizing protoclusters with line of sight Lyman- α absorption. *Monthly Notices of the Royal Astronomical Society*, 489(4):5381–5397.

Momose, R., Shimasaku, K., Kashikawa, N., Nagamine, K., Shimizu, I., Nakajima, K., Terao, Y., Kusakabe, H., Ando, M., Motohara, K., and Spitler, L. (2021). Environmental dependence of galactic properties traced by Ly α forest absorption (II): Diversity among galaxy populations. *The Astrophysical Journal*, 909(2):117.

Monaghan, J. J. and Lattanzio, J. C. (1985). A refined particle method for astrophysical problems. *Astronomy & Astrophysics*, 149:135–143.

- Morales, M. F., Stuart, J., and Wyithe, B. (2010). Reionization and Cosmology with 21-cm Fluctuations. *Annu. Rev. Astron. Astrophys*, 48:127–171.
- Morselli, L., Mignoli, M., Gilli, R., Vignali, C., Comastri, A., Sani, E., Cappelluti, N., Zamorani, G., Brusa, M., Gallozzi, S., and Vanzella, E. (2014). Primordial environment of super massive black holes: large-scale galaxy overdensities around $z \sim 6$ quasars with LBT. *Astronomy & Astrophysics*, 568:A1.
- Mortlock, D. J., Warren, S. J., Venemans, B. P., Patel, M., Hewett, P. C., McMahon, R. G., Simpson, C., Theuns, T., González-Solares, E. A., Adamson, A., Dye, S., Hambly, N. C., Hirst, P., Irwin, M. J., Kuiper, E., Lawrence, A., and Röttgering, H. J. (2011). A luminous quasar at a redshift of $z = 7.085$. *Nature*, 474(7353):616–619.
- Mukae, S., Ouchi, M., Cai, Z., Lee, K.-G., Prochaska, J. X., Cantalupo, S., Ono, Y., Zheng, Z., Nagamine, K., Suzuki, N., Silverman, J. D., Misawa, T., Inoue, A. K., Hennawi, J. F., Matsuda, Y., Mawatari, K., Sugahara, Y., Kojima, T., Shibuya, T., Harikane, Y., Fujimoto, S., Chiang, Y.-K., Zhang, H., and Kakuma, R. (2020a). 3D Distribution Map of HI Gas and Galaxies around an Enormous Ly α Nebula and Three QSOs at $z = 2.3$ Revealed by the HI Tomographic Mapping Technique. *The Astrophysical Journal*, 896(1).
- Mukae, S., Ouchi, M., Hill, G. J., Gebhardt, K., Cooper, E. M., Jeong, D., Saito, S., Fabricius, M., Gawiser, E., Ciardullo, R., Farrow, D., Davis, D., Zeimann, G., Finkelstein, S. L., Gronwall, C., Liu, C., Zhang, Y., Byrohl, C., Ono, Y., Schneider, D. P., Jarvis, M. J., Casey, C. M., and Mawatari, K. (2020b). Cosmological 3D hi gas map with HETDEX Ly α emitters and eBOSS QSOs at $z = 2$: IGM-Galaxy/QSO connection and a ~ 40 -Mpc scale giant H II bubble candidate. *The Astrophysical Journal*, 903(1):24.
- Mukae, S., Ouchi, M., Kakiichi, K., Suzuki, N., Ono, Y., Cai, Z., Inoue, A. K., Chiang, Y.-K., Shibuya, T., and Matsuda, Y. (2017). Cosmic Galaxy-IGM HI Relation at $z \sim 2-3$ Probed in the COSMOS/UltraVISTA $1.6^\circ \times 2^\circ$ Field. *The Astrophysical Journal*, 835(2):281.
- Mukhanov, V. F. and Chibisov, G. V. (1981). Quantum fluctuations and a non-singular universe. *ZhPmR*, 33:549–553.
- Muldrew, S. I., Hatch, N. A., and Cooke, E. A. (2015). What are protoclusters? – Defining high-redshift galaxy clusters and protoclusters. *Monthly Notices of the Royal Astronomical Society*, 452(3):2528–2539.

