The Intergalactic Medium near High Redshift Galaxies

The Intergalactic Medium near High Redshift Galaxies

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 7 februari 2012
klokke 13.45 uur

door

Olivera Rakic
geboren te Zrenjanin
in 1981

Promotiecommissie

Promotores:	Prof. dr. J. Schaye
	Prof. dr. C.C. Steidel (Caltech, USA)
	Prof. dr. P.T. de Zeeuw
Overige leden:	Prof. dr. S. Morris (Durham University, UK)
	Prof. dr. M. Franx
	Dr. J. Brinchmann
	Prof. dr. K. Kuijken

TABLE OF CONTENTS

1

Introduction

OVER the past 15 years, the field of extragalactic astronomy has pushed to high redshift and our knowledge of natal galaxies has grown dramatically. Galaxies are now routinely detected at redshifts $z \approx 6-7$ (e.g. Bouwens et al. 2010; Labbé et al. 2010; Oesch et al. 2010; Bouwens et al. 2011; McLure et al. 2011). However, there are fundamental concepts that are still poorly understood. How do galaxies get their gas? How does galactic feedback affect galaxy evolution?

1.1   Intergalactic Medium

Galaxies form out of the intergalactic medium (IGM), and there was a time, before the formation of galaxies, when all the matter in the Universe was in the IGM. Most of what we know about the IGM comes from studies of the absorption spectra of bright objects, such as QSOs. The intervening gas imprints its composition in their spectra through selective absorption of the continuum radiation. The most prominent pattern in such spectra is the Lyman-α forest of the most abundant element in the Universe, hydrogen, blueward of the QSO's Lyα emission line. It consists of a number of absorption lines, produced by neutral hydrogen clouds along the line of sight (LOS) to the QSO, where they absorb whatever continuum radiation was redshifted to the wavelength of H I Lyα (1216 Å).

The "modern" understanding of the Lyα forest dates to the early 90s, when H. Bi and collaborators realized that the linear density fluctuations in the IGM yield a realistic representation of the forest (Bi et al. 1992; Bi 1993; Bi & Davidsen 1997). This was later supported by hydrodynamical cosmological simulations (e.g. Cen et al. 1994; Zhang et al. 1995; Miralda-Escudé et al. 1996; Hernquist et al. 1996; Theuns et al. 1998). In its simplest

form, the current notion is that intergalactic gas, pulled into dark matter dominated gravitational potential wells, forms sheets and filaments, and ultimately galaxies inside dark matter haloes, and these structures give rise to Lyα forest lines; sheets and filaments constitute weak absorbers, and haloes give rise to strong absorption systems. This in turn has the implication that the Lyα forest provides a map of matter distribution along the LOS to QSOs: baryons to a large extent follow the dark matter, and the hydrogen makes up for most of the baryons.

Lyα absorbers, from the weakest originating in underdense gas to the strongest arising in galaxy disks, present different stages in the odyssey of baryons, from diffuse gas to collapsed structures. Understanding their nature and their relation to galaxies is an essential step in unraveling how galaxies form and evolve.

In addition to being the reservoir of gas for galaxies, the IGM also tells a story of galactic feedback. Feedback constitutes all the processes where the current star formation and an active nucleus in a galaxy (AGN) have an impact on its evolution, e.g. through heating or removal of star-forming gas. Radiation from galaxies ionizes and heats the IGM, and supernova (SN) and AGN winds can shock heat the surrounding IGM and enrich it with elements heavier than helium (i.e. "metals" in the usual astronomical parlance), produced in stars. Such intergalactic metals also leave an imprint in the spectra of background objects, and their distribution—as revealed by absorption spectra—provides important constraints on models of galactic feedback.

1.2   Models of the Lyα forest

We choose to describe two popular models linking the underlying gas density to the absorption signature in the spectra of background objects. The fluctuating Gunn-Peterson approximation (e.g. Rauch et al. 1997), presented in §1.2.1, is appropriate for gas at and below the mean density of the Universe, while the "Jeans" approximation (Schaye 2001), presented in §1.2.2, provides a suitable description of absorption by overdense gas.

1.2.1   The fluctuating Gunn-Peterson approximation

As the radiation from background objects travels trough space, it can get scattered out of the LOS when it encounters intervening atoms of neutral hydrogen. The expected change of the background flux, F_ν, at the frequency ν, for radiation with the mean free path $\lambda_{\mathrm{mfp},\nu}$, is:

$$dF_\nu = -F_\nu \frac{dl}{\lambda_{\mathrm{mfp},\nu}}, \qquad (1.1)$$

where dl is the proper path length. The mean free path can be expressed in terms of the absorption cross-section, σ_ν, and the number density of absorbing atoms, n_{HI}, and so:

$$\frac{dF_\nu}{F_\nu} = -n_{\mathrm{HI}}\sigma_\nu dl. \tag{1.2}$$

Integration of this differential equation results in:

$$F_\nu = F_{\nu,\mathrm{c}}e^{-\int n_{\mathrm{HI}}\sigma_\nu dl} = F_{\nu,\mathrm{c}}e^{-\tau_\nu}, \tag{1.3}$$

where τ_ν is the optical depth.

The absorption cross-section for the Lyα transition ($\lambda_0 = 1215.67$ Å, $h\nu_0 = 10.2$ eV) is a function of frequency, $\sigma_\nu = \sigma_0\phi(\nu - \nu_0)$, where $\phi(\nu - \nu_0)$ is the function describing the line profile, and in the absence of line broadening it takes a form of the delta function, $\delta_{\mathrm{D}}(\nu - \nu_0)$. To estimate the optical depth as a function of observed frequency, we must integrate the expression for τ_ν along the radiation path length, i.e. from the radiation source at redshift z_{em} to $z = 0$:

$$\tau_{\nu,\mathrm{obs}} = \int_0^l n_{\mathrm{HI}}\sigma_\nu dl = \int_0^{z_{\mathrm{em}}} n_{\mathrm{HI}}(z)\sigma_\nu \left|\frac{dl}{dz}\right| dz. \tag{1.4}$$

By taking into account that $\nu = \nu_{\mathrm{obs}}(1+z)$, which implies that $dz = d\nu/\nu_{\mathrm{obs}}$, the integral takes the following form:

$$\tau(z) = \int_{\nu_{\mathrm{obs}}}^{\nu_{\mathrm{obs}}(1+z_{\mathrm{em}})} n_{\mathrm{HI}}(z)\sigma_\nu \left|\frac{dl}{dz}\right| \frac{d\nu}{\nu_{\mathrm{obs}}}. \tag{1.5}$$

In the absence of line broadening with σ_ν being a delta function, this integral becomes:

$$\tau(z) = \frac{\sigma_0\nu_0}{\nu_{\mathrm{obs}}}n_{\mathrm{HI}}(z) \left|\frac{dl}{dz}\right|. \tag{1.6}$$

Neglecting peculiar velocities, we can transform dl/dz:

$$dl = cdt = c\frac{dt}{da}\frac{da}{dz}dz = \frac{c}{\dot{a}}(-a^2)dz = -\frac{c}{H(z)}\frac{dz}{1+z}, \tag{1.7}$$

where c is the speed of light, dt is the time interval that radiation takes to travel the path length dl, $a = 1/(1+z)$ is the expansion factor of the Universe, and $H(z) = \dot{a}/a$ is the Hubble parameter. Taking this into account, and that $\nu_0 = \nu_{\mathrm{obs}}(1+z)$, the observed optical depth is:

$$\tau(z) = \sigma_0 n_{\mathrm{HI}}(z)\frac{c}{H(z)}. \tag{1.8}$$

The neutral hydrogen number density can be expressed in terms of the total hydrogen number density:

$$n_{\mathrm{HI}} = \frac{n_{\mathrm{HI}}}{n_{\mathrm{H}}}\frac{n_{\mathrm{H}}}{\bar{n}_{\mathrm{H}}}\bar{n}_{\mathrm{H}},$$

where $\bar{n}_{\mathrm{H}}$ is the mean hydrogen number density in the Universe:

$$\bar{n}_{\mathrm{H}} = \frac{\bar{\rho}_{\mathrm{b}} X}{m_{\mathrm{H}}} = \frac{X}{m_{\mathrm{H}}} \Omega_{\mathrm{b}} \rho_{0,\mathrm{crit}} (1+z)^3 \approx$$

$$\approx 7.15 \times 10^{-6}\,\mathrm{cm}^{-3} \left(\frac{X}{0.75}\right) \left(\frac{\Omega_{\mathrm{b}} h^2}{0.022}\right) \left(\frac{1+z}{3.4}\right)^3,$$

and $\bar{\rho}_{\mathrm{b}}$ is the mean baryonic density, X is hydrogen mass fraction, m_{H} is the mass of the hydrogen atom, Ω_{b} is the density parameter for baryons, and $\rho_{0,\mathrm{crit}}$ is the critical density of the Universe at $z = 0$. Substituting this into equation 1.8, we get:

$$\tau(z) \approx 1.3 \times 10^5 \left(\frac{\Omega_{\mathrm{m}} h^2}{0.13}\right)^{-1/2} \left(\frac{\Omega_{\mathrm{b}} h^2}{0.022}\right) \left(\frac{X}{0.75}\right) \left(\frac{1+z}{3.4}\right)^{3/2} \frac{n_{\mathrm{HI}}}{n_{\mathrm{H}}} \frac{n_{\mathrm{H}}}{\bar{n}_{\mathrm{H}}}. \quad (1.9)$$

An interesting conclusion about the state of the IGM can be drawn from this equation. Namely, because the observed optical depth at e.g. $z = 2.4$ is $\bar{\tau}(z = 2.4) < 1$, we have $\langle n_{\mathrm{HI}}/n_{\mathrm{H}} \rangle \lesssim 7.7 \times 10^{-6}$, i.e. the IGM is highly ionized at this redshift.

Equation 1.9 can be easily related to the underlying density field if we assume that the gas is in photo-ionization equilibrium, $n_{\mathrm{HII}} n_{\mathrm{e}} \beta = n_{\mathrm{HI}} \Gamma$. In this equation β is the hydrogen recombination rate, which depends on gas temperature as $\beta \approx 4 \times 10^{-13} T_4^{-0.76}\,\mathrm{cm}^{-3}\,\mathrm{s}^{-1}$ ($T \equiv T_4 * 10^4$ K); $\Gamma \equiv \Gamma_{12} \times 10^{-12}\,\mathrm{s}^{-1}$ is the hydrogen photoionization rate due to the background UV radiation with mean intensity J_ν:

$$\Gamma = \int_{\nu_{\mathrm{L}}}^{\infty} \frac{4\pi J_\nu \sigma_{\mathrm{i}}(\nu)}{h\nu} d\nu, \quad (1.10)$$

where $\sigma_{\mathrm{i}}(\nu)$ is the cross-section for photoionization, and ν_{L} is the frequency at the Lyman limit (i.e. 912 Å). Finally, n_{e} is the number density of free electrons. For highly ionized gas we can make the following approximations:

$$n_{\mathrm{HII}} \approx n_{\mathrm{H}}, \text{ and } n_{\mathrm{e}} = n_{\mathrm{H}} + 2n_{\mathrm{He}} = \frac{\rho}{m_{\mathrm{H}}} X + \frac{2\rho}{4m_{\mathrm{H}}}(1-X) = n_{\mathrm{H}}\frac{1+X}{2X},$$

after which we get:

$$\tau(z) \approx 0.45 \left(\frac{X}{0.75}\right) \left(\frac{1+X}{1.75}\right) \left(\frac{\Omega_{\mathrm{m}} h^2}{0.13}\right)^{-1/2} \times$$

$$\times \left(\frac{\Omega_{\mathrm{b}} h^2}{0.022}\right)^2 \left(\frac{1+z}{3.4}\right)^{9/2} T_4^{-0.76} \Gamma_{12}^{-1} \Delta^2, \quad (1.11)$$

where $\Delta = n_H / \bar{n}_H$ is gas overdensity.

In reality, there is some scatter in the relation between the optical depth and overdensity due to gas peculiar velocities and thermal broadening. The approximation is not appropriate in regimes where collisional ionization becomes important (e.g. in regions with $T > 10^5$ K), nor at high density because the line broadening is no longer dominated by the differential Hubble flow.

1.2.2   Jeans approximation

Schaye (2001) argued that overdense (with respect to the mean density of the Universe) Lyα absorbers are typically close to local hydrostatic equilibrium, i.e. that their characteristic size is of order of the local Jeans length. Starting from this assumption, the typical size and mass of absorbers with a given column density may be calculated. We consider these results relevant for building intuition about the physical properties of Lyα absorbers, and thus we repeat the derivation below.

The dynamical time in a cloud with characteristic density $n_{\rm H}$ is:

$$t_{\rm dyn} \equiv \frac{1}{\sqrt{G\rho}} \sim 1.0 \times 10^{15}\,{\rm s} \left(\frac{n_{\rm H}}{1\,{\rm cm}^{-3}}\right)^{-1/2} \left(\frac{X}{0.75}\right)^{1/2} \left(\frac{f_{\rm g}}{0.16}\right)^{1/2}, \quad (1.12)$$

where $f_{\rm g}$ is the fraction of cloud mass in gas (i.e. the baryonic fraction). For the absorbers that Schaye considers in the paper, $f_{\rm g}$ has a value close to the universal baryon fraction ($\approx \Omega_b/\Omega_m$). The sound crossing time in such an absorber, where L is the "characteristic size" of the cloud (i.e. the scale over which the typical density is of order the characteristic density), is:

$$t_{\rm sc} \equiv \frac{L}{c_{\rm s}} \sim 2.0 \times 10^{15}\,{\rm s} \left(\frac{L}{1\,{\rm kpc}}\right) T_4^{-1/2} \left(\frac{\mu}{0.59}\right)^{1/2}. \quad (1.13)$$

In this expression $c_{\rm s}$ is the sound speed in an ideal monoatomic gas, with $\gamma = 5/3$, and μ is the mean molecular weight, and for the rest of the derivation it is set to the value suitable for a fully ionized plasma with primordial abundances, $\mu \approx 0.59$.

The pressure is $P \sim c_{\rm s}^2\rho$, which in hydrostatic equilibrium:

$$\frac{dP}{dr} = -\frac{G\rho M}{r^2}, \quad (1.14)$$

where M is the cloud mass within radius r. It follows that $c_{\rm s}^2\rho/L \sim G\rho^2 L$, which in turn leads to $t_{\rm sc} \sim t_{\rm dyn}$, i.e. the sound crossing time is of order the dynamical time. Consequently, the characteristic size of such a cloud is of order the Jeans length:

$$L_{\rm J} \equiv \frac{c_{\rm s}}{\sqrt{G\rho}} \sim 0.52\,{\rm kpc}\, n_{\rm H}^{-1/2} T_4^{1/2} \left(\frac{f_{\rm g}}{0.16}\right)^{1/2} \quad (1.15)$$

The Lyα absorbers are typically discussed in terms of their column density, and for a cloud satisfying the Jeans condition, it is straightforward to relate total gas column density of a cloud to its characteristic density:

$$N_H \equiv n_{\rm H} L_{\rm J} \sim 1.6 \times 10^{21}\,{\rm cm}^{-2}\, n_{\rm H}^{1/2} T_4^{1/2} \left(\frac{f_{\rm g}}{0.16}\right)^{1/2}. \quad (1.16)$$

For clouds that are optically thin to the ionizing radiation (i.e. neutral gas column density $N_{\mathrm{HI}} \leq 10^{17.2}\,\mathrm{cm}^{-2}$), it is relatively easy to calculate the neutral gas column density from the total gas column density.

The neutral fraction of gas in photoionization equilibrium is:

$$\frac{n_{\mathrm{HI}}}{n_{\mathrm{H}}} = n_{\mathrm{e}}\beta_{\mathrm{HII}}\Gamma^{-1} \sim 0.46 n_{\mathrm{H}} T_4^{-0.76}\Gamma_{12}^{-1}, \tag{1.17}$$

For the rest of the calculations, the adopted value for Γ is $10^{-12}\,\mathrm{s}^{-1}$ (e.g. Haardt & Madau 2001), which is appropriate for $z \approx 2-4$, and the adopted value for the temperature of Lyα absorbers is $T \sim 10^4$ K (e.g. Schaye et al. 2000).

The neutral hydrogen column density can be obtained by combining equations (1.16) and (1.17):

$$N_{\mathrm{HI}} \sim 2.3 \times 10^{13}\,\mathrm{cm}^{-2} \left(\frac{n_{\mathrm{H}}}{10^{-5}\,\mathrm{cm}^{-3}}\right)^{3/2} T_4^{-0.26}\Gamma_{12}^{-1} \left(\frac{f_{\mathrm{g}}}{0.16}\right)^{1/2}. \tag{1.18}$$

One can also express the density in terms of the overdensity, i.e. $\Delta \equiv n_{\mathrm{H}}/\bar{n}_{\mathrm{H}}$, where:

$$\bar{n}_H \approx 7.3 \times 10^{-6}\,\mathrm{cm}^{-3} \left(\frac{1+z}{3.4}\right)^3 \left(\frac{\Omega_{\mathrm{b}}h^2}{0.022}\right), \tag{1.19}$$

from which follows:

$$N_{\mathrm{HI}} \sim 1.5 \times 10^{13}\,\mathrm{cm}^{-2}\Delta^{3/2}T_4^{-0.26}\Gamma_{12}^{-1} \left(\frac{1+z}{3.4}\right)^{9/2} \left(\frac{\Omega_{\mathrm{b}}h^2}{0.022}\right)^{3/2} \left(\frac{f_{\mathrm{g}}}{0.16}\right)^{1/2} \tag{1.20}$$

We can now also relate the size of an absorber to its neutral hydrogen column density, by combining equations (1.15), (1.17), and (1.18):

$$L \sim 1.0 \times 10^2\,\mathrm{kpc} \left(\frac{N_{\mathrm{HI}}}{10^{14}\,\mathrm{cm}^{-2}}\right)^{-1/3} T_4^{0.41}\Gamma_{12}^{-1/3} \left(\frac{f_{\mathrm{g}}}{0.16}\right)^{2/3}, \tag{1.21}$$

and so absorbers with e.g. $N_{\mathrm{HI}} = 10^{15}$, 10^{16}, and $10^{17}\,\mathrm{cm}^{-2}$ are expected to have sizes ~ 50, 20, and 10 physical kpc, respectively.

For spherical absorbers, the characteristic mass can be estimated as $M \sim \rho L^3$, and so the gas mass is:

$$M_{\mathrm{g}} \sim 8.8 \times 10^8\,\mathrm{M}_{\odot} \left(\frac{N_{\mathrm{HI}}}{10^{14}\,\mathrm{cm}^{-2}}\right)^{-1/3} T_4^{1.41}\Gamma_{12}^{-1/3} \left(\frac{f_{\mathrm{g}}}{0.16}\right)^{5/3} \tag{1.22}$$

Of course, absorbers are often not spherical in which case their mass cannot be estimated using this approximation.

The "Jeans" approximation breaks down for underdense absorbers because the sound-crossing time is then longer than the Hubble time, $t_{\mathrm{sc}} > t_{\mathrm{H}}$, i.e. they do not satisfy the Jeans condition.

1.3   The IGM near galaxies

The physics governing the Lyα forest is relatively simple, and thus the fluctuating Gunn-Peterson approximation is likely to describe reality reasonably well. However, the baryon physics becomes much more complicated when the matter fluctuations reach the strongly non-linear regime, where the IGM and galaxies meet. When it comes to observations, a common way of studying this interface is by identifying galaxies close to the line of sight (LOS) of background objects, and examining absorption in the spectra of background objects that coincides in redshift with foreground galaxy positions. Such observations of the galaxy surroundings are challenging.

At high redshift typical star-forming galaxies are faint, and measuring their redshifts requires significant time investment even with 8m class telescopes. Due to their faintness, most studies use galaxies only as foreground objects, while bright QSOs are used as background objects. The number of suitable bright QSOs decreases with redshift, and the number of close QSO-galaxy pairs gets even smaller. Using star-forming galaxies as background objects (e.g. Adelberger et al. 2005; Steidel et al. 2010) yields a higher number of close galaxy-galaxy pairs allowing valuable studies of the immediate galaxy surroundings, but the quality of background spectra is poorer, generally requiring stacking analyses, which in turn limits the type of studies that can be done. Progress in that respect will be possible when a new generation of 30m telescopes becomes available and the quality of galaxy spectra improves significantly (see Steidel et al. 2009, US Decadal Survey White Paper).

At low redshift observations of galaxies are easier and the number of suitable QSO candidates is higher, but studies of many astronomically interesting transitions, such as Lyα, are possible only with space-based facilities because they lie in the rest-frame UV. Significant progress is being made, however, with the UV sensitive Cosmic Origins Spectrograph (COS) on the Hubble Space Telescope, that became available in 2009.

Needless to say, ideally we want studies of the galaxy-IGM interface performed at different redshifts because the Universal "circumstances" change dramatically from e.g. $z \approx 2$ to $z = 0$; for example, at high-z, the universe is denser, galaxies are forming stars more rapidly, active galactic nuclei are more common, and the metagalactic background radiation is more intense.

Theoretical studies of the galaxy-IGM interface are also challenging. Numerous processes that are important in this regime, such as (non-equilibrium) cooling of gas in the presence of metals, collisional ionization, gravitational shock-heating, shock-heating by galactic winds, photoionization by local sources of radiation, and self-shielding, require numerical treatment. Unfortunately, numerical simulations are not at the stage where all the relevant processes can be simulated from first principles. This is because simulating galaxy formation requires a huge dynamic range, with some processes regu-

lated on the atomic level, and others on the scale of galaxy clusters. This is why the use of "subgrid prescriptions" for processes not captured at the current resolution level is unavoidable in simulations of galaxy formation. These subgrid prescriptions are motivated by observations, and the simulation output must be carefully scrutinized in the context of a different set of observational results. Through such a process it became clear that galactic feedback is necessary for, e.g., preventing galaxies from forming too many stars, and growing disks of realistic size (e.g. Springel & Hernquist 2003; Weil et al. 1998).

1.4   This thesis

This thesis presents research on the IGM, as revealed through Lyα absorption, in the vicinity of galaxies at $z \approx 2.4$ in the Keck Baryonic Structure Survey (KBSS, Steidel et al., 2012, in preparation), and the implications of the observed relations through comparison with the hydrodynamical cosmological simulations from the OverWhelmingly Large Simulations (OWLS) set of models (Schaye et al. 2010). This redshift is particularly suitable for simultaneous studies of the Lyα forest and galaxies, because both can be easily observed with ground-based facilities. The Lyα forest is redshifted from the rest-frame UV into the optical part of the spectrum, and the rest-frame UV spectrum of star-forming galaxies is also redshifted into the optical, allowing their efficient identification with optical filters through the Lyman Break technique (e.g. Steidel et al. 2004; Adelberger et al. 2004). The Lyα forest lines at $z \gg 2.4$ are mostly saturated, and at $z \ll 2.4$ they are quite rare, which also makes $z \sim 2.4$ exceptional. In addition, the universal star-formation rate density peaked at $z \sim 2 - 3$, which makes this epoch particularly informative as any interaction between galaxies and their environments should be at its peak during this time as well.

Chapter 2 presents a novel method for calibrating galaxy redshifts using absorption by the surrounding IGM. The most common way of measuring redshifts of high-z galaxies is from rest-frame UV absorption and emission lines originating in the ISM of galaxies. However, they are usually offset from the systemic redshifts due to the combination of radiative transfer effects and galactic outflows. An established way to correct for this is to calibrate the redshifts through near-IR observations of the rest-frame optical nebular emission lines, originating in the H II regions of galaxies, but such observations are currently costly, and not feasible for large samples of faint galaxies. Using KBSS galaxies and background QSOs, we have shown that it is possible to calibrate galaxy redshifts measured from rest-frame UV lines by utilizing the fact that the mean H I Lyα absorption profiles around the galaxies, as seen in spectra of background objects, must be symmetric with respect to the true galaxy redshifts if the galaxies are oriented randomly with respect to the lines of sight (LOS) to background objects.

In **Chapter 3** we present the observations of the H I optical depth near galaxies at $z \approx 2.4$ in the KBSS survey, using the pixel optical depth method to analyze the QSO spectra. We find the Lyα absorption to be enhanced out to at least 2.8 Mpc proper. We present the first two-dimensional maps of the absorption around galaxies, plotting the median Lyα pixel optical depth as a function of transverse and LOS separation from galaxies. We detect two types of redshift space anisotropies. On scales < 200 km s^{-1}, or < 1 Mpc proper, the absorption is stronger along the LOS than in the transverse direction. This "finger of God" effect may be partly due to redshift errors, but is probably dominated by gas motions within or very close to the halos. On the other hand, on scales of 1.4 - 2.0 Mpc proper the absorption is compressed along the LOS, an effect that we attribute to large-scale infall (i.e. the Kaiser effect). We measured the galaxy overdensity within a given volume as a function of pixel optical depth, and we show the covering fraction of absorbers with a given strength within 200 proper kpc from galaxies.

In **Chapter 4** we demonstrate that the observed Lyα absorption distribution near galaxies from Chapter 2 can be used to measure the masses of halos of that galaxy population. We match the observed absorption distribution to the absorption around haloes above a given mass in the cosmological simulations from the OWLS suite of models. The implied minimum halo mass is consistent with the results from the galaxy clustering analysis, and the results are robust to changes in cosmological parameters and feedback prescriptions in models. We also show that this method can be used in narrow field galaxy-QSO surveys, i.e. 30×30 arcseconds.

Inspired by recent theoretical results that imply that most of the fuel for star-formation comes into galaxies through cold accretion, i.e. without getting heated to the virial temperature of the host haloes, we examine in **Chapter 5** how much Lyα absorption near galaxies at $z = 2.25$ is produced in such cold flows. We use OWLS models and study absorption in gas selected based on its thermal history, halo membership, kinematics with respect to the galaxy, and likelihood of becoming part of the interstellar medium by $z = 0$. We also look into the physical properties of the Lyα absorbing gas, i.e. its temperature and density, as a function of distance from galaxies in OWLS models with and without SN and AGN feedback.

References

Adelberger, K. L., Steidel, C. C., Shapley, A. E., et al. 2004, ApJ, 607, 226

Adelberger, K. L., Shapley, A. E., Steidel, C. C., et al., 2005, ApJ, 629, 636

Bi, H. G., Boerner, G., & Chu, Y., 1992, A&A, 266, 1

Bi, H., 1993, ApJ, 405, 479

Bi, H., & Davidsen, A. F., 1997, ApJ, 479, 523

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2010, ApJ, 709, L133

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al., 2011, ApJ, 737, 90

Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M., 1994, ApJ, 437, L9

Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays Modelling the UV/X-ray cosmic background with CUBA, arXiv: 0106018

Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J., 1996, ApJ, 457, L51

Labbé, I., González, V., Bouwens, R. J., et al., 2010, ApJ, 716, L103

McLure, R. J., Dunlop, J. S., de Ravel, L., et al. 2011, arXiv:1102.4881, MNRAS, in press

Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M., 1996, ApJ, 471, 582

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2010, ApJ, 709, L16

Rauch, M., Miralda-Escude, J., Sargent, W. L. W., et al., 1997, ApJ, 489, 7

Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., & Sargent, W. L. W., 2000, MNRAS, 318, 817

Schaye, J. 2001, ApJ, 559, 507

Springel, V., & Hernquist, L. 2003, MNRAS, 339, 289

Steidel, C. C., Shapley, A. E., Pettini, M., et al. 2004, ApJ, 604, 534

Steidel, C., Martin, C., Prochaska, J. X., Pettini, M., Schaye, J., & Rakic, O., 2009, astro2010: The Astronomy and Astrophysics Decadal Survey, 2010, 286

Steidel, C. C., Erb, D. K., Shapley, A. E., Pettini, M., Reddy, N., Bogosavljević, M., Rudie, G. C., Rakic, O., 2010, ApJ, 717, 289

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

Zhang, Y., Anninos, P., & Norman, M. L. 1995, ApJ, 453, L57

Weil, M. L., Eke, V. R., & Efstathiou, G., 1998, MNRAS, 300, 773

2

Calibrating Galaxy Redshifts Using Absorption by the Surrounding Intergalactic Medium

Rest-frame UV spectral lines of star-forming galaxies are systematically offset from the galaxies' systemic redshifts, probably because of large-scale outflows. We calibrate galaxy redshifts measured from rest-frame UV lines by utilizing the fact that the mean H I Lyα absorption profiles around the galaxies, as seen in spectra of background objects, must be symmetric with respect to the true galaxy redshifts if the galaxies are oriented randomly with respect to the lines of sight to the background objects. We use 15 bright QSOs at $z \approx 2.5 - 3$ and more than 600 foreground galaxies with spectroscopic redshifts at $z \approx 1.9 - 2.5$. All galaxies are within 2 Mpc proper from the lines of sight to the background QSOs. We find that Lyα emission and ISM absorption redshifts require systematic shifts of $\Delta v_{\mathrm{Ly}\alpha} = -295^{+35}_{-35}\,\mathrm{km\,s^{-1}}$ and $\Delta v_{\mathrm{ISM}} = 145^{+70}_{-35}\,\mathrm{km\,s^{-1}}$, respectively. Assuming a Gaussian distribution, we put 1σ upper limits on possible random redshift offsets of $< 220\,\mathrm{km\,s^{-1}}$ for Lyα and $< 420\,\mathrm{km\,s^{-1}}$ for ISM redshifts. For the small subset (< 10 percent) of galaxies for which near-IR spectra have been obtained, we can compare our results to direct measurements based on rest-frame optical, nebular emission lines, which we confirm to mark the systemic redshifts. While our Δv_{ISM} agrees with the direct measurements, our $\Delta v_{\mathrm{Ly}\alpha}$ is significantly smaller. However, when we apply our method to the near-IR subsample which is characterized by slightly different selection effects, the best-fit velocity offset comes into agreement with the direct measurement. This confirms the validity of our approach, and implies that no single number appropriately describes the whole population of galaxies, in line with the observation that the line offset depends on galaxy spectral morphology. This

method provides accurate redshift calibrations and will enable studies of circumgalactic matter around galaxies for which rest-frame optical observations are not available.

Olivera Rakic, Joop Schaye, Charles C. Steidel, and Gwen C. Rudie

2.1 Introduction

Two questions that are key to understanding galaxy formation and evolution are: How do galaxies get their gas? and Where does the gas that is ejected via galactic feedback processes end up? A powerful method to study the interface between galaxies and the intergalactic medium (IGM) is to examine the spectra of bright background objects for absorption features at the redshifts of foreground galaxies. Different studies have looked into the relation between Lyα absorbing gas and galaxies (e.g., Lanzetta et al. 1995; Chen et al. 1998; Bowen et al. 2002; Penton et al. 2002; Adelberger et al. 2003, 2005; Crighton et al. 2010; Steidel et al. 2010), as well as metals and galaxies (e.g., Lanzetta & Bowen 1990; Bergeron & Boissé 1991; Steidel et al. 1994, 1997; Chen et al. 2001; Adelberger et al. 2003, 2005; Pieri et al. 2006; Steidel et al. 2010). A crucial requirement for the success of such studies is the availability of accurate redshifts for the foreground galaxies. To illustrate this point, we note that observationally inferred outflow velocities correspond to a change in redshift of order $\Delta z/(1+z) = \Delta v/c \sim 10^{-3}$ and that at $z = 2$ the difference in Hubble velocity across a proper distance of 1 Mpc, which exceeds the virial radius of the typical star-forming galaxy at that redshift by more than an order of magnitude, corresponds to $\Delta z/(1+z) \approx 7 \times 10^{-4}$. Clearly, the systematic errors on the redshifts of foreground objects need to be $\ll 0.1\%$ in order to map the physical properties of the gas in and around the haloes of galaxies.

The most accurate galaxy redshifts are measured from absorption lines arising in stellar atmospheres. Rest-frame optical stellar absorption features are routinely detected in spectra of galaxies in the local Universe. For high-z galaxies rest-frame UV stellar absorption lines are detectable in stacks of ≈ 100 high-quality spectra and have been used to verify the accuracy of redshift calibrations (Shapley et al. 2003; Steidel et al. 2010). One can also measure accurate redshifts from nebular emission lines from stellar HII regions. Lines such as Hα ($\lambda6563$), Hβ ($\lambda4861$), and [O III] ($\lambda\lambda4959, 5007$) are strong and easily accessible for galaxies in the nearby Universe because they are in the rest-frame optical. Measuring redshifts from these lines is, however, more difficult for high-z galaxies. Nebular lines are redshifted into the observed frame near-infrared, and at these wavelengths spectroscopy with ground based instruments is complicated by the bright night-sky, as well as by strong absorption features produced by molecules in the Earth's atmosphere. The prospects for near-IR spectroscopy, at least for bright objects, will improve when multi-object near-IR spectrographs come online (e.g. Keck I/MOSFIRE; McLean et al. 2008), but near-IR spectroscopy will remain challenging for fainter objects.

Alternatively, one may resort to measuring redshifts using spectral lines that lie in the rest-frame UV and can thus be observed with optical spectrographs. Such lines include the H I Lyα emission line at 1216Å, when

present, and UV absorption lines arising in the interstellar medium, of which the strongest are C II ($\lambda 1335$), Si IV ($\lambda 1394$), and Si II ($\lambda 1527$). However, these lines are not ideal as they do not provide accurate redshift measurements. Lyα photons are resonantly scattered, and the observed line profile depends on the kinematics and the amount of gas that is scattering the photons, as well as on the dust content (e.g., Verhamme et al. 2006, 2008; Schaerer & Verhamme 2008; Hansen & Oh 2006; Dijkstra et al. 2006a,b; Zheng & Miralda-Escudé 2002). The observed velocities of interstellar absorption lines also depend on the kinematics of the gas that produces them. Most high-redshift galaxies observable with current facilities have sufficiently high star formation rates to drive galactic scale winds. These winds carry some interstellar material out of the galaxies and this causes both the Lyα emission line and interstellar absorption lines to be systematically offset by a few hundred $\mathrm{km\,s^{-1}}$ from the galaxy systemic redshifts, as measured from nebular lines (e.g., Pettini et al. 2001; Adelberger et al. 2003, 2005; Steidel et al. 2010).

In surveys of high-redshift ($z \approx 2 - 3$) star-forming galaxies, Adelberger et al. (2003, 2005) and Steidel et al. (2010) dealt with these issues by measuring redshifts from rest-frame UV lines, obtaining near-IR spectroscopy for a subset ($< 10\,\%$) of galaxies, and then calibrating the redshifts measured from rest-frame UV lines using the rest-frame optical nebular lines. This calibration was then applied to the rest of the sample that lacked near-IR spectroscopy. However, galaxy populations other than Lyman Break Galaxies (LBGs) may require their own calibrations. This is a problem as calibration using nebular lines is often not possible. For example, most of the galaxies found by future surveys with integral field spectrographs, such as the Multi Unit Spectroscopic Explorer (MUSE, Bacon et al. 2010), will only be detected in Lyα. Also, even for the LBGs at $z \approx 2 - 3$ it is possible that the subsample that is observed in the near-IR is not quite representative of the whole sample, for example because galaxies need to be sufficiently bright to do near-IR spectroscopy.

In this paper we present and apply a method for calibrating redshifts measured from rest-frame UV lines using the absorption features from the surrounding IGM as seen in the spectra of background QSOs. Several studies have found that the mean strength of H I Lyα absorption increases as the distance to the nearest galaxy decreases (e.g., Chen et al. 1998; Adelberger et al. 2003, 2005; Pieri et al. 2006; Ryan-Weber 2006; Morris & Jannuzi 2006; Wilman et al. 2007; Steidel et al. 2010). By using the fact that the mean absorption profiles must be symmetric with respect to the systemic galaxy redshifts if the galaxies are oriented randomly with respect to the lines of sight to the background QSOs, we determine the optimal average velocity offset for which the observed mean absorption profile is symmetric around the positions of the galaxies along the line of sight (LOS). Although the same principle was used to verify the accuracy of the calibrations proposed

in Adelberger et al. (2003) and Steidel et al. (2010), this is the first time that the IGM absorption is used for calibration itself. The use of absorption by the IGM makes the method independent of the galaxies' luminosities, and does not require rest-frame optical observations. We find that the required offsets for redshifts based on ISM absorption lines and Lyα emission lines agree with the calibrations inferred by direct comparison with nebular line redshifts. Using this method we also confirm that different galaxy subsamples require different calibrations depending on the galaxies' spectral morphology.

Throughout this work we use $\Omega_{\mathrm{m}} = 0.258$, $\Omega_{\Lambda} = 0.742$, and $h = 0.719$ (Komatsu et al. 2009). Unless stated otherwise, all distances are proper.

2.2 Data

The data sample used in this study is partially (3 out of 15 fields) described in Steidel et al. (2003; 2004), Adelberger et al. (2005), and Steidel et al. (2010), while the full survey will be presented elsewhere. Here we give only a brief description. This redshift survey was constructed to select galaxies whose U_nGR colors are consistent with star-forming galaxies at redshifts $z \approx 1.9 - 2.7$. The survey was conducted in 15 fields with typical size 5×7 arcmin, centred on bright background QSOs for which there were high-resolution spectra suitable for probing Lyman-α absorbing gas in the same redshift range. The technique for optically selecting galaxies is described in detail in Adelberger et al. (2004) (their BX/BM sample). All 659 galaxies used here have been followed up spectroscopically with the optical spectrograph Keck I/LRIS-B (FWHM ≈ 370 km s^{-1}; Steidel et al. 2004). A subset of 49 galaxies has been observed with the near-IR spectrograph Keck II/NIRSPEC (McLean et al. 1998) at a higher resolution (FWHM ≈ 240 km s^{-1}) than achieved by LRIS-B (see Erb et al. 2006, for details on this galaxy sample). The sample used here is part of a larger sample, with a few thousand galaxies, from which we selected all galaxies within 2 Mpc from the LOS to the background QSOs, and with redshifts that fall in the Lyα forest redshift ranges of the QSOs in their fields. We define the Lyα forest range as the part of the spectrum between the quasar's Lyα and Lyβ emission lines, excluding the region < 5000 km s^{-1} from the quasar's redshift (to avoid the proximity zone).

The QSO observations were conducted between 1996 and 2009 with the Keck I/HIRES echelle spectrograph. The spectra have a resolution of $\Delta v \approx 7.5$ km s^{-1} (R $\simeq 40,000$) and were rebinned into pixels of 0.04 Å. They cover the wavelength range $\approx 3100 - 6000$ Å. The reduction was done using T. Barlow's MAKEE package, and the spectra were normalized using low order spline fits. The typical S/N in the Lyα forest region lies between 40 and 160.

Repeated observations of the same galaxies suggest that the typical statistical measurement uncertainties are ≈ 100 km s^{-1} for redshifts measured from ISM absorption lines (z_{ISM}), ≈ 50 km s^{-1} for redshifts measured from

Lyα emission lines ($z_{\mathrm{Ly}\alpha}$), and ≈ 60 km s^{-1} for redshifts measured from neb-ular emission lines (z_{neb}). The systematic offsets between these redshifts and the true (i.e. systemic) redshifts, which is the subject of this paper, may of course be larger.

2.3   Method

Absorption spectra of background objects (QSOs, GRBs, or star-forming galaxies) can be used to study the intervening IGM. Observations generally show an increase in the mean Lyα and metal absorption as the distance to the nearest galaxy decreases. One expects to find an average absorption profile that is symmetric around the galaxies' positions along the LOS. This is true even if the IGM is not distributed isotropically around galaxies, which can for example occur if the mean absorption at a given distance from a galaxy correlates with the galaxy's inclination. Assuming that the galaxies are oriented randomly with respect to the LOS, we still expect to see a symmetric absorption profile when averaged over an ensemble of galaxies[1]. If the measured redshifts of galaxies are systematically offset from their systemic redshifts, then the mean profile will be symmetric around some other velocity point, offset from the observed redshift by an amount equal to the systematic shift in galaxy redshifts (see Fig. 2.1 for an illustration). In this paper we use this principle to calibrate the systematic offsets between the observed and systemic (i.e. true) redshifts for the galaxies in our sample.

One could imagine that instead of measuring the offset from the average absorption profile, we could measure offsets for individual galaxies by ob-serving the IGM absorption features around them, and then find the average of such estimated offsets. However, near the redshift of a galaxy one often finds several lines at different velocities, of which some are associated with the galaxy in question, but the others are produced by, for example, a cloud that is associated with a neighboring galaxy or an intervening cloud. If the absorption line associated with such a cloud were stronger than that due to the gas associated with the galaxy, then we would infer the wrong offset. Also, at small impact parameters even a cloud that is associated with the galaxy in question could potentially have kinematics that is not dominated by gravity (e.g., it could be a part of the galactic outflow); and at any point in the forest there could be absorption completely unrelated to the observed redshift. This is not a problem when using the average absorption profile, given that it is equally likely for such a line to be at a positive as at a negative velocity with respect to the galaxy.

---

[1]As the blue side corresponds to a lower Lyα redshift than the red side, and the mean Lyα absorption at a random location is known to increase with redshift, we do expect the intergalactic Lyα absorption to be slightly weaker on the blue side. However, the magnitude of this effect (the mean transmission varies by 0.002 over 10^3 km s^{-1} at $z = 2.3$; e.g. Schaye et al. 2003) is negligible compared to the enhancement in the mean absorption near galaxies.
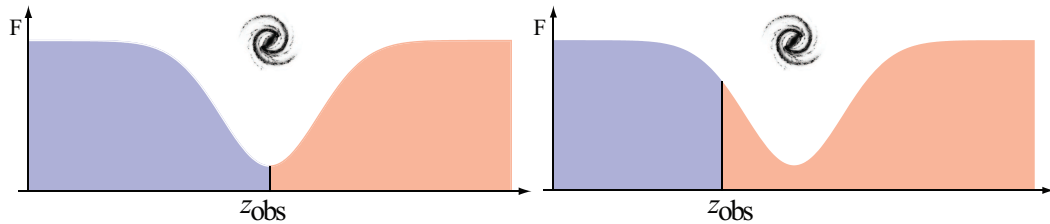
Figure 2.1 – The mean flux profile is symmetric around galaxy positions if the observed redshifts (z_{obs}, vertical line) are equal to the true redshifts (left panel). If the observed redshifts are systematically offset from the true redshifts, the mean profile will be symmetric around some other redshift point, offset from the observed redshifts (right panel).

We use Lyα rather than metal lines, because the metal-line absorption is usually only enhanced within a couple of hundred kpc from galaxies, and our sample does not include a sufficiently large number of galaxies with such small impact parameters (see the discussion below on the required number of galaxies). However, at higher redshifts, where the Lyα forest starts to saturate, it might be advantageous to use metal lines, as long as the survey contains sufficient galaxies.

We compare the "blue" and "red" sides of the median Lyα absorption profile, within $\pm 500\,\mathrm{km\,s^{-1}}$ of the galaxy redshifts of all the foreground galaxies with impact parameters, b, smaller than 2 Mpc. We try a large number of velocity offsets, $z_{\mathrm{obs}} \to z_{\mathrm{obs}} + \frac{\Delta v}{c}(1 + z_{\mathrm{obs}})$, where z_{obs} is the observed redshift, and c is the speed of light. Note that a positive Δv indicates that the observed redshifts systematically underestimate the true redshifts. We verified that there is no justification for using a redshift dependent offset – either because our redshift range is too small for evolution to matter ($z \approx 1.9 - 2.6$), or because there is no dependence on redshift in this galaxy population.

The analysis involves the following steps: *i)* we assume that the observed redshifts are systematically offset from the systemic redshifts by Δv and assign each galaxy a new redshift $z'_{\mathrm{obs}} = z_{\mathrm{obs}} + \frac{\Delta v}{c}(1 + z_{\mathrm{obs}})$; *ii)* for each galaxy we shift the QSO spectrum into the galaxy's rest-frame; *iii)* for each galaxy we interpolate the flux in the QSO spectrum within $\pm 500\,\mathrm{km\,s^{-1}}$ from z'_{obs} in velocity bins of $10\,\mathrm{km\,s^{-1}}$ to get an absorption profile; *iv)* we find the median[2] flux profile as a function of velocity separation from galaxies; *v)* we compare the "blue" ($-500 - 0\,\mathrm{km\,s^{-1}}$) and the "red" ($0 - 500\,\mathrm{km\,s^{-1}}$) side of the median flux profile using χ^2 statistics:

$$\chi^2 = \sum_{i=1}^{N} \frac{(\mathrm{B_i} - \mathrm{R_i})^2}{\sigma^2(\mathrm{B_i}) + \sigma^2(\mathrm{R_i})}$$

---

[2] We verified that using the mean rather than the median gives nearly identical results for samples with enough galaxies, but that using the median is more robust for small samples.

where B_i and R_i are the blue and red median flux arrays, $\sigma(B_i)$ and $\sigma(R_i)$ are the errors on the median flux, and the sum extends over all velocity bins within $\pm 500\,\mathrm{km\,s^{-1}}$ from the assumed redshifts. The errors on the median flux profiles are estimated by bootstrap resampling. We recreated the galaxy sample 1000 times by dividing each QSO field into redshift bins of $1000\,\mathrm{km\,s^{-1}}$ and then randomly drawing redshift intervals to recreate each field, allowing for individual intervals to be selected multiple times. Each bootstrapped galaxy sample consists of galaxies whose redshifts fall in the selected redshift intervals. We chose an interval of $1000\,\mathrm{km\,s^{-1}}$ because this length is larger than the correlation length of LBGs (Adelberger et al. 2005a). These five steps are repeated for a large number of assumed redshift offsets, yielding a reduced χ^2 curve (see Fig. 2.2). We expect the χ^2 to reach a (local) minimum for the offset that brings the measured redshifts closest to the true redshifts. Note that the χ^2 will also have a small value if the applied redshift offsets are very large, because in that case the absorption will no longer be correlated with the assumed locations of galaxies, the median flux profile will be flat, and the blue and red sides will agree. We therefore search for χ^2 minima only between the inferred χ^2 maxima (see Fig. 2.2 for the characteristic shape of the χ^2 curves). In practice we do this as follows: *i)* we smooth the χ^2 curve with a boxcar average with $50\,\mathrm{km\,s^{-1}}$ width; *ii)* we find the global χ^2 maximum and determine the FWHM of the χ^2 peak it belongs to; *iii)* we search for the second (local) χ^2 maximum that is at least one FWHM away from the first one; *iv)* we find the χ^2 minimum between these two peaks, in the un-smoothed curve.

The value that we chose for the maximum impact parameter, $b = 2\,\mathrm{Mpc}$, is a compromise between the need for a strong absorption signal (the mean absorption is stronger for galaxies with smaller impact parameters), and the need for a large number of galaxies (there are more galaxies with larger impact parameters). We verified that using smaller maximum impact parameters gives consistent results. Any maximum impact parameter in the range[3] 0.5-2 Mpc yields redshift calibrations that agree with those obtained for the $b < 2\,\mathrm{Mpc}$ sample within the estimated errors. However, the required number of galaxies does depend on the adopted maximum impact parameter. For $b < 1$ and $b < 2\,\mathrm{Mpc}$, we obtain converged results if we use $\gtrsim 100$ and $\gtrsim 200$ galaxies, respectively, although the error bar on the inferred systematic shift continues to decrease slowly if more galaxies are included. We also varied the velocity range around each galaxy. In this case there is a tradeoff between the strength of the signal and the number of pixels available for statistics. We found that the velocity range of $\pm 500\,\mathrm{km\,s^{-1}}$ generally gives the tightest confidence intervals around the estimated offsets, but the results are insensitive to the precise velocity range chosen.

We apply the above method to our sample and estimate systematic veloc-

---

[3] There are insufficient galaxies with $b < 0.5\,\mathrm{Mpc}$ to test smaller maximum impact parameters.

ity shifts for the subsamples listed in Fig. 2.3. We make a distinction between the whole sample and the NIRSPEC subsample due to the possibility that this subset is not representative of the whole sample (see Erb et al. 2006, for details on the sample selection and possible biases).

2.4   Results

In Fig. 2.2 we plot the reduced χ^2 as a function of the assumed velocity offset. We find that the best offset for redshifts measured from Lyα emission is $\Delta v_{\mathrm{Ly}\alpha} = -295^{+35}_{-35}\,\mathrm{km\,s^{-1}}$ (i.e., the Lyα emission line is systematically redshifted with respect to the systemic velocity), while for redshifts measured from ISM absorption lines we find $\Delta v_{\mathrm{ISM}} = 145^{+70}_{-35}\,\mathrm{km\,s^{-1}}$ (i.e., the ISM absorption lines are systematically blueshifted).

We estimated the confidence intervals around the best estimates for Δv_{ISM} or $\Delta v_{\mathrm{Ly}\alpha}$ using the bootstrap method, described in Section 2.3. We perform all the steps described in Section 2.3 on samples created by bootstrapping the original sample, resulting in a best-fit Δv for each bootstrap realization of the galaxy sample. The error on the estimate of Δv is the 1σ confidence interval around the median Δv from the 1000 bootstrap realizations. These 1σ confidence intervals are shown as the light blue shaded regions in Fig. 2.2 together with the χ^2 curve.

We prefer to determine confidence intervals by using bootstrap resampling rather than $\Delta\chi^2$ because the former is more robust. Indeed, we found that for small galaxy samples the χ^2 curves become too noisy for $\Delta\chi^2$ estimates to work, while for large samples they yield errors slightly larger than those estimated from bootstrap resampling. Contrary to $\Delta\chi^2$, bootstrap resampling does not require the errors on the median flux profile to be Gaussian. It also does not require the different velocity bins to provide independent measurements, which they in fact do not because the individual absorption lines are broader than the bins in Δv. Bootstrap resampling merely requires that the redshift regions into which we divide the fields are independent, which is true since, as already mentioned, they are larger than the galaxies' correlation length. We thus only use the χ^2 curve to find the Δv that minimizes the difference between the blue and red sides.