- Murdoch, H. S., Hunstead, R. W., Pettini, M., and Blades, J. C. (1986). Absorption spectrum of the $z = 3.78$ QSO 2000-330. II - The redshift and equivalent width distributions of primordial hydrogen clouds. *The Astrophysical Journal*, 309:19.
- Nagamine, K., Shimizu, I., Fujita, K., Suzuki, N., Lee, K.-G., Momose, R., Mukae, S., Liang, Y., Kashikawa, N., Ouchi, M., and Silverman, J. D. (2021). Probing Feedback via IGM tomography and the Ly α Forest with Subaru PFS, TMT/ELT, and JWST. *The Astrophysical Journal*, 914(1):66.
- Naiman, J. P., Pillepich, A., Springel, V., Ramirez-Ruiz, E., Torrey, P., Vogelsberger, M., Pakmor, R., Nelson, D., Marinacci, F., Hernquist, L., Weinberger, R., and Genel, S. (2018). First results from the IllustrisTNG simulations: A tale of two elements - Chemical evolution of magnesium and europium. *Monthly Notices of the Royal Astronomical Society*, 477(1):1206–1224.
- Nelson, D., Pillepich, A., Genel, S., Vogelsberger, M., Springel, V., Torrey, P., Rodriguez-Gomez, V., Sijacki, D., Snyder, G. F., Griffen, B., Marinacci, F., Blecha, L., Sales, L., Xu, D., and Hernquist, L. (2015). The Illustris simulation: Public data release. *Astronomy and Computing*, 13:12–37.
- Nelson, D., Pillepich, A., Springel, V., Weinberger, R., Hernquist, L., Pakmor, R., Genel, S., Torrey, P., Vogelsberger, M., Kauffmann, G., Marinacci, F., and Naiman, J. (2018). First results from the IllustrisTNG simulations: the galaxy colour bimodality. *Monthly Notices of the Royal Astronomical Society*, 475(1):624–647.
- Nelson, D., Springel, V., Pillepich, A., Rodriguez-Gomez, V., Torrey, P., Genel, S., Vogelsberger, M., Pakmor, R., Marinacci, F., Weinberger, R., Kelley, L., Lovell, M., Diemer, B., and Hernquist, L. (2019). The IllustrisTNG simulations: Public data release. *Computational Astrophysics and Cosmology*, 6:2.
- Newman, A., Bezanson, R., Johnson, S., Rudie, G., Greene, J., and Hummels, C. (2019). Resolving Galaxy Formation at Cosmic Noon. *BAAS*, 51(3):145.
- Newman, A. B., Rudie, G. C., Blanc, G. A., Kelson, D. D., Rhoades, S., Hare, T., Pérez, V., Benson, A. J., Dressler, A., Gonzalez, V., Kollmeier, J. A., Konidaris, N. P., Mulchaey, J. S., Rauch, M., Fèvre, O. L., Lemaux, B. C., Cucciati, O., and Lilly, S. J. (2020). LATIS: The Ly α Tomography IMACS Survey. *The Astrophysical Journal*, 891(2):147.

- Noterdaeme, P., Petitjean, P., Carithers, W. C., Pâris, I., Font-Ribera, A., Bailey, S., Aubourg, E., Bizyaev, D., Ebelke, G., Finley, H., Ge, J., Malanushenko, E., Malanushenko, V., Miralda-Escudé, J., Myers, A. D., Oravetz, D., Pan, K., Pieri, M. M., Ross, N. P., Schneider, D. P., Simmons, A., and York, D. G. (2012). Column density distribution and cosmological mass density of neutral gas: Sloan Digital Sky Survey-III Data Release 9. *Astronomy & Astrophysics*, 547:L1.
- Oppenheimer, B. D. and Schaye, J. (2013). AGN proximity zone fossils and the delayed recombination of metal lines. *Monthly Notices of the Royal Astronomical Society*, 434(2):1063–1078.
- Orsi, Á. A., Fanidakis, N., Lacey, C. G., and Baugh, C. M. (2016). The environments of high-redshift radio galaxies and quasars: probes of protoclusters. *Monthly Notices of the Royal Astronomical Society*, 456(4):3827–3839.
- Overzier, R. A. (2016). The realm of the galaxy protoclusters. *The Astronomy and Astrophysics Review*, 24(1):14.
- Pakmor, R., Bauer, A., and Springel, V. (2011). Magnetohydrodynamics on an unstructured moving grid. *Monthly Notices of the Royal Astronomical Society*, 418(2):1392–1401.
- Pâris, I., Petitjean, P., Aubourg, É., Ross, N. P., Myers, A. D., Streblyanska, A., Bailey, S., Hall, P. B., Strauss, M. A., Anderson, S. F., Bizyaev, D., Borde, A., Brinkmann, J., Bovy, J., Brandt, W. N., Brewington, H., Brownstein, J. R., Cook, B. A., Ebelke, G., Fan, X., Ak, N. F., Finley, H., Font-Ribera, A., Ge, J., Hamann, F., Ho, S., Jiang, L., Kinemuchi, K., Malanushenko, E., Malanushenko, V., Marchante, M., McGreer, I. D., McMahon, R. G., Miralda-Escudé, J., Muna, D., Noterdaeme, P., Oravetz, D., Palanque-Delabrouille, N., Pan, K., Perez-Fournon, I., Pieri, M., Riffel, R., Schlegel, D. J., Schneider, D. P., Simmons, A., Viel, M., Weaver, B. A., Wood-Vasey, W. M., Yèche, C., and York, D. G. (2014). The Sloan Digital Sky Survey quasar catalog: tenth data release. *Astronomy & Astrophysics*, 563:54.
- Peebles, P. J. E. and Yu, J. T. (1970). Primeval Adiabatic Perturbation in an Expanding Universe. *The Astrophysical Journal*, 162:815.
- Peirani, S., Weinberg, D. H., Colombi, S., Blaizot, J., Dubois, Y., and Pichon, C. (2014). LyMAS: Predicting Large-Scale Ly α Forest Statistics From the Dark Matter Density Field. *The Astrophysical Journal*, 784(1):11.