For the subset of galaxies with near-IR data, the nebular lines provide accurate markers of the systemic redshifts of the individual galaxies. Comparing these with the rest-frame UV lines, we find median offsets of $\Delta v_{\mathrm{Ly}\alpha} = -406^{+22}_{-30}\,\mathrm{km\,s^{-1}}$ and $\Delta v_{\mathrm{ISM}} = 166^{+9}_{-35}\,\mathrm{km\,s^{-1}}$ (based on 42 and 86 galaxies, respectively, including objects that are not within 2 Mpc of the LOS to a QSO). The quoted uncertainties correspond to the 1σ confidence interval around the median, estimated by bootstrap resampling the galaxy sample 1000 times. These results are shown as red polygons in Fig. 2.2. For reference, we note
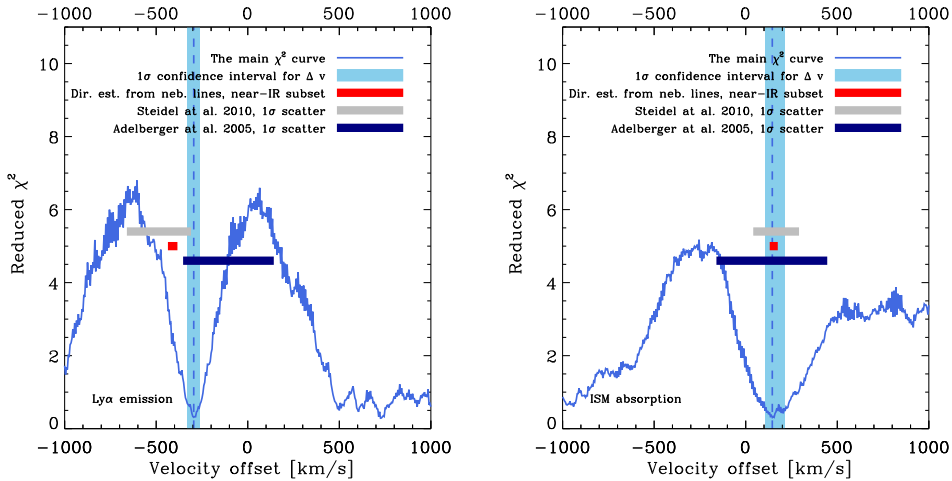
Figure 2.2 – The reduced χ^2 obtained by comparing the red and blue sides of the median Lyα absorption profile, as a function of assumed velocity offset $\Delta v = c(z_{\mathrm{gal}} - z_{\mathrm{obs}})/(1 + z_{\mathrm{obs}})$, where z_{gal} and z_{obs} are the true systemic and observed redshifts, respectively. The red and the blue sides are compared within $\pm 500 \, \mathrm{km \, s^{-1}}$ of all galaxies with impact parameters smaller than 2 Mpc. The *left panel* shows the results for redshifts measured from *Lyα emission lines* (321 galaxies), and the *right panel* for redshifts measured from *ISM absorption lines* (590 galaxies). The dashed vertical line shows the (local) minimum of the χ^2 curve. The vertical, light blue regions show the 1σ confidence intervals, as estimated by bootstrap resampling the galaxy samples. The χ^2 becomes small for very large offsets because in that case we are comparing "no signal with no signal" since we completely miss the galaxies. The horizontal red polygons show the 1σ confidence interval for the median offset for the near-IR subsample, obtained by direct comparison with nebular lines (see also Steidel et al. 2010). The grey and dark blue horizontal lines show the estimates for the mean offset from Steidel et al. (2010) and Adelberger et al. (2005) obtained from nebular lines, respectively, with the size of the bars indicating the 1σ scatter among the galaxies. Our measurements agree well with direct estimates based on comparison with nebular lines for ISM absorption lines, but for Lyα emission lines they differ significantly.

that Steidel et al. (2010) quote[4] mean offsets of $\Delta v_{\mathrm{Ly}\alpha} = -445 \pm 27 \, \mathrm{km \, s^{-1}}$ and $\Delta v_{\mathrm{ISM}} = 164 \pm 16 \, \mathrm{km \, s^{-1}}$. While our result for Δv_{ISM} agrees with the direct estimate, the best estimates for $\Delta v_{\mathrm{Ly}\alpha}$ differ significantly.

We therefore employ our IGM calibration method for the subsample with nebular redshifts. For Lyα the resulting offset is $\Delta v_{\mathrm{Ly}\alpha} = -455^{+45}_{-65} \, \mathrm{km \, s^{-1}}$. This is in good agreement with the direct estimate ($\Delta v_{\mathrm{Ly}\alpha} = -406^{+22}_{-30} \, \mathrm{km \, s^{-1}}$) from nebular lines for the same galaxies. Both disagree with our measurement for the full sample ($\Delta v_{\mathrm{Ly}\alpha} = -295^{+35}_{-35} \, \mathrm{km \, s^{-1}}$).

We emphasize that our measurement of $\Delta v_{\mathrm{Ly}\alpha}$ for the near-IR sample is based on only 26 galaxies, which is generally insufficient to obtain a converged result. Indeed, we find that our method fails for the subsample of 48 galaxies with both ISM absorption and nebular emission lines. The reduced χ^2 curve is noisy and does not show two clear peaks, which causes the best-fit velocity

---

[4]Steidel et al. (2010) also quote $\Delta v_{\mathrm{Ly}\alpha} = -485 \pm 175 \, \mathrm{km \, s^{-1}}$ and $\Delta v_{\mathrm{ISM}} = 166 \pm 125 \, \mathrm{km \, s^{-1}}$, but these errors are standard deviations around the best offset estimates.

offset to vary strongly with the velocity interval over which we measure the median flux. Although the χ^2 curve for the subsample of near-IR galaxies with Lyα emission does look normal and we found the result to be insensitive to the precise velocity interval used, the number of galaxies may well be too small to obtain a robust measurement.

Galaxies whose redshifts have been measured only from their Lyα emission feature tend to be those with poorer continuum S/N, due either to faintness or to poor observing conditions. Such galaxies are under-represented in the near-IR subsample because their redshifts are considered less secure and therefore were selected against for near-IR spectroscopic follow-up (Erb et al. 2006). It has also been observed that the velocity offset of Lyα emission is anti-correlated with the Lyα equivalent width (Shapley et al. 2003) – stronger Lyα lines are expected to have velocities closer to the galaxy systemic redshift. These two effects are likely responsible for the differences between the full sample and the smaller near-IR subsample.

This difference sends a warning that describing a galaxy population with a single number is not the best strategy given that the line offsets depend on galaxy spectral morphology. Thus, conclusions drawn from the near-IR subsample may not be generally applicable and separate redshift calibrations are needed for each sample of interest.

However, we do confirm the finding from Steidel et al. (2010) that absorption line offsets are less sensitive to galaxy spectral morphology than Lyα emission line offsets. This justifies the strategy taken in Adelberger et al. (2003, 2005) and Steidel et al. (2010) to use redshifts measured from ISM absorption lines when they are available ($\approx 90\%$ of the sample).

As an additional test of our IGM calibration method, we measure the systematic velocity offset for redshifts measured from nebular emission lines, which are known to be close to systemic positions. The resulting offset, which is based on a sample of 49 galaxies, is $\Delta v_{\mathrm{nebular}} = -75^{+85}_{-30}\,\mathrm{km\,s^{-1}}$, i.e. consistent with zero, as expected (Fig. 2.3).

2.4.1 Random offsets

We have demonstrated that one can use the fact that Lyα absorption by the IGM is correlated with proximity to galaxies to calibrate the redshifts inferred from Lyα emission and ISM absorption lines. In other words, we measured the systematic offsets in the observed redshifts. We will now use the same idea to constrain the statistical (i.e. random) offsets. Statistical scatter in individual galaxy redshift offsets may be caused by measurement and instrumental errors, but in our case it is probably dominated by the variations in the intrinsic galaxy properties. By comparing redshifts inferred from rest-frame UV lines with those from nebular lines for the subset of 89 galaxies that have been observed in the near-IR, one not only finds a large systematic offset, but there is also a significant scatter of $\approx 180\,\mathrm{km\,s^{-1}}$ for
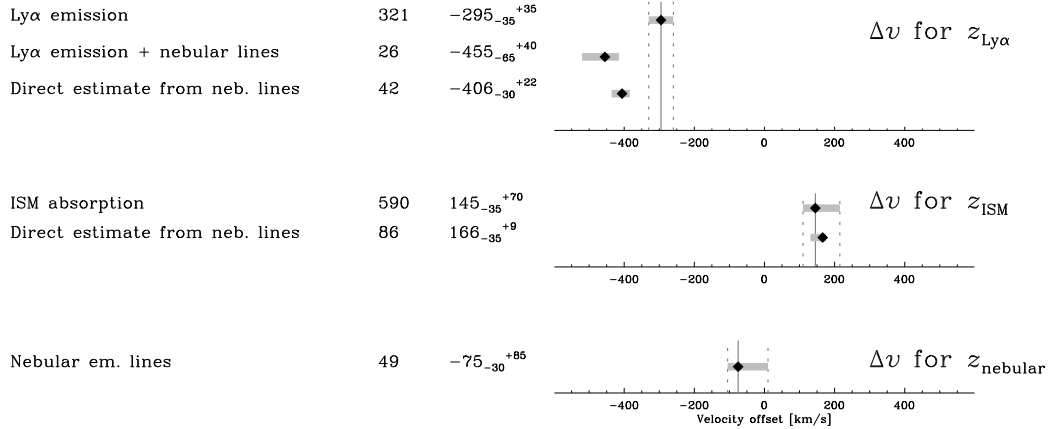
Lyα emission	321	-295_{-35}^{+35}
Lyα emission + nebular lines	26	-455_{-65}^{+40}
Direct estimate from neb. lines	42	-406_{-30}^{+22}

Δv for $z_{\mathrm{Ly\alpha}}$

ISM absorption	590	145_{-35}^{+70}
Direct estimate from neb. lines	86	166_{-35}^{+9}

Δv for z_{ISM}

Nebular em. lines	49	-75_{-30}^{+85}

Δv for z_{nebular}

Velocity offset [km/s]

Figure 2.3 – Different subsamples (first column) together with the number of galaxies in the sample (second column) and the best-fit velocity offsets, Δv, and 1σ confidence intervals (third column). These best offsets (black diamonds) together with the 1σ confidence intervals (grey horizontal stripes) are shown graphically on the right. The results for the full samples are shown as vertical gray lines.

Lyα redshifts, and $\approx 160\,\mathrm{km\,s^{-1}}$ for ISM redshifts.

We use all galaxies with impact parameters smaller than 2 Mpc, and measure the median absorption around them. We then repeat the procedure after applying a random redshift offset to the redshift of each galaxy, drawn from a Gaussian distribution with mean zero and standard deviation σ. We do this many times for increasing values of σ. We then compare the resulting median flux profiles with the original (i.e. $\sigma = 0$) median absorption profile. We expect that the two profiles will be consistent with each other as long as the added random offsets are small compared with the true random offsets. Using this method we obtain 1σ upper limits on the random redshift offsets of $< 220\,\mathrm{km\,s^{-1}}$ ($< 300\,\mathrm{km\,s^{-1}}$) for Lyα redshifts, using mean (median) statistics, and $< 420\,\mathrm{km\,s^{-1}}$ ($430\,\mathrm{km\,s^{-1}}$) for ISM redshifts, which is consistent with the independent estimates based on the subsample with redshifts from nebular lines. The quoted values are upper limits because we measure the amount of scatter needed to cause a significant difference with respect to the original absorption profile, i.e. the value we find exceeds the actual scatter that is present in the sample.

2.5   Summary & Conclusions

Measuring accurate redshifts for high-z galaxies is a daunting task. Distant objects are faint, making even ground-based rest-frame UV observations challenging. Added to this is the complexity of the gas kinematics and radiation transport, which may cause absorption and emission features to be system-

atically offset from the true systemic redshifts. Here we have used a sample of 15 absorption spectra of bright background QSOs, with more than 600 foreground galaxies within 2 Mpc from the LOS to these QSOs, to probe the IGM close to galaxies at $z \approx 2 - 2.5$. All galaxies have spectroscopic redshifts measured from either the H I Lyα emission line and/or rest-frame UV ISM absorption lines. In addition, for a subset of 49 galaxies redshifts have been measured from rest-frame optical nebular emission lines, which are thought to be unbiased with respect to the systemic redshifts. By utilizing the fact that the mean absorption profiles must be symmetric with respect to the true galaxy redshifts if the galaxies are oriented randomly with respect to the LOS, we calibrated the redshifts measured from rest-frame UV spectral lines.

Our results are summarized in Fig. 2.3. We found that observed Lyα emission redshifts require a systematic shift of $\Delta v_{\mathrm{Ly}\alpha} = -295^{+35}_{-35}\,\mathrm{km\,s^{-1}}$ (i.e. the Lyα lines are redshifted with respect to the systemic velocity) and that ISM absorption redshifts need an offset of $\Delta v_{\mathrm{ISM}} = 145^{+70}_{-35}\,\mathrm{km\,s^{-1}}$ (i.e. the ISM lines are blueshifted). For the nebular lines we found $\Delta v_{\mathrm{nebular}} = -75^{+85}_{-30}\,\mathrm{km\,s^{-1}}$, which is consistent with zero, as expected.

We compared our calibrations to those obtained by direct comparison with nebular emission lines, which should be close to systemic, for a subsample observed with NIRSPEC/Keck. While our measurements for Δv_{ISM} agree, the direct measurement for the systematic offset of the Lyα emission line, $\Delta v_{\mathrm{Ly}\alpha} = -406^{+22}_{-30}\,\mathrm{km\,s^{-1}}$, differs significantly. However, when we applied our method to the NIRSPEC subsample we found $\Delta v_{\mathrm{Ly}\alpha} = -455^{+45}_{-65}\,\mathrm{km\,s^{-1}}$, in agreement with the direct measurement. The discrepancy between the Lyα offsets between the near-IR subsample and the full sample reflects the fact that the line offset depends on galaxy spectral morphology, which is different for the NIRSPEC subsample due to the way in which these galaxies were selected.

After having demonstrated that one can use IGM absorption to measure systematic galaxy redshift offsets, we also estimated upper limits on the random offsets for individual galaxies. In our case these random offsets are likely dominated by intrinsic scatter rather than by measurement errors. We applied random velocity shifts drawn from a Gaussian distribution of varying width to the redshift of each galaxy until the median Lyα absorption profile close to galaxies was significantly affected. This procedure yielded 1σ upper limits on the random offsets of $< 220\,\mathrm{km\,s^{-1}}$ ($< 300\,\mathrm{km\,s^{-1}}$) for Ly$\alpha$ redshifts, using mean (median) statistics, and $< 420\,\mathrm{km\,s^{-1}}$ ($< 430\,\mathrm{km\,s^{-1}}$) for ISM redshifts, which are consistent with direct measurements for the subsample with nebular redshifts.

These results will for example be of interest for future Lyα emitter surveys for which follow-up near-IR spectroscopy will not be available. Such surveys would otherwise have to rely on previous calibrations for LBGs at $z \approx 2 - 3$,

which might not be appropriate for different galaxy populations or redshifts. Without applying our self-calibration technique, the redshifts based on Lyα would likely be systematically wrong by a few hundred $\mathrm{km\,s^{-1}}$, which would make it very difficult to study the galaxies' environments through absorption line spectroscopy of nearby background QSOs or GRBs.

In future papers we will also use the calibrations presented here to study the IGM near the galaxies in our sample.

Acknowledgments

References

Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, ApJ, 584, 45

Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005a, ApJ, 619, 697

Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Reddy N. A., 2005b, ApJ, 629, 636

Bacon R. et al., 2010, Proc. SPIE, 7735, 7

Bergeron J., Boissé P., 1991, A&A, 243, 344

Bowen D. V., Pettini M., Blades J. C., 2002, ApJ, 580, 169

Chen H.-W., Lanzetta K. M., Webb J. K., Barcons X., 1998, ApJ, 498, 77C

Chen H.-W., Lanzetta K. M., Webb J. K., 2001, ApJ, 556, 158

Crighton N. et al., 2010 MNRAS, in press (astro-ph/1006.4385)

Dijkstra M., Haiman Z., Spaans M., 2006a, ApJ, 649, 14

Dijkstra M., Haiman Z., Spaans M., 2006b, ApJ, 649, 37

Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, ApJ, 646, 107

Hansen M., Oh S. P., 2006, MNRAS, 367, 979

Komatsu E. et al., 2009, ApJS, 180, 330

Lanzetta K. M., Bowen D., 1998, ApJ, 357, 321

Lanzetta K. M., Bowen D. V., Tytler D., Webb J. K., 1995, ApJ, 442, 538

McLean I. S. et al., 1998, Proc. SPIE, 3354, 566

McLean I. S., Steidel C. C., Matthews K., Epps H., Adkins S. M., 2008, Proc. SPIE, 7014, 99

Morris S. L., Jannuzi B. T., 2006, MNRAS, 367,1261

Penton S. V., Stocke J. T., Shull J. M., 2002, ApJ, 565, 720

Pettini M., Shapley A. E., Steidel C. C., Cuby J.-G., Dickinson M., Moorwood A. F. M., Adelberger K. L., Giavalisco M., 2001, ApJ, 554, 981

Pieri M. M., Schaye J., Aguirre A., 2006, ApJ, 638, 45

Reddy N. A., Erb D. K., Steidel C. C., Shapley A. E., Adelberger K. L., Pettini M., 2005, ApJ, 633, 748

Ryan-Weber E. V., 2006, MNRAS, 367, 1251

Schaerer D., Verhamme A., 2008, A&A, 480, 369

Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2003, ApJ, 596, 768

Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65

Steidel C. C., Dickinson M., Persson S. E., 1994, ApJ, 437, 75

Steidel C. C., Dickinson M., Meyer D. M., Adelberger K. L., Sembach K. R., 1997, ApJ, 480, 568

Steidel C. C., Shapley A. E., Pettini M., Adelberger K. L., Erb D. K., Reddy N. A., Hunt M. P., 2004, ApJ, 604, 534

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289

Verhamme A., Schaerer D., Maselli A., 2006, A&A, 460, 397

Verhamme A., Schaerer D., Atek H., Tapken C., 2008, A&A, 491, 89

Wilman R. J., Morris S. L., Jannuzi B. T., Davé R., Shone A. M., 2007, MNRAS, 375, 735

Zheng Z., Miralda-Escudé J., 2002, ApJ, 578, 33

3

Neutral hydrogen optical depth near star-forming galaxies at $z \approx 2.4$ in the Keck Baryonic Structure Survey

We study the interface between galaxies and the intergalactic medium by measuring the absorption by neutral hydrogen in the vicinity of star-forming galaxies at $z \approx 2.4$. Our sample consists of 679 rest-frame-UV selected galaxies with spectroscopic redshifts that have impact parameters < 2 (proper) Mpc to the line of sight of one of 15 bright, background QSOs and that fall within the redshift range of its Lyα forest. We present the first 2-D maps of the absorption around galaxies, plotting the median Lyα pixel optical depth as a function of transverse and line of sight separation from galaxies. The Lyα optical depths are measured using an automatic algorithm that takes advantage of all available Lyman series lines. The median optical depth, and hence the median density of atomic hydrogen, drops by more than an order of magnitude around 100 kpc, which is similar to the virial radius of the halos thought to host the galaxies. The median remains enhanced, at the $> 3\sigma$ level, out to at least 2.8 Mpc (i.e. > 9 comoving Mpc), but the scatter at a given distance is large compared with the median excess optical depth, suggesting that the gas is clumpy. We clearly detect two types of redshift space anisotropies. On scales < 200 km s^{-1}, or < 1 Mpc, the absorption is stronger along the line of sight than in the transverse direction. This "finger of God" effect may be partly due to redshift errors, but is probably dominated by gas motions within or very close to the halos. On the other hand, on scales of 1.4 - 2.0 Mpc the absorption is compressed along the line of sight (with $> 3\sigma$ significance), an effect that we attribute to large-scale infall

(i.e. the Kaiser effect). Within 100 (200) kpc, and over $\pm 165\,\mathrm{km\,s^{-1}}$, the covering fraction of gas with Lyα optical depth greater than unity is $100^{+0}_{-32}\%$ ($86^{+14}_{-18}\%$). Absorbers with $\tau_{\mathrm{Ly}\alpha} > 0.1$ are typically closer to galaxies than random. The mean galaxy overdensity around absorbers increases with the optical depth and also as the length scale over which the galaxy overdensity is evaluated is decreased. Absorbers with $\tau_{\mathrm{Ly}\alpha} \sim 1$ reside in regions where the galaxy number density is close to the cosmic mean on scales ≥ 0.25 Mpc.

Olivera Rakic, Joop Schaye, Charles C. Steidel, and Gwen C. Rudie

3.1 Introduction

Gas accretion and galactic winds are two of the most important and poorly understood ingredients of models for the formation and evolution of galaxies. One way to constrain how galaxies get their gas, and to learn about the extent of galactic feedback, is to study the intergalactic medium (IGM) in the galaxies' vicinity. The interface between galaxies and the IGM can be studied either in emission (e.g. Bland & Tully 1988; Lehnert et al. 1999; Ryan-Weber 2006; Borthakur et al. 2010; Steidel et al. 2011) or in absorption against the continuum of background objects such as QSOs (e.g., Lanzetta & Bowen 1990; Bergeron & Boissé 1991; Steidel & Sargent 1992; Steidel et al. 1994; Lanzetta et al. 1995; Steidel et al. 1997; Chen et al. 1998, 2001; Bowen et al. 2002; Penton et al. 2002; Frank et al. 2003; Adelberger et al. 2003, 2005; Pieri et al. 2006; Simcoe et al. 2006; Steidel et al. 2010; Crighton et al. 2011; Prochaska et al. 2011; Kacprzak et al. 2011; Bouche et al. 2011) or galaxies (e.g. Adelberger et al. 2005; Rubin et al. 2010; Steidel et al. 2010; Bordoloi et al. 2011).

Emission from intergalactic gas is very faint and observations are currently mostly limited to low redshifts. At $z \approx 2.7$ Lyα emission was recently seen out to ~ 80 kpc from star-forming galaxies in a stacking analysis by Steidel et al. (2011), and the origin of this light seems to be radiation of the central object scattered by galactic halo gas. However, these observations are limited to the immediate vicinity of galaxies, i.e. ≈ 100 kpc, and are currently feasible only for the H I Lyα transition. On the other hand, studying this gas in absorption is viable at all redshifts as long as there are sufficiently bright background objects. Most importantly, absorption studies are sensitive to gas with several orders of magnitude lower density than emission studies.

The background sources used for absorption line probes of the IGM have traditionally been QSOs, but one may also use background galaxies or gamma ray bursts. While the surface density of background galaxies is much higher than that of QSOs, the quality of the individual spectra is much lower, since at the redshifts discussed in this paper the typical background galaxy is $> 10^3$ times fainter than the brightest QSOs. Studies using galaxies as background sources are therefore confined to analyzing strong absorption lines and generally require stacking many lines of sight. On the other hand, QSOs sufficiently bright for high-resolution, high S/N spectroscopy using 8m-class telescopes are exceedingly rare, but the information obtained from a single line of sight is of exceptional quality, well beyond what could be obtained with even the very brightest galaxy at comparable redshift. We note also that galaxies and QSOs do not provide identical information, as the former are much more extended than the latter. This is particularly relevant for metal lines, which often arise in absorbers with sizes that are comparable or smaller than the half-light radii of galaxies (e.g. Schaye et al. 2007).

In this paper we focus on studying the IGM near star-forming galaxies

at $2.0 < z < 2.8$ using absorption spectra of background QSOs. At present we focus on H I Lyα in the vicinity of galaxies, while a future paper will study the relation between metals and galaxies. Star-forming galaxies in this redshift range can be detected very efficiently based on their rest-frame UV colors (Steidel et al. 2003, 2004) and the same redshift range is ideal for studying many astrophysically interesting lines that lie in the rest-frame UV part of spectrum (e.g. Lyα at 1215.67Å, CIV $\lambda\lambda$ 1548.19,1550.77 Å, OVI $\lambda\lambda$1031.93,1037.616 Å, where the CIV and OVI lines are doublets). At these redshifts the Lyα forest is not as saturated as at $z \gtrsim 4$, and at the same time the absorption systems are not as rare as they are at low redshifts. In addition, this redshift range is exceptionally well-suited for studying the galaxy-IGM interface given that this is when the universal star formation rate was at its peak (e.g. Reddy et al. 2008), and hence we expect the interaction between galaxies and their surroundings to be most vigorous.

The QSO-galaxy data samples used by Adelberger et al. (2003, 2005) are the only surveys comparable in size and scope to the survey presented in this paper. Adelberger et al. (2003) studied the IGM close to 431 Lyman Break Galaxies (LBGs) at $z \approx 3 - 4$ in 8 QSO fields. They found enhanced Lyα absorption within $\approx 1 - 6\,h^{-1}$ comoving Mpc ($\approx 0.3 - 2.1$ proper Mpc) from star-forming galaxies. On the other hand, they found that the region within $\lesssim 0.5\,h^{-1}$ comoving Mpc (≈ 0.2 proper Mpc) contains less neutral hydrogen than the global average. This last result was, however, based on only 3 galaxies. Adelberger et al. (2005) studied the IGM at $1.8 \lesssim z \lesssim 3.3$ with an enlarged sample: 23 QSOs in 12 fields containing 1044 galaxies. They confirmed the earlier result of enhanced absorption within $\lesssim 7h^{-1}$ comoving Mpc (≈ 2.4 proper Mpc) and found that even though most galaxies show enhanced absorption within $\lesssim 0.5\,h^{-1}$ comoving Mpc, a third of their galaxies did not have significant associated Lyα absorption.

In comparison with Adelberger et al. (2003, 2005) we use uniformly excellent QSO spectra taken with the HIRES echelle spectrograph, covering from $\simeq$3100 Å to at least the QSO's CIV emission line. Adelberger et al. (2003) also used HIRES spectra, but at $z > 3$ the surface density of galaxies with $R_{\rm s} < 25.5$ is $\sim 4 - 5$ times smaller than at $z = 2.4$, and the QSO spectra did not cover higher order Lyman lines, which are important for the recovery of optical depths in saturated Lyα lines. Adelberger et al. (2005) used HIRES spectra as well as lower-resolution spectra of fainter QSOs obtained using LRIS and ESI.

This paper is organized as follows. We describe our data sample in Section 2.2. In Section 3.3 we discuss the so-called pixel optical depth method that we use to analyze the QSO spectra. The distribution of Lyα absorption as a function of transverse and line of sight (LOS) separation from galaxies is presented in Section 3.4, while the distribution of galaxies around absorbers is described in Section 3.6. Finally, we conclude in Section 4.5.

Throughout this work we use $\Omega_{\rm m} = 0.258$, $\Omega_{\Lambda} = 0.742$, $\Omega_{\rm b} = 0.0418$,

and $h = 0.719$ (Komatsu et al. 2009). When referring to distances in proper (comoving) units we denote them as pMpc (cMpc).

3.2 Data

3.2.1 The Keck Baryonic Structure Survey

The Keck Baryonic Structure Survey (KBSS; Steidel et al 2012) is a new survey which combines high precision studies of the IGM with targeted galaxy redshift surveys of the surrounding volumes, expressly designed to establish the galaxy/IGM connection in the redshift range $1.8 \lesssim z \lesssim 3$. The KBSS fields are centered on 15 background QSOs (see Table 3.1) that are among the most luminous known ($L_{bol} \gtrsim 10^{14}\ L_\odot$) in the redshift range $2.5 \lesssim z \lesssim 2.85$. The QSO redshifts were chosen to maximize the information content and redshift path sampled by their absorption line spectra. Toward that end, the QSO spectra used in the KBSS are of unprecedented quality, combining archival high resolution spectra from Keck/HIRES (and, in 3 cases, archival VLT/UVES spectra as well) with new Keck/HIRES observations to produce the final co-added spectra summarized in Table 3.1.

Within each KBSS field, UV color selection techniques (Steidel et al. 2003; Adelberger et al. 2004) were used to tune the galaxy redshift selection function so as to cover the same optimal range of redshifts probed by the QSO spectra. The spectroscopic follow-up using Keck/LRIS was carried out over relatively small solid angles surrounding the QSO sightlines, but typically 8 to 10 slit masks with nearly identical footprint were obtained in each field, leading to a high level of spectroscopic completeness and a very dense sampling of the survey volumes surrounding the QSO sightlines. The full KBSS galaxy sample contains 2188 spectroscopically identified galaxies in the redshift range $1.5 \leq z \leq 3.5$, with $\langle z \rangle = 2.36 \pm 0.42$. The spectroscopic sample was limited to those with apparent magnitude $\mathcal{R} \leq 25.5$, where galaxies in the range $23 \leq \mathcal{R} \leq 24.5$ and those within 1 arcminute of a QSO sightline were given highest priority in designing slit masks.

The effective survey area is $720\ \mathrm{arcmin}^2$ (i.e. 0.2 square degrees), and the average surface density of galaxies with spectroscopic redshifts is $3.1\ \mathrm{arcmin}^{-2}$ (within 2.5 pMpc from QSO sightlines).

In this paper, we focus on a subset of 679 galaxies that satisfy the following criteria for redshift and projected distance from the relevant QSO sightline:

1) their redshifts are within the Lyα forest range, defined as:

$$(1 + z_{\mathrm{QSO}})\frac{\lambda_{\mathrm{Ly}\beta}}{\lambda_{\mathrm{Ly}\alpha}} - 1 < z < z_{\mathrm{QSO}} - (1 + z_{\mathrm{QSO}})\frac{3000\ \mathrm{km\,s}^{-1}}{c}, \quad (3.1)$$

where z_{QSO} is the QSO redshift, and $\lambda_{\mathrm{Ly}\alpha}$ and $\lambda_{\mathrm{Ly}\beta}$ are the rest-frame wavelengths of the hydrogen Lyα (1215.67Å) and Lyβ (1025.72Å) lines, respectively (for a discussion of the limits in this equation, see §3.2.2);

Table 3.1 – QSO Details

QSO[a]	z[b]	G^c_{AB}	$z^d_{\min}$	$z^e_{\max}$	$\lambda^f_{\min}$[Å]	$\lambda^g_{\max}$[Å]	S/N$^h_{HI}$	$\Delta\Omega$[i] $\times$	N$^j_{gal}$
Q0100+13	2.7210	16.6	2.196	2.654	3125	8595	67	$5.6' \times 7.6'$	35
Q0105+1619	2.6518	16.9	2.102	2.603	3225	6041	112	$5.4' \times 7.4'$	59
Q0142-09	2.7434	16.9	2.210	2.671	3100	6156	65	$5.4' \times 7.4'$	45
Q0207-003	2.8726	16.7	2.264	2.848	3100	9220	71	$5.4' \times 7.0'$	38
Q0449-1645	2.6840	17.0	2.094	2.651	3150	5966	67	$5.0' \times 6.5'$	51
Q0821+3107	2.6156	17.3	2.150	2.584	3235	5967	42	$5.4' \times 7.4'$	32
Q1009+29	2.6520	16.0	2.095	2.593	3160	7032	87	$5.2' \times 7.2'$	29
Q1217+499	2.7040	17.1	2.122	2.680	3075	7000	64	$5.1' \times 6.9'$	37
Q1442+2931	2.6603	17.0	2.078	2.638	3155	6150	83	$5.4' \times 7.5'$	44
Q1549+1933	2.8443	16.3	2.238	2.814	3160	9762	164	$5.2' \times 7.1'$	44
HS1603+3820	2.5510	15.9	1.974	2.454	3185	9762	87	$5.4' \times 7.2'$	41
Q1623-KP77	2.5353	17.4	1.983	2.505	3125	6075	43	$16.1' \times 11.6'$	54
Q1700+64	2.7513	16.1	2.168	2.709	3145	9981	92	$11.5' \times 11.0'$	64
Q2206-199	2.5730	17.5	2.009	2.541	3047	10088	73	$5.4' \times 7.5'$	46
Q2343+12	2.5730	17.0	2.012	2.546	3160	10087	57	$22.5' \times 8.5'$	60

[a]QSO name.
[b]QSO redshift.
[c]QSO AB magnitude.
[d]Minimum galaxy redshift considered in this field.
[e]Maximum galaxy redshift considered in this field.
[f]The lowest available wavelength in each spectrum.
[g]The highest available wavelength in each spectrum.
[h]Median S/N per pixel in the analyzed Lyα region of each spectrum.
[i]Spectroscopically observed area.
[j]Number of galaxies with spectroscopic redshifts in the field.

2) they are within 2 pMpc ($\sim 4'$ or $\sim 5h^{-1}$ cMpc at $z = 2.4$), the transverse distance to which there is significant coverage in all 15 KBSS fields.

Figure 3.1 shows the number of galaxies as a function of their (proper) distance from the QSO sightlines.

We note that 3 of the 15 KBSS fields (1623+268, 1700+64, 2343+12) were also included in the study by Adelberger et al (2005). However, since 2005 the data have been increased in both quantity and quality (for both the QSO spectra and the galaxy surveys) in all 3 of these fields. Further details on the KBSS survey and its data products will be described by Steidel et al (2012).

The smallest impact parameter in the present sample is 55 pkpc, with 29, 106, and 267 galaxies having impact parameters smaller than 200, 500, and 1000 pkpc, respectively.

3.2.1.1   Redshifts

The redshifts of the vast majority (608 out of 679) of the galaxies in our sample are derived from interstellar absorption lines and/or the Lyα emission line.
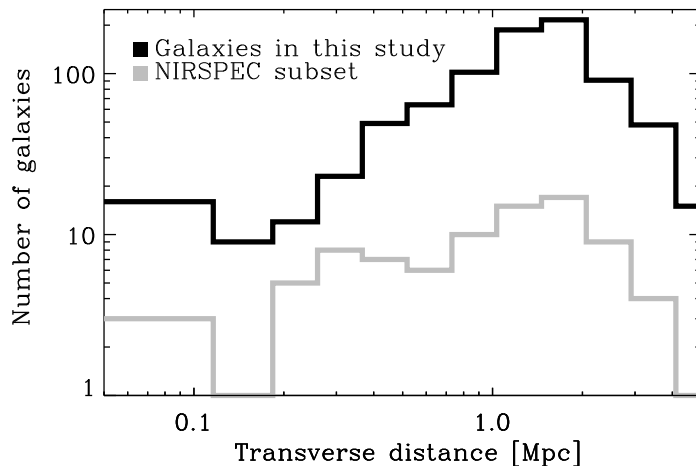
Figure 3.1 − Number of galaxies as a function of proper distance from the line of sight to the background QSO. We do not use galaxies with impact parameters larger than 2 pMpc since such galaxies were only targeted for a fraction of our fields.

These lines were measured from low-resolution (FWHM $\approx 370\,\mathrm{km\,s^{-1}}$) multislit spectra taken between 2000 and 2010 with LRIS-B on the Keck I and II telescopes. However, ideally one would want to measure the galaxy redshifts from the nebular emission lines ([O 2] $\lambda 3727$, Hα, Hβ, [O 3] $\lambda\lambda 4959, 5007$) since those originate in stellar H 2 regions and are more likely to correspond to the galaxies' systemic redshifts. Erb et al. (2006b) have used NIRSPEC, a near-IR instrument on Keck II, to obtain higher resolution (FWHM $\approx 240\ \mathrm{km\,s^{-1}}$) spectra than achieved by LRIS-B and have used these to measure nebular redshifts for 110 galaxies. Our sample contains 71 of those galaxies and we take their systemic redshifts to be equal to the nebular redshifts, i.e. $z_{\mathrm{gal}} = z_{\mathrm{neb}}$. Marginally resolved lines have $\sigma \approx 160\ \mathrm{km\,s^{-1}}$ for LRIS-B and $\sigma \approx 100\ \mathrm{km\,s^{-1}}$ for NIRSPEC.

For the galaxies without near-IR observations we estimate the systemic redshifts using the empirical relations of Rakic et al. (2011). Using the same data as analyzed here, they calibrated galaxy redshifts measured from rest-frame UV lines (Lyman-α emission, $z_{\mathrm{Ly}\alpha}$, and interstellar absorption, z_{ISM}) by utilizing the fact that the mean Lyα absorption profiles around the galaxies, as seen in spectra of background QSOs, must be symmetric with respect to the true galaxy redshifts if the galaxies are oriented randomly with respect to the lines of sight to the background objects. The following values represent their best fits to the data:

a) for galaxies without available interstellar absorption lines (70 objects):

$$z_{\mathrm{gal}} = z_{\mathrm{Ly}\alpha} - 295^{+35}_{-35}\,\mathrm{km\,s^{-1}} \tag{3.2}$$

b) for galaxies without detected Lyα emission lines (346 objects):

$$z_{\mathrm{gal}} = z_{\mathrm{ISM}} + 145^{+70}_{-35}\,\mathrm{km\,s^{-1}} \qquad (3.3)$$

c) for galaxies for which both interstellar absorption and Lyα emission are detected (263 objects) we use the arithmetic mean of the above expressions.

These offsets yield redshift estimates free of velocity systematics. Repeated observations of the same galaxies suggest that the typical measurement uncertainties are $\approx 60\,\mathrm{km\,s^{-1}}$ for z_{neb}. Comparison of redshifts inferred from rest-frame UV lines with those from nebular lines for the subset of 89 galaxies that have been observed in the near IR, shows that the rest-frame UV inferred redshifts have a scatter of $\approx 130\,\mathrm{km\,s^{-1}}$.

Figure 3.2 shows a histogram of the galaxy redshifts, both for the full sample and for the subsample with redshifts measured from near-IR lines. The median redshifts of the full and near-IR samples are 2.36 and 2.29, respectively. The mean and standard deviations of the redshift distributions for the two samples are 2.36 ± 0.17 and 2.28 ± 0.13, respectively.

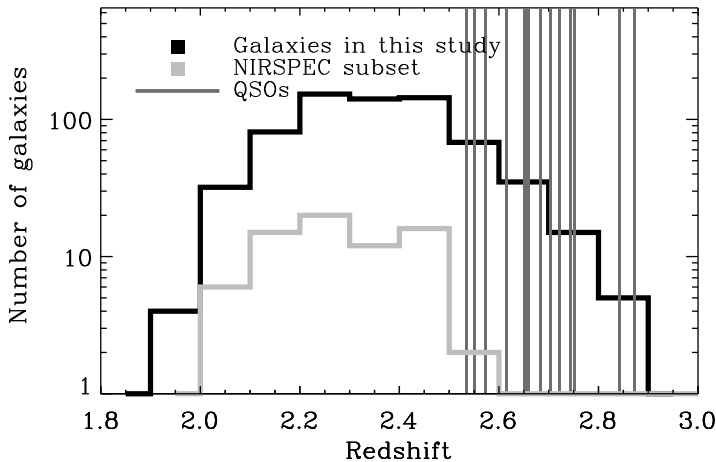**Figure 3.2** – Redshift distribution of QSOs (vertical lines) and galaxies that are used in our study, i.e. of galaxies that are in the Lyman-α forest redshift ranges of QSOs in their fields and that have impact parameters < 2 pMpc. The black histogram shows the redshifts of all galaxies (679 objects in total), and the grey histogram just those with redshifts measured from near-IR spectra (71 galaxies).

3.2.1.2   *Physical properties*

These UV-selected star-forming galaxies have stellar masses $\langle \log_{10}\mathrm{M_*/M_\odot} \rangle = 10.08 \pm 0.51$ (Shapley et al. 2005). The typical star formation rates (SFRs) are $\sim 30\,\mathrm{M_\odot\,yr^{-1}}$, where the SFRs of individual objects vary from ≈ 7 to $\approx 200\,\mathrm{M_\odot\,yr^{-1}}$, and the mean SFR surface density is $\langle \Sigma_{\mathrm{SFR}} \rangle = 2.9$

$M_\odot$ yr^{-1} kpc^{-2} (Erb et al. 2006). These stellar mass and SFR estimates assume a Chabrier (2003) IMF. The galaxies show a correlation between their stellar mass and metallicity, but the relation is offset by 0.3 dex as compared to the local relation, with the same stellar mass galaxies having lower metallicity at $z \approx 2.4$ (Erb et al. 2006c). Typical metallicities range from $\approx 0.3\, Z_\odot$ for galaxies with $\langle M_* \rangle = 2.7 \times 10^9 M_\odot$ to $\approx Z_\odot$ for galaxies with $\langle M_* \rangle = 1 \times 10^{11} M_\odot$.

As discussed in section 3.2.1.1, ISM absorption lines are almost always blue-shifted with respect to the galaxy systemic redshift, and the Lyα emission line is always redshifted. These observed velocity offsets suggest that galaxy-scale outflows, with velocities of hundreds of km s^{-1}, are the norm in these star-forming galaxies.

Trainor et al. (2011, in preparation) use the Millennium simulation (Springel 2005), together with a clustering analysis, to connect galaxies from KBSS to dark matter halos. They find that this type of galaxy resides in halos with masses above $10^{11.75}\, h^{-1} M_\odot$, with a median halo mass of $\sim 10^{12}\, h^{-1} M_\odot$. The corresponding virial radii are ≈ 85 and ≈ 106 pkpc, respectively, with circular velocities ≈ 197 km s^{-1} and ≈ 244 km s^{-1}.

3.2.2 QSO spectra

The typical resolution of the QSO spectra is R $\approx 36,000$, and they were rebinned to pixels of 2.8 km s^{-1}. The spectra were reduced using T. Barlow's MAKEE package and the continua were normalized using low-order spline fits. Further details about the QSO observations will be given in Steidel et al. (in preparation). The QSO redshift distribution can be seen in Figure 3.2 and a summary of the properties of the final spectra is given in Table 3.1.

The redshift range that we consider when studying Lyα absorption in the spectrum of a QSO at redshift $z_{\rm QSO}$ is given by equation 3.1). The lower limit ensures that only Lyα redwards of the QSO's Lyβ emission line is considered, thus avoiding any confusion with the Lyβ forest from gas at higher redshifts. The upper limit is set to avoid contamination of the Lyα forest by material associated with the QSO and to avoid the QSO proximity effect. We verified that excluding 5,000 rather than 3,000 km s^{-1} gives nearly identical results.

The median S/N in the Lyα forest regions of the spectra ranges from 42 to 163 (Table 3.1).

A number of QSO spectra (5 out of 15) contain damped Lyman-α systems (DLAs) and sub-DLAs within the considered redshift range. We divided out the damping wings after fitting the Voigt profiles (see Rudie et al., in preparation, for more details on this procedure). The saturated cores of the (sub-)DLAs were, however, flagged and excluded from the analysis, but this does not have a significant effect on our results.

3.3   Pixel Optical Depths

The goal of the pixel optical depth (POD) method is to use automatic algorithms (as opposed to manual fits) to find the best estimate of the optical depth of the H I Lyα absorption line and those of various metal transitions as a function of redshift. Obtaining accurate pixel optical depths can be challenging due to the presence of noise, errors in the continuum fit, contamination and saturation.

It is more instructive to plot optical depths than normalized fluxes because the optical depth is proportional to the column density of the absorbing species, while the flux is exponentially sensitive to this density. Hence, a power-law column density profile translates into a power-law optical depth profile and a lognormal column density distribution becomes a lognormal optical depth distribution. The shapes of the profiles and distributions common in nature are therefore preserved, which is not the case if we work with the flux. Of course, if the spectral resolution is coarser than the typical line width, or if the wavelength coverage does not allow for the use of weaker components of multiplets to measure the optical depth of saturated lines, then the interpretation of the recovered optical depths is more complicated. Our spectral resolution is, however, sufficient to resolve the Lyα lines and our wavelength coverage is sufficient to recover the optical depth of most saturated lines.

Another advantage of optical depths is that their dynamic range is unconstrained, whereas the normalized flux is confined to vary between zero and one. The large dynamic range does imply that one must use median rather than mean statistics, because the mean optical depth merely reflects the high optical depth tail of the distribution. While this problem could also be solved by using the mean logarithm of the optical depth, that would not resolve another issue resulting from the large intrinsic dynamic range: the observed optical depth distribution is cut off at the low absorption end by noise and at the high end by line saturation (and thus ultimately also by noise). However, errors in the tails of the distribution do not matter if we use median statistics (or other percentiles, provided they stay away from the noise).

The POD statistics method was introduced by Cowie & Songaila (1998) (see also Songaila 1998), and further improved by Ellison et al. (2000), Schaye et al. (2000b), and Aguirre et al. (2002). The POD method used in this study is the one developed and tested by Aguirre et al. (2002), and is identical to that of Ellison et al. (2000) for the case of Lyα.

The H I Lyα optical depth in each pixel is calculated from the normalized flux $F(\lambda)$,

$$\tau_{\mathrm{Ly}\alpha}(\lambda) = -\ln(F(\lambda)). \tag{3.4}$$

Pixels with $\tau_{\mathrm{Ly}\alpha} < 0$ (which can occur in the presence of noise) are assigned an optical depth of $\tau_{\mathrm{min}} = 10^{-5}$, which is smaller than any recovered optical

depth.

Pixels that have $F(\lambda) \leq N_\sigma \sigma_\lambda$ are considered saturated, where σ_λ is the rms noise amplitude at the given pixel and N_σ is a parameter that we set to 3 (see Aguirre et al. 2002, for more details). One advantage of the POD method is that it is easy to recover a good estimate of the Lyα optical depth in these saturated pixels by using the available higher order Lyman lines. The recovered Lyα optical depth is then given by

$$\tau_{\mathrm{Ly}\alpha}^{\mathrm{rec}} = \min\{\tau_{\mathrm{Ly}n} f_{\mathrm{Ly}\alpha} \lambda_{\mathrm{Ly}\alpha} / f_{\mathrm{Ly}n} \lambda_{\mathrm{Ly}n}\}, \tag{3.5}$$

where $f_{\mathrm{Ly}n}$ is the oscillator strength of the nth order Lyman line and $\lambda_{\mathrm{Ly}n}$ is its rest wavelength ($n = 1$ corresponds to α, $n = 2$ to β, etc.). Taking the minimum optical depth (Equation 3.5) minimizes the effect of contamination by other lines. Higher order lines used for Lyα optical depth recovery are those that lie in the wavelength range of the spectrum, and for which $N_\sigma \sigma_n \leq F(\lambda_{\mathrm{Ly}n}) \leq 1 - N_\sigma \sigma_n$, where σ_n is the noise at $\lambda_{\mathrm{Ly}n}$. If none of the available higher order lines satisfy this criterion, or if none of the higher order lines are available, then the pixel is set to a high optical depth (we use $\tau_{\max} = 10^4$). An example of the recovered optical depth in one QSO spectrum is shown in Figure 3.3, together with the positions of galaxies in the same field.

Setting pixels to $\tau_{\min}$ and $\tau_{\max}$ allows them to correctly influence the median and other percentiles. The actual values they are set to are unimportant as long as $\tau_{\min} < \tau_{\mathrm{Ly}\alpha}^{\mathrm{rec}} < \tau_{\max}$ for the percentile of interest.

We use the POD method because: a) It is fast, robust, and automatic, which means it can deal with large amounts of data, both observed and simulated; b) It makes it straightforward to exploit the full dynamic range measurable from our spectra; c) The fact that the optical depth is directly proportional to the column density of neutral hydrogen makes it easy to interpret the results (see §3.4.5). While the POD method has clear advantages, it is important to note that it does not replace the more traditional method of decomposing the absorption spectra into Voigt profile components. For example, in its simplest form, the POD method does not use any information about the line widths, which are clearly of interest as they contain important information about the temperature and the small-scale velocity structure of the absorbing gas. We will present an analysis of the KBSS data based on Voigt profile decompositions in Rudie et al. (in preparation).

3.4 Lyα Absorption near Galaxies

After recovering the optical depths in each pixel of a QSO spectrum, we can build a map of the galaxies' average surroundings, as a function of the transverse distance from the line of sight (LOS), i.e. the impact parameter, and the velocity separation from the galaxy along the LOS. We sometimes convert velocity separation into distance by assuming that it is due to the Hubble flow. However, as we will show in sections 3.4.1 and 3.4.2, this is a
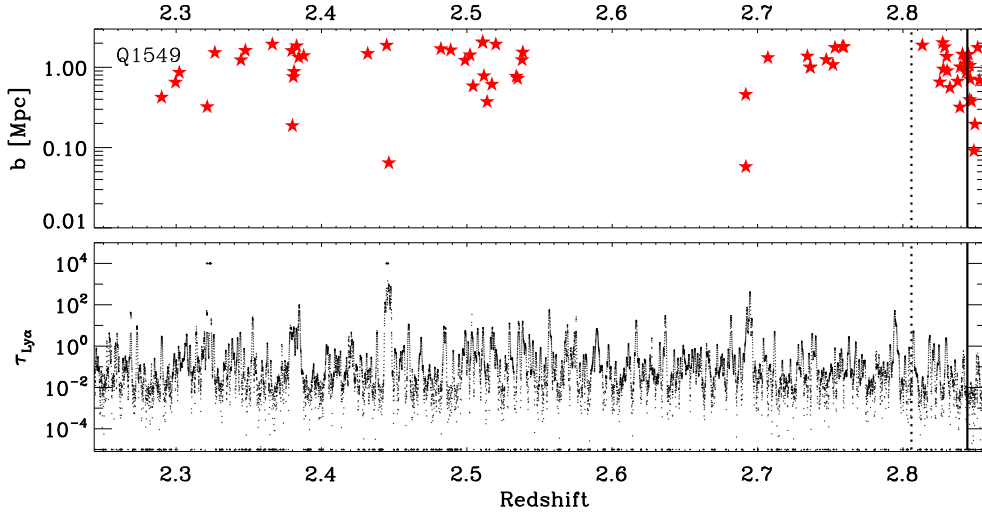
Figure 3.3 − *Lower panel:* Recovered H I Lyα pixel optical depth (black dots) as a function of redshift for the spectrum of Q1549+1933. The black dots at $\tau_{\mathrm{HI}} = 10^{-5}$ are pixels with normalised flux greater than unity that have their values set to $\tau_{\min}$ (see the text for details). The points at 10^4 show pixels that are saturated and that have insufficient available higher order lines for the recovery of their optical depths. The solid vertical line indicates the redshift of the QSO. The dashed vertical line, which is separated from the solid line by 3,000 km s^{-1}, indicates the maximum redshift we consider for this spectrum. *Upper panel:* Impact parameters of galaxies in this field (stars) as a function of spectroscopic redshift.

poor approximation at small velocity differences from galaxies. We will refer to distances computed under the assumption of zero peculiar velocities (and zero redshift errors) as "Hubble distances".