- Pentericci, L., Kurk, J. D., Carilli, C. L., Harris, D. E., Miley, G. K., and Röttgering, H. J. (2002). A Chandra study of x-ray sources in the field of the $z = 2.16$ radio galaxy MRC 1138-262. *Astronomy & Astrophysics*, 396(1):109–115.
- Penzias, A. A. and Wilson, R. W. (1965). A Measurement of Excess Antenna Temperature at 4080 Mc/s. *The Astrophysical Journal*, 142:419.
- Pichon, C., Vergely, J. L., Rollinde, E., Colombi, S., and Petitjean, P. (2001). Inversion of the Lyman α forest: Three-dimensional investigation of the intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 326(2):597–620.
- Pieri, M. M., Bonoli, S., Chaves-Montero, J., Pâris, I., Fumagalli, M., Bolton, J. S., Viel, M., Noterdaeme, P., Miralda-Escudé, J., Busca, N. G., Rahmani, H., Peroux, C., Font-Ribera, A., and Trager, S. C. (2016). WEAVE-QSO: A Massive Intergalactic Medium Survey for the William Herschel Telescope. *arXiv:1611.09388*.
- Pillepich, A., Nelson, D., Hernquist, L., Springel, V., Pakmor, R., Torrey, P., Weinberger, R., Genel, S., Naiman, J. P., Marinacci, F., and Vogelsberger, M. (2018). First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 475(1):648–675.
- Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Battaner, E., Benabed, K., Benoît, A., Benoit-Lévy, A., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bobin, J., Bock, J. J., Bonaldi, A., Bond, J. R., Borrill, J., Bouchet, F. R., Bridges, M., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cappellini, B., Cardoso, J. F., Catalano, A., Challinor, A., Chamballu, A., Chary, R. R., Chen, X., Chiang, H. C., Chiang, L. Y., Christensen, P. R., Church, S., Clements, D. L., Colombi, S., Colombo, L. P. L., Couchot, F., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Danese, L., Davies, R. D., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Delouis, J. M., Désert, F. X., Dickinson, C., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Doré, O., Douspis, M., Dunkley, J., Dupac, X., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Finelli, F., Forni, O., Frailis, M., Fraisse, A. A., Franceschi, E., Gaier, T. C., Galeotta, S., Galli, S., Ganga, K., Giard, M., Giardino, G., Giraud-Héraud, Y., Gjerløw, E., González-Nuevo, J., Górski, K. M.,

Gratton, S., Gregorio, A., Gruppuso, A., Gudmundsson, J. E., Haissinski, J., Hamann, J., Hansen, F. K., Hanson, D., Harrison, D., Henrot-Versillé, S., Hernández-Monteagudo, C., Herranz, D., Hildebrandt, S. R., Hivon, E., Hobson, M., Holmes, W. A., Hornstrup, A., Hou, Z., Hovest, W., Hufferberger, K. M., Jaffe, A. H., Jaffe, T. R., Jewell, J., Jones, W. C., Juvela, M., Keihänen, E., Keskitalo, R., Kisner, T. S., Kneissl, R., Knoche, J., Knox, L., Kunz, M., Kurki-Suonio, H., Lagache, G., Lähteenmäki, A., Lamarre, J. M., Lasenby, A., Lattanzi, M., Laureijs, R. J., Lawrence, C. R., Leach, S., Leahy, J. P., Leonardi, R., León-Tavares, J., Lesgourgues, J., Lewis, A., Liguori, M., Lilje, P. B., Linden-Vørnle, M., López-Cañiego, M., Lubin, P. M., Macías-Pérez, J. F., Maffei, B., Maino, D., Mandolesi, N., Maris, M., Marshall, D. J., Martin, P. G., Martínez-González, E., Masi, S., Massardi, M., Matarrese, S., Matthai, F., Mazzotta, P., Meinhold, P. R., Melchiorri, A., Melin, J. B., Mendes, L., Menegoni, E., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M. A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nati, F., Natoli, P., Netterfield, C. B., Nørgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., O'Dwyer, I. J., Osborne, S., Oxborrow, C. A., Paci, F., Pagano, L., Pajot, F., Paoletti, D., Partridge, B., Pasian, F., Patanchon, G., Pearson, D., Pearson, T. J., Peiris, H. V., Perdureau, O., Perotto, L., Perrotta, F., Pettorino, V., Piacentini, F., Piat, M., Pierpaoli, E., Pietrobon, D., Plaszczynski, S., Platania, P., Pointecouteau, E., Polenta, G., Ponthieu, N., Popa, L., Poutanen, T., Pratt, G. W., Prézeau, G., Prunet, S., Puget, J. L., Rachen, J. P., Reach, W. T., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Ricciardi, S., Riller, T., Ristorcelli, I., Rocha, G., Rosset, C., Roudier, G., Rowan-Robinson, M., Rubiño-Martín, J. A., Rusholme, B., Sandri, M., Santos, D., Savelainen, M., Savini, G., Scott, D., Seiffert, M. D., Shellard, E. P. S., Spencer, L. D., Starck, J. L., Stolyarov, V., Stompor, R., Sudiwala, R., Sunyaev, R., Sureau, F., Sutton, D., Suur-Uski, A. S., Sygnet, J. F., Tauber, J. A., Tavagnacco, D., Terenzi, L., Toffolatti, L., Tomasi, M., Tristram, M., Tucci, M., Tuovinen, J., Türler, M., Umana, G., Valenziano, L., Valiviita, J., Van Tent, B., Vielva, P., Villa, F., Vittorio, N., Wade, L. A., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Wilkinson, A., Yvon, D., Zacchei, A., and Zonca, A. (2014). Planck 2013 results. XVI. Cosmological parameters. *Astronomy & Astrophysics*, 571:A16.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Bartolo, N., Battaner, E., Battye, R., Benabed, K., Benoit, A., Benoit-Levy,

A., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bonaldi, A., Bonavera, L., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Catalano, A., Challinor, A., Chamballu, A., Chary, R. R., Chiang, H. C., Chluba, J., Christensen, P. R., Church, S., Clements, D. L., Colombi, S., Colombo, L. P. L., Combet, C., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Danese, L., Davies, R. D., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Desert, F. X., Di Valentino, E., Dickinson, C., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Dore, O., Douspis, M., Ducout, A., Dunkley, J., Dupac, X., Efstathiou, G., Elsner, F., Ensslin, T. A., Eriksen, H. K., Farhang, M., Fergusson, J., Finelli, F., Forni, O., Frailis, M., Fraisse, A. A., Franceschi, E., Frejsel, A., Galeotta, S., Galli, S., Ganga, K., Gauthier, C., Gerbino, M., Ghosh, T., Giard, M., Giraud-Heraud, Y., Giusarma, E., Gjerlow, E., Gonzalez-Nuevo, J., Gorski, K. M., Gratton, S., Gregorio, A., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Hansen, F. K., Hanson, D., Harrison, D. L., Helou, G., Henrot-Versille, S., Hernandez-Monteagudo, C., Herranz, D., Hildebrandt, S. R., Hivon, E., Hobson, M., Holmes, W. A., Hornstrup, A., Hovest, W., Huang, Z., Huppenberger, K. M., Hurier, G., Jaffe, A. H., Jaffe, T. R., Jones, W. C., Juvela, M., Keihanen, E., Keskitalo, R., Kisner, T. S., Kneissl, R., Knoche, J., Knox, L., Kunz, M., Kurki-Suonio, H., Lagache, G., Lahteenmaki, A., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Leahy, J. P., Leonardi, R., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Linden-Vornle, M., Lopez-Caniego, M., Lubin, P. M., Macias-Perez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marchini, A., Martin, P. G., Martinelli, M., Martinez-Gonzalez, E., Masi, S., Matarrese, S., Mazzotta, P., McGehee, P., Meinhold, P. R., Melchiorri, A., Melin, J. B., Mendes, L., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschenes, M. A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nati, F., Natoli, P., Netterfield, C. B., Norgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., Oxborrow, C. A., Paci, F., Pagano, L., Pajot, F., Paladini, R., Paoletti, D., Partridge, B., Pasian, F., Patanchon, G., Pearson, T. J., Perdureau, O., Perotto, L., Perrotta, F., Pettorino, V., Piacentini, F., Piat, M., Pierpaoli, E., Pietrobon, D., Plaszczynski, S., Pointecouteau, E., Polenta, G., Popa, L., Pratt, G. W., Prezeau, G., Prunet, S., Puget, J. L., Rachen, J. P., Reach, W. T., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Renzi, A., Ristorcelli, I., Rocha, G., Rosset, C., Rossetti, M., Roudier, G., D'Orfeuil, B. R., Rowan-Robinson, M., Rubino-Martin, J. A., Rusholme, B., Said, N., Salvatelli, V., Salvati, L., Sandri, M., Santos, D.,

Savelainen, M., Savini, G., Scott, D., Seiffert, M. D., Serra, P., Shellard, E. P. S., Spencer, L. D., Spinelli, M., Stolyarov, V., Stompor, R., Sudiwala, R., Sunyaev, R., Sutton, D., Suur-Uski, A. S., Sygnet, J. F., Tauber, J. A., Terenzi, L., Toffolatti, L., Tomasi, M., Tristram, M., Trombetti, T., Tucci, M., Tuovinen, J., Turler, M., Umana, G., Valenziano, L., Valiviita, J., Van Tent, B., Vielva, P., Villa, F., Wade, L. A., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Wilkinson, A., Yvon, D., Zacchei, A., and Zonca, A. (2015). Planck 2015 results. XIII. Cosmological parameters. *Astronomy & Astrophysics*, 594:A13.

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Carron, J., Challinor, A., Chiang, H. C., Chluba, J., Colombo, L. P., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., De Bernardis, P., De Zotti, G., Delabrouille, J., Delouis, J. M., Di Valentino, E., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Dusini, S., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Fantaye, Y., Farhang, M., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Hansen, F. K., Herranz, D., Hildebrandt, S. R., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karakci, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Lemos, P., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Lilley, M., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y. Z., Maciás-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M. A., Molinari, D., Montier, L., Morgante, G., Moss, A., Natoli, P., Nørgaard-Nielsen, H. U., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Peiris, H. V., Perrotta, F., Pettorino, V., Piacentini, F., Polastri, L., Polenta, G., Puget, J. L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rosset, C., Roudier, G., Rubiño-Martín, J. A., Ruiz-

- Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Shellard, E. P., Sirignano, C., Sirri, G., Spencer, L. D., Sunyaev, R., Suur-Uski, A. S., Tauber, J. A., Tavagnacco, D., Tenti, M., Toffolatti, L., Tomasi, M., Trombetti, T., Valenziano, L., Valiviita, J., Van Tent, B., Vibert, L., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., White, S. D., Zacchei, A., and Zonca, A. (2020). Planck 2018 results: VI. Cosmological parameters. *Astronomy & Astrophysics*, 641:A6.
- Pontzen, A., Governato, F., Pettini, M., Booth, C. M., Stinson, G., Wadsley, J., Brooks, A., Quinn, T., and Haehnelt, M. (2008). Damped Lyman α systems in galaxy formation simulations. *Monthly Notices of the Royal Astronomical Society*, 390(4):1349–1371.
- Porqueres, N., Hahn, O., Jasche, J., and Lavaux, G. (2020). A hierarchical field-level inference approach to reconstruction from sparse Lyman- α forest data. *Astronomy & Astrophysics*, 642:A139.
- Pospelov, M. and Pradler, J. (2010). Big Bang Nucleosynthesis as a Probe of New Physics. *Annual Review of Nuclear and Particle Science*, 60(1):539–568.
- Potter, D., Stadel, J., and Teyssier, R. (2017). PKDGRAV3: beyond trillion particle cosmological simulations for the next era of galaxy surveys. *Computational Astrophysics and Cosmology*, 4(1):2.
- Prochaska, J. X. (1999). The Physical Nature of the Lyman-Limit Systems. *The Astrophysical Journal*, 511(2):L71–L74.
- Prochaska, J. X., Hennawi, J. F., Lee, K. G., Cantalupo, S., Bovy, J., Djorgovski, S. G., Ellison, S. L., Lau, M. W., Martin, C. L., Myers, A., Rubin, K. H., and Simcoe, R. A. (2013). Quasars probing quasars. VI. Excess HI absorption within one proper Mpc of $z \sim 2$ quasars. *The Astrophysical Journal*, 776(2):136.
- Prochaska, J. X. and Wolfe, A. M. (2009). On the (non)evolution of HI gas in galaxies over cosmic time. *The Astrophysical Journal*, 696(2):1543–1547.
- Puchwein, E. and Springel, V. (2013). Shaping the galaxy stellar mass function with supernova- and AGN-driven winds. *Monthly Notices of the Royal Astronomical Society*, 428:2966–2979.
- Rahmati, A., Pawlik, A. H., Raičević, M., and Schaye, J. (2013). On the evolution of the H I column density distribution in cosmological simulations. *Monthly Notices of the Royal Astronomical Society*, 430(3):2427–2445.

- Rahmati, A., Schaye, J., Bower, R. G., Crain, R. A., Furlong, M., Schaller, M., and Theuns, T. (2015). The distribution of neutral hydrogen around high-redshift galaxies and quasars in the EAGLE simulation. *Monthly Notices of the Royal Astronomical Society*, 452(2):2034–2056.
- Rauch, M. (1998). The Lyman Alpha Forest in the Spectra of Quasistellar Objects. *Annual Review of Astronomy and Astrophysics*, 36(1):267–316.
- Rauch, M., Haehnelt, M., Bunker, A., Becker, G., Marleau, F., Graham, J., Cristiani, S., Jarvis, M., Lacey, C., Morris, S., Peroux, C., Röttgering, H., and Theuns, T. (2008). A population of faint extended line emitters and the host galaxies of optically thick QSO absorption systems. *The Astrophysical Journal*, 681(2):856–880.
- Rauch, M., Miralda-Escude, J., Sargent, W. L. W., Barlow, T. A., Weinberg, D. H., Hernquist, L., Katz, N., Cen, R., and Ostriker, J. P. (1997). The Opacity of the Ly α Forest and Implications for Ω_b and the Ionizing Background. *The Astrophysical Journal*, 489(1):7–20.
- Ravoux, C., Armengaud, E., Walther, M., Etourneau, T., Pomaredé, D., Palanque-Delabrouille, N., Yèche, C., Bautista, J., du Mas des Bourboux, H., Chabanier, S., Dawson, K., Le Goff, J.-M., Lyke, B., Myers, A., Petitjean, P., Pieri, M., Rich, J., Rossi, G., and Schneider, D. (2020). A tomographic map of the large-scale matter distribution using the eBOSS-Stripe 82 Ly α forest. Recent citations TARDIS. II. Synergistic Density Reconstruction from Ly Forest and Spectroscopic Galaxy Surveys with Applications to Protoclusters and the Cosmic. *Journal of Cosmology and Astroparticle Physics*.
- Rees, M. J., Setti, G., Rees, M. J., and Setti, G. (1970). Absorption and Scattering of Ultraviolet and X Ray Photons by Intergalactic Gas. *Astronomy & Astrophysics*, 8:410.
- Reverdy, V., Alimi, J. M., Bouillot, V., Rasera, Y., Corasaniti, P. S., Balmès, I., Requena, S., Delaruelle, X., and Richet, J. N. (2015). DEUS full observable universe simulations: Numerical challenge and outlooks. *International Journal of High Performance Computing Applications*, 29(3):249–260.
- Ricotti, M., Gnedin, N. Y., and Shull, J. M. (1999). The Evolution of the Effective Equation of State of the IGM. *The Astrophysical Journal*, 534(1):41–56.
- Rudie, G. C., Chen, H.-W., Newman, A. B., and Johnson, S. D. (2019). Observing Galaxies and Dissecting their Baryon Cycle at Cosmic Noon. *BAAS*, 51(3):148.

- Salpeter, E. E. (1955). The Luminosity Function and Stellar Evolution. *The Astrophysical Journal*, 121:161.
- Schaye, J. (2001). On the relation between high-redshift starburst galaxies and damped Ly α systems. *The Astrophysical Journal*, 559(1):L1–L4.
- Schaye, J., Crain, R. A., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., Jenkins, A., Rosas-Guevara, Y. M., White, S. D., Baes, M., Booth, C. M., Camps, P., Navarro, J. F., Qu, Y., Rahmati, A., Sawala, T., Thomas, P. A., and Trayford, J. (2015). The EAGLE project: Simulating the evolution and assembly of galaxies and their environments. *Monthly Notices of the Royal Astronomical Society*, 446(1):521–554.
- Schaye, J. and Dalla Vecchia, C. (2008). On the relation between the Schmidt and Kennicutt-Schmidt star formation laws and its implications for numerical simulations. *Monthly Notices of the Royal Astronomical Society*, 383(3):1210–1222.
- Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., and Sargent, W. L. W. (2000). The thermal history of the intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 318(3):817–826.
- Scheuer, P. A. (1965). A Sensitive Test for the Presence of Atomic Hydrogen in Intergalactic Space. *Nature*, 207:963.
- Schmidt, T. M., Hennawi, J. F., Lee, K. G., Lukić, Z., Oñorbe, J., and White, M. (2019). Mapping quasar light echoes in 3d with Ly α forest tomography. *The Astrophysical Journal*, 882(2):165.
- Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C. M., Elvis, M., Gialalisco, M., Guzzo, L., Hasinger, G., Impey, C., Kneib, J., LeFevre, O., Lilly, S. J., Mobasher, B., Renzini, A., Rich, R. M., Sanders, D. B., Schinnerer, E., Schminovich, D., Shopbell, P., Taniguchi, Y., and Tyson, N. D. (2007). The Cosmic Evolution Survey (COSMOS): Overview. *The Astrophysical Journal Supplement Series*, 172(1):1–8.
- Shankar, F., Weinberg, D. H., and Shen, Y. (2010). Constraints on black hole duty cycles and the black hole-halo relation from SDSS quasar clustering. *Monthly Notices of the Royal Astronomical Society*, 406(3):1959–1966.
- Shi, D. D., Cai, Z., Fan, X., Zheng, X. Z., Huang, Y.-H., and Xu, J. (2021). Spectroscopic confirmation of two extremely massive protoclusters, BOSS1244 and BOSS1542, at $z = 2.24$. *The Astrophysical Journal*, 915(1):32.

- Shi, K., Toshikawa, J., Cai, Z., Lee, K.-S., and Fang, T. (2020). A Detailed Study of Massive Galaxies in a Protocluster at $z = 3.13$. *The Astrophysical Journal*, 899(1):79.
- Shimakawa, R., Kodama, T., Shibuya, T., Kashikawa, N., Tanaka, I., Matsuda, Y., Tadaki, K. i., Koyama, Y., Hayashi, M., Suzuki, T. L., and Yamamoto, M. (2017). Similarities and uniqueness of Ly α emitters among star-forming galaxies at $z = 2.5$. *Monthly Notices of the Royal Astronomical Society*, 468(1):1123–1141.
- Shklovskii, I. S. (1965). Physical Conditions in the Gaseous Envelope of 3c-273. *Soviet Astronomy*, 8(5):638–642.
- Simpson, C., Mortlock, D., Warren, S., Cantalupo, S., Hewett, P., McLure, R., McMahon, R., and Venemans, B. (2014). No excess of bright galaxies around the redshift 7.1 quasar ULAS J1120+0641. *Monthly Notices of the Royal Astronomical Society*, 442(4):3454–3461.
- Sobral, D. and Matthee, J. (2019). Predicting Ly α escape fractions with a simple observable. *Astronomy & Astrophysics*, 623:A157.
- Sobral, D., Matthee, J., Best, P., Stroe, A., Röttgering, H., Oteo, I., Smail, I., Morabito, L., and Paulino-Afonso, A. (2016). The CALYMHA survey: Ly α luminosity function and global escape fraction of Ly α photons at $z = 2.23$. *Monthly Notices of the Royal Astronomical Society*, 466(1):1242–1258.
- Somerville, R. S., Gilmore, R. C., Primack, J. R., and Dominguez, A. (2012). Galaxy Properties from the Ultra-violet to the Far-Infrared: Lambda-CDM models confront observations. *Monthly Notices of the Royal Astronomical Society*, 423(3):1992–2015.
- Sorini, D., Davé, R., and Anglés-Alcázar, D. (2020). simba: the average properties of the circumgalactic medium of $2 \approx z \approx 3$ quasars are determined primarily by stellar feedback. *Monthly Notices of the Royal Astronomical Society*, 499(2):2760–2784.
- Sorini, D., Oñorbe, J., Hennawi, J. F., and Lukić, Z. (2018). A Fundamental Test for Galaxy Formation Models: Matching the Lyman- α Absorption Profiles of Galactic Halos Over Three Decades in Distance. *The Astrophysical Journal*, 859:125.
- Sousbie, T. (2011). The persistent cosmic web and its filamentary structure - I. Theory and implementation. *Monthly Notices of the Royal Astronomical Society*.

- Springel, V. (2005). The cosmological simulation code GADGET-2. *Monthly Notices of the Royal Astronomical Society*, 364(4):1105–1134.
- Springel, V. (2010). E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *Monthly Notices of the Royal Astronomical Society*, 401(2):791–851.
- Springel, V. and Hernquist, L. (2003). Cosmological smoothed particle hydrodynamics simulations: A hybrid multiphase model for star formation. *Monthly Notices of the Royal Astronomical Society*, 339(2):289–311.
- Springel, V., Pakmor, R., Pillepich, A., Weinberger, R., Nelson, D., Hernquist, L., Vogelsberger, M., Genel, S., Torrey, P., Marinacci, F., and Naiman, J. (2018). First results from the IllustrisTNG simulations: matter and galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, 475(1):676–698.
- Springel, V., White, S. D., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435(7042):629–636.
- Stark, C. W., White, M., Lee, K. G., and Hennawi, J. F. (2015). Protocluster discovery in tomographic Ly α forest flux maps. *Monthly Notices of the Royal Astronomical Society*, 453(1):311–327.
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., Erb, D. K., Reddy, N. A., and Pettini, M. (2005). Spectroscopic Identification of a Protocluster at $z = 2.300$: Environmental Dependence of Galaxy Properties at High Redshift. *The Astrophysical Journal*, 626(1):44–50.
- Stockton, A. N. and Lynds, C. R. (1966). The Remarkable Absorption Spectrum of 3c 191. *The Astrophysical Journal*, 144:451.
- Strazzullo, V., Pannella, M., Mohr, J. J., Saro, A., Ashby, M. L., Bayliss, M. B., Bocquet, S., Bulbul, E., Khullar, G., Mantz, A. B., Stanford, S. A., Benson, B. A., Bleem, L. E., Brodwin, M., Canning, R. E., Capasso, R., Chiu, I., Gonzalez, A. H., Gupta, N., Hlavacek-Larrondo, J., Klein, M., McDonald, M., Noordeh, E., Rapetti, D., Reichardt, C. L., Schrabback, T., Sharon, K., and Stalder, B. (2019). Galaxy populations in the most distant SPT-SZ clusters: I. Environmental quenching in massive clusters at $1.4 \lesssim z \lesssim 1.7$. *Astronomy & Astrophysics*, 622:117.