Different galaxies and QSO fields are combined, meaning that the median POD for a given impact parameter and velocity difference is estimated over all galaxies irrespective of which field they come from, without applying any weighting. Each galaxy therefore provides an array of pixels with varying velocity difference but fixed impact parameter.

For reference, we note that 0.2, 2.1, and 9.9 percent of the pixels in our sample are at 3-D Hubble distances smaller than 200, 500, and 1000 pkpc, respectively.

3.4.1   2-D Map of Lyα Absorption

The left panel of Figure 3.4 shows the logarithm of the median $\tau_{\mathrm{Ly}\alpha}$ as a function of the transverse and the LOS separations from galaxies. The distance bin size in this plot is 200 by 200 pkpc (which corresponds to ≈ 46 km s^{-1} along the LOS), and the image has been smoothed with a 2-D Gaussian with FWHM equal to the bin size. Note that we take a galaxy-centered approach in constructing this image. Each galaxy contributes a column of pixels whose position along the x-axis corresponds to the galaxy's impact parameter. A
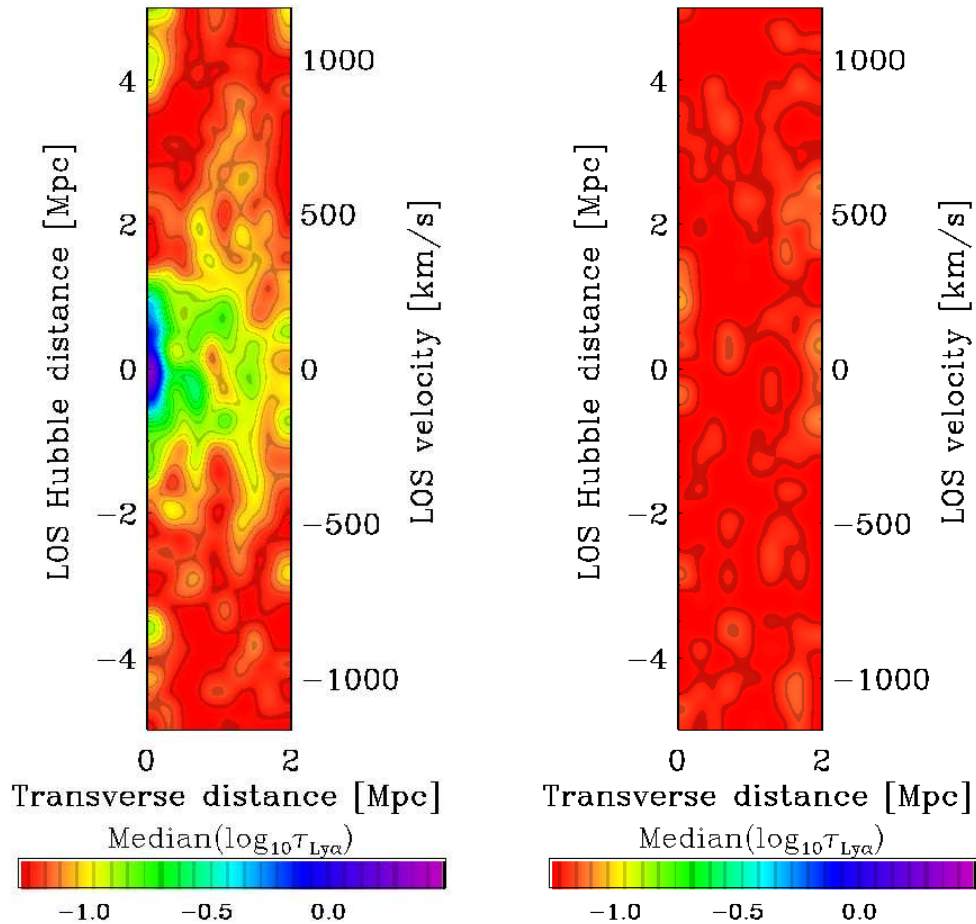
Figure 3.4 − *Left:* Median H I Lyα absorption as a function of transverse and LOS distance from $\langle z \rangle \approx 2.4$ star-forming galaxies. The bin size is 200 pkpc and the images have been smoothed by a 2-D Gaussian with FWHM equal to the bin size. The lower limit of the color scale corresponds to the median optical depth of all pixels, while the upper limit corresponds to the highest optical depth in the map. Absorption is clearly enhanced close to galaxies, out to at least 2 pMpc in the transverse direction, but only out to ≈ 1.5 pMpc along the LOS. This anisotropy suggests large-scale infall of gas. On the other hand, on small scales the absorption declines more rapidly in the transverse direction than in the LOS direction. *Right:* Results after randomizing the galaxy redshifts while keeping their impact parameters fixed. The fact that the correlation does not vary systematically with distance indicates that the features in the left panel are real.

single pixel can be used multiple times: once for each galaxy whose separation from the pixel falls within the range plotted. This is the first published 2-D absorption map around galaxies. The map shows a strong correlation between the Lyα absorption strength and the distance to the galaxies.

The right panel of Figure 3.4 shows the same data, after randomizing the galaxy redshifts. We randomize the galaxy redshifts within the Lyα
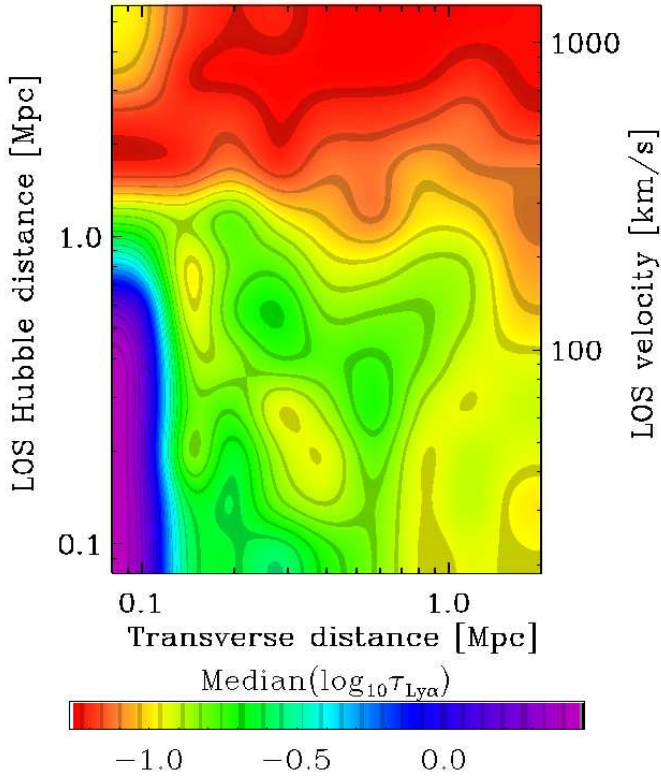
Figure 3.5 – Similar to Fig. 3.4, but we do not separate positive and negative velocity differences between absorbers and galaxies and we use logarithmic bins and axes (bins are separated by 0.15 dex). The map is smoothed by a 2-D Gaussian with FWHM equal to the bin size. The color scale is saturated at the high optical depth end (at the smallest impact parameters the optical depth is $\log_{10} \approx 1.5$).

forest region of each QSO spectrum while keeping their impact parameters unchanged in order to preserve the number of pixels per galaxy. In this way we can estimate the magnitude of the fluctuations due to noise, i.e., in the absence of correlations between the locations of galaxies and absorbers. We can see that the signal is lost, which implies that the features seen in the left panel are real. Finally, we note that because a single galaxy contributes a full column of pixels at its impact parameter, bins along the LOS are somewhat correlated. On the other hand, bins in the transverse direction are independent. In Appendix 3.B we demonstrate that along the LOS the errors are significantly correlated for scales $\lesssim 10^2$ km s^{-1}.

In Figure 3.4 we kept the positive and negative velocity differences between absorbers and galaxies separated, which gives insight into the amount of noise and sample variance. However, given that the Universe is statistically isotropic (and our observations are consistent with this assumption), we can increase the S/N by considering only absolute velocity differences. Figure 3.5 shows a map of gas around galaxies by taking into account only absolute velocity separations. We use logarithmic axes, which brings out the

small-scale anisotropy more clearly.

The most prominent feature is the region of strongly enhanced absorption that extends to $\sim 10^2$ pkpc in the transverse direction, but to ~ 1 pMpc ($\sim 200\,\mathrm{km\,s^{-1}}$) along the LOS. This redshift space distortion (called the "finger of God" effect in galaxy redshift surveys) could be due to redshift errors and/or peculiar velocities. A more subtle anisotropy is visible on large scales. In the transverse direction the correlation between absorption strength and galaxy position persists out to the maximum impact parameter we consider, 2 pMpc, whereas it becomes very weak beyond $\approx 1.5\,\mathrm{pMpc}$ ($\approx 300\,\mathrm{km\,s^{-1}}$) along the LOS, as is most clearly visible in Figure 3.4. If real, such a feature would imply infall of gas on large scales, i.e. the Kaiser (1987) effect. We will examine the significance of these anisotropies in the next section.

3.4.2 Redshift space distortions

Figure 3.6 shows "cuts" along the first 9 rows and columns of the 2-D map shown in Figure 3.5, spanning 0–2 pMpc. The red circles and grey squares show the profiles along the transverse and LOS, respectively. As indicated in the insets, these cuts correspond, respectively, to the horizontal and vertical strips in Fig. 3.5. Observe that the transverse and LOS directions are identical for the nth data point in the nth panel, which corresponds to the intersection of the horizontal and vertical strips shown in the insets. In other words, where the horizontal and vertical strips meet, the data point is replicated as both a red and a black symbol. Note also that the nth red circle (black square) of the mth panel is identical to the mth black square (red circle) of the nth panel. The figure thus contains redundant information. For example, the LOS direction (i.e. black squares) in the 1st panel shows all the 1st data points appearing in the transverse direction (i.e. red circles) of the other panels. Similarly, the 2nd panel shows all the 2nd transverse data points, etc.

As LOS separations have been computed under the assumption of pure Hubble flow, any significant difference between the red and black curves must be due to redshift space distortions (assuming the Universe is isotropic in a statistical sense). By comparing the two curves in each panel, we can therefore identify redshift space distortions and assess their significance.

The error bars in Figure 3.6, as well as those in subsequent plots, were computed by bootstrapping the galaxy sample 1,000 times. That is, for each bootstrap realization we randomly select galaxies, where each galaxy could be selected multiple times, until the number of galaxies from the original set is reached. We calculate the results for each realization and the errors show the 1σ confidence interval. As demonstrated in Appendix 3.B, along the LOS the errors are correlated for separations $\lesssim 10^2$ km s^{-1} (i.e. black squares in each panel are correlated on these scales), but not in the transverse direction (i.e. red circles in each panel are independent).
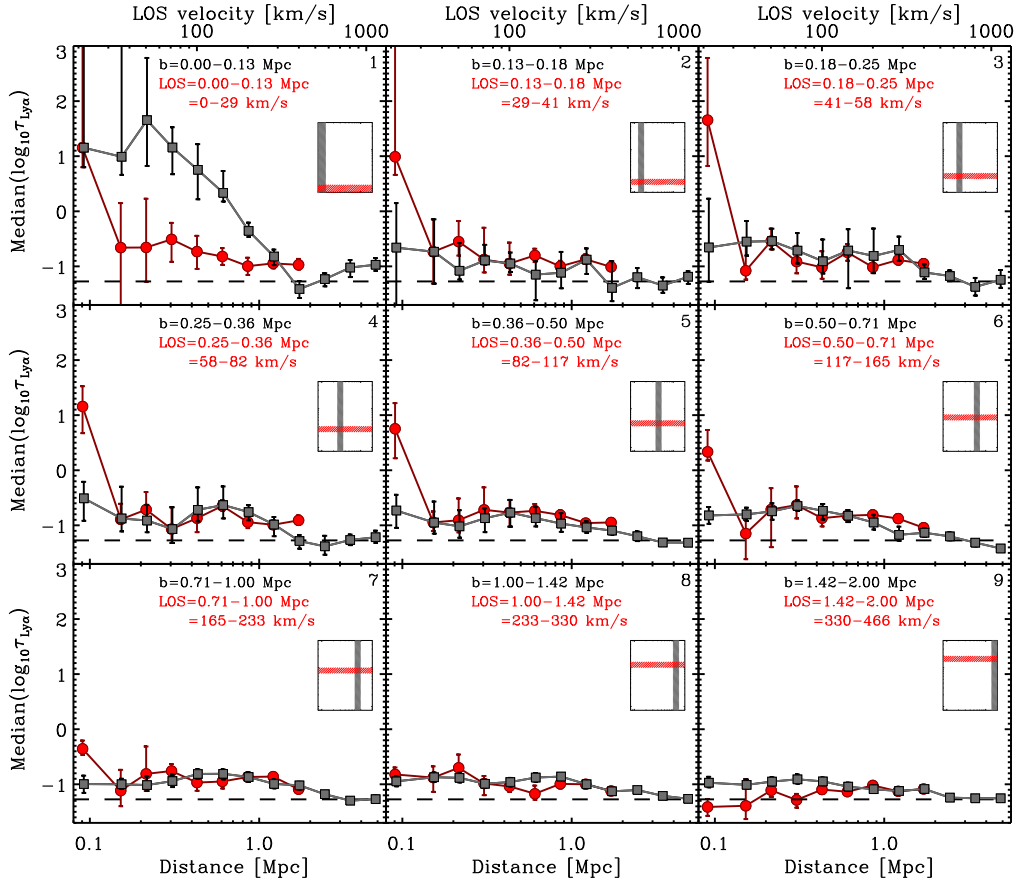
Figure 3.6 – Cuts through the 2-D map from Fig. 3.5, where red circles show median profiles in the transverse direction (i.e. parallel to the x-axis of Fig. 3.5), and grey squares show cuts along the LOS (i.e. parallel to the y-axis of Fig. 3.5). The range of LOS separations resp. impact parameters included in the strips along the y- resp. x-axis of Fig. 3.5 is indicated in each panel. Distances along the LOS have been computed from the corresponding velocity interval under the assumption of pure Hubble flow. Hence, differences between the red and black curves indicate redshift space distortions. The dashed horizontal line shows the median Lyα optical depth of all the pixels in the spectra. With the exception of the last red circles and the corresponding black squares, which correspond to $1.42 - 2.00$ pMpc, the differences between the red and black curves in the first 8 panels indicate that the signal has been smoothed in the LOS direction, probably as a result of redshift errors and, most importantly, small-scale peculiar velocities. On the other hand, the differences revealed by the most distant points in the first six panels or, equivalently (see text), by the first six points in the last panel, imply compression along the LOS. This compression indicates the presence of large-scale infall onto the galaxies.

The number of galaxies per transverse bin depends slightly on the velocity difference, because galaxies separated by Δv km s^{-1} from the Lyα forest region (as defined in section 3.2.2) will still contribute pixels to bins with velocity separations greater than Δv. From small to large impact parameters, the number of galaxies contributing to the 1st (9th) velocity bin is 14 (16), 8 (8), 11 (12), 22 (23), 47 (48), 58 (62), 96 (99), 175 (180), 202 (211). Thus,

the black squares in the first panel, as well as the first red circle in all panels, are based on 14 – 16 galaxies. Similarly, the black squares in the 9th panel, as well as the last red circle in all panels, are based on 202 – 211 galaxies.

The first panel, which shows cuts near the axes of Figure 3.5, clearly shows that out to ~ 1 pMpc, which translates into $\lesssim 200\,\mathrm{km\,s^{-1}}$ along the LOS (see the top x-axis), the absorption is stronger along the LOS (black squares) than in the transverse direction (red circles). In panels 2 – 7, which correspond to strips offset by 0.13 – 1.00 pMpc from the axes in Figure 3.5, the smearing in the LOS direction manifests itself as enhanced absorption at small impact parameters relative to the signal at small LOS separations. The confidence level associated with the detected discrepancy between the two directions, which we estimate as the fraction of bootstrap realizations in which the sign of the discrepancy is reversed, is at least 99% for each data point out to 1 pMpc (i.e. points 2–7 of the first panel).

The differences on scales $\lesssim 1$ pMpc ($\lesssim 233\mathrm{km\,s^{-1}}$) can be explained by two effects. Firstly, we expect gas in and around galaxy halos to have peculiar velocities comparable to the circular velocity of the halos, i.e. $\simeq 200\,\mathrm{km\,s^{-1}}$ (and possibly significantly higher for outflowing gas). Secondly, as discussed in Section 3.2.1.1, there are random errors in the redshift measurements of $\approx 130\,\mathrm{km\,s^{-1}}$ for LRIS redshifts, and $\approx 60\,\mathrm{km\,s^{-1}}$ for the NIRSPEC subsample. These two effects smooth the signal in velocity space on the scale that is a result of the combination of these velocities. In Appendix 3.A we show that redshift errors may be able to account for the observed smoothing. However, this does not prove the observed elongation along the LOS is due to redshift errors rather than peculiar velocities. In fact, since the elongation is more extended than the typical redshift errors ($\sim 200\,\mathrm{km\,s^{-1}}$ vs. $\approx 125\,\mathrm{km\,s^{-1}}$), but similar to the expected circular velocities ($\approx 200\,\mathrm{km\,s^{-1}}$), it is likely that the finger of God effect is dominated by small-scale peculiar velocity gradients due to virial motions, infall, and/or outflows.

On the other hand, at distances > 1.42 pMpc the situation is reversed: the absorption is compressed in the LOS direction. For impact parameters 1.42 – 2 pMpc and LOS separations 0– 0.71 pMpc ($0 - 165\,\mathrm{km\,s^{-1}}$; 9th (i.e. last) red circles in panels 1–6) the absorption is stronger than for impact parameters 0– 0.71 pMpc and LOS separations 1.42 – 2 pMpc (9th black square in panels 1–6). The same information is collected in the last panel, which corresponds to strips separated by 1.42 – 2.00 pMpc from the axes. The first six black squares are higher than the corresponding red circles, which implies that the absorption is enhanced along the LOS relative to the direction transverse to the LOS. Observe that this enhancement is absent from all other panels. The confidence level with which this discrepancy is detected is, from small to large scales, 99.4%, 77.9%, 83.2%, 99.8%, 91.6% and 82.6% for points 1–6, respectively.

This compression along the LOS can be explained by large-scale infall, i.e. the Kaiser effect. As the absorbing gas must be cool, i.e. $T \sim 10^4$ K $\ll \mathrm{T_{vir}}$,

to be visible in HI, this form of gas accretion could be called "cold accretion" (e.g. Kereš et al. 2005), although we note that this term is most often used in the context of cold streams within the virial radii of halos hosting galaxies. We conclude that we have an unambiguous and highly significant detection of cool gas falling towards star-forming galaxies at $z \approx 2.4$.

The dashed, horizontal lines in Figure 3.6 indicate the median Lyα optical depth of all pixels in the Lyα forest regions of the QSO spectra. In the transverse direction we do not probe sufficiently large distances to see the signal disappear: for all but the last panel the red curves stay above the dashed lines out to impact parameters of 2 pMpc. In the LOS direction (black curves) we do see convergence for separations $\gtrsim$ 3 pMpc. The last black square in the first panel is an outlier, but note that it is based on only 16 galaxies, whereas the same points in panels 5-9 are based on 49–217 galaxies. Indeed, according to the redshift randomization method described in the next section, the last black square in the first panel is only a 1.8σ outlier, whereas the last red circles of panels 1-6 (or, equivalently, the first 6 black squares in the last panel) represent detections of excess absorption with significance varying between 2.6σ and $> 3\sigma$.

3.4.3   Lyα absorption as a function of 3-D Hubble distance

Figure 3.7 shows the median Lyα optical depth in radial bins around the galaxy positions, where we assumed that velocity differences between absorbers and galaxies are entirely due to the Hubble flow. We emphasize that Figures 3.5 and 3.6 show that this is not a good approximation, particularly for distances $\lesssim 1\,$pMpc. It does, however, provide us with a compact way to present a lot of information.

The 3-D Hubble distance, as we call it, is therefore just $\sqrt{b^2 + (H(z)\Delta v)^2}$, where b is the galaxy's impact parameter, $H(z)$ is the Hubble parameter, and Δv is the velocity separation between an absorber and the galaxy. As mentioned in Section 3.2.1, we use only galaxies with impact parameters smaller than 2 pMpc, even when making Figure 3.7. Hence, distances $\gtrsim$ 2 pMpc reflect mostly LOS separations.

The horizontal dashed line shows the median level of absorption in the Lyα forest pixels. The significance of the excess absorption can be estimated by comparing the error bars, which indicate the 1σ confidence intervals determined by bootstrap resampling the galaxies, to the difference between the data points and the horizontal dashed line. More precisely, we can estimate the confidence level associated with the detection of excess absorption as $1 - 2f_{\mathrm{b,low}}$, where $f_{\mathrm{b,low}}$ is the fraction of 1,000 bootstrap realizations for which the data point falls below the dashed line. This method indicates that within 2.8 pMpc excess absorption is detected with greater than 99.7% significance (i.e. $> 3\sigma$). For $2.8 - 4.0$ pMpc the significance is 87% (i.e. 1.5σ), while the absorption is consistent with random beyond 4 pMpc.

The dotted curve and the right y-axis show the number of galaxies contributing to each bin. Since the inner few bins contain only a few tens of galaxies each (14 for the first bin), the bootstrap errors may not be reliable for these bins. The significance of the excess of absorption can be estimated more robustly by making use of the fact that each QSO spectrum provides many independent spectral regions at the impact parameter of each galaxy. We can do this by comparing the excess absorption to the grey region, which indicates the 1σ detection threshold and which was determined by re-measuring the median $\log_{10}\tau_{\mathrm{Ly}\alpha}$ after randomizing the galaxy redshifts (while keeping the impact parameters fixed). The grey shaded region shows the 1σ confidence interval obtained after doing this 1,000 times. For each distance bin, the confidence level of the detection is then given by $1 - 2f_{z,\mathrm{high}}$, where $f_{z,\mathrm{high}}$ is the fraction of realizations resulting in a median optical depth that is higher than actually observed. In agreement with the errors estimated by bootstrap resampling the galaxies, we find that the significance is $> 99.7\%$ within 2.8 pMpc and that there is no evidence for excess absorption beyond 4 pMpc. For 2.8–4.0 pMpc the significance of the detection is, however, larger than before: 99.2% (i.e. 2.7σ). We conclude that the absorption is significantly enhanced out to at least 3 pMpc proper, which is $7h^{-1}$ cMpc.

The fact that the absorption is enhanced out to several pMpc is in good agreement with Adelberger et al. (2005), who measured the mean flux as a function of 3-D Hubble distance. However, the profile measured by Adelberger et al. (2005) is much flatter. Converting their data into optical depths, they measure $\log_{10}\tau_{\mathrm{Ly}\alpha} \approx -0.1$ in their innermost bin, which extends to about 200 pkpc. This is about an order of magnitude lower than our median recovered optical depth at this distance. Conversely, at large distances their mean flux asymptotes to 0.765, or $\log_{10}\tau_{\mathrm{Ly}\alpha} \approx -0.57$, which is much higher than our asymptotic median optical depth of $\log_{10}\tau_{\mathrm{Ly}\alpha} \approx -1.27$, even though we measure a similar mean flux of 0.806, or $\log_{10}\tau_{\mathrm{Ly}\alpha} \approx -0.67$. Thus, our dynamic range is about two orders of magnitude larger than that of Adelberger et al. (2005). This difference arises because we use median optical depth rather than mean flux statistics and because we use higher order Lyman lines to recover the optical depth in saturated lines.

The dashed curves show the 15.9% and 84.1% percentiles, indicating the 1σ scatter in the PODs (which is obviously much larger than the error in the median). It is important to note that, except on the smallest scales ($\lesssim 200$ pkpc), the scatter is similar to or larger than the median excess absorption. Hence, there will be a wide range of PODs for all separations probed here.

Finally, the solid curve shows the best-fit power-law through the data points,

$$\mathrm{median}(\log_{10}\tau_{\mathrm{Ly}\alpha}) = (0.32 \pm 0.08)d^{-0.92\pm0.17} - 1.27, \qquad (3.6)$$

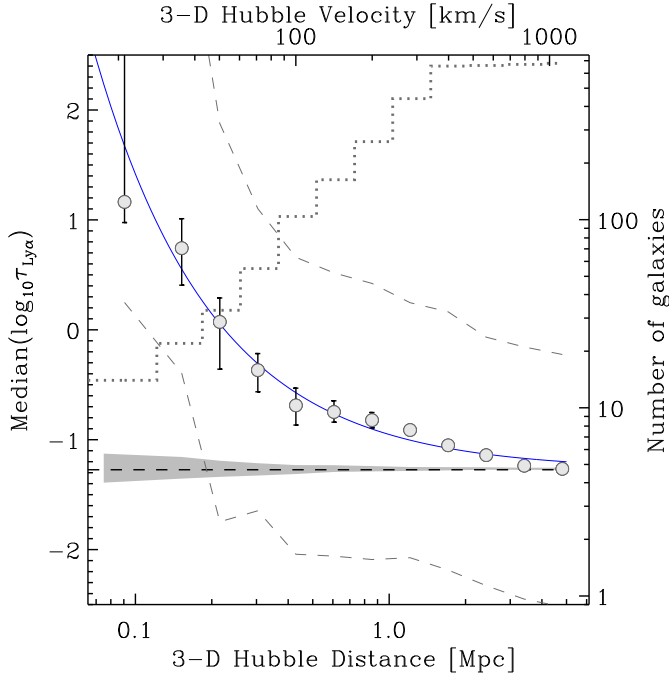where we required the fit to asymptote to the median of all pixels (horizontal, dashed line in Figure 3.7).

Figure 3.7 – Median $\log_{10}\tau_{\mathrm{Ly}\alpha}$ as a function of 3-D Hubble distance from galaxies. Distance bins are separated by 0.15 dex. The dashed lines show the 15.9% and 84.1% percentiles, i.e. the 1σ scatter of the PODs around the median. Note that the errors are correlated over scales $\lesssim 10^2$ km s^{-1} (see Appendix 3.B). The horizontal dashed line shows the median of all pixels in the spectra. The grey shaded region shows the 1σ detection threshold. It shows the 1σ confidence interval for the median($\log_{10}\tau_{\mathrm{Ly}\alpha}$) that we obtain after randomizing the redshifts. The dotted line shows the number of galaxies in each distance bin (right y-axis). The solid curve shows the best-fit power-law: $\log_{10}\tau_{\mathrm{Ly}\alpha} = 0.32d^{-0.92} - 1.27$, where d is the 3-D Hubble distance.

3.4.3.1   Testing the robustness

In this section we will show that our results are robust to changes in the S/N ratio of the QSO spectra, to the omission of NIRSPEC or non-NIRSPEC redshifts, to the exact redshift calibration, and, for optical depths $\ll 10$, to the use of higher order Lyman lines. We have chosen to demonstrate this using the plots of median absorption versus 3-D Hubble distance, because these offer a compact summary of the data.

In the left panel of Figure 3.8 we compare the median Lyα absorption as a function of 3-D Hubble distance for the lower and higher S/N subsamples of the QSO spectra (with median S/N ratios of ≈ 65 and ≈ 85 respectively). It appears that better data yields slightly more absorption at 3-D Hubble distances of ≈ 0.3 pMpc, but the differences are not significant.

The middle panel of Figure 3.8 compares the subset of 71 galaxies for which redshifts have been measured from nebular emission lines using the NIRSPEC instrument (red circles) with the default sample (grey curve) as well as with the result obtained when we ignore the NIRSPEC redshifts and
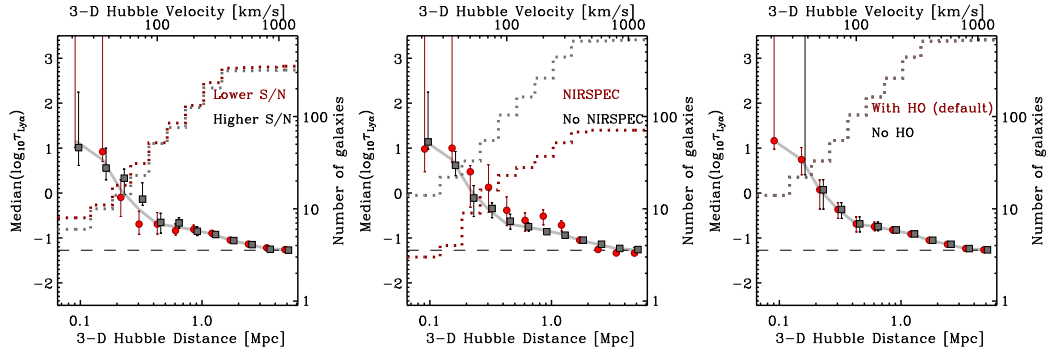
Figure 3.8 – Similar to Figure 3.7 but: *the left panel* shows the differences arising when using only higher or lower S/N QSO spectra; *the middle panel* shows the results for the galaxy subsample with near-IR redshifts, and for the whole sample when not using redshifts measured from near-IR lines; *the right panel* shows the effect of relying solely on Lyα and neglecting higher order transitions, which otherwise allow the recovery of the optical depth in saturated Lyα pixels. In all panels the black points have been slightly offset for clarity and the grey curve shows the result for the default sample and method.

instead use Lyα emission and/or interstellar absorption redshifts measured from LRIS spectra for all galaxies (black squares). As discussed in Sections 3.2.1.1 and appendix 3.A, the redshifts estimated from the NIRSPEC spectra have errors of $\Delta v \approx 60\,\mathrm{km\,s^{-1}}$, while the redshifts estimated from LRIS spectra typically have $\Delta v \approx 130\,\mathrm{km\,s^{-1}}$. In Figure 3.8 we see that the signal appears to drop slightly more steeply for the NIRSPEC subsample, as would be expected given the smaller redshift errors, but both the NIRSPEC and pure-LRIS samples are consistent with the default sample.

We note that for Lyα emission and interstellar absorption we also tried using the redshift calibrations from Adelberger et al. (2005) and Steidel et al. (2010) instead of the one from Rakic et al. (2011). For Rakic et al. (2011) the signal tends to be slightly stronger and the bootstrap errors slightly smaller, but the differences are small compared with the errors (not shown).

One of the advantages of the POD method is the possibility to recover the optical depth in the saturated Lyα pixels by using higher order Lyman lines. The right panel of Figure 3.8 shows the effect of omitting this feature of the POD method. The two curves are nearly identical except for the first two bins, $0-0.18$ pMpc, where the median optical depths increase from ~ 10 when we make use of higher order lines to 10^4 when we do not. The latter value is not meaningful as it is the optical depth that we assign to pixels for which saturation prevents recovery of the optical depth (see §3.3). Without higher order lines, we cannot constrain the flux to be much smaller than the S/N ratio, which in our case corresponds to optical depths of about 4–5.

Hence, measuring the median optical depth in the circumgalactic region requires the use of higher order lines, but the recovery of the optical depth in saturated pixels appears to be unimportant at large distances. This is, however, only true if we restrict ourselves to median statistics. As we will

see in Sections 3.5.1.1 and 3.5.1.2, higher order lines are also crucial at large distances if we are interested in the PDF of pixel optical depths.

3.4.4   Lyα absorption as a function of transverse distance

Figure 3.5 shows that absorption signal is smeared out over ~ 1 pMpc (i.e. ~ 200 km s^{-1}) in the LOS direction. Indeed, for all panels of Figure 3.6 the first 6 bins along the LOS (black squares), which correspond to velocities ≤ 165 km s^{-1}, are consistent with each other at the 1σ level. This velocity is slightly larger than the expected redshift errors and similar to the expected circular velocities of the halos hosting our galaxies. Because there is no evidence for structure in the velocity direction on scales ≤ 165 km s^{-1} and because the errors are strongly correlated for smaller velocity differences (see Appendix 3.B), it makes sense to group these first 6 LOS bins together and measure the absorption as a function of transverse distance. The result is shown as the blue squares in Figure 3.9, which shows the 3-D Hubble distance results for comparison (grey circles).

Estimating the errors by bootstrapping galaxies, we find that excess absorption is detected with $\geq 2\sigma$ confidence over the full range of impact parameters. Using the more robust method of randomizing galaxy redshifts, we find that the significance is at least 2σ for all but the second point (0.13–0.18 pMpc), which is based on only 8 galaxies and for which the significance is only 1.6σ.

As expected, the Hubble and transverse results converge at large distances, where redshift space distortions are small[1]. They also agree at very small distances (first bin; < 0.13 pMpc), because a small 3-D Hubble distance implies that the transverse distance must also be small (note that the reverse is not true since galaxies with small impact parameters contribute pixels with large LOS separations). On intermediate scales, however, the two measures of distance yield significantly different results, with the absorption at fixed 3-D Hubble distance being stronger than that at the same fixed transverse distance. In particular, while the transverse direction shows a rapid fall off at $b \sim 0.1$ pMpc followed by a constant excess absorption out to 2 pMpc, the absorption decreases smoothly with 3-D Hubble distance.

Because of redshift space distortions, it is not possible to measure the absorption as a function of the true, 3-D distance. However, we expect the true result to be in between the transverse and Hubble results shown in Figure 3.9. The 3-D Hubble distance may overestimate the true distance on intermediate scales, because it assumes all velocity differences to be due to the Hubble flow, whereas in reality redshift errors and peculiar velocities will contribute. Note, however, that infall of the right magnitude could lead the 3-

---

[1]Note that in the last bin the absorption is slightly stronger in the transverse direction. This is a weaker version of the distortion that we attributed to the Kaiser effect when we compared the transverse and LOS directions (see Fig. 3.6).
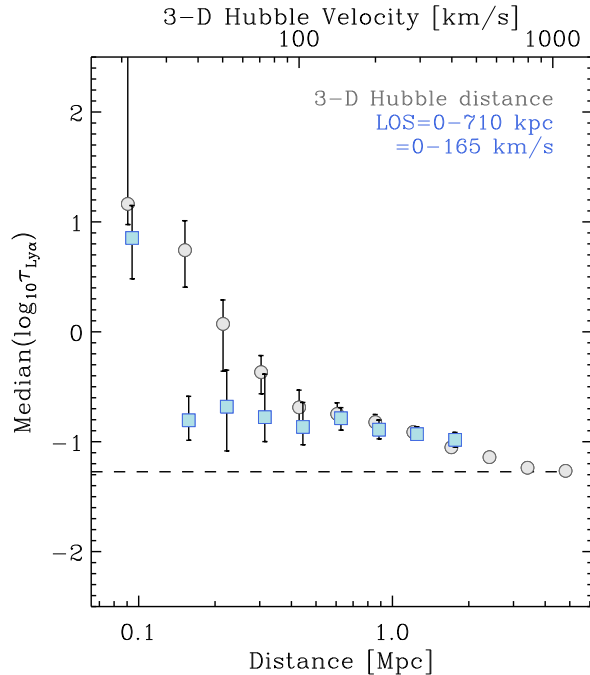
Figure 3.9 – Median ($\log_{10}\tau_{\mathrm{Ly}\alpha}$) as a function of two different measures of the distance from galaxies. The blue squares, which have been offset slightly for clarity, show the result as a function of transverse distance for LOS separations < 165 km s^{-1} (i.e. 0.71 pMpc for pure Hubble flow), which is about the scale over which trends in the LOS direction are smoothed out (see Fig. 3.6) and over which errors are correlated (see Appendix 3.B). For comparison, the grey circles repeat the Hubble distance results shown in Fig. 3.7. The results for the true, 3-D distance are not measurable, but are expected to lie in between the transverse and Hubble results shown here.

D Hubble distances to underestimate the true distances. On the other hand, because we group LOS separations of $0 - 165$ km s^{-1}, the transverse distance will typically underestimate the true distance on intermediate scales because it implicitly assumes that the contribution of the Hubble flow is negligible up to velocity differences of 165 km s^{-1}.

However, we chose to average over this velocity difference with good reason: within 165 km s^{-1} the signal is independent of velocity (see Fig. 3.6) and the errors are strongly correlated (Appendix 3.B), which suggests that the velocities are not dominated by the Hubble flow. We therefore expect the transverse results to be closer to the truth and the excess absorption for 3-D Hubble distances ~ 200 pkpc to be due mostly to the inclusion of absorption around galaxies with smaller impact parameters. Indeed, we find that around these distances the scatter in the PODs is much greater if we bin in terms of 3-D Hubble distance than if we bin in terms of transverse distance (even though a velocity interval of 165 km s^{-1} corresponds to a relatively large Hubble distance of 710 pkpc).

3.4.4.1   Scatter

We have so far focused on the median absorption. However, the dashed curves in Figure 3.7 demonstrate that there is a large degree of scatter in the distribution of PODs at a fixed distance. In this section we will therefore investigate how the median absorption around individual galaxies varies from galaxy to galaxy[2].
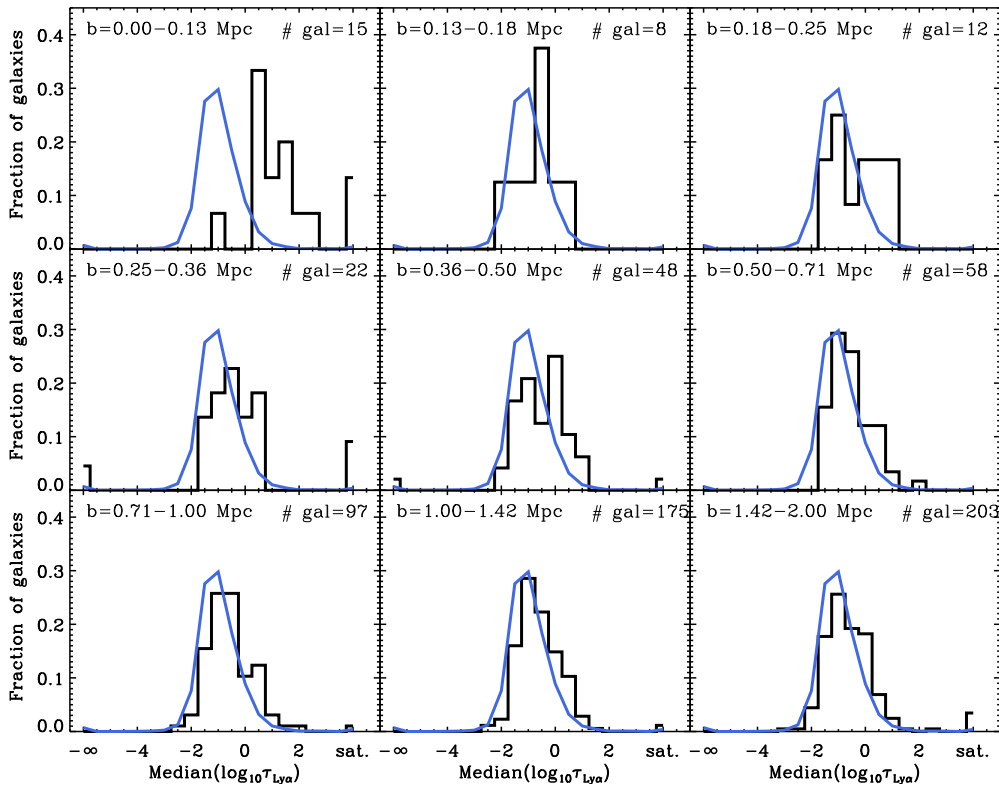

Figure 3.10 − The histograms show the distribution of the median HI Lyα optical depth within 165 km s^{-1} from galaxies. Different panels correspond to different impact parameters in the range 0–2 pMpc, as indicated above each panel. The grey solid curve, which is repeated in each panel, shows the distribution after randomizing the galaxy redshifts, i.e., the distribution expected in the absence of absorption correlated with galaxy positions. There is strong ($\gtrsim 2\sigma$) evidence for excess absorption at all impact parameters, but there is no evidence for a trend with impact parameter beyond 0.13 pMpc.

Figure 3.10 shows histograms of the distribution of the median optical

---

[2]We could also have studied the distribution of individual PODs, but we prefer to consider only one data point per galaxy (i.e. the median POD within 165 km s^{-1}) because galaxies are independent, whereas pixels are not.

depth in velocity intervals of ± 165 km s^{-1} centered on galaxies. The different panels correspond to different impact parameters. The x-axis has $-\infty$ as its lower limit. This indicates pixels set to $\tau = 10^{-5}$, because they had $F > 1$ (see Section 3.3 for details). We expect the true optical depths of these pixels to be similar to the inverse of the S/N ratio, which is $\sim 10^{-2}$. The label "saturated" at the high-absorption end of the x-axis is for saturated pixels whose optical depth could not be recovered using higher order Lyman lines (because they are either also saturated or unavailable). These pixels do not necessarily have optical depths that are higher than the highest recovered values ($\sim 10^2$), because higher order lines can be either contaminated or unavailable. In that case we cannot constrain the flux to be smaller than the S/N ratio, which corresponds to optical depths of about 4–5.

Comparing the different panels with each other, it seems that the distribution is shifted to higher optical depths for $b < 0.13$ pMpc, but that the results are similar for impact parameters $0.13 < b < 2.0$ pMpc. Indeed, Kolmogorov-Smirnov (K-S) tests show that, except for the first panel, panels 2–8 are consistent with the last panel (at the 1σ level). For $b < 0.13$ pMpc, however, the distribution of median optical depths differs at the 99.98% confidence level (3.7σ). These results are consistent with our findings for the medians of the distributions (Fig. 3.9, blue squares).

Ignoring the outliers for which the median optical depth is saturated or $-\infty$, the scatter is about 0.75 dex. The median absorption near galaxies with a fixed impact parameter is thus highly variable, suggesting that the gas is clumpy. This is true at all impact parameters and therefore not a distinguishing feature of the circumgalactic medium.

The grey curve, which is repeated in every panel, indicates the distribution after randomizing the galaxy redshifts. It therefore gives the distribution that we would expect to measure if the absorption were uncorrelated with galaxy positions. Using the K-S test to compare each histogram with the grey curve, we find that except for the second panel (which contains only 8 galaxies), all are discrepant at the $> 2\sigma$ level. For $b > 0.5$ pMpc the excess absorption appears small, but the difference is in each case significant at the $\gtrsim 4\sigma$ level. For the second panel ($b = 0.13 - 0.18$ pMpc) the significance of the detection of excess absorption is $\lesssim 1\sigma$. These results are in near perfect agreement with those based on median statistics. Thus, examination of the full distribution of median optical depths around galaxies confirms the result obtained for the medians. There is a sharp drop in the absorption around 10^2 pkpc, which is similar to the virial radius, followed by a near constant, small, but highly significant, excess absorption out to transverse separations of at least 2 pMpc.

3.4.5 Interpreting PODs

The central optical depth is related to the absorbing gas column density, N, through the following approximate relation:

$$\tau_0 \approx \left(\frac{N}{3.43 \times 10^{13}\,\text{cm}^{-2}} \right) \left(\frac{f}{0.4164} \right)$$

$$\times \left(\frac{\lambda_0}{1215.67\,\text{Å}} \right) \left(\frac{b}{26\,\text{km s}^{-1}} \right)^{-1} \tag{3.7}$$

where f is the oscillator strength, λ_0 is the transition's rest wavelength, $b = \sqrt{2}v_{\text{RMS}} = (\text{FWHM})/2\sqrt{\ln 2}$ is the line width, and v_{RMS} is the line of sight velocity dispersion (e.g. Padmanabhan 2002§9.5). In our study we use statistics based on either median optical depths or (in §3.5.1.2) on the maximum optical depth within a given distance range from a galaxy. Measurements based on the maximum optical depth in a given region are relatively easy to interpret using Equation (3.7). However, given that most of our analysis is based on median optical depths, all the conversions that we make to column densities of the absorbing gas are likely to be underestimates if most of the lines are not blended, and could be overestimates in regions where the lines are highly blended.

The left panel of Figure 3.11 shows the median neutral hydrogen column density as a function of both transverse distance for LOS velocities $0 - 165\,\text{km s}^{-1}$ (solid curve) and 3-D Hubble distance (dashed curve), which we obtained from Figure 3.9 and equation (3.7). The median column density decreases from $\sim 10^{14.5}\,\text{cm}^{-2}$ at $d \sim 10^2$ pkpc to $\sim 10^{12.7}\,\text{cm}^{-2}$ at ~ 1 pMpc, which is in excellent agreement with results from Rudie et al. (in preparation) based on Voigt profile decompositions. Note that if we had selected the strongest system within a given distance of each galaxy, instead of taking into account all pixels within that distance, the value of the median column density (i.e., the median of the maximum column density) would have been significantly higher.

To gain intuition about what the observed absorption represents, we can convert the optical depths into overdensities. Combining Equation (3.7) with Equation (10) of Schaye (2001), who treats Lyα absorbers as gravitationally confined gas clouds with sizes of order the local Jeans length, we obtain:

$$\Delta \approx 2.1\tau_{0,\text{Ly}\alpha}^{2/3}\Gamma_{12}^{2/3} \left(\frac{1+z}{3.36} \right)^{-3} \left(\frac{T}{2 \times 10^4\,\text{K}} \right)^{0.17}$$

$$\times \left(\frac{f_{\text{g}}}{0.162} \right)^{-1/3} \left(\frac{b}{26\,\text{km s}^{-1}} \right)^{2/3} \tag{3.8}$$

where $\Delta \equiv \rho/\langle\rho\rangle$ is the density of gas in units of the mean baryon density of the Universe, Γ_{12} is the photo-ionization rate in units of $10^{12}\,\text{s}^{-1}$, and f_{g}

denotes the fraction of the cloud mass in gas. We have assumed a temperature typical for the moderately overdense IGM (e.g. Schaye et al. 2000a; Lidz et al. 2010; Becker et al. 2011), a line width consistent with the median value measured by Rudie et al. (in preparation) for our data, and a photo-ionization rate appropriate for ionization by the ultraviolet background radiation (e.g. Bolton et al. 2005; Faucher-Giguère et al. 2008). In collapsed gas clouds f_g could be close to unity, but far away from galaxies gas will not be in dense clumps and f_g should be close to its universal value of Ω_b/Ω_m. Note that these densities are effectively evaluated on the local Jeans scale, which is typically $\sim 10^2$ pkpc for the densities of interest here (Schaye 2001).

The above expression is a good approximation for overdense absorbers. Absorbers with densities around or below the cosmic mean have not had sufficient time to reach local hydrostatic equilibrium (Schaye 2001) and will be better described by the fluctuating Gunn-Peterson approximation (e.g. Rauch et al. 1997), which assumes smoothly varying density fluctuations and pure Hubble flow, yielding

$$\Delta \approx 2.02\tau_{\mathrm{Ly}\alpha}^{1/2}\Gamma_{12}^{1/2}\left(\frac{1+z}{3.36}\right)^{-9/4}\left(\frac{T}{2\times 10^4\,\mathrm{K}}\right)^{0.38}. \qquad (3.9)$$

Both equations (3.8) and (3.9) assume primordial abundances, highly ionized gas, and photo-ionization equilibrium. However, close to galaxies, UV radiation from local sources may dominate over the background and the gas may be shock-heated to temperatures sufficiently high for collisional ionization to dominate. Both of these effects would cause us to underestimate the gas density, possibly by a large factor.

The middle and right panels of Figure 3.11 show the median overdensity profiles obtained after applying the above equations to the median optical depth measured as a function of 3-D Hubble and transverse distance (for LOS separations $0-165$ km s^{-1}), respectively. The "Jeans" and fluctuating Gunn-Peterson approximation are generally in good agreement, although the former yields steeper density profiles at small distances, where the gas is highly overdense. However, in this regime we do not expect the fluctuating Gunn-Peterson approximation to hold. The dashed curve shows the result if we convert the recovered optical depths into overdensities using the fit to hydrodynamical simulations given in Aguirre et al. (2002) for $z = 2.5$. The relation provided by those authors was obtained by producing mock Lyα absorption spectra for sight lines through a hydrodynamical simulation, scaling the simulated spectra to fit the observed mean flux decrement, and then fitting the relation between the Lyα optical depthweighted overdensity of gas responsible for the absorption in each pixel and the recovered Lyα POD in that pixel. The agreement between the hydrodynamical simulation and the Jeans method is clearly excellent.

We conclude that typical gas overdensities decrease from $\gtrsim 10$ at $d \lesssim 10^2$ pkpc, which is similar to the virial radii of the halos hosting our galaxies,
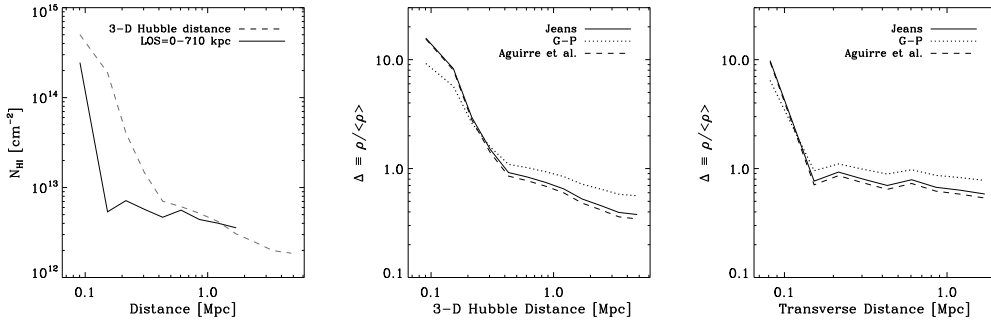
Figure 3.11 − Inferred median neutral hydrogen column density (left panel) and inferred median overdensity (middle and right panels) as a function of distance from galaxies. Column densities were obtained from the measured, median Lyα optical depths and equation (3.7), assuming $b = 30 \, \mathrm{kms}^{-1}$. Overdensities were obtained by converting the median optical depths using the Schaye (2001) "Jeans" approximation (solid curves), the fluctuating Gunn-Peterson approximation (dotted lines), or the fit to the results of the hydrodynamic simulation of Aguirre et al. (2002) (dashed curves). Overdensities inferred from observed optical depths are effectively volume weighted (and are lower than mass-weighted overdensities). The right panel and the solid curve in the left panel show results as a function of transverse distance for velocity separations $< 165 \, \mathrm{km \, s}^{-1}$, while the middle panel and the dashed curve in the left panel are for 3-D Hubble distances. Typical column densities (overdensities) decrease from $\sim 10^{14.5} \, \mathrm{cm}^{-2}$ (~ 10) at $d \sim 10^{-1}$ pMpc to $\lesssim 10^{13} \, \mathrm{cm}^{-2}$ ($\lesssim 1$) at $d \gtrsim 1$ pMpc.

to ~ 1 at $d \sim 1$ pMpc. The steepness of the drop at intermediate scales is uncertain due to redshift space distortions, but it is likely to be bracketed by the the middle and right panels of Figure 3.11. Observe that on large scales the overdensity asymptotes to values smaller than unity. This is expected, as the median optical depth is a volume weighted measure rather than a mass-weighted measure. While the mass-weighted mean overdensity is by definition unity, the volume-weighted mean overdensity is smaller because underdense regions dominate the volume. The difference between volume and mass-weighted quantities probably also accounts for the relatively low overdensity measured around the virial radius ($\sim 10^{-1}$ pMpc).

3.5   Circum-Galactic Matter

The smallest impact parameter in our sample is ≈ 55 pkpc. It is difficult to probe scales smaller than this with QSO-galaxy pairs owing to the small number of bright QSOs and the rarity of close pairs, as well as to the difficulty of observing objects right in front of a bright QSO. For such small scales it is therefore more efficient to resort to using galaxies as background objects (e.g. Adelberger et al. 2005; Rubin et al. 2010; Steidel et al. 2010). Nevertheless, in this section we will use our sample of background QSOs to study the gas within transverse distances of 200 pkpc of galaxies. But before doing so, we will compare our results with the small-scale data from Steidel et al. (2010).

3.5.1 Comparison with results for galaxy-galaxy pairs

Steidel et al. (2010) measured Lyα rest-frame equivalent width (EW) as a function of distance from the same sample of galaxies that we study here. Since they used background galaxy spectra for probing foreground galaxies' circumgalactic matter, they were able to study smaller scales than we can probe with the QSO spectra. While they could measure the EW in a stack of background galaxy spectra, the spectral resolution of the galaxy spectra was too low to measure column densities or resolved pixel optical depths and the S/N was too low to obtain EW measurements for individual galaxies. The red circles in Figure 3.12 show their measurements at small impact parameters and the blue circles show their measurements using QSO spectra, together with our EW measurements from QSO spectra at larger impact parameters (black squares). Although we used the same data, their technique for measuring EW differs from ours (see Steidel et al. 2010, for more details).

We computed the EWs as follows. We shifted the spectra into the rest-frame of each galaxy within a given impact parameter bin, found the mean flux profile, divided it by the mean flux level of all pixels in the spectra (i.e. 0.804) in order to mimic the effect of continuum fitting low-quality spectra, and integrated the flux decrement over $\pm 500\,\mathrm{km\,s^{-1}}$ centered on the galaxies' positions. We verified that using a larger velocity interval ($\leq 1000\,\mathrm{km\,s^{-1}}$) changes the results within the plotted error bars.

It appears that the absorption strength falls off according to a power law, out to ~ 800 pkpc. Beyond this impact parameter the relation flattens off at EW$\sim 0.2 - 0.3$ Å.

3.5.1.1 Scatter

Adelberger et al. (2005) found that the absorption near galaxies follows a bimodal distribution. While the flux decrement is usually large, $\approx 1/3$ of galaxies are not associated with strong Lyα absorption within a 3-D Hubble distance of 1 h^{-1} pMpc (i.e. ≈ 400 pkpc).

As shown in Figure 3.10, we also find that the absorption varies strongly from galaxy to galaxy. However, we showed that this is not only true near galaxies, but also if we look in random places of the spectrum. Moreover, the distribution of median optical depths is not bimodal.

As Adelberger et al. (2005) did not measure optical depths, we show the distribution of the median flux decrement for impact parameters < 200 pkpc and velocity differences $< 165\,\mathrm{km\,s^{-1}}$ from galaxies in the left panel of Figure 3.13. For comparison, the right panel shows the corresponding median optical depths. As before, we chose this velocity interval because there is little structure in the median profiles for velocity separations smaller than this value (see Figs. 3.5 and 3.6), which is not surprising given that $165\,\mathrm{km\,s^{-1}}$ is similar to both the redshift errors and the circular velocities of the halos
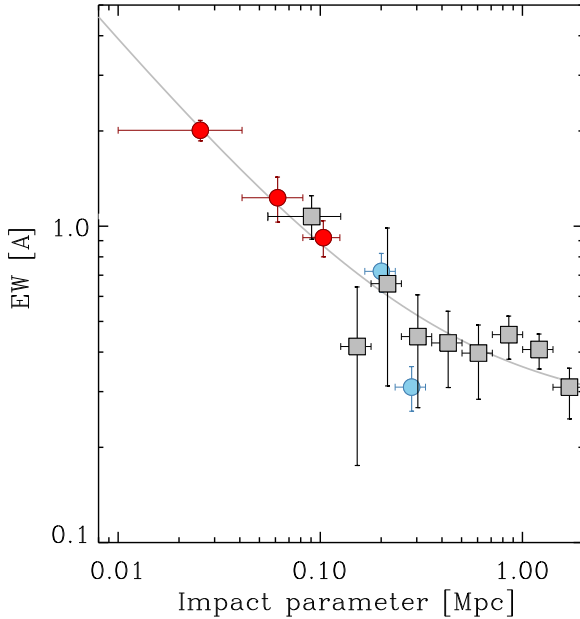
Figure 3.12 – Rest-frame equivalent width of Lyα absorption as a function of impact parameter, for the galaxy-galaxy pair samples from Steidel et al. (2010) shown as red circles, and as measured from the data set used in this paper, QSO-galaxy pairs, as grey squares. The blue circles show QSO-galaxy pair results from Steidel et al. (2010), using the same KBSS data, but a different way of measuring EW and different bin sizes.) The vertical error bars are estimated by bootstrap resampling the galaxy sample. The horizontal error bars indicate bin sizes. The grey line shows the best fit: EW= $0.11d^{-0.76} + 0.25$, where d is impact parameter of galaxies.

thought to host the galaxies. The grey curves show the results for randomized galaxy redshifts, where we assign random redshifts (within the Lyα forest redshift range) to the galaxies in our sample. The curves are the result of estimating a histogram for randomized redshifts 1,000 times, and taking the mean median in each bin. A K-S test shows that the two distributions are discrepant at the $> 6\sigma$ level, so the absorption is clearly enhanced near galaxies.

It can be seen in the left panel that galaxies either show very little transmission, or relatively high flux. The flux distribution is thus bimodal, in agreement with Adelberger et al. (2005). However, the right panel shows that there is no evidence for bimodality of the optical depth distribution. Hence, the bimodality seen in the left panel is a consequence of the mapping $F = \exp(-\tau)$ which bunches the low (high) optical depth tail of the distribution together at a flux decrement of zero (one). Note that the shape of the optical depth distribution (which appears to be approximately lognormal) is physically more relevant than that of the flux, because the former is proportional to the neutral hydrogen column density.
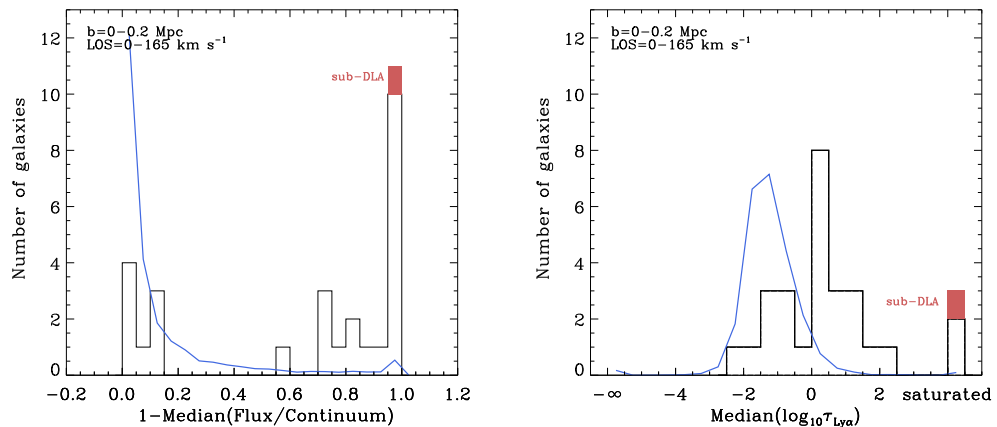
Figure 3.13 − Number of galaxies vs. median flux (left panel) and median $\log_{10}\tau_{\mathrm{Ly}\alpha}$ (right panel) within $165\,\mathrm{km\,s^{-1}}$ from galaxies with impact parameters $b < 200\,\mathrm{pkpc}$ from galaxies. The grey solid curves show the results expected if absorbers and galaxy positions were uncorrelated, and were obtained after randomising the galaxy redshifts 1,000 times. In the right panel $-\infty$ indicates the pixels with $F > 1$, and "saturated" indicates saturated pixels for which the optical depth could not be recovered (see text for more details). The observed absorption is enhanced at the $> 6\sigma$ level.

We will demonstrate next that the scatter is not random as the optical depth is strongly anti-correlated with the impact parameter.

3.5.1.2 Cold flows

In recent years cosmological simulations and theoretical models have converged on the idea that gas accretion is bimodal. While the gas accreting onto massive halos is shock-heated to the virial temperature, in lower mass halos most of the gas falls in cold ($T \sim 10^4$ K). The cold mode feeds the galaxy through filaments (cold streams) whereas the hot mode results in the formation of a hydrostatic halo which fuels the galaxy through a cooling flow (e.g. Kereš et al. 2005; Ocvirk et al. 2008; Dekel et al. 2009; Crain et al. 2010; van de Voort et al. 2011a,b; Faucher-Giguère et al. 2011b). At $z = 2$ halo masses of $\sim 10^{12}$ M$_\odot$, i.e. similar to those hosting our galaxies (Trainor et al. 2011, in preparation), are especially interesting in that they mark the transition between the cold- and hot-mode accretion regimes. For example, van de Voort et al. (2011a) predict that the hot mode contributes on average $\approx 50\%$ and $\approx 30\%$ to the growth of 10^{12} M$_\odot$ halos and their central galaxies, respectively. The simulations predict that individual galaxies in such halos are fed simultaneously through both modes with cold streams penetrating hot, hydrostatic halos.

Despite the theoretical consensus, there is no direct observational evidence for this galaxy formation picture as of yet. Absorption by neutral hydrogen is a promising way to observe the cold streams. Based on a high-resolution

simulation of a single halo with mass 3×10^{11} M$_\odot$ at $z = 2$, Faucher-Giguere et al. (2010) predict that the covering fraction of cold flows with Lyman Limit (LLS; $10^{17.2} \leq N_{HI} \leq 2 \times 10^{20}$ cm^{-2}) and DLA column densities ($N_{HI} > 2 \times 10^{20}$ cm^{-2}), within the virial radius (2 virial radii) is 10–15% (4%) and 3–4% (1–2%), respectively, where they quote 74 pkpc for the virial radius. Kimm et al. (2011) use cosmological simulations to predict that the covering fraction of systems with $N_{HI} > 10^{20}$ cm^{-2} within 100 pkpc of halos of $\sim 10^{12}$ M$_\odot$ at $z = 2.5$ is $\approx 5\%$. Stewart et al. (2011) simulate two halos with masses of $\sim 10^{11.5}$ M$_\odot$ at $z = 2$ and predict covering fractions for $N_{HI} > 10^{16}$ cm^{-2} of about 15% at 33 pkpc.
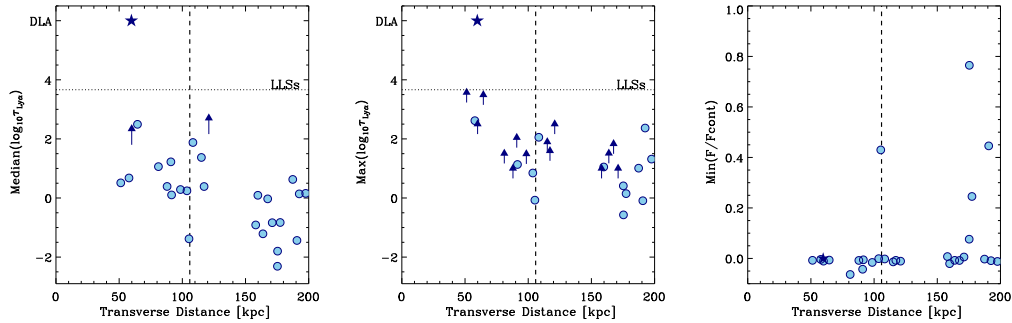

Figure 3.14 $-$ *The left and middle panels* show, respectively, the median and the maximum optical depth within 165 km s^{-1} from each galaxy that is separated from the LOS by less than 200 pkpc as a function of its impact parameter. In cases where the median/maximum corresponded with a saturated pixel without an optical depth estimate we plot the largest recovered value of the POD within 165 km s^{-1} as an upward pointing arrow - these are lower limits to the actual optical depths. The dotted horizontal line indicates the approximate threshold for Lyman limit absorbers. Stars indicate a sub-DLA, identified via its damping wings (and with line centers < 165 km s^{-1} from the nearest galaxy). *The right panel* shows the minimum flux within 165 km s^{-1} from each galaxy. The vertical dashed lines indicate the virial radius for typical halos hosting the galaxies in our sample, with $M_h = 10^{12} h^{-1}$M$_\odot$ (Adelberger et al. 2005; Conroy et al. 2008, Trainor et al., in preparation) at $z = 2.4$.

The left panel of Figure 3.14 shows the median optical depth within ± 165 km s^{-1} from galaxies, as a function of the impact parameter. The results are insensitive to the exact velocity interval chosen, although a galaxy with $b = 155$ pkpc would have been a DLA system if we had used a maximum velocity separation ≥ 200 km s^{-1}. Each circle represents a galaxy. Sub-DLA systems are shown as star symbols. In two non-DLA cases the median optical depth could not be measured because more than half of the pixels were saturated and there were insufficient higher order lines to recover the optical depth. For those cases we plot the largest recovered optical depth and show it as an upwards-pointing arrow to indicate that it is a lower limit. Note, however, that there is no evidence for DLA absorption associated with any of the galaxies plotted as lower limits. The vertical dashed line shows the virial radius for halos with mass M $= 10^{12}$M$_\odot$ (Trainor et al. 2011, in preparation).

To compare with predictions from the literature, it is more appropriate to look at the maximum pixel optical depth in the $\pm 165 \ \mathrm{km \, s^{-1}}$ interval around each galaxy, because we can convert this into a neutral hydrogen column density using equation (3.7). This results in significantly more lower limits, as can be seen in the middle panel of Figure 3.14. Finally, the right panel shows the minimum flux.

Figure 3.14 shows a clear trend of increasing absorption strength with decreasing distance. While the scatter in the optical depth is large at all impact parameters, in terms of flux there is actually little variation within 100 pkpc. In particular, it is striking that all 10 galaxies with impact parameters smaller than the virial radius are associated with saturated absorbers ($F \approx 0$).

Within 100 pkpc 10 out of 10 galaxies have median $\tau_{\mathrm{Ly}\alpha} > 1$, which corresponds to a covering fraction of $100 \pm 32\%$. Within 200 pkpc this decreases to $66 \pm 16\%$. If we use the maximum optical depth than the covering fraction of $\tau_{\mathrm{Ly}\alpha} > 1$ absorbers within 200 pkpc increases to $89 \pm 18\%$. Note that the central optical depth is unity for lines with column density $N_{\mathrm{HI}} \approx 4 \times 10^{13} \ \mathrm{cm^{-2}}$ (see equation 3.7).

It appears that 1–8 out of 10 galaxies with impact parameters smaller than the virial radius of their host halos are associated with LLSs (the horizontal dashed line shows the central optical depth for a $N_{\mathrm{HI}} = 10^{17.2} \ \mathrm{cm^{-2}}$ absorber assuming a line width of 26 $\mathrm{km \, s^{-1}}$). In addition, 1 out of 10 galaxies has a sub-DLA within R_{vir}. Within two virial radii 1–15 out of 27 galaxies are associated with LL systems and 1 out of 27 galaxies is associated with a sub-DLA (in addition there would have been 1 DLA if we had used a maximum velocity separation $\geq 200 \ \mathrm{km \, s^{-1}}$). These numbers are higher than the predictions from Faucher-Giguere et al. (2010) and Kimm et al. (2011) but the errors are too large for the discrepancy to be significant. We note that when it comes to measuring the column densities of strongly saturated systems (e.g. Figure 3.14 where our method yields a number of lower limits), it would be more appropriate to use Voigt profile decompositions (see Rudie et al., in preparation).

Even though these observations are consistent with what is seen in simulations, it is unclear whether the observed absorbing gas is inflowing or outflowing. Steidel et al. (2010) found, using 500 galaxy pairs sampling angular scales of 1–15", that the circumgalactic medium within $\approx 80 - 90$ pkpc shows strong H I and low-ion metallic absorption, which is consistent with the findings in this paper. However, they also find that galaxy spectra exhibit kinematics consistent with radial flows with velocity increasing outward.

3.6   The distribution of galaxies around absorbers

In the previous sections we investigated the typical IGM environment of the galaxies in our sample. Here we address the complementary question: what is the galaxy environment for a pixel of a given optical depth?
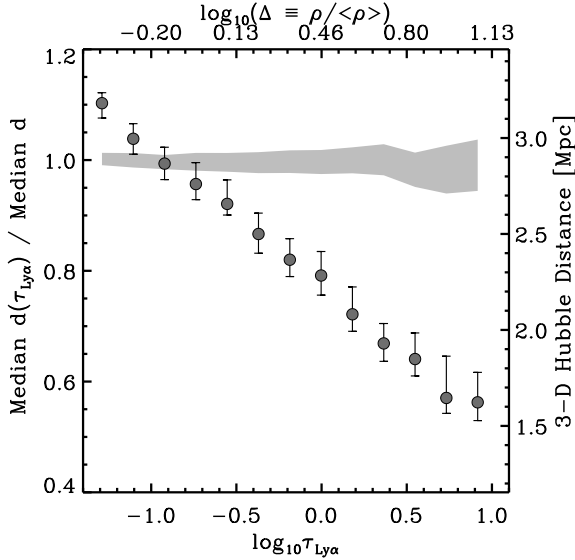

Figure 3.15 – The median 3-D Hubble distance to the nearest galaxy, normalized by the median 3-D Hubble distance to the nearest galaxy for all pixels, as a function of the pixel's Lyα optical depth. We normalize by this distance in order to get results that do not depend on the galaxy sample's completeness. The unnormalized distance scale is shown on the right y-axis. The grey shaded area shows the 1σ confidence interval around the results obtained when using randomized galaxy redshifts. Absorbers with $\log_{10}(\tau_{\mathrm{Ly}\alpha}) \gtrsim -1.0$ tend to be closer to galaxies than a random place in the Universe is, while lower optical depth pixels tend to be further away.

Figure 3.15 shows the median 3-D Hubble distance to the nearest galaxy as a function of a pixel's Lyα optical depth, normalized by the median 3-D Hubble distance to the nearest galaxy for all pixels (2.88 pMpc), irrespective of their optical depth (the unnormalized distance is shown on the right y-axis). Note that because we divided by the median distance for all pixels, the result is not directly dependent on the completeness of our sample (the actual, unnormalized distances are, however, highly sensitive to the completeness). The error bars were determined by bootstrap resampling the QSO spectra using chunks of 500 pixels. The grey shaded area shows the 1σ confidence interval of results obtained for randomized galaxy samples. It shows what would be expected if absorbers and galaxies were distributed randomly with respect to each other.

We cut the plot at the low optical depth end at the value corresponding to a 2σ detection of absorption for a S/N of 50. Noise prevents us from distinguishing lower optical depths from each other, causing the trend to flatten. At the high optical depth end we cut the plot at $\tau_{\mathrm{Ly}\alpha} = 10$. Although

we can recover higher optical depths, we find that pixels with $\tau_{\mathrm{Ly}\alpha} > 10$ are biased: while the median redshift stays close to 2.36 for optical depths in the plotted range, it increases to about 2.5 for higher optical depths. Because of our selection function, the galaxy number density is lower for such high redshifts, causing an upturn of the distance to the nearest galaxy for $\tau > 10$.

The optical depth is strongly anti-correlated with the distance to the nearest galaxy. Pixels with $\tau > 0.1$, are closer to galaxies than a random place in the Universe is, while pixels with $\tau < 0.1$ are on average farther from galaxies than a random location is.

To provide some physical interpretation of this observed trend we use the Jeans approximation described in Section 3.4.5 to convert optical depths into estimates for the gas overdensities (implicitly smoothed on the scale of the absorbers, i.e. $\sim 10^2$ pkpc) (top x-axis). Gas at the mean density of the Universe would produce $\log_{10}\tau_{\mathrm{Ly}\alpha} \approx -0.65$ and is thus typically closer to galaxies than a random place in the Universe. Although it is unclear whether we can estimate the overdensity sufficiently accurately for this statement to be reliable, it would in fact not be a surprising result. Because voids take up most of the volume, a random place in the universe will be underdense.

If we cube the normalized distance to the nearest galaxy, then we get something close to the inverse of the overdensity of galaxies. However, in that case the scale over which this number density is measured, would be the distance to the nearest galaxy and would thus vary with the density itself. It is more instructive to measure the overdensity of galaxies on a fixed length scale, which is something we can also measure from our data.

In Figure 3.16 we show the mean galaxy overdensity as a function of pixel optical depth, evaluated on (3-D Hubble) scales of 0.25, 0.5, 1, 2, 4, and 8 pMpc around them. Note that these galaxy number densities cannot be directly compared with the gas overdensities that we can estimate from the optical depths because they are evaluated over different length scales.

The galaxy number density increases with increasing optical depth, and the density contrast is stronger when it is measured on smaller scales. For example, pixels with $\tau_{\mathrm{Ly}\alpha} \sim 10$ (at $z \approx 2.4$) see on average a galaxy overdensity ≈ 2 within 8 pMpc, but an overdensity ~ 10 within 0.25 pMpc. Pixels with $\tau_{\mathrm{Ly}\alpha} \sim 1$ typically reside in regions where the galaxy number density is close to the cosmic mean on scales ≥ 0.25 pMpc. Because galaxies are clustered, this implies that for such pixels the distance to the nearest galaxy will typically be smaller than it is for a random place in the Universe (see Fig. 3.15).

Given that we have both measurements of the galaxy number density and estimates of the gas overdensities, we could attempt to estimate the bias of these two components relative to each other by computing the ratio of the two densities. However, the length scale over which gas densities corresponding to Lyα absorbers are implicitly smoothed (i.e. the local Jeans scale in case

of overdense gas; Schaye 2001), varies with the density and hence the optical depth, while galaxy and gas density have to be evaluated on the same scale in order to calculate the relative bias. We could in principle estimate the galaxy overdensity on scales that vary with the optical depth, but for overdense gas the relevant scales are somewhat too small to get a robust estimate of the galaxy number density.
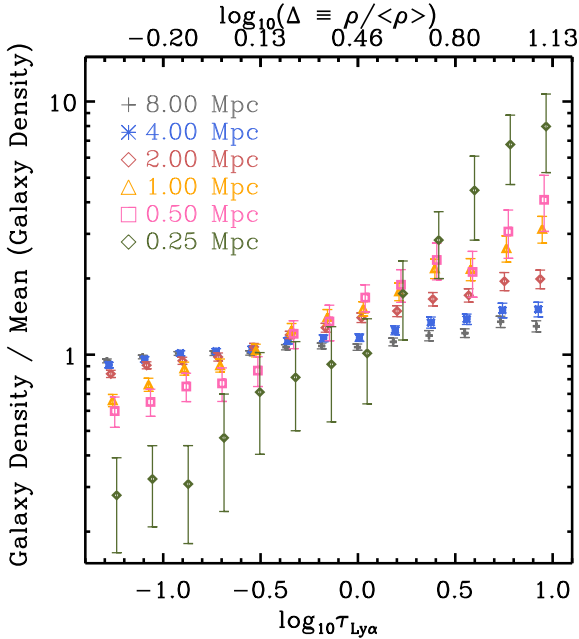

Figure 3.16 – The mean galaxy overdensity within 0.25 (green diamonds), 0.5 (pink squares), 1 (orange triangles), 2 (red diamonds), 4 (blue asterisks), and 8 (black crosses) 3-D Hubble pMpc from pixels of a given Lyα optical depth. The mean galaxy density is higher near stronger absorbers, and the density contrast is more pronounced when estimated on smaller scales.

3.7   Summary & Conclusions

In this paper we studied the galaxy-IGM interface at $z \approx 2-2.8$ in the 15 fields of the Keck Baryonic Structure Survey. Each KBSS field is centered on a very bright background QSO of which high S/N, high resolution spectra have been obtained, and includes a densely-sampled galaxy survey optimized to sample the same range of redshifts probed by the IGM measurements. Our analysis has included a total of 679 galaxies, the subset of the larger KBSS sample that lies in the appropriate range of redshifts within 2pMpc of the corresponding QSO sightline. The redshifts of 71 of our galaxies were measured directly from nebular lines observed in the near-IR, while the remaining redshifts were derived from ISM absorption and/or Lyα emission lines and the redshift calibration of Rakic et al. (2011). The median redshift of our galaxies is 2.36.

We analyzed the spectra using the pixel optical depth method, which uses

an automatic algorithm to recover the Lyα optical depth in each pixel, making use of higher order Lyman lines to recover the optical depth in pixels with saturated Lyα absorption. Our analysis relied almost entirely on median statistics. Because the Lyα optical depth is proportional to the column density of neutral hydrogen, its distribution is physically more relevant than that of the flux.

Our results can be summarized as follows (as before, all distances are proper):

- We presented the first 2-D map of absorption around galaxies, showing the median Lyα optical depth as a function of transverse and LOS separations. By converting velocity differences along the LOS into distances assuming pure Hubble flow, redshift space distortions show up as anisotropies in the absorption maps.

- Hydrogen Lyα absorption is strongly enhanced near galaxies. Within $\sim 10^2$ pkpc the median optical depth is $\gtrsim 2$ orders of magnitude higher than in a random place ($\tau_{\mathrm{Ly}\alpha} \gtrsim 10$ versus $< 10^{-1}$).

- Two types of anisotropies are clearly detected:

  1. On scales $\sim 30 - 200$ km s^{-1}, or $0.1 - 1$ pMpc, the median absorption is significantly stronger, by $\sim 1 - 2$ orders of magnitude, along the LOS than in the transverse direction. We showed that this "finger of God" effect can be partially explained by galaxy redshift errors, but argued that it is probably dominated by peculiar velocities of gas within and around the halos.

  2. On scales of $\sim 1.4 - 2.0$ pMpc the absorption is significantly stronger (with $> 3\sigma$ significance) in the transverse direction than along the LOS. We interpret this compression in the LOS direction as large-scale infall towards the potential wells occupied by the galaxies, i.e. the Kaiser (1987) effect. To our knowledge, this is the first time that this signature of gravitational collapse has been detected for gas.

- The median optical depth as a function of impact parameter (over ± 165 km s^{-1} intervals around galaxies) drops by more than an order of magnitude going from $b < 0.13$ pMpc to $0.13 < b < 0.18$ pMpc and remains approximately constant thereafter out to at least 2 pMpc, where it is still a factor of two higher than in a random part of the spectrum. The sharp transition occurs at a radius that is similar to the virial radius of the halos that are thought to host the galaxies.

- The median optical depth as a function of 3-D Hubble distance (i.e. the 3-D distance computed under the assumption of pure Hubble flow) shows no steep drop at 10^2 pkpc. Instead, it decreases smoothly away from galaxies. The lack of a steep drop reflects the fact that 3-D Hubble distances of a few times 10^2 pkpc mix strong absorption arising at

impact parameters $< 10^2$ pkpc with the mostly weaker absorption at larger transverse separations.

- Excess absorption is detected with $> 3\sigma$ significance out to 2.0 pMpc in the transverse direction. In terms of the 3-D Hubble distance the significance is $> 3\sigma$ out to 2.8 pMpc and 1.5σ for 2.8 to 4 pMpc.

- The median optical depth near galaxies with a fixed impact parameter is highly variable, suggesting that the gas is clumpy. The scatter in the median optical depth at fixed impact parameter is ~ 0.75 dex. Beyond 100 pkpc this is large compared with the excess median absorption.

- Converting optical depths to column densities, we estimate that the median neutral hydrogen column density decreases from $\gtrsim 10^{14.5}$ cm^{-2} within 10^2 pkpc to $< 10^{13}$ cm^{-2} for larger impact parameters. We note that if we had selected the strongest absorber within 10^2 pkpc of each galaxy, instead of taking all the pixels within that distance into account, the median column density of so selected systems would have been substantially higher.

- Converting optical depths into overdensities (implicitly smoothed on the scale of the absorbers, i.e. $\sim 10^2$ pkpc) by applying the "Jeans" approximation (Schaye 2001), the fluctuating Gunn-Peterson approximation (e.g. Rauch et al. 1997), or a fit to a hydrodynamical simulations (Aguirre et al. 2002) all yield similar results, with the Jeans approximation and the hydro simulation agreeing particularly well. We estimate that median gas overdensities decrease from $\gtrsim 10$ within 10^2 pkpc to ~ 1 at 1 pMpc. Beyond this impact parameter the overdensity asymptotes to values smaller than unity, which is expected because underdense regions dominate the volume and gas densities inferred from optical depths are volume rather than mass weighted.

- The covering fraction of absorbers with median optical depth $\tau_{\mathrm{Ly}\alpha} > 1$ within 165 km s^{-1} is $100 \pm 32\%$ for $b < 100$ pkpc and $66 \pm 16\%$ for $b < 200$ pkpc. If we use the maximum rather than the median optical depth than the latter fraction increases to $89 \pm 18\%$.

- By turning to a pixel-centric view, i.e. estimating the median distance to the nearest galaxy as a function of the pixel's optical depth, we show that absorbers with $\tau_{\mathrm{Ly}\alpha} > 0.1$ are typically closer to galaxies than a random location in the Universe, while weaker absorbers are farther from galaxies compared to random.

- The average number density of galaxies around absorbers increases with increasing absorption strength, and the galaxy density contrast is higher when it is evaluated on smaller scales. For example, on scales of 0.25 pMpc the galaxy overdensity increases from $\sim 10^{-0.5}$ at $\tau_{\mathrm{Ly}\alpha} = 10^{-1}$ to ~ 1 at $\tau_{\mathrm{Ly}\alpha} = 1$ and to ~ 10 at $\tau_{\mathrm{Ly}\alpha} = 10$. When evaluated in spheres with radius of 2 pMpc the galaxy overdensity increases from ≈ 0.9 at $\tau_{\mathrm{Ly}\alpha} \sim 10^{-1}$ to ≈ 2 for $\tau_{\mathrm{Ly}\alpha} \sim 10$. Absorbers with $\tau_{\mathrm{Ly}\alpha} \sim 1$ reside in regions where the galaxy number density is close to the cosmic mean

on scales ≥ 0.25 pMpc. As most of the volume is underdense, such absorbers are nearer to galaxies than a random place in the Universe is.

This work could be improved considerably by reducing the redshift errors, which will be possible through observations of nebular emission lines with the new multi-object near-IR spectrograph Keck I/MOSFIRE (McLean et al. 2010). Such observations would also allow us to investigate whether the absorption varies with the properties of the galaxies such as their stellar masses or star formation rates.

In a future paper we will present a comparison of our observational findings with cosmological, hydrodynamical simulations, which will aid the interpretation of the absorption profiles that we see around galaxies. In addition, Rudie et al. (in preparation) will present an analysis of the same data based on Voigt profile decompositions.

Acknowledgments

References

Adelberger K. L., Steidel C. C., Shapley A. E., & Pettini M., 2003, ApJ, 584, 45

Adelberger K. L., Steidel C. C., Shapley A. E., Hunt M. P., Erb D. K., Reddy N. A., & Pettini M., 2004, ApJ, 607, 226

Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., & Reddy N. A., 2005, ApJ, 629, 636

Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., & Erb D. K. 2005, ApJ, 619, 697

Aguirre A., Schaye J., & Theuns T., 2002, ApJ, 576, 1

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

Bergeron J. & Boissé P., 1991, A&A, 243, 344

Bland J., & Tully B., 1988, Nature, 334, 43

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

Bordoloi R., et al., 2011, ApJ, in press (arXiv:1106.0616)

Borthakur S., Tripp T. M., Yun M. S., Momjian E., Meiring J. D., Bowen D. V., & York D. G., 2010, ApJ, 713, 131

Bouche N., et al., 2011, MNRAS, in press (arXiv:1107.4618)

Bowen D. V., Pettini M., & Blades J. C., 2002, ApJ, 580, 169

Brooks A. M., Governato F., Quinn T., Brook C. B., & Wadsley J., 2009, ApJ, 694, 396

Chabrier, G., 2003, PASP, 115, 763

Chen H.-W., Lanzetta K. M., Webb J. K., & Barcons X., 1998, ApJ, 498, 77C

Chen H.-W., Lanzetta K. M., & Webb J. K., 2001, ApJ, 556, 158

Conroy C., Shapley A. E., Tinker J. L., Santos M. R., & Lemson, G. 2008, ApJ, 679, 1192

Cowie L. L., & Songaila A., 1998, Nature, 394, 44

Crain, R. A., McCarthy, I. G., Frenk, C. S., Theuns, T., & Schaye, J. 2010, MNRAS, 407, 1403

Crighton N. H. M., et al., 2011, MNRAS, 414, 28

Danforth C. W., & Shull J. M., 2008, ApJ, 679, 194

Dekel A., et al., 2009, Nature, 457, 451

Ellison S. L., Songaila A., Schaye J., & Pettini M., 2000, AJ, 120, 1175

Erb, D. K., Shapley, A. E., Pettini, M., Steidel, C. C., Reddy, N. A., & Adelberger, K. L., 2006, ApJ, 644, 813

Erb, D. K., Steidel, C. C., Shapley, A. E., Pettini, M., Reddy, N. A., & Adelberger, K. L., 2006, ApJ, 646, 107

Erb, D. K., Steidel, C. C., Shapley, A. E., Pettini, M., Reddy, N. A., & Adelberger, K. L., 2006, ApJ, 647, 128

Faucher-Giguère C.-A., Lidz A., Hernquist L., & Zaldarriaga M., 2008, ApJ, 682, L9

Faucher-Giguère, C.-A., & Kereš D., 2011, MNRAS, 412, L118

Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, in press (arXiv:1103.0001)

Frank, S., Appenzeller, I., Noll, S., & Stahl, O., 2003, A&A, 407, 473

Kaiser N., 1987, MNRAS, 227, 1

Kacprzak G. G., Churchill C. W., Evans J. L., Murphy M. T., & Steidel C. C. 2011, MNRAS, in press (arXiv:1106.3068)

Kereš D., Katz N., Weinberg D. H., & Davé R., 2005, MNRAS, 363, 2

Kimm T., Slyz A., Devriendt J., & Pichon C., 2011, MNRAS, 413, L51

Komatsu E. et al., 2009, ApJS, 180, 330

Lanzetta K. M., & Bowen D., 1998, ApJ, 357, 321

Lanzetta K. M., Bowen D. V., Tytler D., & Webb J. K., 1995, ApJ, 442, 538

Lehnert M. D., Heckman T. M., & Weaver K. A., 1999, ApJ, 523, 575

Lidz A., Faucher-Giguère C.-A., Dall'Aglio A., McQuinn M., Fechner C., Zaldarriaga M., Hernquist L., & Dutta S., 2010, ApJ, 718, 199

McLean, I. S., Steidel, C. C., Epps, H., et al. 2010, Proc. SPIE, 7735, 47

Ocvirk P., Pichon C., & Teyssier R., 2008, MNRAS, 390, 1326

Padmanabhan, T. 2002, Theoretical Astrophysics - Volume 3, Galaxies and Cosmology, by T. Padmanabhan, pp. 638. Cambridge University Press, December 2002. ISBN-10: 0521562422. ISBN-13: 9780521562423

Penton S. V., Stocke J. T., & Shull J. M., 2002, ApJ, 565, 720

Pieri M. M., Schaye J., & Aguirre A., 2006, ApJ, 638, 45

Prochaska J. X., Weiner B., Chen H. -W., Mulchaey J. S., & Cooksey K. L., 2011, arXiv, arXiv:1103.1891

Rakic, O., Schaye, J., Steidel, C. C., & Rudie, G. C., 2011, MNRAS, 414, 3265

Rauch M., et al., 1997, ApJ, 489, 7

Reddy, N. A., Steidel, C. C., Pettini, M., Adelberger, K. L., Shapley, A. E., Erb, D. K., & Dickinson, M., 1998, ApJS, 175, 48

Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., & Weiner B. J., 2010, ApJ, 712, 574

Ryan-Weber E. V., 2006, MNRAS, 367, 1251

Schaye J., Theuns T., Rauch M., Efstathiou G., & Sargent W. L. W., 2000a, MNRAS, 318, 817

Schaye J., Rauch M., Sargent W. L. W., & Kim T.-S., 2000b, ApJ, 541, 1

Schaye J., 2001, ApJ, 559, 507

Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., & Sargent W. L. W., 2003, ApJ, 596, 768

Schaye, J., Carswell, R. F., & Kim, T.-S. 2007, MNRAS, 379, 1169

Shapley A. E., Steidel C. C., Erb D. K., Reddy N. A., Adelberger K. L., Pettini M., Barmby P., & Huang, J., 2005, ApJ, 626, 698

Simcoe, R. A., Sargent, W. L. W., Rauch, M., & Becker, G., 2006, ApJ, 637, 648

Songaila A., 1998, AJ, 115, 2184

Springel V., White S. D. M., 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., & Pearce F., 2005, Nature, 435, 629S

Steidel C. C., & Sargent W. L. W., 1992, ApJS, 80, 1

Steidel C. C., Dickinson M., & Persson S. E., 1994, ApJ, 437, 75

Steidel C. C., Dickinson M., Meyer D. M., Adelberger K. L., & Sembach K. R., 1997, ApJ, 480, 568

Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., & Giavalisco M., 2003, ApJ, 592, 728

Steidel C. C., Shapley A. E., Pettini M., Adelberger K. L., Erb D. K., Reddy N. A., & Hunt M. P., 2004, ApJ, 604, 534

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., & Rakic O., 2010, ApJ, 717, 289

Steidel, C. C., Bogosavljević, M., Shapley, A. E., Kollmeier, J. A., Reddy, N. A.,

Erb, D. K., & Pettini, M., 2011, ApJ, 736, 160

Stewart K. R., Kaufmann T., Bullock J. S., Barton E. J., Maller A. H., Diemand J., & Wadsley J., 2011, ApJ, 735, L1

van de Voort F., Schaye J., Booth C. M., Haas M. R., & Dalla Vecchia C., 2011a, MNRAS, 414, 2458

van de Voort F., Schaye J., Booth C. M., & Dalla Vecchia C., 2011b, MNRAS, 415, 2782

Vogt S. S., Allen S. L., Bigelow B. C., Bresee L., Brown B., Cantrall T., Conrad A., & Couture M., 1994, Proc. SPIE, 2198, 362

3.A   The effect of redshift errors

As discussed in Section 3.2.1.1, there are random errors in the redshift measurements of $\approx 130\,\mathrm{km\,s^{-1}}$ for those obtained from LRIS spectra, and $\approx 60\,\mathrm{km\,s^{-1}}$ for the NIRSPEC subsample. We expect that adding errors that are small compared to the redshift errors that are already present in the data will have very little effect, whereas adding errors that are large compared to the actual redshift errors should strongly degrade the signal.

Figure 3.17 illustrates what effect the addition of random, Gaussian redshift errors has. The left and right panels show median optical depth profiles in cuts along the x- and y-axes of Figure 3.5 corresponding to the nearest LOS and transverse bins, respectively. The data points in the left (right) panel are thus identical to the red circles (black squares) of the first panel of Figure 3.6. The blue, red, and orange curves show the results after adding redshift errors with $\sigma = 55$, 165, and 495 $\mathrm{km\,s^{-1}}$, respectively. Redshift errors smear out the signal along the LOS direction, thereby strongly reducing its amplitude. Adding 55 $\mathrm{km\,s^{-1}}$ errors hardly changes anything, but for 165 $\mathrm{km\,s^{-1}}$ significant differences start to appear. Errors of 495 $\mathrm{km\,s^{-1}}$ completely wash out the signal. We conclude that the true redshift errors are $\lesssim 165\,\mathrm{km\,s^{-1}}$.

3.B   Correlations between data points

Because a single galaxy contributes pixels to multiple distance bins along the LOS, we expect the errors to be correlated in the LOS direction. Conversely, since a single galaxy only has one impact parameter, we expect the errors to be uncorrelated in the direction transverse to the LOS. In this section we will show that this is indeed the case and that the errors along the LOS are correlated over scales of $\sim 10^2$ $\mathrm{km\,s^{-1}}$.

To study the correlation between points, we made 1,000 bootstrap realisations of our galaxy sample, found the median $\tau_{\mathrm{Ly}\alpha}$ in each of the distance bins for each of the realisations, and then computed the Spearman rank correlation coefficient between different distance bins. In other words, if we created a scatter plot of the median optical depth in bins a and b where each point is a bootstrap realisation, then a high rank correlation would show up as
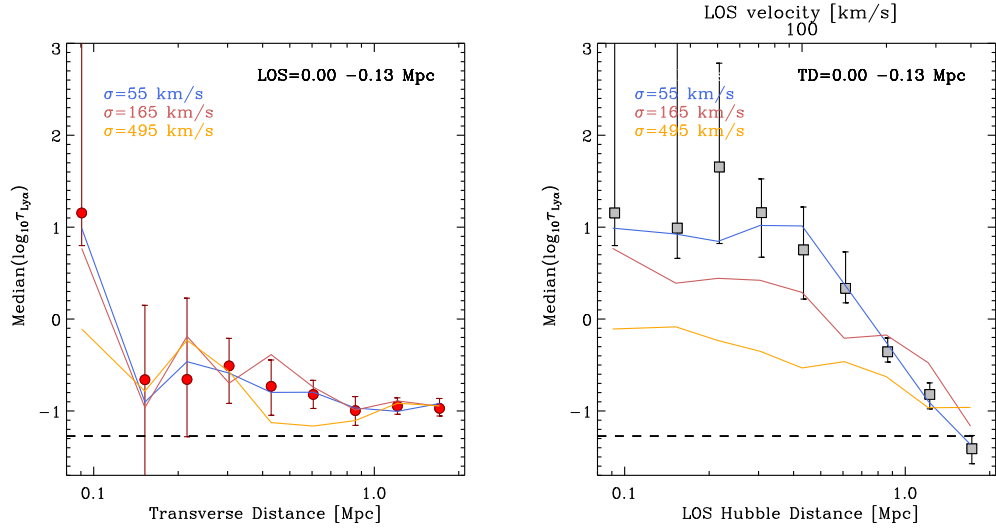
Figure 3.17 – The effect of redshift errors. The data points with error bars in left (right) panels shows the median $\log_{10}\tau_{\mathrm{Ly}\alpha}$ as a function of transverse (LOS) separation for the velocity (impact parameter) range indicated in the panels. These correspond to the first panel of Fig. 3.6. The blue, red, and orange curves show the result after adding random, Gaussian redshift errors with $\sigma = 55$, 165, and 495 $\mathrm{km\,s^{-1}}$, respectively. Significant differences start to appear for 165 $\mathrm{km\,s^{-1}}$, while three times greater redshift errors nearly completely wash out the signal. This suggests that the actual errors on the redshifts are $\lesssim 165\,\mathrm{km\,s^{-1}}$.
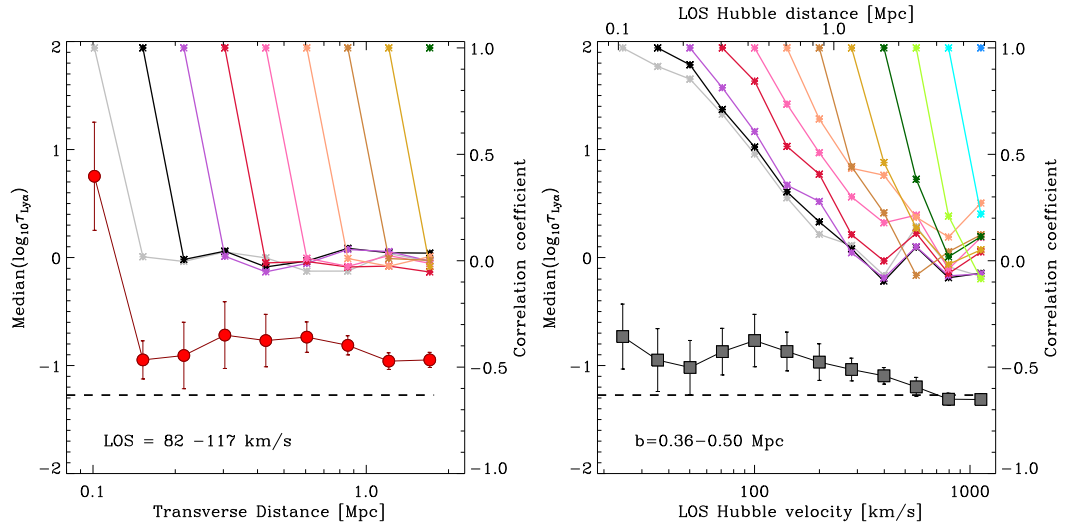

Figure 3.18 – Correlations between errors of different data points. The data points with error bars in left (right) panels shows the median $\log_{10}\tau_{\mathrm{Ly}\alpha}$ as a function of transverse (LOS) separation for the velocity (impact parameter) range indicated in the panels. These correspond to the central panel of Fig. 3.6. The different coloured curves and the right y-axis show the Spearman rank correlation coefficient between the errors of different points, computed by correlating 1,000 different bootstrap samples (see text). Data points are uncorrelated in the transverse direction. Along the LOS they are strongly correlated for separations < 100 $\mathrm{km\,s^{-1}}$ but become uncorrelated for distances $\gg 10^2$ $\mathrm{km\,s^{-1}}$.

points clustered along a line, whereas a correlation coefficient of zero would correspond to a random cloud of points in $\tau_a - \tau_b$ space.

We use rank correlation tests instead of the covariance matrix because some of the pixels are set to very high optical depths due to saturation, or to very low optical depths because they have an apparent positive flux (see Section 3.3). By using a non-parametric measure of the correlation strength, we are not susceptible to such problems (we use medians rather than averages for the same reason).

The red circles in the left panel of Figure 3.18 shows the median optical depth for LOS separations $80 - 114$ km s^{-1} (i.e. 0.36 – 0.50 pMpc for pure Hubble flow) as a function of transverse distance. Similarly, the black squares in the right panel show the median optical depth for transverse distances 0.36 – 0.50 pMpc as a function of LOS separation. These points are thus identical to those shown in the central panel of Figure 3.6.

The rest of the coloured curves in Figure 3.18 show the rank correlation coefficient between different points, with the corresponding scale indicated on the right y-axis. Values of the correlation coefficient from 0.5 to 1 imply a strong correlation, between 0.2 and 0.5 a weak correlation, values around 0 imply there is no correlation, and negative values imply that there is an anti-correlation.

For example, in the left panel the red curve shows the correlation between the 4th point and all the ones that lie at larger distances. The rank correlation coefficient for the 4th point with itself is 1, but the correlation with the subsequent points is about zero. This confirms that the points are uncorrelated in the transverse direction.

On the other hand, the right panel shows that the points are strongly correlated for LOS separations < 100 km s^{-1} and become uncorrelated for distances $\gg 10^2$ km s^{-1}. This is consistent with the fact that we observe no evidence for structure along the LOS for separations $< 10^2$ km s^{-1} (see Figs. 3.4 and 3.6).