- Sunyaev, R. A. and Zeldovich, Y. B. (1970). Small-scale fluctuations of relic radiation. *Astrophysics and Space Science*, 7(1):3–19.
- Sunyaev, R. A., Zeldovich, Y. B., Sunyaev, R. A., and Zeldovich, Y. B. (1972). The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies. *CoASP*, 4:173.
- Takada, M., Ellis, R. S., Chiba, M., Greene, J. E., Aihara, H., Arimoto, N., Bundy, K., Cohen, J., Doré, O., Graves, G., Gunn, J. E., Heckman, T., Hirata, C. M., Ho, P., Kneib, J.-P., Fèvre, O. L., Lin, L., More, S., Murayama, H., Nagao, T., Ouchi, M., Seiffert, M., Silverman, J. D., Sodr e, Laerte, J., Spergel, D. N., Strauss, M. A., Sugai, H., Suto, Y., Takami, H., and Wyse, R. (2014). Extragalactic science, cosmology, and Galactic archaeology with the Subaru Prime Focus Spectrograph. *Publications of the Astronomical Society of Japan*, 66(1). R1.
- Tanaka, M., Goto, T., Okamura, S., Shimasaku, K., and Brinkman, J. (2004). The Environmental Dependence of Galaxy Properties in the Local Universe: Dependence on Luminosity, Local Density, and System Richness. *The Astronomical Journal*, 128(6):2677–2695.
- Tepper-García, T. (2006). Voigt profile fitting to quasar absorption lines: An analytic approximation to the Voigt-Hjerting function. *Monthly Notices of the Royal Astronomical Society*, 369(4):2025–2035.
- Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., and Thomas, P. A. (1998). P3M-SPH simulations of the Ly α forest. *Monthly Notices of the Royal Astronomical Society*, 301(2):478–502.
- Theuns, T., Viel, M., Kay, S., Schaye, J., Carswell, R. F., and Tzanavaris, P. (2002). Galactic winds in the intergalactic medium. *The Astrophysical Journal*, 578(1):L5–L8.
- Toshikawa, J., Uchiyama, H., Kashikawa, N., Ouchi, M., Overzier, R., Ono, Y., Harikane, Y., Ishikawa, S., Kodama, T., Matsuda, Y., Lin, Y. T., Onoue, M., Tanaka, M., Nagao, T., Akiyama, M., Komiyama, Y., Goto, T., and Lee, C. H. (2018). GOLDRUSH. III. A systematic search for protoclusters at $z \sim 4$ based on the $> 100 \text{ deg}^2$ area. *Publications of the Astronomical Society of Japan*, 70(SP1):12–13.
- van deVoort, F., Springel, V., Mandelker, N., van denBosch, F. C., and Pakmor, R. (2018). Cosmological simulations of the circumgalactic medium with 1 kpc

- resolution: enhanced H I column densities. *Monthly Notices of the Royal Astronomical Society: Letters*, 482(1):L85–L89.
- Vargas-Magana, M., Brooks, D. D., Levi, M. M., and Tarle, G. G. (2019). Unravelling the Universe with DESI. *arXiv:1901.01581*.
- Venemans, B. P., Röttgering, H. J. A., Miley, G. K., Van Breugel, W. J. M., De Breuck, C., Kurk, J. D., Pentericci, L., Stanford, S. A., Overzier, R. A., Croft, S., and Ford, H. (2007). I. Characteristics of high redshift protoclusters. *Astronomy & Astrophysics*, 461:823–845.
- Verner, D. A. and Ferland, G. J. (1996). Atomic Data for Astrophysics. I. Radiative Recombination Rates for H-like, He-like, Li-like, and Na-like Ions over a Broad Range of Temperature. *The Astrophysical Journal Supplement Series*, 103:467.
- Verner, D. A., Ferland, G. J., Korista, K. T., and Yakovlev, D. G. (1996). Atomic Data for Astrophysics. II. New Analytic FITS for Photoionization Cross Sections of Atoms and Ions. *The Astrophysical Journal*, 465:487.
- Viel, M., Haehnelt, M. G., and Springel, V. (2004). Inferring the dark matter power spectrum from the Lyman α forest in high-resolution QSO absorption spectra. *Monthly Notices of the Royal Astronomical Society*, 354(3):684–694.
- Viel, M., Schaye, J., and Booth, C. M. (2013). The impact of feedback from galaxy formation on the Lyman α transmitted flux. *Monthly Notices of the Royal Astronomical Society*, 429(2):1734–1736.
- Villaescusa-Navarro, F., Genel, S., Castorina, E., Obuljen, A., Spergel, D. N., Hernquist, L., Nelson, D., Carucci, I. P., Pillepich, A., Marinacci, F., Diemer, B., Vogelsberger, M., Weinberger, R., and Pakmor, R. (2018). Ingredients for 21 cm Intensity Mapping. *The Astrophysical Journal*, 866:135.
- Vogelsberger, M., Genel, S., Sijacki, D., Torrey, P., Springel, V., and Hernquist, L. (2013). A model for cosmological simulations of galaxy formation physics. *Monthly Notices of the Royal Astronomical Society*, 436:3031–3067.
- Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Nelson, D., and Hernquist, L. (2014). Introducing the Illustris project: Simulating the coevolution of dark and visible matter in the universe. *Monthly Notices of the Royal Astronomical Society*, 444(2):1518–1547.
- Vogelsberger, M., Marinacci, F., Torrey, P., and Puchwein, E. (2020). Cosmological simulations of galaxy formation. *Nature Reviews Physics*, 2(1):42–66.