Finally, Figure 3.19 shows the extent of the correlation between the errors of different points when we bin in terms of the 3-D Hubble distance. As the 3-D Hubble distance combines uncorrelated transverse with correlated LOS separations, the result must be intermediate between that for the two directions. Figure 3.19 demonstrates that this is indeed the case. Adjacent points are weakly correlated, while more distant points are essentially uncorrelated.
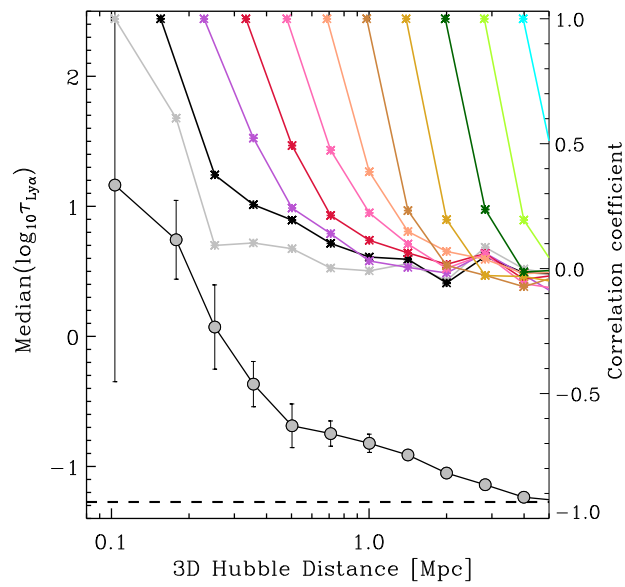
Figure 3.19 − Similar to Figure 3.18 but showing the median optical depth and the rank correlation coefficient as a function of 3-D Hubble distance. The correlations are weaker than along the LOS (right panel of Fig. 3.18) and are only substantial for adjacent bins.

4

Measuring Galaxy Halo Masses from the Surrounding Lyα Absorption Distribution

We measure dark matter halo masses of $\langle z \rangle \approx 2.36$ star-forming galaxies by matching the observed H I Lyα absorption distribution around them, as observed in the spectra of background QSOs, to the absorption around haloes above a given mass in the SPH cosmological simulations. The implied minimum halo mass, $\log_{10} M_{\mathrm{min}} / \, \mathrm{M}_\odot = 11.6^{+0.20+0.50}_{-0.20-0.60}$, is in good agreement with the mass estimates from clustering analysis. To test the robustness of the method, we employ a simulation with WMAP1 cosmological parameters, as opposed to the WMAP3 parameters used for the rest of the simulations in our study, and find that it gives results consistent with our reference model. A simulation with a highly efficient AGN feedback implementation implies masses slightly higher, but consistent with those inferred from simulations with weaker feedback. A particularly strong point of this method is that it can be used in narrow field galaxy-QSO surveys, i.e. $\approx 30 \times 30$ arcseconds. In addition, we find that the recently observed redshift space distortions, presented in Chapter 3, in the 2-D H I Lyα absorption distribution are a consequence of large-scale gas infall into potential wells occupied by galaxies.

4.1   Introduction

THE mass of the host dark matter halo is believed to be the most important factor that determines galaxy evolution. When it comes to the most massive structures in the Universe, the employed techniques for measuring halo mass include measurements of Sunyaev-Zeldovich effect (e.g. Sunyaev & Zeldovich 1972; Motl et al. 2005), cluster X-ray temperature and luminosity (e.g. Reiprich & Böhringer 2002), strong lensing (e.g. Gavazzi et al. 2007), weak lensing (also called galaxy-galaxy lensing; e.g. Kaiser & Squires 1993), and kinematics of galaxies (e.g. Zwicky 1937). Measuring dark matter halo masses of individual galaxies is more difficult and usually requires statistical methods such as already mentioned weak lensing (e.g. Mandelbaum et al. 2006), clustering analysis (e.g. Kaiser 1984), satellite kinematics (e.g. More et al. 2011), and abundance matching (e.g. Conroy et al. 2006), although Haas et al. (2011) found that the most popular environmental indicators (e.g. number of galaxies within a given distance, or distance to the Nth nearest neighbour) also strongly correlate with halo mass even for single galaxies but with a large scatter.

Weak lensing studies rely on measuring gravitationally induced tangential shear distortions of background galaxies around foreground galaxies. Even though the individual distortions are small, averaging over a large ensemble of galaxies allows statistical reconstruction of dark matter density distributions around a given galaxy population. This method has been very successful at estimating dark matter halo masses for low-redshift galaxies. However, it is not a viable method for high-redshift objects since lensing cross-section peaks half way towards background galaxy, and is thus very small at $z \geq 2$. Furthermore, the number of both foreground and background galaxies becomes too low for getting a sufficient statistical sample, and measuring shapes of galaxies becomes increasingly hard.

As known from structure formation studies, more massive dark matter haloes cluster more strongly, and on scales greater than the haloes the galaxies inhabiting such haloes cluster similarly. Matching the observed clustering properties of a given galaxy population with the clustering of haloes in N-body simulations leads to estimates of the most likely host halo masses. This method has facilitated some important insights into the properties of galaxies, and their likely descendants (e.g. Adelberger et al. 1998, 2005b; Quadri et al. 2007; Conroy et al. 2008).

In this chapter we measure galaxy halo masses from H I Lyα absorption profiles. Numerous studies have shown that the absorption of the light from a background object (e.g. QSO) is increased when the sightline passes near a foreground galaxy (e.g. Chen et al. 1998, 2001; Penton et al. 2002; Adelberger et al. 2003, 2005; Crighton et al. 2010; Prochaska et al. 2011; Rakic et al. 2011), and at high-redshift this enhancement can be statistically detected out to distances of several proper Mpc (pMpc). Kim & Croft (2008) found that

such absorption profiles can be used to measure halo masses of foreground QSOs and galaxies by comparing them to mock observations from smoothed particle hydrodynamics (SPH) simulations. Somewhat related to this method is also work of Faucher-Giguère et al. (2008) who found that the Lyα optical depth statistics in the QSO proximity zones depends on dark matter halo masses of the QSO hosts. In comparison with previous studies we use higher resolution SPH simulations, and more models with varied galactic feedback prescriptions.

We apply the method to the observed Lyα absorption distribution around galaxies at $\langle z \rangle \approx 2.36$ from Rakic et al. (2011). They study the intergalactic matter distribution around rest-frame UV selected star-forming galaxies through the analysis of the absorption spectra of background QSOs, and find enhanced Lyα absorption within ≈ 3 pMpc from galaxies. They also present a 2-D (i.e. along the line of sight, LOS, and in the transverse direction) absorption distribution where they measure a slight compression of signal along the LOS on scales $\gtrsim 1.5$ pMpc. This compression suggests large-scale infall of gas into the potential wells occupied by the galaxies.

This observational study is particularly suitable for testing the method because the dark matter halo masses of their galaxies are already known from clustering analyses (e.g. Adelberger et al. 2005b; Conroy et al. 2008, Trainor et al., 2011, in preparation). Conroy et al. (2008) found that every halo at $z \sim 2$ with $M_{\rm h} > 10^{11.4} h^{-1} M_\odot$ hosts a star-forming galaxy from the photometric sample out of which the galaxies used in Rakic et al. (2011) were drawn. This result was obtained using N-body simulations with the WMAP1 cosmology. Using the WMAP3 cosmology (which is what the simulations in this paper are based on) they found a 0.3-0.4 dex smaller minimum halo mass. Using a spectroscopic sample, almost identical to the one used by Rakic et al. (2011), Trainor et al. (2011, in preparation) found that these galaxies reside in halos with $M_{\rm h} > 10^{11.75} h^{-1} M_\odot$, using WMAP1 cosmological parameters.

We describe the simulations in Section 4.2. Our halo mass measuring method is presented in Section 4.3. We explain the observed redshift space distortions seen in the 2-D absorption maps from Rakic et al. (2011) in Section 4.4. We conclude in Section 4.5. Proper distances throughout the text are denoted as pMpc, and comoving as cMpc. When a quantity is not expressed in terms of h (e.g. when stating distances), the value of h from the appropriate set of cosmological parameters is assumed.

4.2 Simulations

We use a subset of simulations from the suite of OverWhelmingly Large Simulations (OWLS; Schaye et al. 2010). The simulations were run using a modified version of GADGET-3 (last described in Springel 2005), which is a Lagrangian SPH code. These simulations include star formation (Schaye & Dalla Vecchia 2008), supernova (SN) feedback (Dalla Vecchia & Schaye

2008), radiative heating and cooling via 11 elements (hydrogen, helium, carbon, nitrogen, oxygen, neon, magnesium, silicon, sulphur, calcium, and iron) in the presence of the cosmic microwave background and the Haardt & Madau (2001) model for the ionizing background radiation from galaxies and quasars (Wiersma et al. 2009a), and chemodynamics (abundances of elements released by type Ia and type II SNe, stellar winds, and asymptotic giant branch stars; Wiersma et al. 2009b). The assumed initial mass function (IMF) is that of Chabrier (2003), with stars ranging in mass from 0.1 to 100 $M_\odot$. Box sizes and (the initial) baryonic and dark matter particle masses of the simulations used here are specified in Table 4.1. The cosmological parameters used in the simulations are from the Wilkinson Microwave Anisotropy Probe (WMAP) 3-year results (Spergel et al. 2007), $\{\Omega_{\rm m}, \Omega_{\rm b}, \Omega_\Lambda, \sigma_8, n_{\rm s}, h\} = \{0.238, 0.0418, 0.762, 0.74, 0.951, 0.73\}$.

We use the 'reference' (REF) model, model with AGN feedback (AGN), model using 'Millennium cosmology' ($MILL$), and a model with a different SN feedback prescription ($WML4$). All of the models are described in detail by Schaye et al. (2010); here we provide only short descriptions.

The REF and the AGN models are almost identical, except that the latter includes prescriptions for the growth of supermassive black holes and for AGN feedback. The AGN feedback implementation is a modified version of that of Springel (2005), and is described in Booth & Schaye (2009). It involves black holes injecting 1.5 % of the rest-mass energy of the accreted gas, in the form of heat, into the surrounding matter. Booth & Schaye (2009, 2010) show that this efficiency results in excellent agreement with the observed, $z = 0$, scaling relations between black hole masses and the properties of their host galaxies. We chose this simulation to test the effect of extreme feedback on our results as we expect that it could potentially influence physical properties (e.g. density and temperature) of the absorbing gas near galaxies.

To test the effect of 'cosmology', and to facilitate comparisons with previous studies that used older estimates of the cosmological parameters, we employ the 'Millennium cosmology' (Springel et al. 2005) in our $MILL$ simulation. The $MILL$ simulation uses the cosmological parameter values $\{\Omega_{\rm m}, \Omega_{\rm b}, \Omega_\Lambda, \sigma_8, n_{\rm s}, h\} = \{0.25, 0.045, 0.75, 0.9, 1.0, 0.73\}$. As found by Conroy et al. (2008), the lower values of σ_8 that are implied by more recent WMAP results cause the lower mass dark matter haloes to cluster more strongly than in models with higher σ_8. Hence the minimum mass of haloes implied by the observed clustering of a galaxy population is lower. The faster growth of structure in the $MILL$ simulation also causes higher predicted global star-formation rate (SFR) densities, and thus the mass loading factor in the $MILL$ model for the SN driven winds is doubled in comparison with the REF model, to $\eta = 4$. This is done in order to match the observed peak in the star-formation history (SFH). We note that the latest, WMAP year 7, estimates of cosmological parameters (Komatsu et al. 2011), $\{\Omega_{\rm m}, \Omega_{\rm b}, \Omega_\Lambda, \sigma_8, n_{\rm s}, h\} = \{0.272, 0.0455, 0.728, 0.81, 0.967, 0.704\}$,

Table 4.1 – Relevant information for 50 and 100 h^{-1} cMpc simulation boxes used here. The columns give the comoving size of the box L, the total number of particles per component N (dark matter and baryons), the initial baryon particle mass m_b, the dark matter particle mass m_{dm}, and the maximum gravitational softening length.

Simulation	L $[h^{-1}\text{cMpc}]$	N	m_b $[h^{-1}\,\text{M}_\odot]$	m_{dm} $[h^{-1}\,\text{M}_\odot]$	Soft. length $[h^{-1}\text{ckpc}]$
L050 N512	50	512^3	1.1×10^7	5.1×10^7	1
L100 N512	100	512^3	8.7×10^7	4.1×10^8	2

have σ_8 value that is in between those from WMAP year 1 and WMAP year 3 results, while the value for Ω_m is closer to that from WMAP1.

WML4 is almost identical to the *REF* model, but the mass loading factor is the same as in the *MILL* simulation, which allows us to separate the effects of cosmological parameters and feedback prescriptions in our models.

4.2.1   Extracting sightlines from the simulations

We use only the $z = 2.25$ snapshot, as this is close to peak in the redshift distribution of the galaxy sample used in Rakic et al. (2011, $z \approx 2.36$), and the snapshots of the simulations are saved every $\Delta z = 0.25$. The observed galaxies cover the redshift range from $z \approx 2-2.8$, and we leave for future work the estimate of uncertainty in our results arising due to possible evolution from $z = 2.8$ to $z = 2$.

The masses and locations of the dark matter haloes in the simulations are determined using a spherical overdensity criterion as implemented in the SubFind algorithm (Dolag et al. 2009). SubFind finds the radius, r_{200}, and mass, m_{200}, of a spherical halo centered on the potential minimum of each identified halo, so that they contain a mean density of 200 times the critical density of the Universe at a given redshift.

We extract 12,500 sightlines within 5 pMpc from the halo centers, drawing haloes above a given mass limit randomly (which ensures that we sample the halo mass function correctly). This is done for 20 cases where we vary the minimum halo mass, $M_{\min}$, in steps of 0.1 dex in the range $M_{\min} = 10^{10.5} - 10^{12.5}\,\text{M}_\odot$. There are 500 sightlines for every 200 pkpc impact parameter bin. Given that there are, e.g. $\sim 20,000$ haloes with mass above $10^{11}\,\text{M}_\odot$, or ~ 700 haloes with mass higher than $10^{12}\,\text{M}_\odot$ in a 100 h^{-1} cMpc box (see Figure 4.1 for number densities of haloes above a given mass), in most distance bins a single halo features only once. In addition, sightlines are always parallel to the box sides, but the axis is chosen randomly.

By setting only the minimum halo mass we implicitly assume that each halo above a given mass limit is a host to a galaxy from the observed population. We do not set the upper limit on halo mass because there is no upper
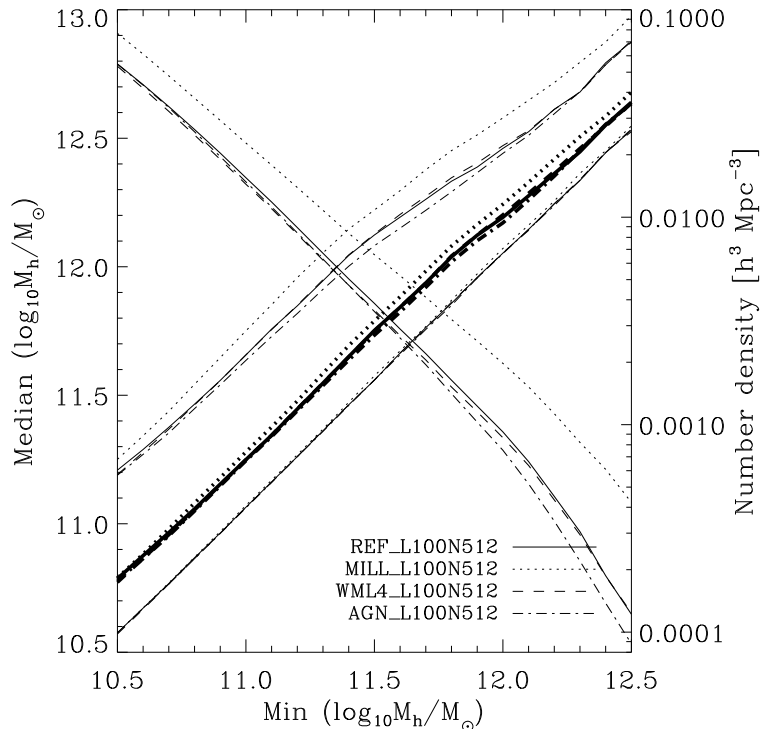
Figure 4.1 − Median halo mass above a given minimum halo mass (rising thick black curves), for different models, as measured from 100 h^{-1} cMpc boxes, with the 15.9th and 84.1th percentile (thin black rising curves above and below the median curves). The declining curves show the corresponding number density of haloes above a given minimum mass (right y-axis).

limit on the galaxy luminosity in the observed sample. Objects that are potentially missed by the selection based on rest-frame UV colors, e.g. massive red galaxies identified through their rest-frame optical colors (e.g. Franx et al. 2003), have space density smaller by more than an order of magnitude than the rest-frame UV selected galaxies used here, and are expected to have little impact on our statistics. In fact, we verified that using overlapping bins in halo mass with the specified minimum and maximum mass spaced by 1 dex yields very similar results to setting only the lower mass limit. This is due to the shape of the halo mass function, where the typical (median) mass above a given mass limit is only 0.2-0.3 dex higher than the set minimum (Figure 4.1).

The procedure for calculating absorption spectra is described in Appendix A4 of Theuns et al. (1998). QSO spectra from Rakic et al. (2011) were taken with Keck I/HIRES (Vogt et al. 1994) and have spectral resolution of FWHM $\approx 8.5\,\mathrm{km\,s^{-1}}$. We convolved our simulated spectra with a Gaussian with a FWHM = $6.6\,\mathrm{km\,s^{-1}}$, and rebin to $1.4\,\mathrm{km\,s^{-1}}$ pixels. We do not expect the slightly higher resolution in the simulated spectra to have impact

on our results. We add Gaussian noise with a signal to noise (S/N) ratio of 100 to the spectra, to match the typical S/N from the observations.

The optical depth of all pixels from simulated spectra, with original median optical depth of $\log_{10}\tau_{\mathrm{Ly}\alpha} = -1.32$ (for randomly extracted sightlines), is scaled to match the median absorption level in the observed spectra, $\log_{10}\tau_{\mathrm{Ly}\alpha} = -1.28$ (Rakic et al. 2011). This is justified because of the large uncertainties in the intensity of the ionizing background radiation.

4.3   Measuring halo masses

In this section we show the results of matching 2-D H I Lyα absorption maps, 1-D cuts through such maps, and radially averaged absorption profiles to those from the simulations.

4.3.1   Resolution tests

The *REF* model was run in 25, 50, and 100 h^{-1}cMpc boxes, while the rest of the simulations are only available in 25 and 100 h^{-1}cMpc boxes. The smallest box size does not have sufficient high-mass haloes and is therefore not suitable for our study. The 50 h^{-1}cMpc simulation box samples the full range of the mass function well, and for the *REF* model we used this box size to test convergence and resolution by comparing it with the results of the 100 h^{-1}cMpc box. The latter samples the high end of the mass function well, but resolves haloes with e.g. $M_{\mathrm{h}} = 10^{10.5}\,\mathrm{M}_\odot$ with less than 10^2 dark matter particles (for the simulation with 512^3 particles), which is insufficient to get reliable results for such objects. However, comparison of the *REF* model with observations suggests that the most relevant mass range is $M_{\mathrm{h}} > 10^{11}\,\mathrm{M}_\odot$, which means that using 100 h^{-1}cMpc simulation boxes is appropriate for our particular problem. We show convergence tests in the Appendix.

4.3.2   Measuring halo mass from 2-D absorption maps

Figure 4.2 shows the observed Lyα 2-D absorption map in the top panel and the maps from simulations for the minimum halo masses of $M_{\mathrm{min}} = 10^{10.5}$, $10^{11.5}$, and $10^{12.5}\,\mathrm{M}_\odot$, in the bottom panels from left to right respectively. The distance bins in these maps are logarithmically spaced by 0.15 dex, both in transverse direction and along the LOS. We added errors to simulated galaxy redshifts, to mimic those from the observations (Steidel et al. 2010): to a random 10% subsample of simulated redshifts we added errors with $\sigma = 60\,\mathrm{km\,s^{-1}}$ (mimicking errors in redshifts measured from nebular emission lines), and to the remaining 90% of redshifts $\sigma = 125\,\mathrm{km\,s^{-1}}$ (imitating errors measured from rest-frame UV absorption and emission lines). It is evident that haloes with $M_{\mathrm{h}} > 10^{10.5}\,\mathrm{M}_\odot$ produce too little absorption to account for the observed absorption maps, while the more massive haloes, $M_{\mathrm{h}} > 10^{12.5}\,\mathrm{M}_\odot$, produce too much absorption at large impact parameters.

The middle panel, showing absorption near haloes with the minimum mass of $M_h > 10^{11.5}$ M$_\odot$, resembles the observed distribution the most. The observed map is noisier, however, which is not surprising given that hundreds of sightlines contribute to each bin in simulated maps, as compared to ≈ 10 in the innermost bins of the observed map. None of the simulated maps shows enough absorption very close to galaxies (impact parameters $b \lesssim 100$ pkpc). Steidel et al. (2010) find that the absorption strength keeps rising closer to galaxies using their galaxy-galaxy pairs observations where the $b < 100$ pkpc region is sampled well, which suggests that there is probably a true deficiency of high column gas close to galaxies in the simulations, and that the discrepancy is not only a result of small number statistics (there are only 15 galaxies in the first, and 8 in the second impact parameter bin of Rakic et al. 2011). We will come back to this question in more detail below.
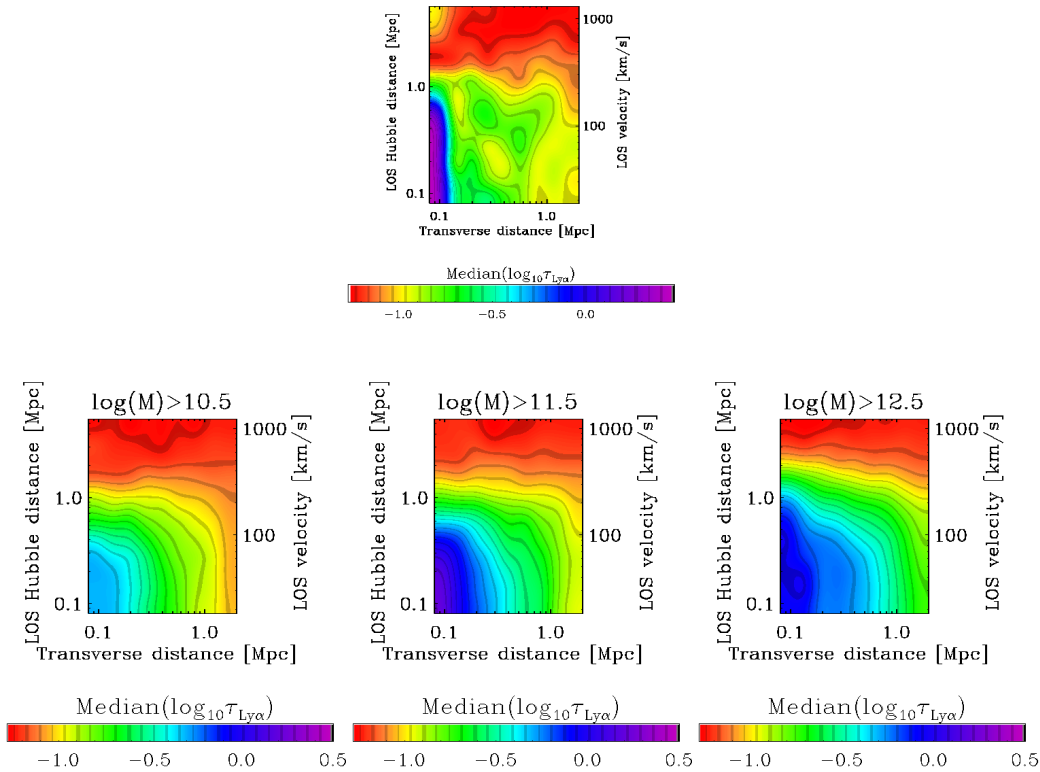

Figure 4.2 – Median 2-D distribution of Lyα absorption signal near haloes with, from left to right, $M_{min} > 10^{10.5}, 10^{11.5}$ and $10^{12.5}$ M$_\odot$ at $z = 2.25$ using the *REF* model. Redshift errors (as described in the text) were included. The observed map from Rakic et al. (2011, their Figure 6) is in the top panel. In comparison with the observed absorption map, the extent of absorption in the first panel is too small, while the third panel shows too much absorption at large impact parameters.

Figure 4.2 indicates the approximate mass range for host haloes to produce absorption comparable to the observations. More quantitatively, we estimate the minimum halo mass by minimizing the reduced χ^2 between the

observed 2-D Lyα absorption maps and maps from the simulations, taking into account errors on data points from observations. We use only regions of maps with LOS separations smaller than $616\,\mathrm{km\,s^{-1}}$ because the absorption signal is consistent with noise beyond this point (see Figure 4.2), and therefore these parts have no constraining power. Different distance bins in 2-D maps are uncorrelated in the transverse direction, but are correlated on scales of $200\,\mathrm{km\,s^{-1}}$ along the LOS (Chapter 3). We therefore use 1000 bootstrap realizations of the observed galaxy sample to estimate the errors on the halo mass, instead of using a $\Delta\chi^2$ criterion. For each bootstrap realization of the data, we find the implied minimum halo mass by comparing the resulting maps to the simulated maps, and then find the 1 and 2σ confidence intervals from the sample of bootstrap realizations. Given a grid of minimum halo masses spaced by 0.1 dex, the errors on halo masses sometimes take discrete values. In order to prevent errors with zero values, in cases where this happens we set the error to half the grid spacing, i.e. 0.05 dex, since we cannot determine errors to better than this value.
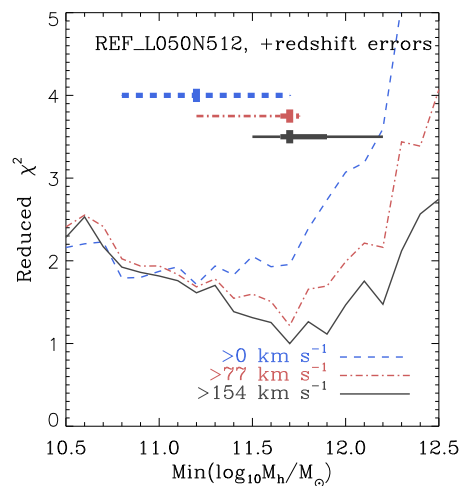

Figure 4.3 – The reduced χ^2 estimated by comparing the observed 2-D Lyα absorption map with the simulated maps for different minimum halo masses (x-axis), for the *REF* model in a 50 h^{-1} cMpc box. Given that regions close to galaxies are most affected by the uncertain baryonic physics and galaxy redshift errors, we show the resulting curves after using only absorbing regions separated by LOS velocity intervals greater than the values indicated in the legend. The curves show the results for the full sample. Their minima correspond to the best-fit minimum halo masses. The vertical rectangles show the median minimum halo masses as estimated from different bootstrap realizations. The thick horizontal lines show the 1σ confidence interval from bootstrap realizations, and the thin horizontal lines show the 2σ confidence intervals.

The dashed curves in Figures 4.3 and 4.4 show the reduced χ^2 from comparing absorption maps as a function of minimum halo mass used for simulated data. For the 50 h^{-1}cMpc *REF* simulation the implied mass is $\log_{10} M_{\min} = 11.2^{+0.5+0.5}_{-0.4-0.4}\,\mathrm{M_\odot}$, and for the 100 h^{-1}cMpc box it is $\log_{10} M_{\min} =$
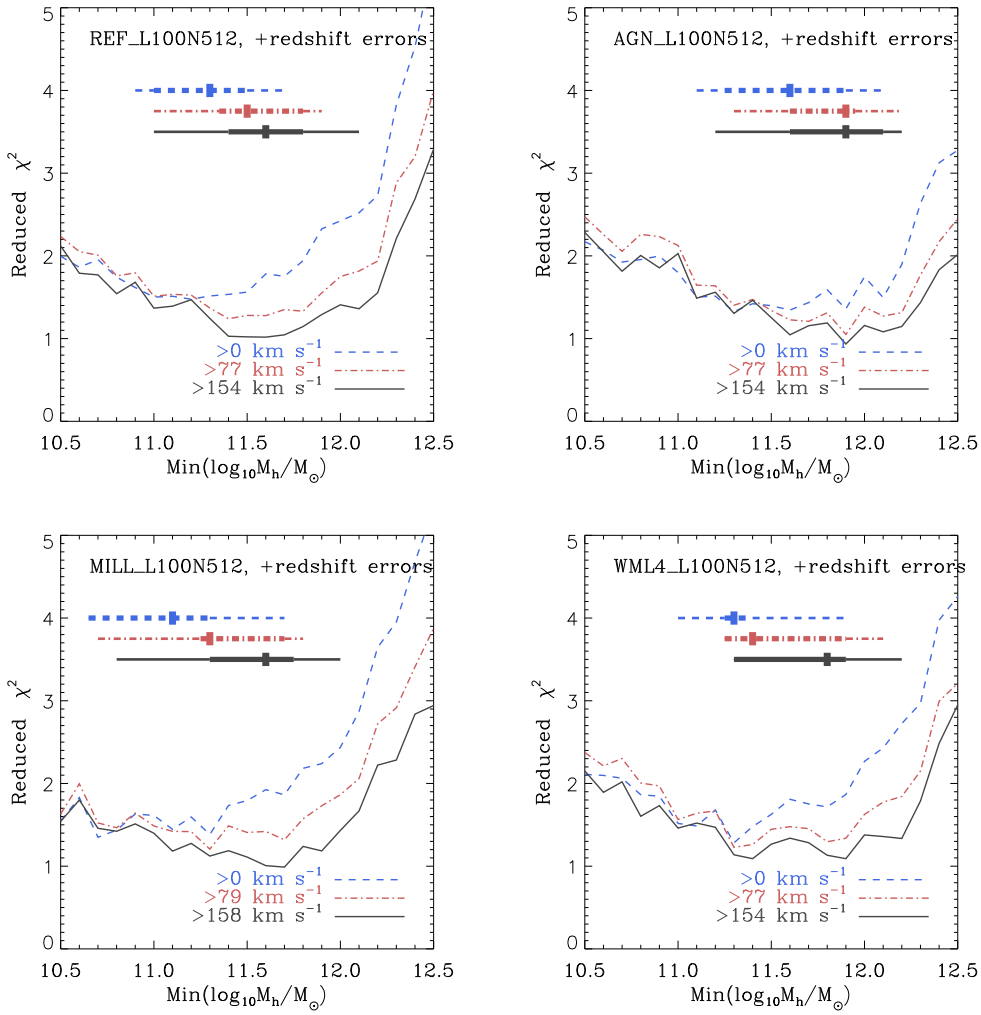
Figure 4.4 − Similar to Figure 4.3 but for more models simulated in 100 h^{-1} cMpc boxes (as opposed to 50 h^{-1} cMpc in the previous figure).

$11.2^{+0.3+0.5}_{-0.3-0.4}$ M$_\odot$ (the first and second pairs of errors indicate 1σ and 2σ confidence intervals, respectively). The two simulations give almost identical results, which justifies using larger boxes for the rest of the models.

Because uncertain baryonic physics and redshift errors could affect the absorption signal close to galaxies, we also compare parts of maps where these uncertainties are expected to have a smaller impact. We do so by excluding regions with $v_{\rm LOS} < 77$, and $154\,{\rm km\,s^{-1}}$ (separated by 0.3 dex in logarithmic space), or $v_{\rm LOS} < 79$ and $158\,{\rm km\,s^{-1}}$ for the *MILL* model (due to different cosmological parameters the same proper distance corresponds to different velocity separation), where $v_{\rm LOS}$ is the line of sight velocity separation between absorbers and galaxies. The resulting minimum halo mass from such conservative comparisons is in all cases consistent with the one inferred by compar-

ing the complete maps. Given that the simulations seem to underproduce the level of absorption very close to galaxies in comparison with the observations (e.g. Figure 4.2), and given the uncertain baryonic physics in the vicinity of galaxies, we choose to compare absorption distributions from observations and simulations only at LOS velocity separations $v_{\rm LOS} > 154\,{\rm km\,s^{-1}}$ (which is comparable to the redshift errors in the galaxy sample) as our default case. This implies a minimum halo mass of $\log_{10}M_{\rm min} = 11.7^{+0.2+0.5}_{-0.05-0.2}\,{\rm M_\odot}$ using the $50\ h^{-1}{\rm cMpc}$ box, and $\log_{10}M_{\rm min} = 11.6^{+0.20+0.50}_{-0.20-0.60}\,{\rm M_\odot}$ using the $100\ h^{-1}{\rm cMpc}$ box.

4.3.2.1   *The AGN, MILL and WML4 models*

The minimum mass implied by comparison of observations with the *AGN* simulation is $\log_{10}M_{\rm min} = 11.9^{+0.20+0.30}_{-0.30-0.70}\,{\rm M_\odot}$. This is somewhat higher than mass inferred from the *REF* model ($\log_{10}M_{\rm min} = 11.6^{+0.20+0.50}_{-0.20-0.60}\,{\rm M_\odot}$), but still consistent within 1σ. McCarthy et al. (2011) found that AGN eject large amounts of gas from galaxies at $z \sim 2-3$ using this simulation, which is why we need to move to higher mass haloes to get high enough columns of neutral gas to account for the observed absorption. Nevertheless, given that results differ by only 0.3 dex even when using a model with such extreme feedback, we conclude that this halo mass measuring method is robust to changes in baryonic physics prescriptions.

To quantify the impact of cosmology on our results, we compare the *MILL* and *WML4* simulations, which differ mostly in terms of σ_8, which is 0.9 for the former, and 0.74 for the latter model. We find minimum masses of $\log_{10}M_{\rm min} = 11.7^{+0.05+0.30}_{-0.40-0.90}\,{\rm M_\odot}$ and $11.4^{+0.50+0.80}_{-0.1-0.10}\,{\rm M_\odot}$, for *MILL* and *WML4*, respectively. These estimates are consistent with the value implied by the *REF* model. The minimum halo mass for the *MILL* model, slightly higher than the one for *WML4*, has a large uncertainty on the lower side. In addition, Figure 4.4 shows that the χ^2 curve for this simulation where we exclude only regions with LOS separations $\le 79\,{\rm km\,s^{-1}}$ gives results that are more consistent with the *WML4* model ($\log_{10}M_{\rm min} = 11.3^{+0.4+0.5}_{-0.05-0.6}\,{\rm M_\odot}$). The fact that *WML4* gives a result consistent with that for the *MILL* model suggests that the exact value of σ_8 does not make a large difference.

For comparison we show 2-D Lyα absorption maps for different models in Figure 4.5, for haloes with $M_{\rm h} > 10^{11.5}\,{\rm M_\odot}$. *AGN* model shows weaker absorption signal within ~ 100 kpc from galaxies in comparison with other models, and *WML4* shows less absorption relative to the *REF* model.

The implied minimum masses from all models are summarized in Table 4.2. It is clear that the statistical errors on the minimum halo mass, which are around 0.2-0.3 dex, are comparable to the systematic errors arising due to uncertainties in the physics important for galaxy formation, as inferred by using simulations with different feedback prescriptions and different cosmologies. This demonstrates the robustness of this method for
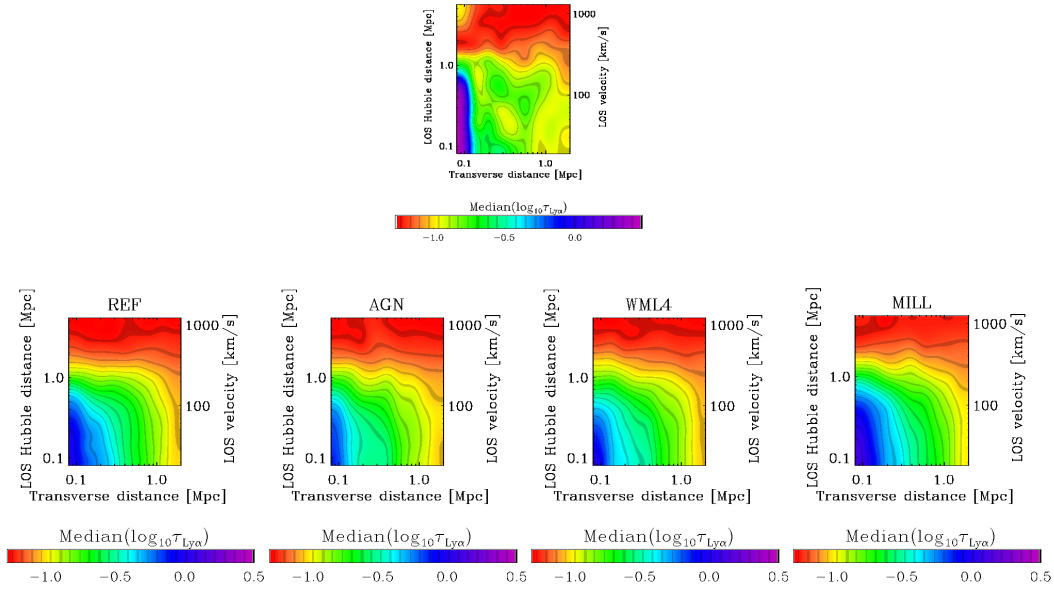
Figure 4.5 − Median 2-D distribution of Lyα absorption signal near haloes with $M_{\min} > 10^{11.5}$ at $z = 2.25$ using the *REF*, *AGN*, *WML4*, and *MILL* model, from left to right respectively. The observed map from Rakic et al. (2011, their Figure 6) is in the top panel.

Table 4.2 − The inferred minimum halo masses for different simulations are shown in the first column (with 1 and 2σ confidence intervals), while the corresponding number densities of haloes (and 1 and 2σ confidence intervals) as measured from the simulations with $100\, h^{-1}$ cMpc boxes are presented in the second column. For reference, the observed number density of galaxies in the photometric sample, from which objects in Rakic et al. (2011) were drawn for spectroscopic follow-up, is $11 \times 10^{-3}\, h^3\, \mathrm{Mpc}^{-3}$, with an uncertainty of 10% (Reddy et al. 2008). Galaxies in the spectroscopic sample are weighted toward brighter objects in comparison, and their number density is lower.

Models	Min. halo mass [$\log_{10} \mathrm{M}_\odot$]	n [$h^3\, \mathrm{Mpc}^{-3}$]
REF	$11.6^{+0.20+0.50}_{-0.20-0.60}$	$2.9^{+2.2+12.6}_{-1.2-2.2} \times 10^{-3}$
AGN	$11.9^{+0.20+0.30}_{-0.30-0.70}$	$1.0^{+1.6+7.3}_{-0.5-0.7} \times 10^{-3}$
MILL	$11.7^{+0.05+0.30}_{-0.40-0.90}$	$4.1^{+6.7+33.3}_{-0.5-2.1} \times 10^{-3}$
WML4	$11.4^{+0.50+0.80}_{-0.10-0.10}$	$4.7^{+1.5+1.5}_{-3.5-4.3} \times 10^{-3}$

measuring halo masses.

4.3.2.2   *Number densities of haloes*

The number density of galaxies in the photometric sample, from which galaxies in Rakic et al. (2011) were drawn, is $11 \times 10^{-3}\, h^3\, \mathrm{Mpc}^{-3}$, with an uncertainty of 10% (Reddy et al. 2008). The largest weights for spectroscopic follow-up were given to objects in the apparent magnitude range $R = 23-24.5$

(Steidel et al. 2004), while the limit for the photometric sample is $R = 25.5$. In addition, objects near the QSO sightlines (i.e. within 1-2 arcminutes) were given extra weight. There are ~ 2.5 times more objects in the photometric sample for the limiting magnitude of $R = 25.5$, than for $R = 24.5$, and so the lower limit on the number density of galaxies in the spectroscopic sample is $\sim 4.5 \times 10^{-3} \, h^3 \, \mathrm{Mpc}^{-3}$. The real number density of objects in the spectroscopic sample is somewhere in between, because a non-negligible fraction of objects with $R > 24.5$ also have a spectroscopic follow-up.

Number densities of haloes above a minimum mass inferred in the previous section using different models, are given in Table 4.2. Their values are $2.9^{+2.2+12.6}_{-1.2-2.2} \times 10^{-3}$, $1.0^{+1.6+7.3}_{-0.5-0.7} \times 10^{-3}$, $4.1^{+6.7+33.3}_{-0.5-2.1} \times 10^{-3}$, and $4.7^{+1.5+1.5}_{-3.5-4.3} \times 10^{-3} \, h^3 \, \mathrm{Mpc}^{-3}$ (quoting 1 and 2 σ confidence intervals as estimated from the errors on the minimum halo mass), for the *REF*, *AGN*, *MILL* and *WML4* model, respectively. All the estimates, apart for *AGN*, are consistent with the observed number densities within 1σ, and *AGN* agrees with the observations within 2σ. The minimum halo mass for the *MILL* model is 0.1 dex higher than that for the *REF* model, and yet the number density is also somewhat higher, while we would naively expect the number density to be lower given the higher halo mass. This is due to higher value of σ_8 in the *MILL* model (0.9 as compared to 0.74), which causes faster growth of structure (see Figure 4.1). Number densities for the *AGN* model although very low in comparison with other models, are consistent with them due to the large uncertainty on the minimum halo mass estimate.

4.3.2.3 Comparison with estimates from clustering measurements

The latest estimates of halo masses of $z \sim 2$ star-forming galaxies were presented by Conroy et al. (2008) and Trainor et al. (2011, in preparation). As already mentioned in the introduction, using a clustering analysis Conroy et al. (2008) found the minimum halo mass to be $10^{11.55} \, \mathrm{M}_\odot$ using the WMAP1 cosmology, and 0.3-0.4 dex smaller (i.e. $10^{11.15} - 10^{11.25} \, \mathrm{M}_\odot$) using the WMAP3 cosmology (which is the default cosmology in our study), and quote an error of ± 0.2 dex. Our results for the models that use these cosmological parameters ($\log_{10} M_{\min} = 11.7^{+0.05+0.30}_{-0.40-0.90} \, \mathrm{M}_\odot$ for the *MILL* model, using WMAP1, and $11.4^{+0.50+0.80}_{-0.10-0.10} \mathrm{M}_\odot$ for the *WML4* model, using WMAP3), although consistent within errors with their results, are both slightly higher. This could be due to the fact that they used the correlation function measured from the photometric galaxy sample, while we use only the brighter spectroscopic sample. Adelberger et al. (2005b) found that brighter galaxies in their photometric sample cluster more strongly, and therefore likely reside in more massive haloes. Trainor et al. (2011, in preparation) measure masses of $10^{11.9} \, \mathrm{M}_\odot$ for the spectroscopic sample using WMAP1 cosmological parameters, i.e. higher by ~ 0.3 dex than result by Conroy et al. (2008) for the photometric sample. Given that our results when using both the WMAP1

and WMAP3 cosmological parameters are ~ 0.2 dex higher than those from Conroy et al. (2008), in line with the expectation for the brighter spectroscopic sample, we conclude that the two methods (i.e. clustering analysis and matching of the Lyα absorption profiles) give consistent results.

4.3.2.4 Relevance for future surveys

We demonstrated that comparisons of observed 2-D Lyα absorption maps around galaxies with maps extracted from cosmological SPH simulations can be used to measure halo masses of the galaxy population in question. Our observed QSO-galaxy fields typically span areas of $\approx 5 \times 7$ arcminutes with the QSO in the middle. Here we test the performance of our method for a limited survey area of only $\approx 30 \times 30$ arcseconds, which corresponds to a cut along the LOS through the first impact parameter bin of the above 2-D maps ($b \lesssim 130$ pkpc). Figure 4.6 shows that the implied minimum mass is $\log_{10} M_{\mathrm{min}} = 11.6^{+0.50+0.50}_{-0.05-0.05}$ M$_\odot$, which agrees very well with that inferred from a comparison of the full maps ($\log_{10} M_{\mathrm{min}} = 11.6^{+0.20+0.50}_{-0.20-0.60}$ M$_\odot$).
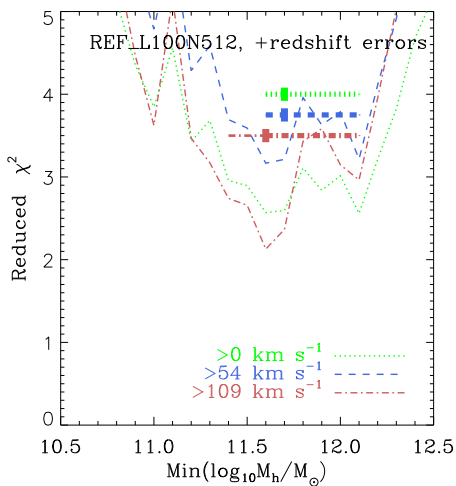

Figure 4.6 – The reduced χ^2 estimated by comparing the observed Lyα absorption profile through the first impact parameter bin of a 2-D map (i.e. cut along the LOS). The implied minimum mass agrees with the one indicated by comparison of full 2-D maps ($\log_{10} M_{\mathrm{min}} = 11.6^{+0.50+0.50}_{-0.05-0.05}$ M$_\odot$ as compared to the $11.6^{+0.20+0.50}_{-0.20-0.60}$ M$_\odot$, respectively).

This finding is relevant for future galaxy-QSO pair surveys such as the one planned with the VLT/MUSE integral field spectrograph (Bacon et al. 2010), whose field of view is 1×1 arcminute. Applying our method to MUSE observations will allow estimates of halo masses of Lyα emitters from a single telescope pointing on each QSO in the survey. This eliminates the need to make a wide field mosaic to measure clustering, which would in any case not be feasible for the faint objects dominating the number density of galaxies detected in the deep QSO fields.

4.3.3   Measuring mass from radially averaged absorption profiles

Comparing radially averaged absorption profiles to the ones from simulations yields results consistent with those implied by 2-D distributions but the confidence intervals are larger. This is not surprising given that here we average over the transverse and LOS signal, in which case we lose some constraining power in comparison with the 2-D map. Figure 4.7 shows the median $\log_{10}(\tau_{\mathrm{Ly}\alpha})$ as a function of distance from galaxy haloes, in 5 mass bins, with the observational results shown as gray symbols. The observed profile follows the one near haloes with $M_\mathrm{h} > 10^{11.5}\,\mathrm{M}_\odot$ the most closely. In this plot we see that the discrepancy at small impact parameters, where simulations underpredict the amount of absorption, is at the 2σ level. As mentioned above, Steidel et al. (2010) find that the absorption within 100 pkpc keeps rising, which suggests a true deficiency of high column density gas at small impact parameters. This could be explained if cold gas physics in the simulations is not captured sufficiently well (e.g. in reality gas at ISM densities can self-shield to form a cold phase with $T \ll 10^4$ K, while in the simulations its temperature stays at $\gtrsim 10^4$ K). Also, the SN feedback prescription in the *REF* model is possibly not fully appropriate and perhaps heats the gas near galaxies to higher temperatures than in reality.
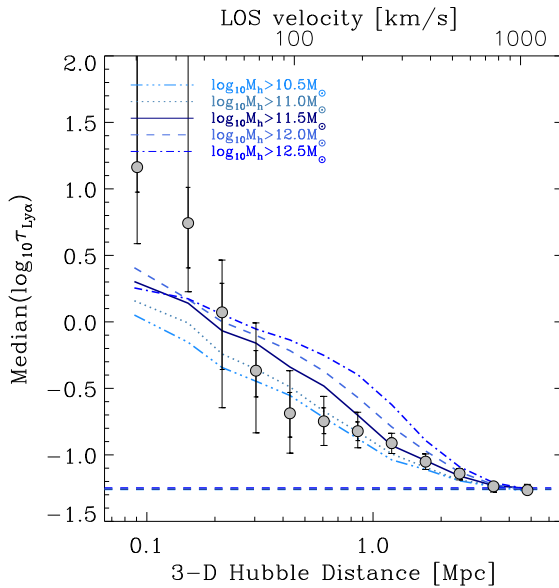

Figure 4.7 – Median $\log_{10}(\tau_{\mathrm{Ly}\alpha})$ as a function of proper distance from galaxy haloes in the 100 h^{-1} cMpc *REF* simulation. Curves represent different mass bins, as indicated in the legend, while black symbols with 1σ and 2σ error bars show the results from the observations. We added redshift errors to galaxy positions in simulations (see text for more details).

4.4   Redshift space anisotropies

Rakic et al. (2011) report on two types of redshift space anisotropies in the observed 2-D Lyα absorption maps. At small impact parameters ($\lesssim 200$ kpc) the absorption signal is elongated along the LOS, which is likely a result of gas peculiar velocities of $\approx 200\,\mathrm{km\,s^{-1}}$, and errors in galaxy redshifts ($\approx 125\,\mathrm{km\,s^{-1}}$). On the other hand, at large separations from galaxies (≈ 1.5 Mpc) the absorption appears compressed along the LOS, which probably reflects large-scale gas infall into the potential wells populated by the observed galaxies.

Here we use simulated observations to examine the origin of the observed redshift space anisotropies. Figure 4.8 shows 2-D absorption maps centered on haloes with masses higher than the minimum mass inferred in the previous section, i.e. $M_h > 10^{11.5}\,M_\odot$, for 4 cases: *i)* default case, i.e. redshift errors of $\sigma = 60\,\mathrm{km\,s^{-1}}$ for 10% of galaxy redshifts, and $\sigma = 125\,\mathrm{km\,s^{-1}}$ for the remaining 90% of redshifts, and taking into account peculiar velocities, *ii)* no redshift errors, *iii)* ignoring peculiar velocities, and *iv)* a number of sightlines per distance bin identical to the observations of Rakic et al. (2011).

Comparing the middle left and right panel of Figure 4.8 we see that the effect of redshift errors is to wash out the signal along the LOS.

The bottom left panel shows the case where we ignore peculiar velocities. The absorption enhancement is more isotropic around the halo positions. At small impact parameters ($\lesssim 200$ pkpc), the absorption is much reduced, demonstrating that peculiar velocities strongly enhance the small-scale absorption excess. In addition, comparison with the middle left panel suggests that the large-scale ($\gtrsim 1.5$ Mpc) compression along the LOS is due to large-scale infall. This is the first observation of gas collapse into large-scale potential wells. We just note here that ignoring both peculiar velocities and redshift errors results in a map with axisymmetric absorption.

In the bottom right panel we limit the number of sightlines per distance bin to match the observations. The resulting map is noisier in comparison with the previous panels with simulated maps, but the redshift space distortions seen in the first panel are also clearly visible in this panel.

A more quantitative comparison is provided by Figures 4.9 and 4.10 which show cuts along the LOS and in transverse direction, respectively, for three different cases. The top left panel of Figure 4.9 makes it very clear that the simulations underpredict the amount of absorption for impact parameters smaller than ≈ 100 kpc. Note, however, that a simple "χ by eye" would strongly overestimate the significance of the discrepancy because the points are strongly correlated in the line of sight direction (Rakic et al. 2011). Both figures show that redshift errors smooth the signal along the LOS, on scales $\lesssim 10^2\,\mathrm{km\,s^{-1}}$. The curves for the case where we do not take peculiar velocities into account lie below the default curves for velocity separations $\lesssim 150\,\mathrm{km\,s^{-1}}$, and above them at larger velocity separations. The increased

optical depth closer to galaxies could be due to large-scale infall of gas enhancing absorption through "filling" in redshift space. Another possibility is that the enhanced absorption is due to lines getting broadened due to large scale galactic outflows, which would also significantly affect our median statistics. While there is some ambiguity at small impact parameters, the evacuation of absorption on large scales is likely due to large scale gas infall.

In Figure 4.11 we add different amounts of scatter to the galaxy redshifts. Large redshift errors smooth the absorption signal more along the LOS. The bottom panels show cases with redshift errors of $\sigma = 300$ and $400\,\mathrm{km\,s^{-1}}$, where instead of compression along the LOS relative to the direction transverse to the LOS, we see elongation. This suggests that the errors in galaxy redshifts have to be $< 200\,\mathrm{km\,s^{-1}}$ to study large scale-gas infall into the potential wells of star-forming galaxies.

4.5   Summary & Conclusions

Several observations have shown that the absorption of the light from background QSOs is enhanced at the wavelength of H I Lyα when a sightline passes within several pMpc from a galaxy. We compared the observed Lyα absorption distribution around star-forming galaxies at $z \approx 2.36$ from Rakic et al. (2011) with cosmological SPH simulations in order to statistically constrain the total halo masses of the observed galaxy population, and to explain the observed redshift space distortions.

The most important conclusions of this study are:

- Using cosmological SPH simulations from the OWLS suite of models we derive the minimum halo mass of $\log_{10}M_{\min} = 11.6^{+0.20+0.50}_{-0.20-0.60}\,\mathrm{M_\odot}$ for the *REF* model, in good agreement with the estimates from the clustering analysis (Conroy et al. 2008) and Trainor et al. (2011, in preparation) performed on the same galaxy population.

- Our method is robust to changes in feedback prescriptions as demonstrated by comparison with the *AGN* model in which galaxies eject large amounts of gas at the redshift of interest here. The implied minimum mass of $\log_{10}M_{\min} = 11.9^{+0.20+0.30}_{-0.30-0.70}\,\mathrm{M_\odot}$, although 0.3 dex higher, is consistent with the *REF* model within 1σ.

- The minimum halo mass inferred from comparison with the *MILL* simulation ($\log_{10}M_{\min} = 11.7^{+0.05+0.30}_{-0.40-0.90}\,\mathrm{M_\odot}$) is consistent with that required by the *WML4* model ($\log_{10}M_{\min} = 11.4^{+0.50+0.80}_{-0.10-0.10}\,\mathrm{M_\odot}$), which suggests that our method is not very sensitive to differences in cosmological parameters (WMAP1 vs. WMAP3).

- If we consider only a cut along the LOS for impact parameters $\lesssim 100$ pkpc, then we infer a mass of $\log_{10}M_{\min} = 11.6^{+0.50+0.50}_{-0.05-0.05}\,\mathrm{M_\odot}$, which is almost identical to the one obtained from the full 2-D map (impact parameters $\lesssim 2$ pMpc). This is very encouraging for future narrow

field QSO-galaxy surveys (e.g. the one planned with the VLT/MUSE), where one will be able to estimate halo masses of Lyα emitters with a single telescope pointing, without the need for a wide field mosaic.

- The observed elongation of the absorption signal of $\approx 200\,\mathrm{km\,s^{-1}}$ at small impact parameters is a product of uncertainties in galaxy redshifts, observed to be $\approx 125\,\mathrm{km\,s^{-1}}$, and the peculiar motions of gas in and around galaxy haloes.
- The compression of the signal on large scales is a result of gas infall into potential wells occupied by galaxies. To observe this infall, the accuracy of galaxy redshifts has to be better than $\sim 200\,\mathrm{km\,s^{-1}}$, as larger redshift errors smooth the absorption signal along the LOS, disguising the compression signature.
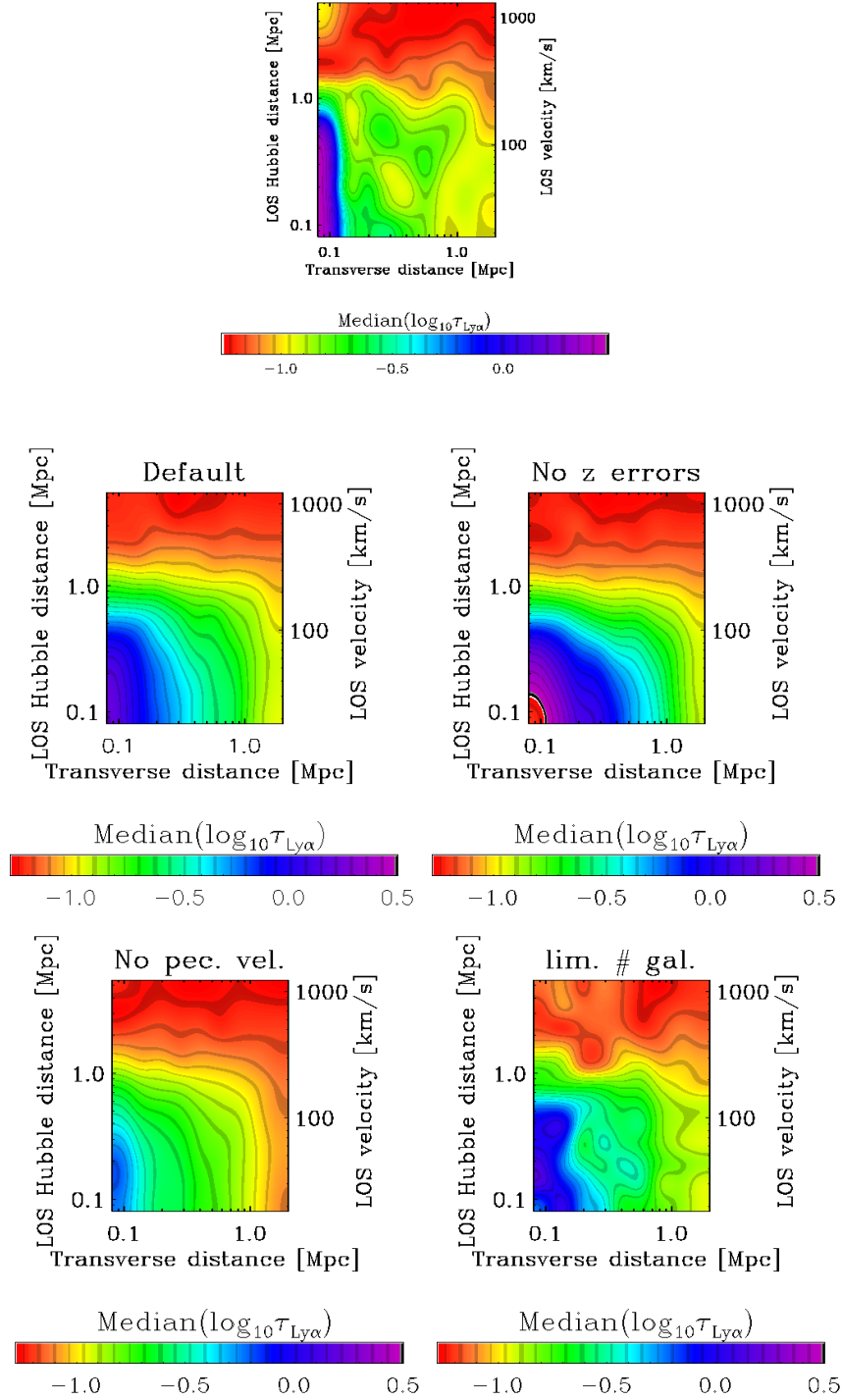
Figure 4.8 − 2-D Lyα absorption maps for haloes with $M_{\rm h} > 10^{11.5}\,{\rm M_\odot}$ at $z = 2.25$ in the 50 h^{-1} cMpc *REF* simulation. Top panel: observations of Rakic et al. (2011). Middle left panel: default case (errors in galaxy redshifts, and taking into account peculiar velocities). Middle right panel: no redshift errors. Lower left panel: ignoring peculiar velocities. Lower right panel: number of galaxies per impact parameter bin matches observations of Rakic et al. (2011). Redshift errors increase redshift space distortions on small scales and decrease the signal, while peculiar velocities act to compress signal along the LOS on large scales, elongate it along the LOS for small impact parameters, and increase the signal.
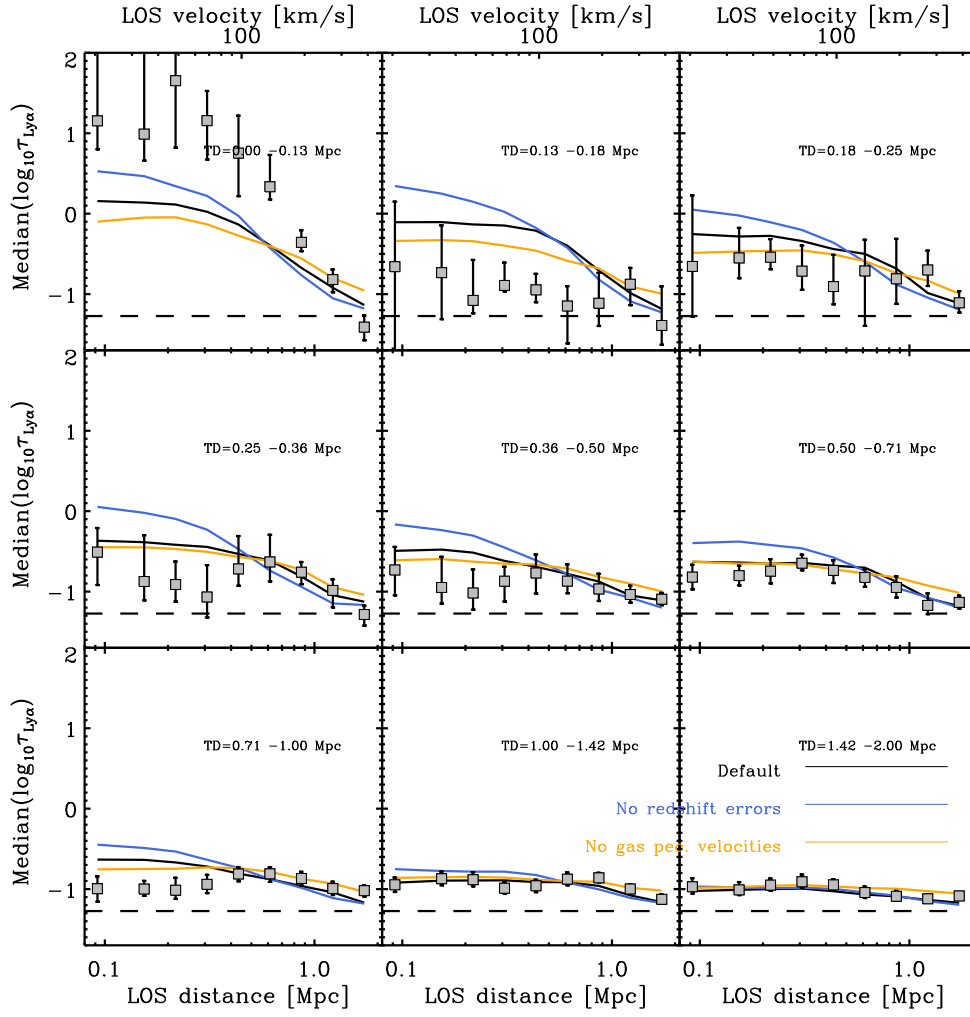
Figure 4.9 − Cuts along the LOS through panels of Figure 4.8 ($M_h > 10^{11.5}$ $M_\odot$). The gray symbols show the observations of Rakic et al. (2011). The horizontal dashed line indicates the median optical depth of all pixels.
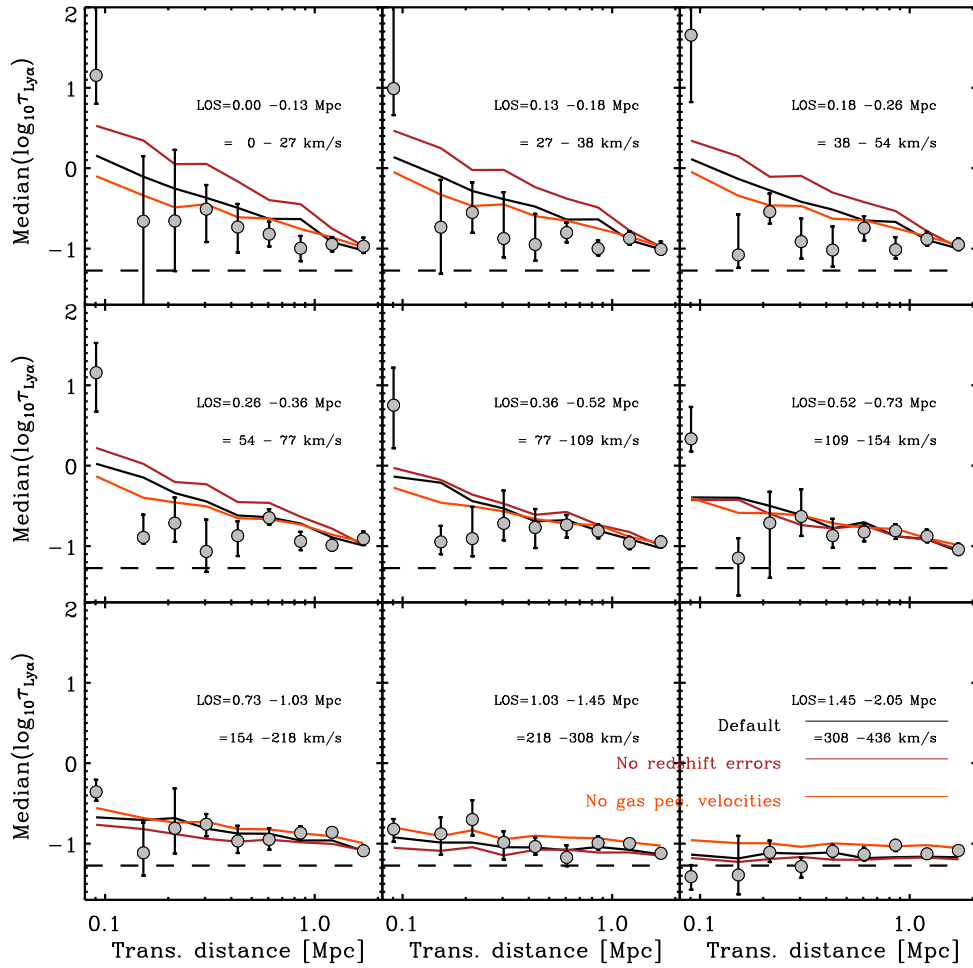
Figure 4.10 − Similar to Figure 4.9, but showing cuts in transverse direction through panels of Figure 4.8 ($M_h > 10^{11.5}$ $M_\odot$).

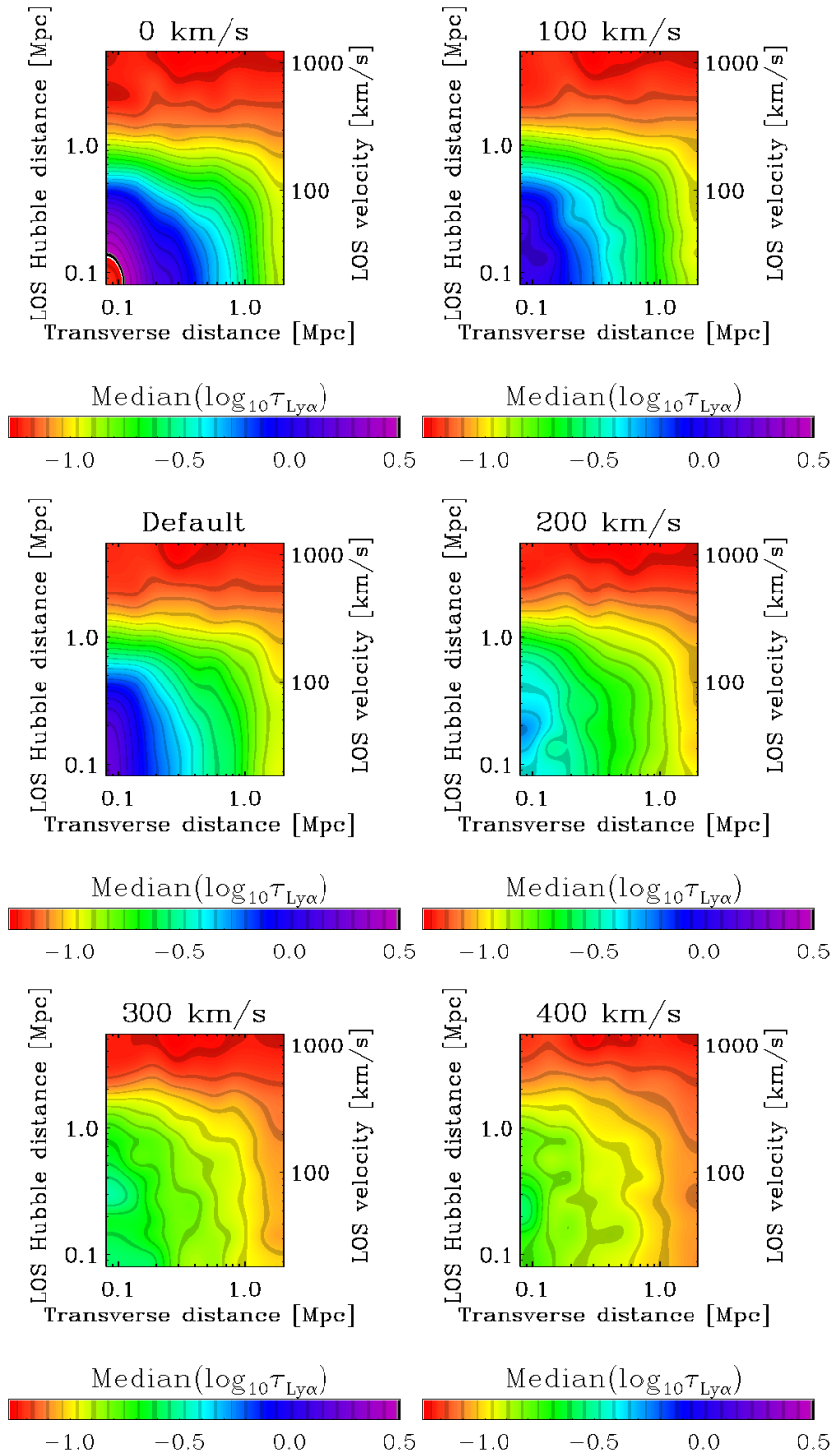**Figure 4.11** − Similar to Figure 4.8 but adding different levels or errors in galaxy redshifts: $\sigma = 0$, 100, default redshift errors (described in the text), 200, 300, and 400 $\mathrm{km\,s^{-1}}$, from left to right, and from top to bottom. Adding errors with $\sigma \gtrsim 200\,\mathrm{km\,s^{-1}}$ masks the compression of the absorption distribution along the LOS relative to that transverse to the LOS.

References

Adelberger K. L., Steidel C. C., Giavalisco M., Dickinson M., Pettini M., Kellogg, M., 1998, ApJ, 505, 18

Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, ApJ, 584, 45

Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Reddy N. A., 2005, ApJ, 629, 636

Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K. 2005, ApJ, 619, 697

Bacon R. et al., 2010, Proc. SPIE, 7735, 7

Booth, C. M., Schaye, J., 2009, MNRAS, 398, 53

Booth, C. M., Schaye, J., 2010, MNRAS, 405, 1

Chabrier, G., 2003, PASP, 115, 763

Chen H.-W., Lanzetta K. M., Webb J. K., Barcons X., 1998, ApJ, 498, 77C

Chen H.-W., Lanzetta K. M., Webb J. K., 2001, ApJ, 556, 158

Croft R. A. C., Hernquist L., Springel V., Westover M., White M., 2002, ApJ, 580, 634

Conroy, C., Wechsler, R. H., Kravtsov, A. V., 2006, ApJ, 668, 826

Conroy C., Shapley A. E., Tinker J. L., Santos M. R., Lemson G.

Crighton N. et al., 2010 MNRAS, 414, 28

Cole, S., Kaiser, N., MNRAS, 237, 1127

Dalla Vecchia, C., Schaye, J., 2008, MNRAS, 387, 1431

Dolag, K., Borgani, S., Murante, G., Springel, V., 2009, MNRAS, 399, 497

Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., Hernquist, L., 2008, ApJ, 673, 39

Franx, M., Labbé, I., Rudnick, G., van Dokkum, P. G., Daddi, E., Förster Schreiber, N. M., Moorwood, A., Rix, H.-W., Röttgering, H., van der Wel, A., van der Werf, P., van Starkenburg, L., 2003, ApJ, 587, 79

Gavazzi, R., Treu, T., Rhodes, J. D., Koopmans, L. V. E., Bolton, A. S., Burles, S., Massey, R. J., Moustakas, L. A., 2007, ApJ, 667, 176

Gunn, J. E., Peterson, B. A., 1965, ApJ, 142, 1633

Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays Modelling the UV/X-ray cosmic background with CUBA, arXiv: 0106018

Haas, M. R., Schaye, J., Jeeson-Daniel, A., 2011, arXiv: 1103.0547

Kaiser, N., 1984 ApJ, 284, 9

Kaiser, N., Squires, G., 1993, ApJ, 404, 441

Kim, Y.-R., Croft, R. A. C., 2008, MNRAS, 387, 377

Komatsu E. et al., 2011, ApJS, 192, 18

Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., Brinkmann, J., 2006, MNRAS, 368, 715

McCarthy, I. G., Schaye, J., Bower, R. G., Ponman, T. J., Booth, C. M., Dalla Vecchia, C., Springel, V., 2011, MNRAS, 412, 1965

More, S., van den Bosch, F. C., Cacciato, M., Skibba, R., Mo, H. J., Yang, X., 2011, MNRAS, 410, 210

Motl, P. M., Hallman, E. J., Burns, J. O., Norman, M. L., 2005, ApJ, 623, 63

Penton S. V., Stocke J. T., Shull J. M., 2002, ApJ, 565, 720

Prochaska, J. X., Weiner, B., Chen, H. -., Mulchaey, J. S., Cooksey, K. L., 2011, arXiv: 1103.1891

Quadri, R., van Dokkum, P., Gawiser, E., Franx, M., Marchesini, D. , Lira, P., Rudnick, G., Herrera, D., Maza, J., Kriek, M., Labbé, I., Francke, H., 2007, ApJ, 654, 138

Rakic et al., 2011, in preparation

Reddy, N. A., Steidel, C. C., Pettini, M., Adelberger, K. L., Shapley, A. E., Erb, D. K., Dickinson, M., 1998, ApJS, 175, 48

Reiprich, T. H., Böhringer, H., ApJ, 567, 716

Schaye, J., 2001, ApJ, 559, 507

Schaye J., Rauch M., Sargent W. L. W., Kim T.-S., 2000, ApJ, 541, 1

Schaye J., Dalla Vecchia, C., 2008, MNRAS, 383.1210

Schaye, J., Dalla Vecchia, C., Booth, C. M., Wiersma, R. P. C., Theuns, T., Haas, M. R., Bertone, S., Duffy, A. R., McCarthy, I. G., van de Voort, F., 2010, MNRAS, 402, 1536

Spergel et al., 2007, ApJS, 170, 377

Springel, V., White, S. D. M., Tormen, G., Kauffmann, G., 2001, MNRAS, 328, 726

Springel V., 2005, MNRAS, 364, 1105

Springel, V. et al., 2005, Nature, 435, 629

Steidel, C. C., Shapley, A. E., Pettini, M., Adelberger, K. L., Erb, D. K., Reddy, N. A., Hunt, M. P., 2004, ApJ, 604, 534

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289

Sunyaev, R. A., Zeldovich, Y. B., CoASP, 4, 173

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

van de Voort, F., Schaye, J., Booth, C. M., Haas, M. R., & Dalla Vecchia, C. 2011, MNRAS, 414, 2458

Vogt S. S., Allen S. L., Bigelow B. C., Bresee L., Brown B., Cantrall T., Conrad A., Couture M., 1994, Proc. SPIE, 2198, 362

Wiersma, R. P. C., Schaye, J., Smith, B. D., 2009. MNRAS, 393, 99

Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., Tornatore, L., 2009, MNRAS, 399, 574

Zwicky, F., 1937, ApJ, 86, 217

4.A Convergence Tests

Figure 4.12 shows the mass resolution test for the 50 h^{-1} cMpc *REF* simulation, performed with 128^3, 256^3, and 512^3 particles. The resulting median $\log_{10}(\tau_{\mathrm{Ly}\alpha})$ curves as a function of proper distance from the galaxies are consistent with each other for all but the highest mass bin. In the highest mass bin the resulting curve for the simulation with the poorest resolution is below the curves for simulations with higher resolution, but the simulations with 256^3, and 512^3 particles give consistent results. Therefore, we can conclude that the simulation results are converged with respect to the particle mass.

Figure 4.13 shows the results for the 50 h^{-1} cMpc *REF* simulation with 256^3, and 100 h^{-1} cMpc *REF* simulation with 512^3 particles. In this case the particle masses in the two simulations are the same, but the simulation

Figure 4.12 – Median $\log_{10}(\tau_{Ly\alpha})$ as a function of proper distance from galaxy haloes in the 50 h^{-1} cMpc *REF* simulations, with 128^3, 256^3, and 512^3 particles. Each panel shows the results for a different mass range, as indicated in the legend, while grey symbols with 1σ and 2σ error bars show the results from the observations. We added errors to galaxy positions in simulations (see text for more details).

volume is changed. The median optical depth curves show that the results are converged with respect to the box size.
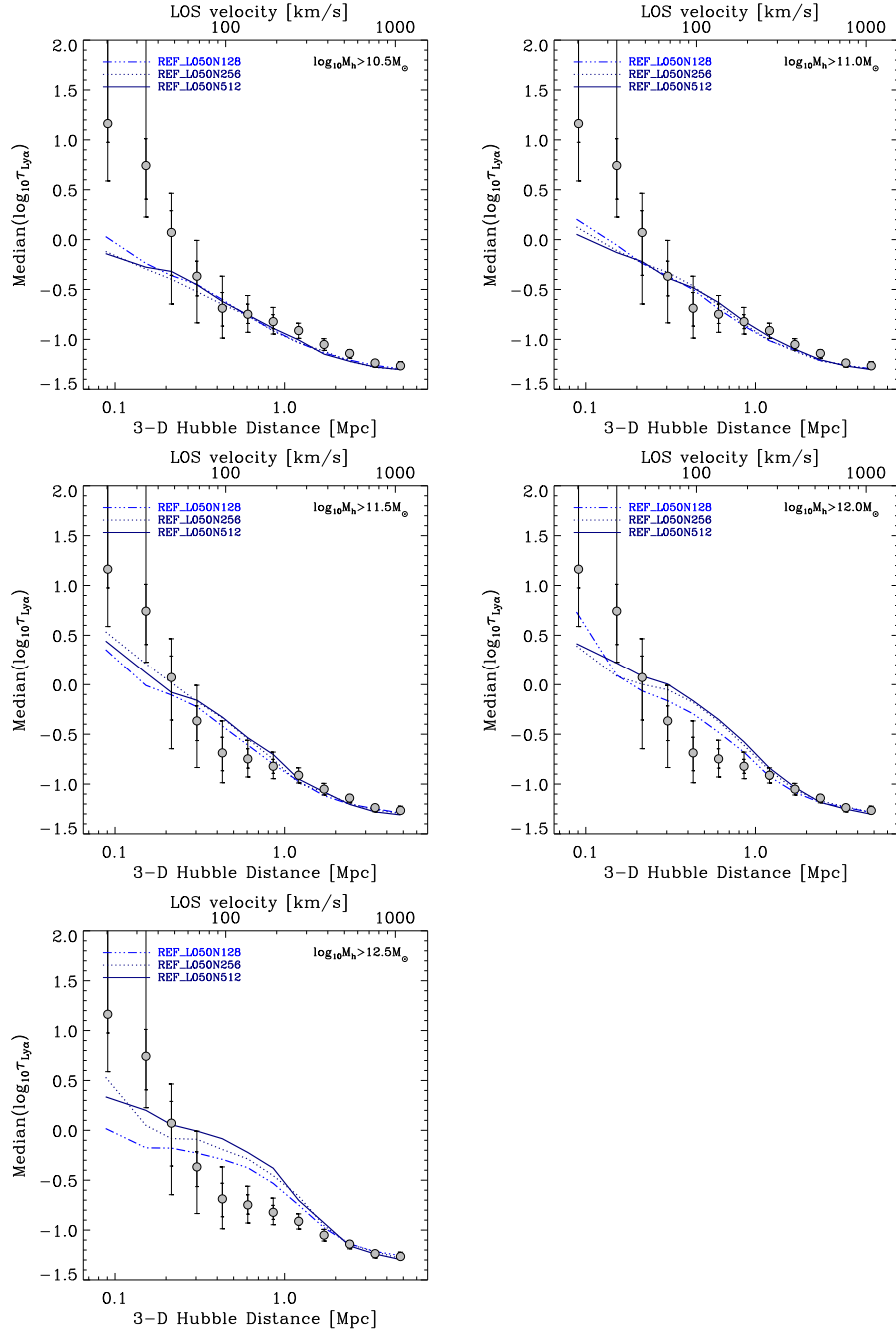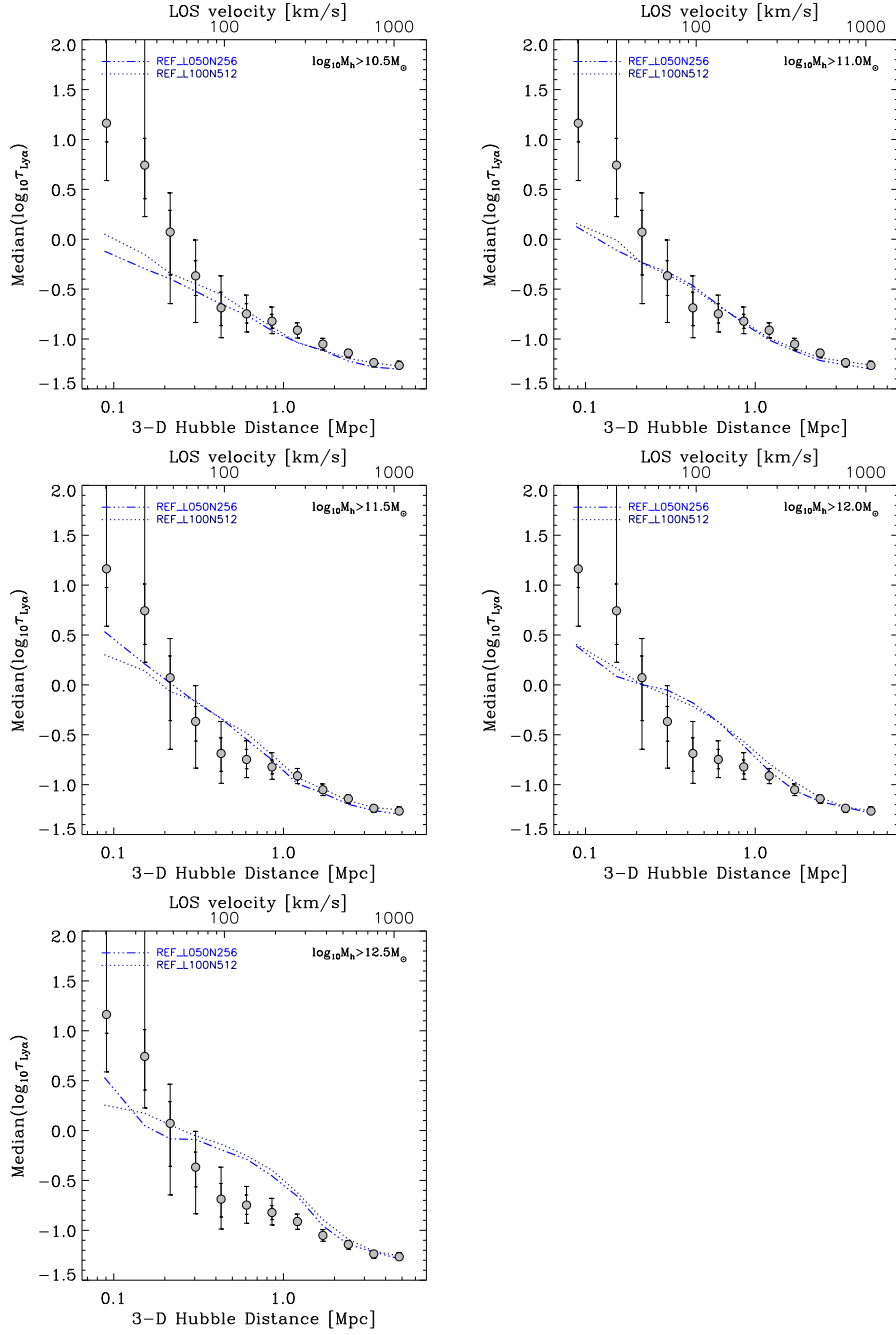
Figure 4.13 – Median $\log_{10}(\tau_{\mathrm{Ly}\alpha})$ as a function of proper distance from galaxy haloes in the 50 h^{-1} cMpc *REF* simulation with 256^3, and 100 h^{-1} cMpc *REF* simulation with 512^3 particles. Each panel shows the results for a different mass range, as indicated in the legend, while grey symbols with 1σ and 2σ error bars show the results from the observations. We added errors to galaxy positions in simulations (see text for more details).

5

How much H I Lyα Absorption near Galaxies at $z \approx 2.4$ is due to Cold Flows?

Inspired by recent studies that suggest that galaxies grow primarily by acquiring gas through cold accretion, we studied the covering fraction of Lyα absorbing gas with different properties within 200 proper kpc from haloes with masses $M_h \geq 10^{11.5}$ M$_\odot$, which are thought to host Lyman Break type galaxies at redshift $z \approx 2.4$. We use a cosmological hydrodynamical simulation with radiative transfer applied in post-processing. We consider the contributions to the absorption of different gas samples, selected on the basis of the maximum past temperature, halo membership, and past, current and future relation to the interstellar medium. We also examine the median optical depth that the different gas samples produce on 3-D distance scales $\lesssim 5$ proper Mpc. We find that gas with a maximum past temperature of $T_{max} < 10^{5.5}$ K accounts for almost all the absorption. The contribution of gas residing in the halo increases with declining distance to galaxies and with the strength of the absorber. The majority of gas, on all scales, is moving towards the haloes. While outflowing and static gas have similar and non-negligible covering fractions close to haloes, the contribution of outflowing gas declines with distance, while absorption by static gas exceeds that of inflowing gas at ~ 5 pMpc from galaxies. Most of the Lyα absorbing gas within 200 pkpc will enter the ISM of galaxies by $z = 0$, and most of it will do so for the first time. Gas that is at $\gtrsim 500$ proper kpc from haloes will not enter the ISM by $z = 0$. Gas with $T_{max} < 10^{5.5}$ K, as well as infalling gas, and gas that will join galaxies by $z = 0$ can account for the observed covering fraction of Lyα absorbing gas. Gas with $T_{max} < 10^{5.5}$ K in addition has a median optical depth consistent with that observed at distances $\gtrsim 200$ pkpc. We also examine the temperature and density of Lyα absorbing gas as a function of distance from the galaxy in simulations with and

without SN and AGN feedback, and find that galactic feedback is the main culprit in heating gas to $T \gtrsim 10^5$ K close to galaxies, although gravitational shock-heating is a non-negligible effect. Strong feedback can affect the state of the IGM out to ~ 2 proper Mpc from galaxies through heating of gas by winds emanating from galaxies in the same general region, causing an inversion of the temperature-density relation, which provides another way to constrain galactic feedback effects through measurements of the spatially resolved IGM temperature.

5.1   Introduction

Recent theoretical efforts imply that most of the fuel for star-formation comes into galaxies through "cold accretion flows", which consist of gas that did not shock-heat to the virial temperature when it accreted onto the host halo (e.g. Birnboim & Dekel 2003; Kereš et al. 2005, 2009; Ocvirk et al. 2008; Brooks et al. 2009; Crain et al. 2010; Van de Voort et al. 2011a,b; Faucher-Giguere et al. 2011). Instead, this gas falls in while maintaining a temperature of $\sim 10^4$ K, although it may shock heat to much higher temperatures if and when it accretes onto the galaxy. Simulations show that this "cold mode" gas typically falls in through filaments which can penetrate haloes of hot, hydrostatic or outflowing gas. In this chapter we examine how much of the H I Lyα absorption close to galaxies is due to this material.

Recent results of Van de Voort et al. (2011a) are based on the simulations from the same suite of OverWhelmingly Large Simulations (OWLS; Schaye et al. 2010) as the work presented in this chapter, so here we provide a short summary of their findings. They found that gas accretion is bimodal, with maximum past temperatures being either of order the virial temperature, or $\lesssim 10^5$ K. This is a consequence of the cooling function having a maximum at $T \approx 10^{5-5.5}$ K (e.g. Wiersma et al. 2009a). Van de Voort et al. (2011a) also find that the rate of accretion onto galaxies is smaller than the accretion rate onto haloes, and that this difference increases with time. It is worth noting that the ratio between the accretion rates onto haloes and galaxies is minimal for $M_{\rm halo} \sim 10^{12}$ M$_\odot$ (and at $z = 2$ this ratio equals 2), which falls in the range of typical masses of dark matter halo hosts of the LBG-type galaxies at $z \approx 2-3$ (Chapter 4). The importance of gas accreted in the hot mode increases with halo mass and towards lower redshifts. When it comes to accretion onto halos, cold accretion dominates for halos with $M_{\rm h} \ll 10^{12}$ M$_\odot$, and hot accretion for $M_{\rm h} \gg 10^{12}$ M$_\odot$, and the relative importance of the two is robust to changes in feedback and cooling prescriptions. On the other hand, when it comes to accretion onto galaxies, the cold mode is always important, and for all halo masses and at all redshifts most of the stars were on average formed from gas accreted in the cold mode. Stars formed from hot-mode gas account for only $\sim 20\%$ of the stellar mass of galaxies in halos with $M_{\rm h} \gtrsim 10^{11.5}M_\odot$ at $z \approx 2$. For accretion onto galaxies, the relative importance of the two modes is more sensitive to feedback and cooling prescriptions than for accretion onto halos.

Several publications have recently made predictions for the absorption signatures of cold flows. They in general either give predictions useful for observations of QSO-galaxy pairs, or for "down the barrel" absorption signatures, i.e. for absorption in spectra of the galaxies onto which gas is being accreted. Faucher-Giguere & Kereš (2011) use Smoothed Particle Hydrodynamics (SPH) simulations to find the covering fraction of Lyman Limit Systems (LLSs; $N_{\rm HI} \geq 10^{17.2}$ cm^{-2}) and Damped Lyman Alpha systems

(DLAs; $N_{\text{HI}} \geq 10^{20.3}\,\text{cm}^{-2}$) within 0.5, 1, and 2 R_{vir} of a galaxy halo ($M_{\text{h}} = 3.2 \times 10^{11}\,\text{M}_\odot$ at $z = 2$) at redshifts $z = 2$, 3, and 4. Their simulations do not include any galactic feedback. They do not track the gas history and can therefore not distinguish between currently cool gas accreted through the hot and cold modes. Kimm et al. (2011) and Fumagalli et al. (2011) used AMR simulations to examine, among other things, the covering fraction of cold gas for haloes of different masses and at different redshifts. They also do not take gas history into account (which would also be impossible due to the Eulerean nature of their simulations). Stewart et al. (2011a,b) used cosmological SPH simulations to study the covering fraction of cold gas (also without looking at the gas history) within 50 comoving kpc of galaxies residing in Milky Way type progenitor haloes ($M_{\text{h}}(z = 0) = 1.4 \times 10^{12}$), from redshift $z \approx 4$ to 0. While their predictions are useful for galaxy-galaxy pair observations, it will be extremely challenging to build a statistical sample of QSO-galaxy pairs with such small separations at high z. Indeed, the smallest separation in the sample used by Rakic et al. (2011) and Rudie et al. (in preparation) is ≈ 55 kpc physical. Using QSOs, as opposed to galaxies, as background sources is, however, important if one wants to study lower columns of gas than is possible with spectra of background galaxies which are generally of much lower quality than those of QSOs.

Van de Voort et al. (2011c) used the reference simulation from the OWLS suite, post-processed with radiative transfer, to study the nature of high column density ($N_{\text{HI}} \geq 10^{16}\,\text{cm}^{-2}$) Ly$\alpha$ absorbers at $z = 3$. They computed the contributions to the H I columns of gas selected in terms of its maximum past temperature, whether it is in haloes, its kinematics with respect to the haloes, and participation in star formation in the past, present and future. They found that gas with a maximum past temperature of $T_{\text{max}} \leq 10^{5.5}$ K accounts for almost all absorption in systems with $N_{\text{HI}} \geq 10^{16}\,\text{cm}^{-2}$, and that gas accreted in the "hot mode" is only important for very strong absorbers, $N_{\text{HI}} \geq 10^{21}\,\text{cm}^{-2}$. Considering this, they concluded that the cold flows have already been detected in the form of high column density systems. The likelihood of halo membership increases with the absorption strength, and $\gtrsim 80\%$ of absorbers with $N_{\text{HI}} \geq 10^{18}\,\text{cm}^{-2}$ reside within haloes. Gas moving towards the nearest (in units of R_{vir}) halo with a velocity, v_{gas}, larger than 1/4 of the circular velocity, v_{circ}, of the halo accounts for $\sim 60\%$ of the absorption, while gas outflowing with $v_{\text{gas}} \geq 0.25 v_{\text{circ}}$ accounts for $\sim 20\%$ of the absorption. The ISM dominates the absorption only for $N_{\text{HI}} \geq 10^{21}\,\text{cm}^{-2}$. The chance of having been part of the ISM before ("recycled gas") or after the moment of observation ($z = 3$) increases with absorption strength. About 70% of the gas in $N_{\text{HI}} \geq 10^{17} - 10^{21}\,\text{cm}^{-2}$ systems enters a galaxy by $z = 2$, $\lesssim 40\%$ has been recycled, and $\lesssim 35\%$ will enter galaxies after $z = 3$ for the first time. They also found that haloes with total mass $< 10^{10}\,\text{M}_\odot$ dominate the absorption for $10^{17} < N_{\text{HI}} < 10^{21}\,\text{cm}^{-2}$.

Although the idea that cold flows dominate the fueling of galaxies is

strongly supported by cosmological hydrodynamical simulations, it has, as of yet, no conclusive backing in the observations. In fact, spectra of star-forming galaxies show only clear signs of outflows, at both high (e.g. Pettini et al. 2001; Adelberger et al. 2003, 2005; Shapiro et al. 2009; Weiner et al. 2009; Steidel et al. 2010) and low redshift (e.g. Martin & Bouche 2009; Chen et al. 2010). At the same time, there is circumstantial evidence that gas replenishment is needed to provide the necessary fuel for the measured build-up of stellar mass (e.g. Papovich et al. 2010) and to explain the relationship between stellar mass and metallicity (Erb et al. 2006; Erb 2008), and the HI gas density (non)evolution from $z \approx 2$ to $z = 0$ (Prochaska et al. 2009).

Ribaudo et al. (2011) discovered a metal-poor Lyman limit system (LLS) at $z = 0.27395$, separated by only 37 kpc from a star-forming galaxy at an almost identical redshift, and they suggest that its low metallicity ([Mg/H] $= -1.71 \pm 0.06$) makes it consistent with being cold accretion. However, association with an undetected dwarf galaxy, although unlikely, cannot be completely ruled out, and a low metallicity does not guarantee that the gas is accreting. Giavalisco et al. (2011) have also recently reported the discovery of a large amount of cold, metal-poor gas in an overdensity of galaxies at $z \approx 1.6$. They suggest, based on the analysis of the co-added absorption spectra of background galaxies, that this gas is potentially undergoing infall onto the background overdensity and its galaxies, which they say is a tentative detection of cold accretion. Dijkstra & Loeb (2009) and Goerdt et al. (2010) claimed that Lyα blobs at $z \sim 3$ (e.g. Steidel et al. 2000, 2011) may be cooling radiation from cold flows. However, simulations by Furlanetto et al. (2005) and Faucher-Giguere et al. (2010) suggest that cooling radiation is insufficient to explain Lyα blob luminosities, which is also what Steidel et al. (2011) concluded based on stacked narrow-band observations of star-forming galaxies at $z \approx 2.65$. Instead, these authors suggested that the blobs are fully consistent with being powered by the light of central objects that is scattered by surrounding gas. Finally, Hayes et al. (2011) measured a high polarization for the Lyα radiation from the blobs which supports the idea that they are powered by a central source.

Motivated by the recent observations of Rakic et al. (2011) who studied the distribution of Lyα absorbing gas around $\langle z \rangle \approx 2.36$ star-forming galaxies, we use SPH simulations to determine how much of the absorption that they measured could be due to cold accretion gas. They found the Lyα absorption to be enhanced out to at least 2.8 Mpc proper from galaxies (3σ confidence). They also found evidence of gas infall into the overdensities occupied by galaxies. The observed infall manifests itself as a compression of the absorption signal along the LOS, i.e. the absorption is enhanced out to at least 2 Mpc in the transverse direction from galaxies, but only out to ≈ 1.5 Mpc along the LOS. These scales are large in comparison with the virial radius of the galaxies ($R_{\rm vir} \approx 100$ kpc), and therefore the observed infall could not be considered "cold accretion", in the sense that separation

into cold and hot mode is possible only on scales $\lesssim R_{\mathrm{vir}}$. Nevertheless, it is the only observation that unequivocally proves that gas is moving towards galaxies. We analyze the OWLS simulations, which are particularly well suited for this type of study because their Lagrangian nature allows one to trace back the history of the gas. The maximum past temperature, $T_{\max}$, and the redshift when it was reached, $z_{\max}$, are saved for each particle. This allows us to distinguish between particles that did and those that did not go through an accretion shock near the virial radius. In comparison with most earlier works, we improve by differentiating between absorption that comes from gas with particular past properties, kinematics with respect to nearby haloes, membership of a halo, and participation in ISM of galaxies. In comparison with Van de Voort et al. (2011c), we complement their work by considering the spatial distribution of gas with respect to galaxies, while they studied global properties of absorbers, and we also study weaker absorbers ($N_{\mathrm{HI}} \ll 10^{16}\,\mathrm{cm}^{-2}$) along with strong ones ($N_{\mathrm{HI}} \geq 10^{16}\,\mathrm{cm}^{-2}$), while they focused only on stronger systems. In addition, we focus on haloes with $M_{\mathrm{h}} > 10^{11.5}\,\mathrm{M}_{\odot}$ whose virial temperatures are $> 10^{5.5}$ K, and are therefore suitable places to look for cold accretion streams penetrating hot haloes.

In the first part of this chapter we will consider the contribution to the absorption of gas with different kinematics, past temperature, halo membership, and participation in the ISM at any time in the history of the Universe. In the second part of the chapter we will study the physical properties of the Lyα absorbing gas, i.e. the temperature and density, as a function of the distance from the haloes. We will also examine how sensitive these properties are to changes in feedback prescriptions in the simulations.

We describe the simulations in Section 5.2. We show how much of the absorption on scales $\sim R_{\mathrm{vir}}$ is produced by gas with various properties in Section 5.3, and we do the same for absorption on larger scales ($\gg R_{\mathrm{vir}}$) in Section 5.4. In Section 5.5 we look into the temperature and density of Lyα absorbing gas in different models. We conclude in Section 5.6.

We denote proper distances as pMpc and pkpc, and comoving distances as cMpc and ckpc.

5.2   Simulations

We rely mostly on the "reference" (*REF*) simulation from the suite of Over-Whelmingly Large Simulations (OWLS; Schaye et al. 2010). The simulation is run in a 50 h^{-1} cMpc box, containing 2×512^3 dark matter and baryonic particles. The simulations were already briefly described in Chapter 4. This is the only model we use in the first part of this chapter where we study absorption by cold flows. In the second part of the chapter, which investigates the physical properties of the Lyα absorbing gas, we also make use of the *AGN* model, which was also described in the previous chapter, to study the effect of very strong feedback. We will use the *NOSN* model, which

is identical to the *REF* model except that it does not contain supernova (SN) feedback, to determine how much gas heating comes from gravitational shock-heating, and how much is due to galactic winds.