- Wang, M. Y., Croft, R. A., Peter, A. H., Zentner, A. R., and Purcell, C. W. (2013). Lyman- α forest constraints on decaying dark matter. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 88(12):123515.
- Weinberg, D. H., Hernquist, L., Katz, N., Croft, R., and Miralda-Escudé, J. (1997). Hubble Flow Broadening of the Lyman-alpha Forest and its Implications. In *Structure and Evolution of the Intergalactic Medium from QSO Absorption Line System, Proceedings of the 13th IAP Astrophysics Colloquium, 1-5 July 1997*, Paris. Editions Frontieres.
- Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4(1):389–396.
- Weymann, R. J., Carswell, R. F., and Smith, M. G. (1981). Absorption Lines in the Spectra of Quasistellar Objects. *Annual Review of Astronomy and Astrophysics*, 19(1):41–76.
- White, M. (2000). The mass of a halo. *Astronomy & Astrophysics*, 367:27–32.
- White, M., Myers, A. D., Ross, N. P., Schlegel, D. J., Hennawi, J. F., Shen, Y., McGreer, I., Strauss, M. A., Bolton, A. S., Bovy, J., Fan, X., Miralda-Escude, J., Palanque-Delabrouille, N., Paris, I., Petitjean, P., Schneider, D. P., Viel, M., Weinberg, D. H., Yèche, C., Zehavi, I., Pan, K., Snedden, S., Bizyaev, D., Brewington, H., Brinkmann, J., Malanushenko, V., Malanushenko, E., Oravetz, D., Simmons, A., Sheldon, A., and Weaver, B. A. (2012). The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey. *Monthly Notices of the Royal Astronomical Society*, 424(2):933–950.
- White, S. D. M. and Frenk, C. S. (1991). Galaxy formation through hierarchical clustering. *The Astrophysical Journal*, 379:52.
- Wold, M., Armus, L., Neugebauer, G., Jarrett, T. H., and Lehnert, M. D. (2003). Overdensities of Extremely Red Objects in the Fields of High-Redshift Radio-Loud Quasars. *The Astronomical Journal*, 126(4):1776–1786.
- Wolfe, A. M., Gawiser, E., and Prochaska, J. X. (2005). Damped Ly α Systems. *Annual Review of Astronomy and Astrophysics*, 43(1):861–918.
- Worseck, G., Prochaska, J. X., O’meara, J. M., Becker, G. D., Ellison, S. L., Lopez, S., Meiksin, A., Ménard, B., Murphy, M. T., and Fumagalli, M.

(2014). The Giant Gemini GMOS survey of $z_{\text{em}} > 4.4$ quasars-I. Measuring the mean free path across cosmic time. *Monthly Notices of the Royal Astronomical Society*, 445:1745–1760.

Wylezalek, D., Galametz, A., Stern, D., Vernet, J., De Breuck, C., Seymour, N., Brodwin, M., Eisenhardt, P. R., Gonzalez, A. H., Hatch, N., Jarvis, M., Rettura, A., Stanford, S. A., and Stevens, J. A. (2013). Galaxy clusters around radio-loud active galactic nuclei at $1.3 < z < 3.2$ as seen by spitzer. *The Astrophysical Journal*, 769(1):79.

Zheng, W., Overzier, R. A., Bouwens, R. J., White, R. L., Ford, H. C., Benitez, N., Blakeslee, J. P., Bradley, L. D., Jee, M. J., Martel, A. R., Mei, S., Zirm, A. W., Illingworth, G. D., Clampin, M., Hartig, G. F., Ardila, D. R., Bartko, F., Broadhurst, T. J., Brown, R. A., Burrows, C. J., Cheng, E. S., Cross, N. J. G., Demarco, R., Feldman, P. D., Franx, M., Golimowski, D. A., Goto, T., Gronwall, C., Holden, B., Homeier, N., Infante, L., Kimble, R. A., Krist, J. E., Lesser, M. P., Menanteau, F., Meurer, G. R., Miley, G. K., Motta, V., Postman, M., Rosati, P., Sirianni, M., Sparks, W. B., Tran, H. D., and Tsvetanov, Z. I. (2006). An Overdensity of Galaxies near the Most Distant Radio-loud Quasar. *The Astrophysical Journal*, 640(2):574–578.