In the first part of the chapter we use the *REF* model with radiative transfer of ionizing radiation applied in post-processing, while no such treatment was applied to the simulations used in the second part of the chapter.

5.2.1   Radiative transfer in post-processing

The gas in the OWLS simulations is exposed to the cosmic microwave background and to the ionizing background radiation from galaxies and quasars according to the model from Haardt & Madau (2001), but this is done in the optically thin limit where each gas particle, irrespective of its density, is subject to the same radiation intensity. This is an excellent approximation for gas with column densities $N_{\mathrm{HI}} \lesssim 10^{17}\,\mathrm{cm}^{-2}$, but gas at higher densities can self-shield against the ionizing radiation, in which case it reaches higher neutral fractions than in the optically thin limit. Given that we are interested in gas, accreting onto dark matter halos and their galaxies and that this gas might reach relatively high densities, we post-process the $z = 2.25$ snapshot from the simulation with radiative transfer, as described in Altay et al. (2011, and in preparation). Note that local sources of radiation were ignored.

Altay et al. (2011) used the *REF* model from the OWLS suite of simulations, which is identical to the one used here, except that their simulation was run with WMAP7 cosmological parameters, as opposed to WMAP3 in our case, and that they used a 25 h^{-1} cMpc rather than a 50 h^{-1} cMpc box. After applying the radiative transfer calculation with self-shielding, they found that the *REF* model reproduces the observed $z = 3$ properties of Lyα absorbers, i.e. the abundance of Lyα forest lines, LLSs, and DLAs, spanning over ten orders of magnitude in H I column density.

5.2.2   Thermal history of gas particles

The maximum temperature ever reached is recorded for every time step for each particle in the OWLS simulations. Given that the simulations have insufficient resolution to properly treat the interstellar medium (ISM), any gas particle with density $n_{\mathrm{H}} \geq 0.1\,\mathrm{cm}^{-3}$ is assumed to be part of the ISM, multi-phase and star-forming, and its pressure is imposed in the form of a polytropic equation of state $P \propto n_{\mathrm{H}}^{4/3}$. The temperature of such gas just reflects the imposed equation of state, and therefore the maximum recorded temperature does not get updated as long as the particle is part of the ISM.

Since we can follow the thermal history of the gas, we can study, for example, how much absorption at $z = 2.25$ is produced by gas that has never reached temperatures higher than some threshold value, that is inside haloes

and infalling onto galaxies at a given velocity, or is found to have been a part of the ISM before, at, or after $z = 2.25$. Table 5.1 summarizes all the selection criteria that we consider in this Chapter.

We chose the maximum past temperature of $T = 10^{5.5}$ K as the threshold temperature separating cold and hot accretion, because this is where the cooling function has its maximum (e.g. Wiersma et al. 2009a), and because Van de Voort et al. (2011a) found that this value best separates the cold and hot accretion modes. The virial temperature of haloes, $T_{\rm vir}$, is defined as follows:

$$T_{\rm vir} = \left(\frac{G^2 H_0^2 \Omega_{\rm m} 18\pi^2}{54} \right)^{1/3} \frac{\mu m_{\rm H}}{k_{\rm B}} M_{\rm halo}^{2/3} (1 + z)$$

$$\approx 10^6 \, {\rm K} \left(\frac{\mu}{0.59} \right) \left(\frac{M_{\rm halo}}{10^{12} {\rm M}_\odot} \right)^{2/3} \left(\frac{1 + z}{3.4} \right) \tag{5.1}$$

where G is the gravitational constant, H_0 the Hubble constant, μ the mean molecular weight, $m_{\rm H}$ the mass of a hydrogen atom, and $k_{\rm B}$ Boltzmann's constant. While the stated definition of hot and cold accretion is not appropriate for haloes with virial temperatures lower than $10^{5.5}$ K, i.e. $M_{\rm h} \lesssim 10^{11.2} \, {\rm M}_\odot$, since in this regime it becomes impossible to determine whether gas has gone through the shock or not (i.e. its maximum past temperature is similar to the virial temperature of such haloes), it is fine for our mass range of interests, $M_{\rm h} \gtrsim 10^{11.5} \, {\rm M}_\odot$, with virial temperatures of $\gtrsim 10^{5.7}$ K at $z = 2.4$.

5.2.3   Observing simulations

Similar to the procedure described in Chapter 4, we compute H I Lyα absorption spectra for 12,500 sightlines randomly drawn within 5 pMpc from haloes with masses $M_{\rm h} > 10^{11.5} \, {\rm M}_\odot$, and within 200 pkpc we draw an additional 6000 sightlines in order to sample the immediate halo surroundings sufficiently well. The distribution of impact parameters is uniform as a function of radius. This mass threshold was chosen because in the same chapter we found haloes with masses $M_{\rm h} > 10^{11.5} \, {\rm M}_\odot$ to be the most likely hosts of the galaxy population used in Rakic et al. (2011). This procedure is repeated for each gas sample listed in Table 5.1, where the density of particles not satisfying the stated conditions is set to zero. Due to the uncertainty in the intensity of the ionizing background radiation, we rescale the optical depth in the default case, i.e. when including all gas, to match the median optical depth in the observations of Rakic et al. (2011), $\log_{10} \tau_{\rm Ly\alpha} = -1.27$. For the rest of the variations listed in Table 5.1 we use the multiplication factor as determined for the default case. We use a subset of selections from Van de Voort et al. (2011c), where more details about the selection of particles can be found.

We use the Friends-of-Friends (FoF) algorithm to identify haloes. Two particles belong to the same group if their separation is less than 20% of the

Table 5.1 – List of applied cuts. The first column states the conditions that gas particles need to satisfy with respect to their maximum past temperature, whether they are inside halos with $M_h > 10^{11.5}$ M$_\odot$, whether they are infalling/outflowing with respect to the nearest such halo, and whether they are part of the ISM at, after, or before $z = 2.25$. In the second column we give global fraction of the gas mass satisfying the stated conditions.

Variations	Global gas fraction
All particles - default case	1.000
$T_{max} \leq 10^{5.5}$ K	0.882
In haloes ($M_h \geq 10^{11.5}$ M$_\odot$)	0.024
In haloes ($M_h \geq 10^{11.5}$ M$_\odot$), infall with $v > 0.25 v_{circ}$	0.012
In haloes ($M_h \geq 10^{11.5}$ M$_\odot$), outflow with $v > 0.25 v_{circ}$	0.004
In haloes ($M_h \geq 10^{11.5}$ M$_\odot$), static, i.e. $v < 0.25 v_{circ}$	0.008
Infall toward the nearest (in units of R_{vir}) halo ($M_h \geq 10^{11.5}$ M$_\odot$), with $v > 0.25 v_{circ}$	0.480
Outflow from the nearest (in units of R_{vir}) halo ($M_h \geq 10^{11.5}$ M$_\odot$), with $v > 0.25 v_{circ}$	0.056
Static w.r.t. the nearest (in units of R_{vir}) halo ($M_h \geq 10^{11.5}$ M$_\odot$), i.e. $v < 0.25 v_{circ}$	0.463
Gas particles that are in the ISM at $z = 2.25$	0.015
Gas particles that became ISM after $z = 2.25$	0.104
Gas particles that became ISM after $z = 2.25$ for the first time	0.069
Gas particles that were in the ISM before, but are not at $z = 2.25$	0.035

average inter-particle separation. Such groups have an average overdensity typical for virialized objects, i.e. $\langle \rho_{halo} \rangle / \langle \rho \rangle \approx 180$ (e.g. Lacey & Cole 1994). A baryonic particle belongs to the same halo as its nearest dark matter particle.

The positions and masses of haloes are determined using a spherical overdensity criterion as implemented in the SUBFIND algorithm (Dolag et al. 2009). The position of a halo is taken to be the position of its most bound particle. The virial radius, R_{vir}, is estimated by requiring that the enclosed density agrees with the top-hat spherical collapse approximation (Bryan & Norman 1998). The virial mass, M_{vir}, is the total mass within R_{vir}.

For some of the gas samples listed in Table 5.1 we consider whether particles are moving with respect to the nearest halo with $M_h > 10^{11.5}$ M$_\odot$, in units of R_{vir}. In other words, distances towards a halo, d, are normalized by R_{vir}, and the nearest halo is one for which d/R_{vir} has minimum. The halo circular velocity is defined as $v_{circ} = \sqrt{GM_{vir}/R_{vir}}$.

To allow comparison with observations of Rakic et al. (2011) and Rudie et al. (2011, in preparation) whose galaxies have measured redshift errors of ± 125 km s^{-1}, we add random errors from a Gaussian distribution, $\sigma = 125$ km s^{-1}, to the line of sight positions of haloes in the simulations (i.e. along only one dimension that is parallel to the line of sight).

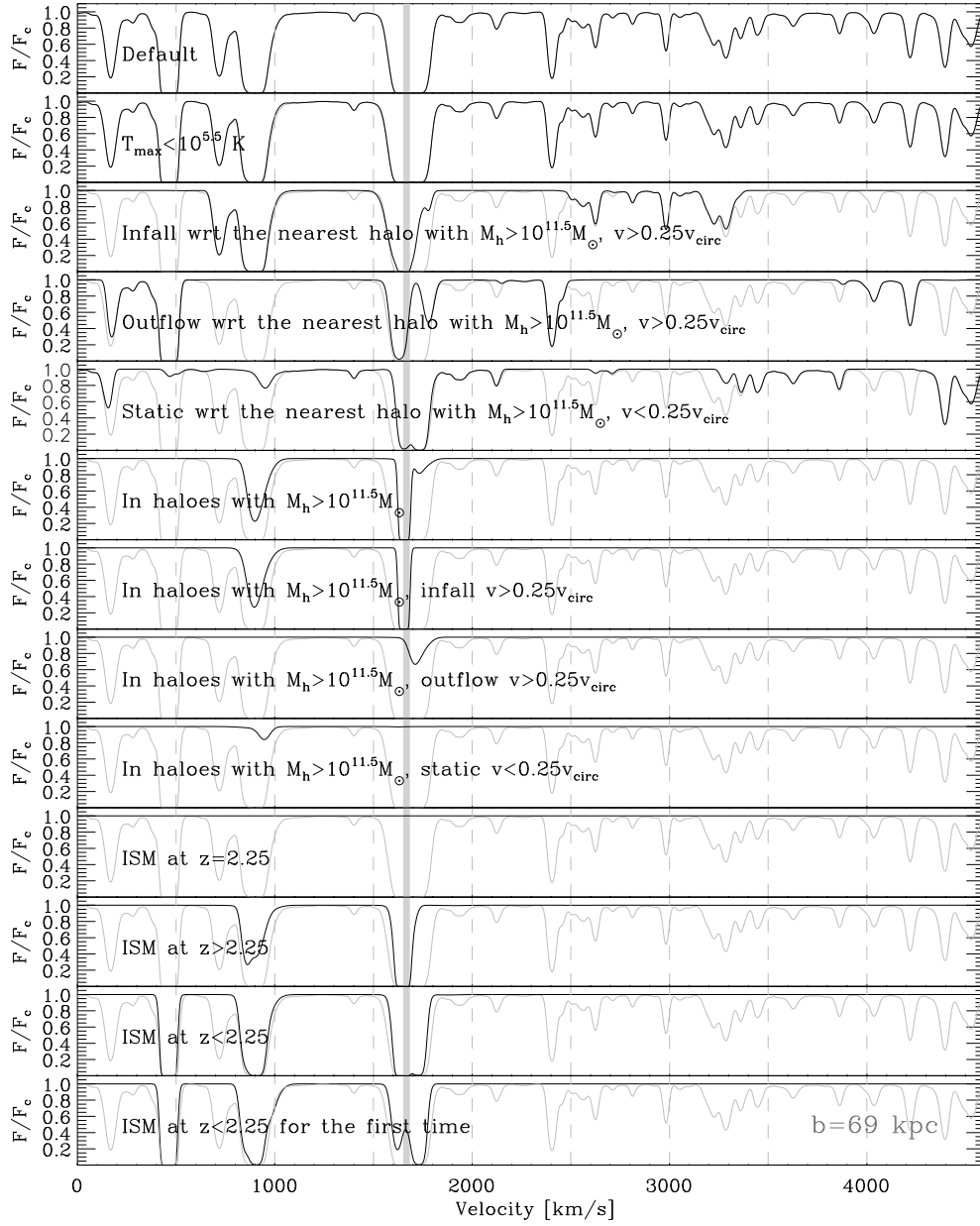Figure 5.1 shows example spectra along a sightline that passes 69 pkpc

Figure 5.1 − Spectra along a LOS that passes 69 kpc from a nearby halo with total mass $M_\mathrm{h} = 10^{11.74}$ M$_\odot$ at $z = 2.25$, whose position is indicated by the grey vertical stripe. Different panels show the spectra for different gas samples. The fourth panel from the bottom shows no absorption, because the ISM gas is concentrated in the centers of haloes, and to have such gas at $b = 69$ pkpc it is necessary to intersect a neighboring galaxy, which is not the case for this sightline. The default case (all gas, top panel) is repeated in each panel in grey.

from a nearby halo of mass $M_\mathrm{h} = 10^{11.74}$ M$_\odot$, for the gas samples listed in Table 5.1. By comparing absorption in different gas samples with the

default case (i.e. where all particles are included), it can be seen that the majority of Lyα forest lines originate in gas that was never shock-heated to temperatures $T > 10^{5.5}$ K, and is not in haloes. In addition, very few lines originate in gas that becomes part of the ISM. In the next sections we will quantify the contributions from these different IGM components to the H I Lyα absorption near galaxies.

5.3   Circum-galactic medium

As already discussed in Chapter 3, when it comes to stronger absorbers, which we expect to be close to galaxies, it is more appropriate to measure their strength in terms of their central optical depth or column density, instead of the median optical depth. Using the median optical depth within a given velocity interval might underestimate the amount of neutral gas, and is very dependent on the size of the region over which the median is estimated. We look both into the covering fraction of sightlines as a function of the maximum optical depth within a given distance from galaxies, and into the covering fraction as a function of the H I column density. We use the former for absorbers that are optically thin to the ionizing radiation ($N_{\mathrm{HI}} < 10^{17.2}\,\mathrm{cm}^{-2}$, or $\log_{10}\tau_{\mathrm{Ly}\alpha} \approx 3.66$ for a b parameter of $26\,\mathrm{km\,s}^{-1}$) and the latter for optically thick absorbers ($N_{\mathrm{HI}} > 10^{17.2}\,\mathrm{cm}^{-2}$).

Figure 5.2 shows the covering fractions of gas with different Lyα absorption strength, within a given distance bin from the centers of haloes with $M_{\mathrm{h}} > 10^{11.5}\,\mathrm{M}_\odot$. The maximum optical depth is measured within $165\,\mathrm{km\,s}^{-1}$ from galaxy positions. Rakic et al. (2011) found that the signal is independent of velocity within $165\,\mathrm{km\,s}^{-1}$, and that the errors are strongly correlated for smaller velocity differences, so they chose this velocity interval for many of their calculations. We use the same interval here to facilitate the comparison with their observations. We note that this velocity interval is larger than the estimated galaxy redshift errors in the observations ($\sigma = 125\,\mathrm{km\,s}^{-1}$). On the other hand, the column density results are based on the total column density along the LOS through the $50\,h^{-1}$cMpc box. We use such projections only for optically thick absorbers, and since they are relatively rare, they will likely reside in the vicinity of galaxies (the maximum impact parameter here is 200 pkpc). Also, it is unlikely that a collection of weak absorbers along the LOS could cumulatively account for such large column densities (see Altay et al. 2011). We use a total column density along the LOS for practical reasons, as finding column densities for individual absorbers would require decomposing the spectra into Voigt profiles, which is impractical for the number of sightlines considered here. We verified that this is a good approximation for e.g. LLSs, where we found the expected maximum optical depth of an absorber with LLS column densities (assuming a b parameter of $26\,\mathrm{km\,s}^{-1}$), and performed the analysis equivalent to the one we apply to optically thin absorbers, and found almost perfect agreement with results based on the to-

tal column density along the LOS. Naturally, redshift errors have no effect on such estimated column density results.
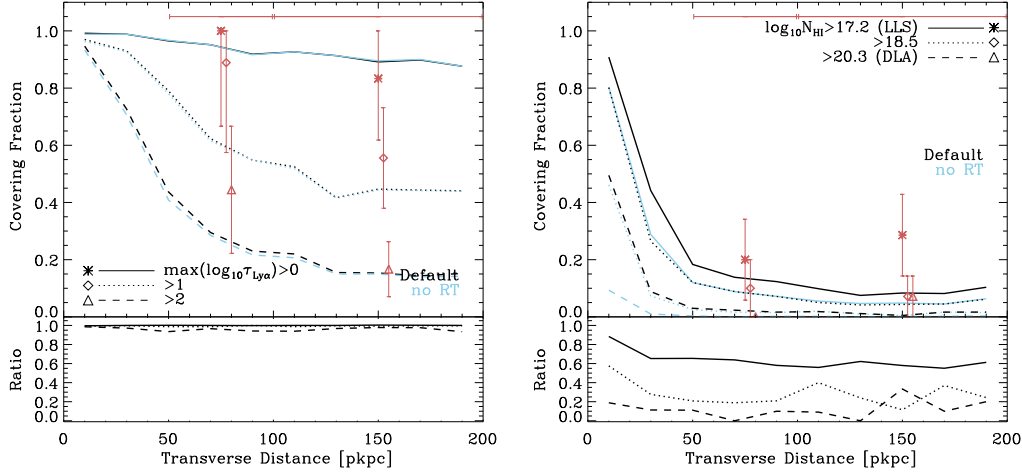

Figure 5.2 – *Top panels*: covering fractions of absorbers with maximum H I Lyα optical depth $\log_{10}\tau_{\text{Ly}\alpha} > 0$, 1, 2 (left panel), and covering fractions of absorbers with $\log_{10}N_{\text{HI}} \geq 17.2$ (LLSs), 18.5, and 20.3 cm^{-2} (DLAs) (right panel), as a function of impact parameter from the center of haloes with $M_{\text{h}} > 10^{11.5}$ M$_\odot$ at z=2.25. Black lines show the results for the default case, i.e. absorption from all gas, in the simulation with radiative transfer applied in post-processing. The blue lines show the results for the simulation without radiative transfer, which fall on top of the black curves in the left panel (radiative transfer does not make a difference for gas optically thin to ionizing radiation). The red symbols show the covering fractions observed by Rakic et al. (2011) in the left panel, and by Rudie et al. (in prep.) in the right panel. The covering fractions are measured in 20 pkpc bins for the simulations, and in 100 pkpc bins for the observations (top horizontal bars show the extent of the bins for the observations, where the first bin effectively starts only at 55 pkpc, which is equal to the smallest impact parameter). *Bottom panels:* the ratios between the blue and black curves, i.e. the effect of the radiative transfer.

The left panel of Figure 5.2 shows the covering fraction of sightlines with gas with maximum H I Lyα optical depth higher than 1, 10, or a 100, within $\pm 165\,\text{km s}^{-1}$ of a halo of mass $M_{\text{h}} > 10^{11.5}$ M$_\odot$ at $z = 2.25$, while the right panel shows the same but for gas with the H I column densities greater than $10^{17.2}$, $10^{18.5}$, and $10^{20.3}$ cm^{-2}. The blue lines show the results for the simulation that was not treated with radiative transfer, and the black lines for the case where it was applied in post-processing. As expected, the radiative transfer does not affect the results significantly for absorbers optically thin to the ionizing radiation (left panel). Although this panel includes all absorbers with optical depth higher than a given value from the optically thin regime, i.e. also absorbers that are optically thick, strong absorbers are relatively rare and the covering fraction is dominated by weaker absorbers. In the discussion below we will refer to absorbers from the left panel as 'optically thin', even though they include optically thick absorbers as well. However, the situation is dramatically different for optically thick absorbers, as can be seen in the

right panel. Without radiative transfer the covering fraction of DLAs is ~ 5 times smaller. This demonstrates the importance of this post-processing step for studying high column density absorbers close to galaxies.

The covering fraction of absorbers with maximum $\tau_{\text{Ly}\alpha} \geq 1$, 10, and 100, changes from unity within 20 pkpc to ≈ 0.90, ≈ 0.45, and ≈ 0.15, respectively, at an impact parameter of 200 pkpc. The covering fraction of optically thick absorbers changes from ~ 0.90 (~ 0.45) for LLSs (DLAs) within 20 pkpc, to ~ 0.15 (~ 0.02) at 200 pkpc. Below we will refer to the covering fraction for the default case (i.e. all gas) as the "total covering fraction".

In the same figure we also show covering fractions of absorbers measured from observations. In the left panel (for the optically thin lower limit) we show the covering fractions measured by Rakic et al. (2011), and in the right panel (for the optically thick regime) we use the results from Rudie et al. (in preparation). These covering fractions were estimated in two bins of 100 pkpc, as compared to 20 pkpc for the simulated data. The observed covering fractions should therefore be compared to the simulated ones by averaging simulated covering fractions over appropriate distances. In addition, the smallest impact parameter in the observations is ≈ 55 pkpc, which should also be taken into account for the comparison. It appears that the simulations match all observed covering fractions, except that for $N_{\text{HI}} > 10^{17.2}\,\text{cm}^{-2}$ from 100 to 200 pkpc, although the errors on the observations are large. Given that Rudie et al. (2011) observe only 1 absorber with $N_{\text{HI}} > 10^{18.5}\text{cm}^{-2}$ in each 100 pkpc bin, we will not use (dis)agreement with observations in the optically thick regime to determine the importance of cold-mode gas for the total covering fraction of Lyα absorbers.
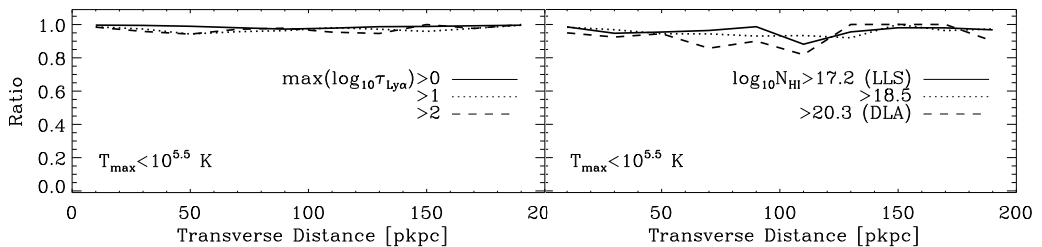

Figure 5.3 – The ratios of covering fractions of absorbers with a given minimum absorption strength including only gas with the maximum past temperature of $\lesssim 10^{5.5}$ K and total covering fractions (i.e. absorption from all gas), as a function of impact parameter. *Left panel* shows the ratios for absorbers with maximum H I Lyα optical depth $\log_{10}\tau_{\text{Ly}\alpha} \geq 0$, 1, and 2. *Right panel:* similar to the left panel, but for optically thick absorbers with $\log_{10}N_{\text{HI}} \geq 17.2$ (LLSs), 18.5, and 20.3 cm^{-2} (DLAs).

Figure 5.3 shows the fractional contribution that gas with a maximum past temperature $\lesssim 10^{5.5}$ K makes to the total covering fraction within 200 pkpc from galaxy haloes, for both optically thin and optically thick absorbers. The covering fraction of such gas is equal to the value on the y-axis multiplied by the covering fraction in Figure 5.2 (black lines). We note that the results
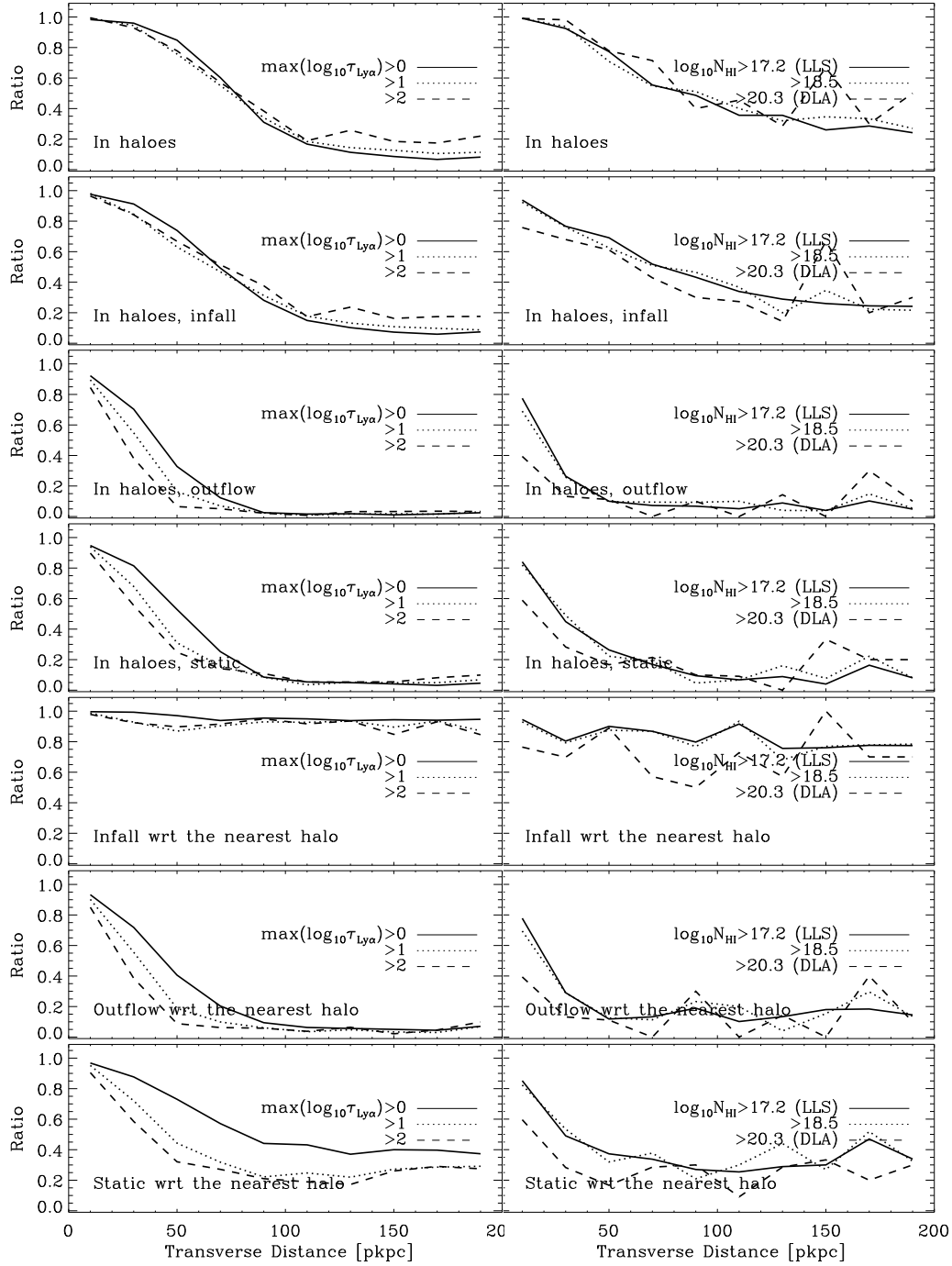
Figure 5.4 – Similar to Figure 5.3, but for more cuts referring to the gas halo membership (in haloes with $M_{\mathrm{h}} > 10^{11.5}$ M$_\odot$) and kinematics with respect to the nearest halo with $M_{\mathrm{h}} > 10^{11.5}$ M$_\odot$ (in units of R_{vir}) at velocities $v > 0.25 v_{\mathrm{circ}}$ for inflowing and outflowing gas, and $v < 0.25 v_{\mathrm{circ}}$ for static gas.

are noisy for DLAs, due to small number statistics. The covering fraction of DLAs at distances $\gtrsim 50$ pkpc is only a few per cent (Figure 5.2) which means

that an estimate in any given 20 pkpc distance bin (with ≈ 200 sightlines per bin) is based on only a few DLAs.

Gas with a maximum past temperature $\lesssim 10^{5.5}$ K has a covering fraction that is almost identical to the total covering fraction. Given that these covering fractions are consistent with the observations, it appears that cold-mode gas can account for Lyα absorption near galaxies from Rakic et al. (2011) and Rudie et al. (2011, in preparation).

We next determine how much absorption near haloes with $M_{\rm h} > 10^{11.5} {\rm M}_\odot$ comes from gas within those haloes. We select all the particles belonging to haloes with $M_{\rm min} = 10^{11.5}$ M$_\odot$. It is likely that some of the haloes in this mass range overlap, and that by studying the distribution near one such halo, we also intersect absorption by gas belonging to a different halo. We expect the contribution from neighboring halos to be small, and leave the precise treatment of this issue for future work.

Figure 5.4 shows that for impact parameters $\lesssim 30$ pkpc, gas in haloes produces a covering fraction that is almost identical to the total for absorbers of all strengths. At larger distances, the contribution of haloes asymptotes to $\sim 10\%$ of the total for gas with $\tau_{\rm Ly\alpha} \geq 1$, and to $\sim 25\%$ of the total for optically thick absorbers. This suggests that optically thick gas is more likely to reside in the halo. Van de Voort et al. (2011c) also found that the likelihood of halo membership increases with the strength of Lyα absorption. We also see that the probability of optically thick systems belonging to the halo in question declines with distance. We investigated (not shown here) the origin of the remainder of optically thick absorbers that contribute to the total covering fraction and found that they reside within smaller haloes. By multiplying the curves in the top left panel of Figure 5.4 with the corresponding covering fractions in Figure 5.2 (black lines), we conclude that gas in haloes with $M_{\rm h} > 10^{11.5}$ M$_\odot$ cannot account for the observed covering fractions. We cannot draw any conclusions about the optically thick gas due to small number statistics.

Gas within the haloes that is moving towards their centers with a radial velocity greater than $1/4$ of the halo circular velocity, $v_{\rm circ} = \sqrt{GM_{\rm vir}/R_{\rm vir}}$, has a covering fraction that is similar to that of all gas that is in haloes (second set of panels from the top). We take a threshold value of $1/4$ of the circular velocity because the contribution of gas that is moving at velocities higher than the circular velocity is negligible. Gas that is outflowing with $v > 0.25 v_{\rm circ}$ and gas that is "static" with respect to the halo center (i.e. $v < 0.25 v_{\rm circ}$), generally has a lower covering fraction than the inflowing gas. The covering fractions of outflowing and static gas are similar, and for optically thin (thick) gas they change from $\sim 90\%$ ($\sim 80\%$) of the total covering fraction within 20 pkpc, to $\sim 5\%$ ($\sim 10\%$) at 200 pkpc.

If we relax the condition about halo membership, and instead only consider whether the gas is falling towards the nearest halo with $M_{\rm h} > 10^{11.5} {\rm M}_\odot$

with a velocity greater than a quarter of the circular velocity of that halo, we find that the covering fraction of such gas is $\gtrsim 85\%$ of the total within 200 pkpc for optically thin gas. This shows that most of the Lyα absorbing gas is accreting onto the haloes, even though it is not yet inside them. The covering fraction of this gas is consistent with the observations, which suggests that (if the *REF* model describes reality sufficiently well) the infalling gas is necessary to account for the observed covering fraction within 200 pkpc. The covering fraction of optically thick absorbers is $\lesssim 20\%$ lower than total at all distances, suggesting that most strong absorbers (and their small haloes) are moving towards larger haloes. Gas that is moving away from the nearest massive halo with $v > 0.25v_{\mathrm{circ}}$, as well as static gas (i.e. $v < 0.25v_{\mathrm{circ}}$), has a covering fraction that is $\lesssim 20\%$ and $\lesssim 40\%$, respectively, of the total at distances $\sim 50 - 200$ pkpc, and therefore also comprises a significant amount of gas. Their covering fractions alone are insufficient to account for the observations.
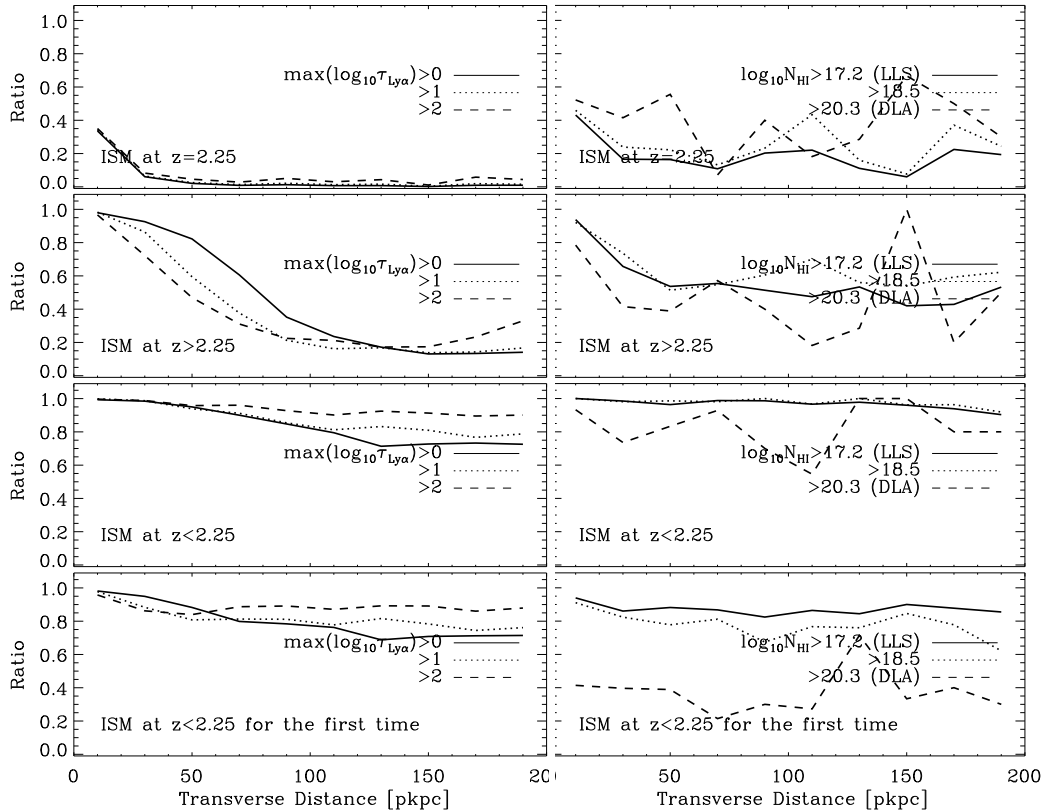

Figure 5.5 − Similar to Figure 5.3, but for more cuts referring to the gas participation in the ISM of galaxies, before, at, and after $z = 2.25$.

Figure 5.5 refers to the gas past, present, and future participation in the ISM of galaxies, which we define as any gas with density $n_{\mathrm{H}} \geq 0.1\,\mathrm{cm}^{-3}$ (see §5.2.2). The contribution of the ISM to the covering fraction of Lyα absorbing gas increases closer to the centers of haloes and with increasing absorption

strength. Within 200 pkpc, $\sim 20-40\%$ of optically thick absorbers are part of the ISM. At 200 pkpc, however, we do not expect them to be the ISM of the central galaxy, but of companion galaxies.

Gas that was part of the ISM at $z > 2.25$, but is not in the ISM at $z = 2.25$, has a covering fraction close to the total within 20 pkpc, and then it asymptotes to $\sim 20\%$ ($\sim 60\%$) of the total covering fraction at 200 pkpc for optically thin (thick) gas. Gas that will become part of the ISM at $z < 2.25$, but is not ISM at $z = 2.25$, has a covering fraction $\gtrsim 80\%$ ($\gtrsim 90\%$) of the total for optically thin (thick) gas within 200 pkpc, and most of that gas will accrete onto a galaxy for the first time. The covering fraction of this gas is consistent with the observations for maximum $\log_{10}\tau_{\text{Ly}\alpha} > 0$ and > 2, but is lower than observed for maximum $\log_{10}\tau_{\text{Ly}\alpha} > 1$ (1σ discrepancy). This suggests that gas that will become the ISM of galaxies by $z = 0$ can be observed in absorption at $z \approx 2.4$. Gas that accretes onto a galaxy after $z = 2.25$ but not for the first time, is the "recycled" gas (Oppenheimer et al. 2010). This is gas that was in the ISM before $z = 2.25$, and then for example got stripped or ejected from the ISM through outflows, but accretes again onto galaxies at $z < 2.25$. In addition, stronger absorbers are more likely to become the ISM of galaxies than weaker absorbers at the same separation from haloes, as can be seen by comparing the left and right panels.

We note that in order to say with certainty whether any of the gas samples is crucial for the observed absorption, it would be necessary to consider absorption by all but gas in a given sample (e.g. all but gas with $T_{\max} < 10^{5.5}$ K). If it turns out that the predicted absorption in such gas is insufficient to account for the observations, we can conclude that that gas sample (e.g. cold-mode gas) plays a major role and had been observed already. Also, it is necessary to determine how sensitive the results are to different feedback prescriptions. We leave treatment of these issues for our future publication.

5.4 Absorption by cold flows on pMpc scales

In the previous Section we considered the distribution of H I around galaxies on scales that are comparable to the sizes of the dark matter haloes hosting the galaxies used in Rakic et al. (2011), i.e. < 200 pkpc, $\lesssim 2R_{\text{vir}}$. In this section we will study H I Lyα absorption as a function of distance from the galaxies on scales up to 5 pMpc ($\gg R_{\text{vir}}$) by analyzing the median optical depth as a function of 3-D Hubble distance from haloes. This distance is equal to $\sqrt{d_{\text{LOS}}^2 + b^2}$, where d_{LOS} is the LOS separation between the halo and the absorber estimated from their velocity separation assuming that it is due to Hubble expansion, and b is the impact parameter. We will compare the simulated absorption distributions with the observations of Rakic et al. (2011), and therefore we set the minimum impact parameter in the simulations to match their observations, i.e. to 55 pkpc.

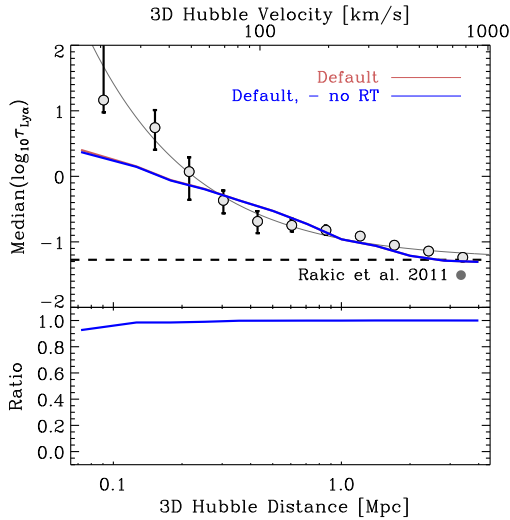Figure 5.6 shows the effect of the improved handling of radiative transfer

Figure 5.6 − *The Top panel* shows the median $\log_{10}(\tau_{\mathrm{Ly\alpha}})$ as a function of physical 3-D Hubble distance from galaxy halos. The circles show the observed points with 1σ error bars from Rakic et al. (2011), the red curve shows the results obtained from simulations where self-shielding was applied in post-processing, and the blue curve shows the case where we apply no self-shielding correction. *The lower panel* shows the ratio of the latter and former. Self-shielding makes negligible difference to the median level of absorption at distances greater than 200 pkpc from galaxies, and $\sim 10\%$ difference for distances $\lesssim 100$ pkpc.

on absorption at large distances from galaxies. Without applying self shielding in post-processing, the median level of absorption within $\sim 55-100$ pkpc from galaxies is $\sim 10\%$ lower, while at larger distances it is nearly identical to the case with radiative transfer. This is consistent with Figure 5.2, which showed that radiative transfer does not change the results for optically thin absorbers, and far away from galaxies we expect gas to be predominantly optically thin. The simulation under-predicts the absorption level within ~ 200 pkpc from galaxies in comparison with the observations from Rakic et al. (2011). This is probably due to cold gas physics in the simulations not being captured sufficiently well (e.g. in reality gas at ISM densities can self-shield to form a cold phase with $T \ll 10^4$ K, while in the simulations its temperature stays at $\gtrsim 10^4$ K), as well as potentially not fully adequate SN feedback prescription in the *REF* model (as we will show below, feedback can raise gas temperature close to haloes).

The top left panel of Figure 5.7 shows that if we only include gas that was never heated to temperatures $T > 10^{5.5}$ K, then the absorption level decreases by $\approx 20\%$ within ≈ 130 pkpc from galaxies, and by only $\lesssim 5\%$ at distances $\gtrsim 1$ pMpc. In other words, gas that was never heated to $T > 10^{5.5}$ K accounts for almost all absorption in the Lyα forest, and only close to galaxies is there a mixture of hotter gas that cooled down sufficiently to absorb in H I Lyα and gas that is considered to be part of cold accretion. At distances $\gtrsim 200$ pkpc, the absorption profile of gas with $T_{\mathrm{max}} < 10^{5.5}$ K is consistent
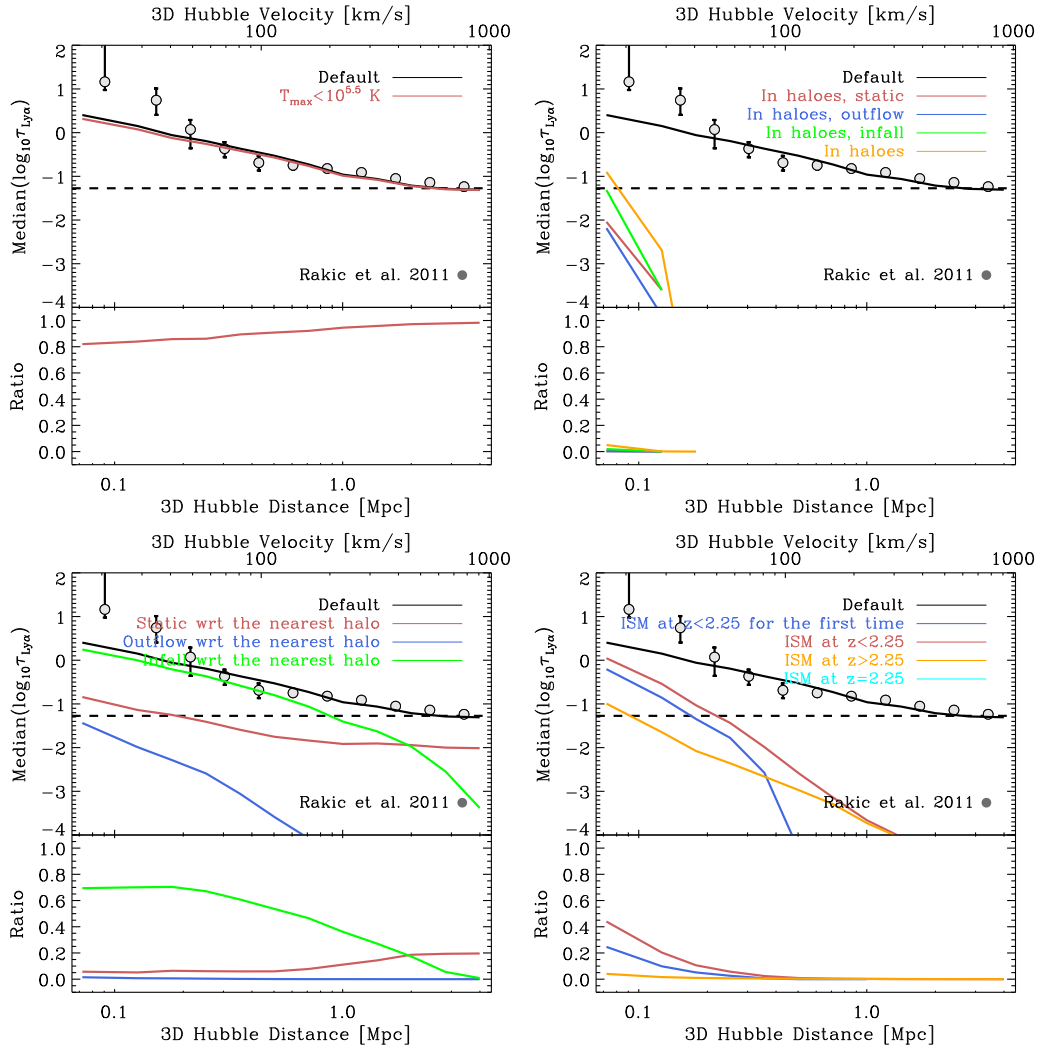
Figure 5.7 – *The top parts of all panels* show the median $\log_{10}(\tau_{\mathrm{Ly}\alpha})$ as a function of physical 3-D Hubble distance from $M_{\mathrm{h}} > 10^{11.5}\,\mathrm{M}_\odot$ galaxy halos at $z = 2.25$, where colored curves represent results when different gas samples are included, as indicated in the legends. *The bottom parts of all panels* show the ratios between the curves for different cuts and the default curve. The dashed line shows the median optical depth of all pixels. *The top right panel* shows the median optical depth of gas that is in haloes with $M_{\mathrm{h}} > 10^{11.5}\,\mathrm{M}_\odot$, and also of gas that is in addition moving with respect to the halo center with radial velocities $v > 0.25 v_{\mathrm{circ}}$ (for infalling and outflowing gas) and $v < 0.25 v_{\mathrm{circ}}$ (static gas).

with the observed profile. A comparison with Figure 5.2, which showed that within 200 pkpc the covering fraction of gas with $T_{\mathrm{max}} < 10^{5.5}\,\mathrm{K}$ is almost identical to the total covering fraction, shows that the effect on the median absorption statistics is larger.

In the top right panel of Figure 5.7 we apply more stringent cuts. It appears that gas that is within the host haloes produces absorption that is $\gtrsim 90\%$ smaller than the total absorption within ≈ 130 pkpc from the galaxies,

while at distances $\gtrsim 130$ pkpc its contribution is negligible. This is consistent with results from the previous Section (5.3), where strong absorbers are confined to haloes, and lower optical depth gas is less likely to be in haloes, which is why contribution of halo gas is negligible at large distances. Considering that 130 pkpc is comparable to the $R_{\rm vir}$ of the haloes, and that halo gas produces very small median optical depth within that 3-D distance, suggests that gas that is outside haloes, but due to peculiar velocities absorbs at halo positions, is a very important contributor.

Halo gas that is moving towards the halo center with $v > 0.25v_{\rm circ}$ produces absorption that is only a few percent of the total within 130 pkpc, and negligible beyond. Outflowing and static gas produce even weaker absorption.

In the bottom left panel of Figure 5.7 we see that if we include all gas that is moving towards the nearest halo (with $M_{\rm min} = 10^{11.5}$ M$_\odot$) with $v > 0.25v_{\rm circ}$ (where $v_{\rm circ}$ is the circular velocity of that halo) rather than only the gas within the haloes, then such gas makes a significant contribution to absorption even far from galaxies. This reflects the ongoing structure formation in the Universe where even low-density gas is undergoing motion towards large-scale overdensities. It produces absorption that is $\sim 30\%$ lower than the total within ~ 0.5 pMpc, and then slowly falls off to being only a few percent of the total at 5 pMpc. Gas that is going away from the galaxies with $v > 0.25v_{\rm circ}$ produces a median absorption level that is $\lesssim 1.5\%$ of the total at all distances from galaxies. Gas that is static relative to the galaxies ($v < 0.25v_{\rm circ}$) produces absorption that is $\lesssim 10\%$ of the total within ~ 1 pMpc from galaxies, but its contribution grows to $\sim 20\%$ of the total at ~ 5 pMpc, where it also surpasses the absorption by inflowing gas. The picture on large scales is quite different from that within 200 pkpc of the galaxies, where the covering fractions of outflowing and static gas are comparable. This suggests that outflowing gas becomes less important further away from galaxies.

In the bottom right panel of Figure 5.7 we select gas that is part of the ISM at $z = 2.25$, at $z > 2.25$, $z < 2.25$, and at $z < 2.25$ for the first time. The ISM contributes a negligible amount of Lyα absorption on the distance scales considered, and its median optical depth is consistent with zero. Absorption by ejected gas, i.e. gas that has been in galaxies before and is not at $z = 2.25$, is $\sim 95\%$ lower than total absorption within ≈ 130 pkpc from galaxies, and its contribution is negligible beyond that distance. Gas that is not in the ISM at $z = 2.25$, but becomes part of it after that redshift, produces significant absorption out to ~ 0.5 pMpc, and its median optical depth is only $\sim 50\%$ lower than total within ~ 130 pkpc from the galaxies. Most of this gas becomes the ISM after $z = 2.25$ for the first time. An interesting result from this panel is that although gas that is closer to galaxies is more likely to join the ISM, most of the Lyα absorbing gas at distances greater than ~ 200 pkpc will not accrete onto the galaxy before $z = 0$.

5.5   Physical properties of the Lyman-α absorbing gas

In this section we show the physical properties of the H I Lyα absorbing gas in the simulations. Given that the predicted median absorption profiles do not exactly match the observations, the relations below might not resemble the true gas densities and temperatures in those regions. However, we consider them useful for building intuition about gas properties, especially farther away from galaxies where the absorption profiles between observations and simulations agree well. We also note that the simulations used below were not post-processed with radiative transfer, unlike all previous results presented in this chapter.

In the figures below we show three different overdensity and temperature profiles of gas as a function of 3-D Hubble distance from the haloes:

   i) mass weighted, in real space;

  ii) $N_{\rm HI}$ weighted, i.e. weighted by the contribution to the optical depth, in real space;

 iii) optical depth weighted, i.e. weighted by the contribution to the optical depth of the pixel. This estimate is almost identical to the $N_{\rm HI}$ weighted one, the only difference being that it is evaluated in velocity space.

Along each sightline used in the previous section, we extract gas density and temperature distributions (that were used for calculating absorption spectra in the first place). Estimates of these quantities in real space were extracted from the simulations by ignoring peculiar velocities, as opposed to those estimated in velocity (i.e. redshift) space. We estimate the profiles below as a function of 3-D distance, where for quantities extracted in redshift space (i.e. the optical depth weighted overdensity and temperature) we assume that the LOS separation between haloes and pixels is due only to the Hubble flow.

Figure 5.8 compares the different median overdensity profiles, for the halo mass range of interest. The $N_{\rm HI}$ weighted overdensity is higher than the mass weighted overdensity suggesting that Lyα absorbing gas is denser than the gas that makes up most of the mass. The optical depth weighted overdensity profile, which only differs from the $N_{\rm HI}$ profile in that it is estimated in redshift space, is also higher than the mass weighted profile. In comparison with the $N_{\rm HI}$ profile, however, it shows a shallower slope with distance, because it is affected by the peculiar motions of the gas. The median optical depth weighted overdensity also appears higher than the median $N_{\rm HI}$ weighted overdensity, which is due to overdense absorbers having relatively small sizes in real space that get smoothed in redshift space due to peculiar velocities. For example, Lyα absorbers with the column density $N_{\rm HI} = 10^{16}\,{\rm cm}^{-2}$ have typical sizes of ~ 20 pkpc (Schaye 2001), while the typical Doppler parameter of such absorbers is $\approx 30\,{\rm km\,s}^{-1}$ (Rudie et al. 2011, in preparation) making them smeared over $\gtrsim 100$ pkpc in redshift space at $z = 2.25$.

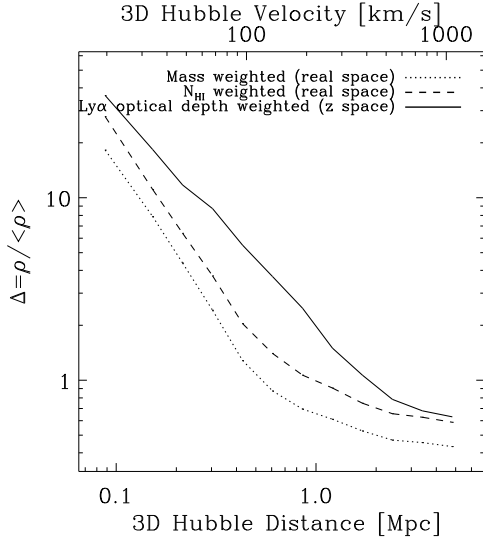The left panel of Figure 5.9 shows the different median temperature pro-

Figure 5.8 − Mass weighted (dotted lines), N_{HI} weighted (dashed lines), and H I Lyα optical depth weighted (solid lines) median overdensity of gas as a function of distance from the haloes, with mass $M_h > 10^{11.5} \, M_\odot$ at $z = 2.25$, in the 50 h^{-1} cMpc *REF* simulation.

Figure 5.9 − The left panel shows the median mass weighted (dotted lines), N_{HI} weighted (dashed lines), and H I Lyα optical depth weighted (solid lines), temperatures as a function of distance from the haloes, with $M_h > 10^{11.5} \, M_\odot$ at $z = 2.25$, in the 50 h^{-1} cMpc *REF* simulation. In the right p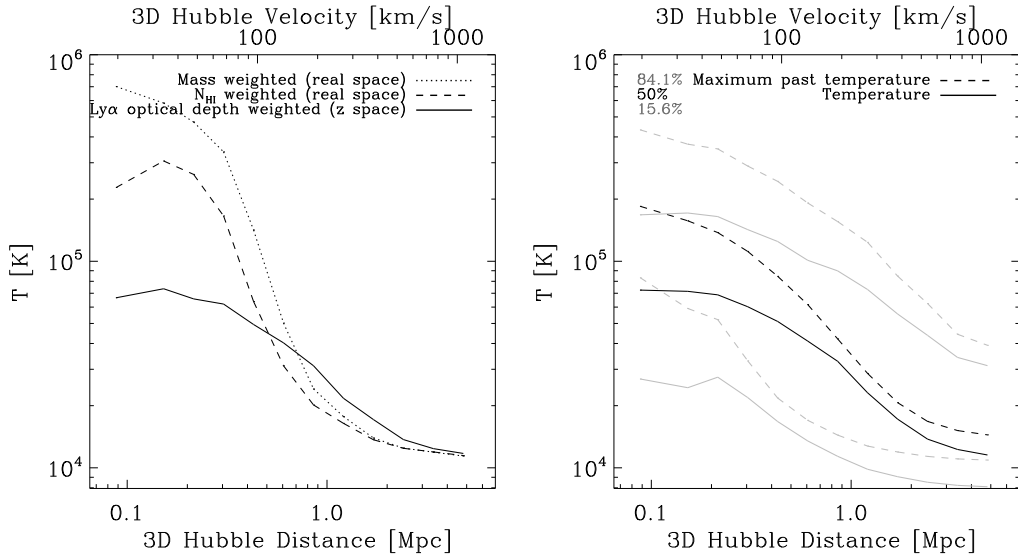anel we repeat the optical depth weighted median temperature of the gas from the left panel (black solid line) and show also the median of the maximum temperature that the Lyα absorbing gas has ever reached (black dashed line). Grey lines show the corresponding 15.1 and 84.1 percentiles (encompassing the 1σ interval around the median).

files. The mass weighted temperature is higher than the N_{HI} weighted temperature, because it traces all gas, including gas that is collisionally ionized

at the virial temperature ($T_{\rm vir} > 10^{5.5}$ K). The Lyα absorption traces denser and colder, neutral gas. The optical depth weighted temperature is also lower than the mass weighted temperature, and the profile is shallower than the $N_{\rm HI}$ weighted temperature due to peculiar velocities. The optical depth weighted temperature is lower than the ion weighted temperature in the innermost regions because peculiar velocities bring the colder gas closer to galaxy redshifts in velocity space (see Chapter 4 where we found that neglecting peculiar velocities reduces the signal close to galaxies).

The right panel of Figure 5.9 compares the median optical depth weighted current gas temperature with the median optical depth weighted maximum temperature that the gas has ever reached, as a function of distance from the halo. The 1σ scatter around the median is also shown. Most of the Lyα absorbing gas is and has always been cold ($T_{\rm max} \lesssim 10^{5.5}$ K), while there is a mixture of hot- and cold-mode gas close to galaxies ($\lesssim 200$ kpc). This is consistent with the findings presented in the previous two sections.

Each panel of Figure 5.10 shows the median optical depth weighted overdensity and temperature of gas as a function of the optical depth for a particular range of impact parameters (estimated over $\pm 165\,{\rm km\,s^{-1}}$ along the LOS from galaxy positions), in the *REF* model. In each panel we also show the results for 1000 random sightlines (grey lines), which is identical in each panel. Looking at the results for random sightlines, we notice that both the overdensity and the temperature increase with the optical depth, for $-2 \lesssim \log_{10}\tau \lesssim 1$. This reflects the well-known temperature-density relation (e.g. Hui & Gnedin 1997), where the temperature of the IGM is set by the competition between adiabatic cooling due to the expansion of the Universe and photoheating. At the low optical depth end ($\log_{10}\tau_{\rm Ly\alpha} < -2$) there is an upturn in both the density and temperature due to gas that is shock-heated to high temperatures and therefore highly ionized. At the high optical depth end ($\log_{10}\tau_{\rm Ly\alpha} > 1$) the temperature decreases with increasing optical depth and density, because this gas is dense enough to cool radiatively.

Even for impact parameters $1.42 < b < 2$ pMpc the temperature for a given optical depth in the range $-1 \lesssim \log_{10}\tau \lesssim 1$ is higher than for a random sightline. Approaching galaxies, the temperature of the Lyα absorbing gas increases further and for $b \lesssim 0.5$ pMpc the density corresponding to a fixed optical depth also becomes noticeably higher than for a random place. The temperature-density relation (for $-2 \lesssim \log_{10}\tau \lesssim 1$) gets gradually inverted with decreasing impact parameter and is almost fully inverted at $b \lesssim 0.25$ pMpc; instead of both gas temperature and overdensity increasing with optical depth, the temperature is higher for low optical depths. Absorbers with optical depth $\tau = 0.1$ are gas clouds with a typical overdensity of $\sim 1-2$ and a temperature $T \sim 2 \times 10^4$ K at ~ 2 Mpc from galaxies, but they are gas clouds with overdensity ~ 10 and temperature $\sim 10^5$ K within ~ 100 pkpc from galaxies. This means that *gas with the same Lyα optical depth has very different properties depending on where it is with respect to the haloes.*
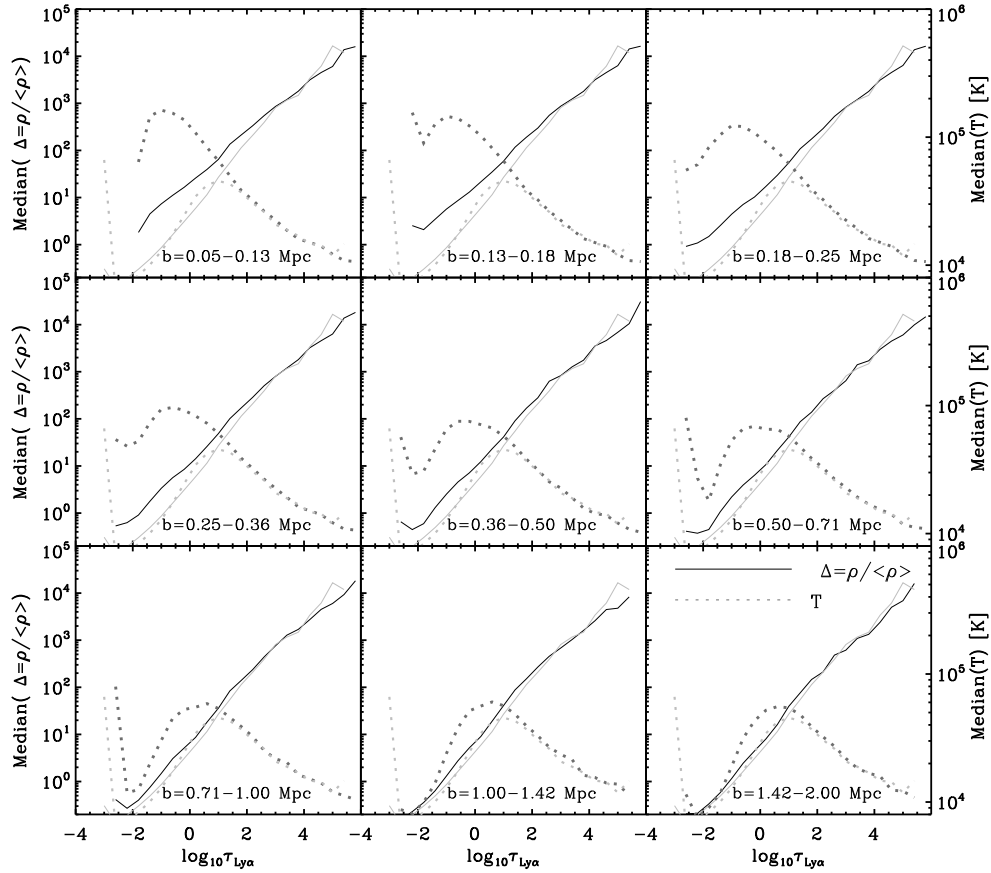
Figure 5.10 – Different panels show the $z = 2.25$ median H I Lyα optical depth weighted overdensity (solid black line, left axis) and temperature (dotted black line, right axis) as a function of the optical depth, with different impact parameter ranges shown in different panels (estimated over $\pm 165\,\mathrm{km\,s^{-1}}$ from halos with $M_\mathrm{h} > 10^{11.5}\,\mathrm{M_\odot}$). The grey lines are repeated in each panel and show the results for sightlines drawn at random positions. We can see that gas with the same optical depth has very different properties depending on where it is with respect to the haloes.

Figure 5.11 is similar to Figure 5.10 but shows the properties of absorbers with a given optical depth for the *AGN* model. With respect to the *REF* model (grey curves), more gas has been shock-heated to high temperatures at the low optical depth end, and the temperature-density relation is already fully inverted at impact parameters $b \lesssim 1$ pMpc.

In Figure 5.12, we show the results for the *NOSN* model. Comparing them with results for the *REF* model (grey curves), it becomes obvious that at least some of the temperature-density inversion, for $-2 \lesssim \log_{10}\tau \lesssim 1$, at small impact parameters is due to SN feedback. Nevertheless, even pure gravitational heating is able to somewhat invert the temperature-density relation.
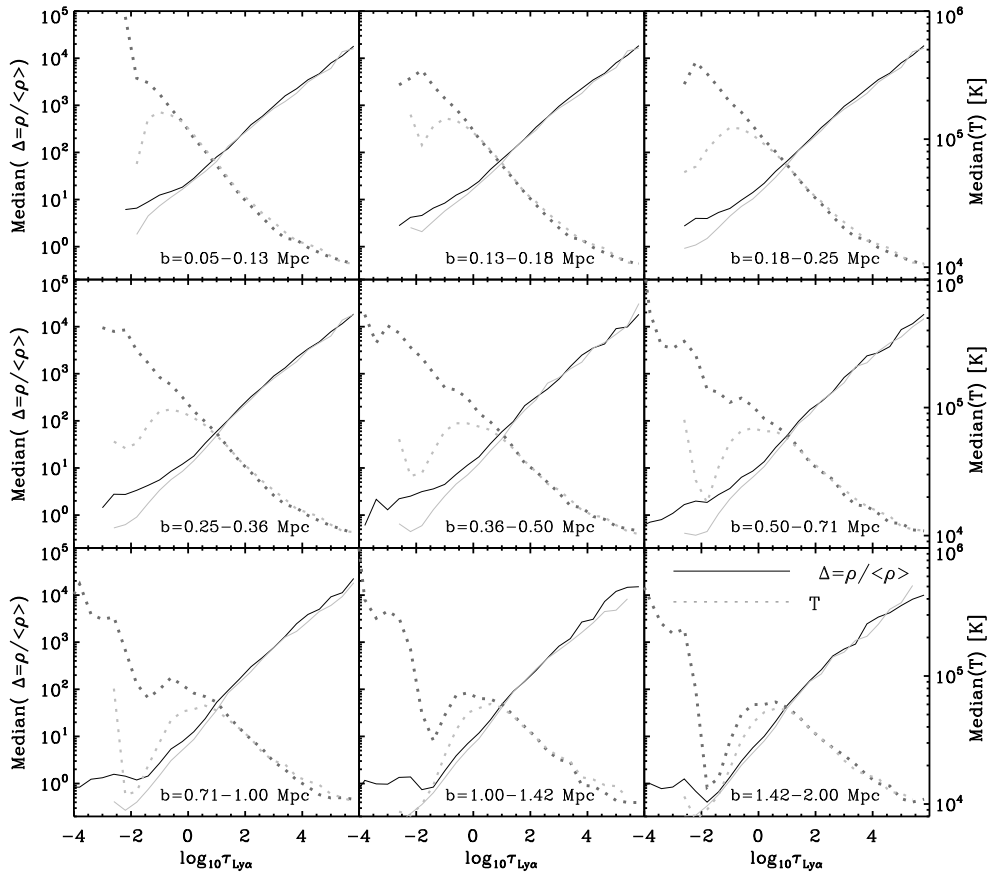
Figure 5.11 – Similar to Figure 5.10 but for the *AGN* model. The grey lines show the results for the *REF* model. In comparison with the *REF* model, gas properties are affected at greater distances from the galaxies.

5.6   Summary & Conclusions

Cold flows are thought to be the dominant process through which galaxies acquire fuel for star-formation, but there is no unequivocal observational evidence for their existence as of yet. Here we explore how much of the gas seen in H I Lyα absorption at $z = 2.25$ is due to cold flows. We use the reference model from the OWLS suite of simulations (Schaye et al. 2010), with radiative transfer applied in post-processing, and study absorption near haloes with $M_\mathrm{h} \geq 10^{11.5}\,\mathrm{M}_\odot$, which are expected to host the galaxies from Rakic et al. (2011). Our results are dominated by haloes with $M_\mathrm{h} \sim 10^{12}\,\mathrm{M}_\odot$, which is where the simulations predict that the cold and hot modes provide comparable contributions to the gas accretion rates onto haloes (Van de Voort et al. 2011a). We found that the improved radiation transfer handling is only important for the part where we focus on the dense halo gas which can self-shield against the ionizing radiation.
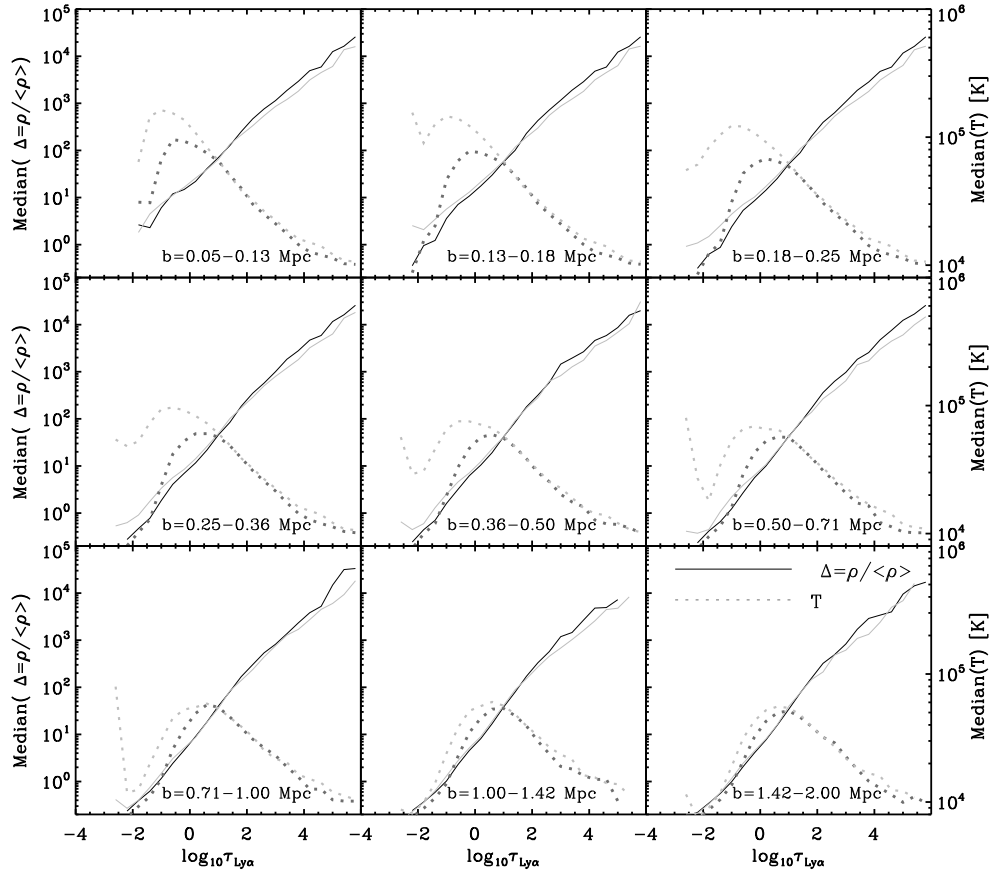
Figure 5.12 – Similar to Figure 5.10, but for the *NOSN* model. In comparison with the *REF* model (grey curves), the temperature density-relation is less inverted at small impact parameters, suggesting that at least some of the heating is due to SN-driven winds.

In the first part of this chapter we studied the nature of the Lyα absorbing gas by examining its past and future properties, and in the second part we studied its temperature and density. We looked into their dependence on the feedback implementation by using simulations with SN feedback (the reference model), with both SN and AGN feedback (model *AGN*), and without feedback (model *NOSN*). Unlike in the first part of the chapter, radiative transfer was not applied to any of these simulations.

We studied the covering fraction of absorbers with a given strength in the immediate surroundings of the halo, i.e. at impact parameters $b \leq 200$ pkpc, as a function of impact parameter, and across $\pm 165 \, \mathrm{km \, s^{-1}}$, for gas that is optically thin to the ionizing radiation. For gas with column densities $N_{\mathrm{HI}} \geq 10^{17.2} \, \mathrm{cm^{-2}}$ we instead considered the total column density along the sightline through the simulation volume. We compared the predicted covering fractions to those observed by Rakic et al. (2011) for maximum $\log_{10} \tau_{\mathrm{Ly}\alpha} > 0$, 1, and 2, and Rudie et al. (2011, in preparation) for $\log_{10} N_{\mathrm{HI}} > 17.2$, 18.5,

and 20.3 cm^{-2}. We found that:

- Most of the Lyα absorbing gas within 200 pkpc from galaxies has not been shock heated to temperatures $T > 10^{5.5}$ K.
- Gas is mostly moving towards the haloes with $v > 0.25v_{\mathrm{circ}}$, although a non-negligible fraction is flowing outwards (also with $v > 0.25v_{\mathrm{circ}}$), or is static with respect to the halo centers (i.e. $v < 0.25v_{\mathrm{circ}}$).
- Most of the Lyα absorbing gas is on its way to joining the ISM of galaxies after the time of observation ($z = 2.25$), and will be accreting for the first time. Optically thick gas is more likely to become part of galaxies before $z = 0$ than weaker absorbers at the same impact parameter.
- Gas with $T_{\mathrm{max}} \lesssim 10^{5.5}$ K has a covering fraction that is consistent with the observed covering fraction of gas with maximum $\log_{10}\tau_{\mathrm{Ly}\alpha} > 0, 1$, and 2. Gas that is infalling towards haloes, but is not necessarily inside them, and gas that will join the ISM of galaxies by $z = 0$, also have covering fractions that are consistent with the observed.

In order to say whether different gas samples have a critical role in accounting for the observed covering fraction, it would be more appropriate to exclude particles belonging to those samples, and compare the absorption produced by the remaining gas to the observed absorption. If the remaining gas exhibits the covering fraction of Lyα absorbers that is lower than observed, we would be able to conclude that the excluded gas plays a major role in absorption. In addition, the complete analysis should be repeated for simulations with different feedback models to examine the sensitivity of the results to particular prescriptions. We will address these issues in our future publication.

We also studied the median H I Lyα optical depth of gas with given properties on relatively large scales in comparison with halo sizes ($\lesssim 5$ pMpc from the haloes), as a function of 3-D distance ($\equiv \sqrt{d_{\mathrm{LOS}}^2 + b^2}$, where d_{LOS} is the LOS separation between the halo and the absorber estimated from their velocity separation assuming that it is due to Hubble expansion). The main findings are:

- At large distances ($\gtrsim 1$ pMpc), gas with a maximum past temperature of $\lesssim 10^{5.5}$ K produces a median optical depth that is only $\lesssim 5\%$ lower than the total median optical depth (i.e. when we include all gas), suggesting that most of absorption in the Lyα forest is due to gas that was never shock-heated to high temperatures. This gas alone can account for the observed median optical depth at distances $\gtrsim 200$ pkpc.
- Gas within the haloes produces a median optical depth that is $\lesssim 20\%$ lower than that produced by all gas within ~ 130 pkpc (3-D distance), and its contribution to the total median optical depth is negligible beyond this distance.
- Gas that is inflowing towards the nearest halo with $M_{\mathrm{h}} > 10^{11.5}\mathrm{M}_{\odot}$ pro-

duces a median optical depth that is $\gtrsim 30\%$ lower than that produced by all particles within ~ 1 pMpc from the haloes; it remains significant even at 2 pMpc, beyond which static gas becomes dominant. This suggests that even at large distances, a non-negligible amount of gas is moving towards the haloes, which is a consequence of the ongoing structure formation. This is consistent with the observations of Rakic et al. (2011) who detected gas infall on scales of 1.4-2 pMpc from galaxies. The outflowing gas produces a median optical depth that is $\lesssim 1.5\%$ of that produced by all gas within ~ 1 pMpc and negligible beyond.

- Gas that will join the ISM of galaxies by $z = 0$ and is not in the ISM at $z = 2.25$, produces a median optical depth that is $\sim 30\%$ lower than that produced by all Lyα absorbing particles within ~ 130 pkpc. Its contribution to the median optical depth declines with distance from haloes, and is negligible beyond ~ 500 pkpc. The majority of this gas will accrete onto a galaxy for the first time. Most of the Lyα forest gas, however, does not come into galaxies before $z = 0$.

Finally, we looked at the overdensity and temperature of gas with a given optical depth, at different impact parameters from the haloes, using simulations where no radiative transfer post-processing was applied. We find:

- The *REF* model predicts that absorbers with optical depth $10^{-2} \lesssim \tau_{\mathrm{Ly}\alpha} \lesssim 10$ are typically produced by gas whose temperature and density follow the temperature-density relation, i.e. the temperature increases with the gas density. This relation changes closer to galaxies, and is almost fully inverted, i.e. gas temperature increases with decreasing density, within ~ 0.25 pMpc from haloes with $M_{\mathrm{h}} > 10^{11.5}$ M$_\odot$. This is mostly due to SN-driven winds heating the gas near galaxies. As a result, the physical properties of gas with the same Lyα optical depth change depending on where it is with respect to galaxies.

- The *AGN* model predicts a temperature-density relation that differs from the reference model even at ~ 2 pMpc from the haloes, with more gas heated to high temperatures. The relation is fully inverted (i.e. gas temperature increases with decreasing density) within ~ 1 pMpc.

- In the *NOSN* model the temperature-density relation appears largely independent of separation from haloes, but is slightly inverted within ~ 200 pkpc. This suggests that the galactic feedback is not solely responsible for the inversion, and that gravitational shock-heating plays a role as well.

The fact that the temperature-density relation appears sensitive to the feedback model, and that it changes as a function of distance from haloes, provides another way to constrain galactic feedback once observations of the temperature-density relation can be performed at different distances from galaxies. Although performed on larger scales, an indirect proof of concept is the measurement of the IGM temperature in the QSO proximity zone by Bolton et al. (2010). Their measurement was conducted across ≈ 30 cMpc in

a single sightline, while we found that strong feedback can impact the physical state of the IGM out to (at least) ~ 2 pMpc. While it will be difficult to test our predictions in a single sightline, a number of sightlines (QSO-galaxy pairs) passing at different distances from galaxies can make up sufficient redshift path for the temperature to be measured as a function of distance from galaxies (see Rudie et al., 2011, in preparation).

References

Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, ApJ, 584, 45

Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Reddy N. A., 2005, ApJ, 629, 636

Altay, G., Theuns, T., Schaye, J., Crighton, N. H. M., Dalla Vecchia, C., 2011, ApJ, 737, L37

Birnboim, Y., Dekel, A., 2003, MNRAS, 345, 349

Bolton, J. S., Becker, G. D., Wyithe, J. S. B., Haehnelt, M. G., & Sargent, W. L. W. 2010, MNRAS, 406, 612

Brooks, A. M., Governato, F., Quinn, T., Brook, C. B., Wadsley, J., 2009, ApJ, 694, 396

Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80

Chen, Y.-M., Tremonti, C. A., Heckman, T. M., Kauffmann, G., Weiner, B. J., Brinchmann, J., Wang, J., 2010, AJ, 140, 445

Chabrier, G., 2003, PASP, 115, 763

Crain, R. A., Theuns, T., Dalla Vecchia, C., Eke, V. R., Frenk, C. S., Jenkins, A., Kay, S. T., Peacock, J. A., Pearce, F. R., Schaye, J., Springel, V., Thomas, P. A., White, S. D. M., Wiersma, R. P. C., 2009, MNRAS, 399.1773

Dalla Vecchia, C., Schaye, J., 2008, MNRAS, 387, 1431

Dijkstra M., Loeb, A., 2009, MNRAS, 400, 1109

Dolag, K., Borgani, S., Murante, G., Springel, V., 2009, MNRAS, 399, 497

Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, ApJ, 644, 813

Erb, D. K. 2008, ApJ, 674, 151

Faucher-Giguère, C.-A., Kereš, D., Dijkstra, M., Hernquist, L., Zaldarriaga, M., 2010, ApJ, 725, 633

Faucher-Giguere C.-A., Kereš D., 2011, MNRAS, 412, 118

Faucher-Giguere, C.-A., Kereš, D., Ma, C.-P., 2011, MNRAS, in press

Fumagalli, M., Prochaska, J. X., Kasen, D., Dekel, A., Ceverino, D., Primack, J. R., 2011, MNRAS, in press

Furlanetto, S. R., Schaye, J., Springel, V., & Hernquist, L. 2005, ApJ, 622, 7

Geach, J. E., Alexander, D. M., Lehmer, B. D., Smail, I., Matsuda, Y., Chapman, S. C., Scharf, C. A., Ivison, R. J., Volonteri, M., Yamada, T., Blain, A. W., Bower, R. G., Bauer, F. E., Basu-Zych, A., 2009, ApJ, 700, 1

Giavalisco, M., Vanzella, E., Salimbeni, S., Tripp, T. M., Dickinson, M., Cassata, P., Renzini, A., Guo, Y., Ferguson, H. C., Nonino, M., Cimatti, A., Kurk, J., Mignoli, M., Tang, Y., 2011, ApJ, in press

Goerdt, T., Dekel, A., Sternberg, A., et al., 2010, MNRAS, 407, 613

Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays Modelling the UV/X-ray cosmic background with CUBA, arXiv: 0106018

Hayes, M., Scarlata, C., & Siana, B. 2011, Nature, 476, 304

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

Kereš, D., Katz, N., Weinberg, D. H., Davé, R., 2005, MNRAS, 363, 2

Kereš, D., Katz, N., Fardal, M., Davé, R., Weinberg, D. H., 2009, MNRAS, 395, 160

Kimm, T., Slyz, A., Devriendt, J., & Pichon, C. 2011, MNRAS, 413, L51

Lacey, C., & Cole, S. 1994, MNRAS, 271, 676

Martin, C. L., Bouché, N., 2009, ApJ, 703, 1394

Nilsson, K. K., Fynbo, J. P. U., Møller, P., Sommer-Larsen, J., Ledoux, C., 2006, A&A, 452, 23

Ocvirk, P., Pichon, C., Teyssier, R., 2008, MNRAS, 390, 1326

Oppenheimer, B. D., Davé, R., Kereš, D., et al. 2010, MNRAS, 406, 2325

Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., Giavalisco, M., MNRAS, 412, 1123

Pettini, M., Shapley, A. E., Steidel, C. C., Cuby, J.-G., Dickinson, M., Moorwood, A. F. M., Adelberger, K. L., Giavalisco, M., 2001, ApJ, 554, 981

Prochaska, J. X., Wolfe, A. M., 2009, ApJ, 696, 1543

Rakic, O., Schaye, J., Steidel, C. C., & Rudie, G. C. 2011, arXiv:1109.4944

Rauch M., et al., 1997, ApJ, 489, 7

Ribaudo, J., Lehner, N., Howk, J. C., Werk, J. K., Tripp, T. M., Prochaska, J. X. , Meiring, J. D., Tumlinson, J., 2011, ApJ, in press

Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., Weiner B. J., 2010, ApJ, 712, 574

Schaye, J. 2001, ApJ, 559, 507

Schaye, J., Dalla Vecchia, C., Booth, C. M., Wiersma, R. P. C., Theuns, T., Haas, M. R., Bertone, S., Duffy, A. R., McCarthy, I. G., van de Voort, F., 2010, MNRAS, 402, 1536

Shapiro, K. L., Genzel, R., Quataert, E., Förster Schreiber, N. M., Davies, R., Tacconi, L., Armus, L., Bouché, N., Buschkamp, P., Cimatti, A., Cresci, G., Daddi, E.,Eisenhauer, F., Erb, D. K., Genel, S., Hicks, E. K. S., Lilly, S. J., Lutz, D., Renzini, A., Shapley, A., Steidel, C. C., Sternberg, A., ApJ, 701, 955

Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., Giavalisco, M., 2000, ApJ, 532, 170

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289

Steidel, C. C., Bogosavljević, M., Shapley, A. E., Kollmeier, J. A., Reddy, N. A., Erb, D. K., Pettini, M., 2011, ApJ, 736, 160

Stewart, K. R., Kaufmann, T., Bullock, J. S., Barton, E. J., Maller, A. H., Diemand, J., Wadsley, J., 2011,ApJ, 735, L1

Stewart, K. R., Kaufmann, T., Bullock, J. S., Barton, E. J., Maller, A. H., Diemand, J., Wadsley, J., 2011, ApJ, 738, 39

Van de Voort, F., Schaye, J., Booth, C. M., Haas, M. R., & Dalla Vecchia, C. 2011, MNRAS, 414, 2458

Van de Voort, F., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, MNRAS, 415, 2782

Van de Voort, F., Schaye, J., Altay, G., & Theuns, T. 2011, arXiv:1109.5700

Weiner et al., 2009, ApJ, 692, 187

Wiersma, R. P. C., Schaye, J., Smith, B. D., 2009. MNRAS, 393, 99

Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., Tornatore, L., 2009, MNRAS, 399, 574

Nederlandse samenvatting

In de laatste 15 jaar is het veld van de extra-galactische astronomie doorgedrongen tot het verre heelal op hoge roodverschuiving en is onze kennis van jonge, afgelegen sterrenstelsels dramatisch gegroeid. Sterrenstelsels worden momenteel routinematig gedetecteerd op $z \approx 6 - 7$ (e.g. Bouwens et al. 2010; Labbé et al. 2010; Oesch et al. 2010; Bouwens et al. 2011; McLure et al. 2011). Toch zijn er fundamentele vragen waarop het antwoord nog niet is gegeven: Hoe komen sterrenstelsels aan hun gas? Hoe beïnvloedt feedback de ontwikkeling van sterrenstelsels?

5.7 Intergalactisch Medium

Sterrenstelsels worden gevormd uit het intergalactisch medium en er was een tijd, voordat er stelsels waren, dat alle materie in het heelal zich in het intergalactisch medium bevond. Het meeste wat we weten van het intergalactisch medium komt van onderzoek naar absorptiespectra van heldere objecten zoals quasars. Het tussenliggende gas laat zijn sporen na op quasarspectra door de selectieve absorptie van continuümstraling. Het meest prominente patroon in zulke spectra is het Lyman-α woud van het meest talrijke element in het heelal, waterstof. Het woud bestaat uit een aantal absorptielijnen die worden veroorzaakt door wolken van neutraal waterstofgas langs de gezichtslijn naar de quasar. Deze gaswolken absorberen alle continuümstraling die roodverschoven is naar een golflengte van H I Lyα (1216 Å).

Het moderne begrip van het Lyα woud stamt uit de vroege jaren 90, toen H. Bi en zijn collegas zich realiseerden dat dichtheidsverschillen in het intergalactisch medium een realistische verklaring voor het woud boden. Dit werd later bevestigd door hydrodynamische kosmologische simulaties. Het idee is simpel gezegd dat intergalactisch gas de ruimtelijke verdeling van donkere materie volgt en een netwerk van vlakken en filamenten vormt. Middenin de halos van donkere materie ontstaan hieruit uiteindelijk sterrenstelsels. Deze structuur veroorzaakt het Lyα woud: de vlakken en filamenten zijn zwakke absorbeerders van de achterliggende straling, de halo absorbeert juist sterk. Dit zorgt ervoor dat het Lyα woud een overzicht biedt van de verdeling van alle materie langs de gezichtslijn naar quasars: de verdeling van baryonen wordt namelijk bepaald door de verdeling van donkere materie en de baryonen bestaan voor het grootste gedeelte uit waterstof. De structuren die

Lyα absorberen, van het zeer zwak absorberende ijle gas tot de sterk absorberende schijven van sterrenstelsels, vertegenwoordigen verschillende stadia in de odyssee van de baryonen, van diffuus gas tot samengeperste materie. Het doorgronden van hun aard en de link met sterrenstelsels is een essentiële stap als we willen begrijpen hoe sterrenstelsels zich vormen en ontwikkelen. Naast het feit dat de intergalactische materie een gasreservoir is voor sterrenstelsels, bevat het ook informatie over feedback. Feedback omhelst alle processen waarbij de huidige stervorming en actieve kernen in een sterrenstelsel de evolutie van het stelsel beïnvloeden, bijvoorbeeld door het verwarmen of verwijderen van gas dat sterren vormt. Straling afkomstig uit sterrenstelsels verwarmt en ioniseert het intergalactisch medium. Supernova-explosies en actieve kernen veroorzaken winden die het intergalactisch medium schokachtig verwarmen en het verrijken met elementen die zwaarder zijn dan helium (metalen in de volksmond van astronomen). Deze intergalactische metalen laten ook hun sporen na op de spectra van achtergelegen objecten. Via de absorptiespectra kan de verdeling van dergelijke metalen achterhaald worden, wat weer meer informatie geeft over feedback in sterrenstelsels.

5.8 Het intergalactisch medium in de buurt van sterrenstelsels

De natuurkunde die het Lyα woud bepaalt is betrekkelijk eenvoudig, maar wordt snel ingewikkeld in het gebied waar het intergalactische medium en de sterrenstelsels samenkomen en de materie niet-lineair fluctueert. Observationeel gezien wordt dit grensgebied vaak bestudeerd door sterrenstelsels te zoeken die dicht langs de gezichtslijn naar de achtergelegen objecten liggen. Vervolgens wordt de absorptie in de spectra van de achtergelegen objecten die overeenkomt met de roodverschuiving van de sterrenstelsels onderzocht. Zulke waarnemingen, van de omgeving van sterrenstelsels, zijn zeer tijdrovend. Op hoge roodverschuiving zijn typische stervormende sterrenstelsels zeer zwak en het bepalen van hun roodverschuiving vergt veel tijd, zelfs met grote telescopen. Doordat ze zo zwak zijn, worden deze sterrenstelsels meestal als object in de voorgrond gebruikt en kiest men voor de achtergelegen objecten vaak voor de veel helderdere QSOs. Het aantal geschikte QSOs daalt met roodverschuiving en het aantal hechte QSO-sterrenstelsel paren is nog kleiner. Door stervormende sterrenstelsels als achtergrondobjecten te gebruiken kan een groter aantal hechte sterrenstelsel-paren gevonden worden, waardoor waardevolle studies naar de omgeving van de stelsels mogelijk wordt gemaakt. Het nadeel hiervan is dat de kwaliteit van de achtergrondspectra veel slechter is, zodat er opeenstapeling van data vereist is, wat beperkend is voor het soort onderzoek dat kan worden uitgevoerd. Pas wanneer de nieuwe generatie van 30m telescopen beschikbaar wordt zal de kwaliteit van dit soort spectra voldoende verbeteren.

Op lage roodverschuiving is het waarnemen van sterrenstelsels makkelij-

ker zijn en het aantal geschikte QSO-kandidaten hoger, maar hiervoor geldt weer dat de interessante overgangen zoals die van Lyα enkel toegankelijk zijn met behulp van ruimte-faciliteiten aangezien ze hun straling uitzenden in het ultraviolet (UV) liggen. Toch wordt hier flinke vooruitgang geboekt dankzij de komst van de UV-gevoelige Cosmic Origins Spectrograph (COS) op de Hubble ruimtetelescoop in 2009.

Het behoeft geen uitleg dat we dit soort studies, in het grensgebied tussen sterrenstelsel en intergalactisch medium, het liefst op verschillende roodverschuivingen verrichten. De omstandigheden in het heelal veranderen dramatisch van roodverschuiving $z \sim 2$ tot $z = 0$. Op hoge roodverschuiving is het heelal dichter, sterrenstelsels vormen sterren op een hoger tempo, actieve kernen komen veel vaker voor en galactische achtergrondstraling is intenser dan in het lokale heelal.

Ook theoretisch onderzoek naar het grensgebied tussen sterrenstelsels en het intergalactisch medium is lastig. Verscheidene processen die belangrijk zijn in dit gebied, zoals het (niet-evenwichtig) afkoelen van gas in nabijheid van metalen, ionisatie door botsingen, gravitationele verwarming door schokken, schokachtige verwarming door galactische winden, foto-ionisatie door lokale stralingsbronnen en zelf-afscherming, vereisen een numerieke benadering. Helaas zijn numerieke simulaties nog niet in het stadium dat ze alle relevante processen mee kunnen nemen. Dat komt doordat het vormen van een sterrenstelsel een groot dynamisch bereik vergt van een simulatie; van processen op atomair niveau tot het gedrag van clusters van sterrenstelsels. Deze voorschriften worden ondersteund door waarnemingen en de uitkomst van de simulaties moet zorgvuldig en kritisch worden bekeken ten opzichte van andere waarnemingen. Door een dergelijk proces werd het duidelijk dat feedback nodig is in sterrenstelsels om bijvoorbeeld te voorkomen dat er te veel sterren worden gevormd of ervoor te zorgen dat er schijven groeien tot een realistische grootte.

5.9   Dit proefschrift

Dit proefschrift presenteert onderzoek naar het intergalactisch medium in de buurt van sterrenstelsels op roodverschuiving $z \sim 2.4$ in de Keck Baryonic Structure Survey (KBSS, Steidel et al. in voorbereiding), zoals het door Lyα absorptie wordt getoond. Deze waarnemingen worden vergeleken met de hydrodynamische kosmologische simulaties van de OverWhelmingly Large Simulations (OWLS). Deze roodverschuiving is zeer geschikt voor studies naar zowel sterrenstelsels als het Lyα woud, aangezien ze beiden goed waargenomen kunnen worden met telescopen op aarde. De typische spectrale kenmerken, uitgezonden in het ultraviolet zijn op deze afstand verschoven naar het optische deel van het centrum wat efficiënte identificatie mogelijk maakt met behulp van de zogenaamde Lyman Break techniek. Bovendien zijn de lijnen van het Lyα woud op $z \gg 2.4$ grotendeels verzadigd en op $z \ll 2.4$ juist zeer

zeldzaam, wat deze roodverschuiving extra bijzonder maakt. Daarbij komt ook nog dat de algehele stervorming in het heelal een hoogtepunt meemaakte op $z \sim 2-3$, wat dat tijdperk bijzonder bruikbaar maakt voor dit soort studies, aangezien de interactie tussen sterrenstelsels en hun omgeving op dat moment maximaal zou moeten zijn.

Hoofdstuk 2 beschrijft een nieuwe methode om roodverschuivingen van sterrenstelsels te kalibreren door gebruik te maken van absorptie-eigenschappen van het omliggende intergalactische medium. De gebruikelijke manier om roodverschuivingen te bepalen van afgelegen sterrenstelsels is door middel van absorptie van de uitgezonden UV-straling en emissielijnen afkomstig van de interstellaire materie in sterrenstelsels. Vanwege een combinatie van effecten die te maken hebben met stralingsoverdracht en galactische stromen, wordt hiermee echter niet precies de roodverschuiving van het systeem bepaald. Een algemeen geaccepteerde manier om deze roodverschuivingen te kalibreren is met behulp van infrarode waarnemingen van emissielijnen afkomstig uit H II-gebieden in sterrenstelsels. Helaas zijn dat soort waarnemingen momenteel zeer kostbaar en niet uit te voeren voor grote verzamelingen van zwakke sterrenstelsels. Met behulp van de sterrenstelsels in het KBSS en de achterliggende, laten we zien dat het ook mogelijk is om de roodverschuivingen van de sterrenstelsels op een andere manier te kalibreren. Namelijk door gebruik te maken van het feit dat de gemiddelde H I Ly-alpha absorptieprofielen rondom de sterrenstelsels, zoals ze zich in de spectra van achtergelegen objecten manifesteren, symmetrisch moeten zijn ten opzichte van de echte roodverschuiving van de sterrenstelsels als je ervan uitgaat dat sterrenstelsels willekeurig verdeeld zijn rond de gezichtslijn naar de achtergelegen objecten.

In Hoofdstuk 3 presenteren we waarnemingen van de H I optische diepte in de buurt van sterrenstelsels op $z \sim 2.4$ in de KBSS data, door de quasarspectra te analyseren met behulp van de pixel optische diepte methode. Hieruit bleek dat Lyα absorptie wordt versterkt tot een afstand van tenminste 2.8 Mpc van het sterrenstelsel. We presenteren de eerste tweedimensionale kaarten van absorptie rond sterrenstelsels, waarbij we de mediaan van de Lyα optische diepte tonen als een functie van afstand van de sterrenstelsels, zowel langs de gezichtslijn naar de achtergelegen bron als er loodrecht op. We vinden twee soorten vervormingen van de roodverschuivingen. Op afstanden < 1 Mpc (of < 200 km s^{-1} in de spectra) is de absorptie sterker langs de gezichtslijn dan dwars erop. Dit effect, ook wel vinger van God genoemd, zou gedeeltelijk kunnen zijn veroorzaakt door onzekerheden in de bepaling van roodverschuivingen, maar wordt waarschijnlijk gedomineerd door de beweging van gas binnenin of dicht bij de halos. Aan de andere kant wordt, op afstanden tussen de 1.4-2.0 Mpc, de absorptie samengedrukt langs de gezichtslijn, een effect wat we associëren met inval op grote schaal (het Kaiser effect). We hebben de dichtheid van de sterrenstelsels binnen een bepaald volume gemeten als een functie van optische diepte van de pixel and tonen de covering fraction van absorberend materiaal met een bepaalde kracht tot

op 200 kpc van de sterrenstelsels.

In Hoofdstuk 4 laten we zien dat de verdeling van Lyα absorptie in de buurt van sterrenstelsels waargenomen in hoofdstuk 2 gebruikt kan worden om de massa te bepalen van de halos van die sterrenstelsels. We koppelen de waargenomen absorptieverdeling aan de absorptie rondom sterrenstelsels in de kosmologische simulaties van de OWLS set. De minimale halomassa die op deze manier bepaald wordt, is consistent met resultaten van analyses van het clustergedrag van sterrenstelsels en de resultaten zijn onafhankelijk van verandering in kosmologische parameters of de beschrijving van feedback in de modellen. We laten ook zien dat deze methode gebruikt kan worden in kleinere velden met sterrenstelsel-QSO waarnemingen, van 30×30 arcseconds.

Geïnspireerd door recente theoretische resultaten, die laten zien dat de meeste brandstof voor stervorming in sterrenstelsels daar terechtkomt door zogenaamde koude accretie, dat wil zeggen zonder te worden verwarmd tot de viriële temperatuur van de gasthalos, onderzoeken we in hoofdstuk 5 hoeveel Lyα absorptie er in de buurt van sterrenstelsels op $z = 2.5$ er in dit soort koude stromen geproduceerd wordt. We gebruiken modellen van OWLS en bestuderen de absorptie in gas dat we selecteren op basis van zijn thermische geschiedenis, gasthalo, beweging ten opzichte van het sterrenstelsel en de kans dat het terecht zal komen in het interstellair medium op $z = 0$. We kijken ook naar de fysische eigenschappen van het Lyα absorberende gas zoals de temperatuur en dichtheid als een functie van afstand tot de sterrenstelsels in de OWLS modellen met en zonder feedback door supernova en actieve kernen.

Curriculum vitae

O LIVERA was born in 1981 in Zrenjanin, in former Yugoslavia, present day Serbia. She attended Dr Boško Vrebalov primary school in Melenci. In Zrenjaninska Gimnazija high-school she followed the natural sciences and mathematics track. Her first contact with the scientific method and contemporary science was in Petnica Science Center - the non-governmental institution for high-school students interested in science, where she participated in both the chemistry (1998) and astronomy program (1999-2000). After finishing high-school she continued visiting Petnica as a junior assistant.

In 2000 Olivera started her bachelor studies at the astrophysics department of the Belgrade University. During her undergraduate years she visited Instituto de Astrofísica de Canarias in 2003 and Durham University in 2004 as a summer student. She was also awarded several fellowships in Serbia, among which The Scholarship of Republic Foundation for Development of Youth in Science and Arts, and The Scholarship of Royal Norwegian Embassy in Belgrade. After finishing her bachelor studies in 2005, she pursued MSc studies in astronomy at Leiden University, supported by the Huygens and Oort scholarships.

In 2007 she started her PhD studies under the supervision of Prof. Dr. Joop Schaye and Prof. Dr. Charles Steidel (Caltech), also at Leiden University. Her PhD involved observing at the Keck Observatory in Hawaii, and Palomar Observatory in California, frequent working visits to California Institute of Technology, and presenting work at international conferences and major research institutions.

In November 2011 she started her postdoc at MPIA in Heidelberg.

Publications

- **Olivera Rakic**, Joop Schaye, Charles C. Steidel, & Gwen C. Rudie, *Neutral hydrogen optical depth near star-forming galaxies at $z \sim 2.4$ in the Keck Baryonic Structure Survey*, 2011, submitted to ApJ, arXiv:1109.4944

- **Olivera Rakic**, Joop Schaye, Charles C. Steidel, & Gwen C. Rudie, *Calibrating Galaxy Redshifts Using Absorption by the Surrounding Intergalactic Medium*, 2011, MNRAS, 414, 3265

- Ryan Cooke, Max Pettini, Charles C. Steidel, King J. Lindsay, Gwen C. Rudie, & **Olivera Rakic**, *A newly discovered DLA and associated Ly-α emission in the spectrum of the gravitationally lensed quasar UM 673*, 2010, MNRAS, 409, 679

- Charles C. Steidel, Dawn K. Erb, Alice E. Shapley, Max Pettini, Naveen Reddy, Milan Bogosavljević, Gwen C. Rudie, & **Olivera Rakic**, *The Structure and Kinematics of the Circum-Galactic Medium from Far-UV Spectra of $z \sim 2 - 3$ Galaxies*, 2010, ApJ, 717, 289

- Charles C. Steidel, Christopher Martin, Xavier J. Prochaska, Max Pettini, Joop Schaye, & **Olivera Rakic**, *Mapping the 'Cosmic Web' During the Peak Epoch of Galaxy Formation*, Astro2010: The Astronomy and Astrophysics Decadal Survey, Science White Papers, no. 286