

THE UNIVERSITY OF CHICAGO

MULTIPHASE GASEOUS HALOS AROUND GALAXIES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF ASTRONOMY AND ASTROPHYSICS

BY
CAMERON JIA LIANG

CHICAGO, ILLINOIS

AUGUST 2018

Copyright © 2018 by Cameron Jia Liang
All Rights Reserved

In memory of my grandfather, 梁邻长 (1930-2018).

To my parents, who have always believed in me,

And my family and friends,

Without whom this work would not have been possible ...

TABLE OF CONTENTS

LIST OF FIGURES	vii
LIST OF TABLES	xx
ACKNOWLEDGMENTS	xxi
PREFACE	xxiii
ABSTRACT	xxiv
 1 A BRIEF HISTORY OF THE GASEOUS HALOS AROUND GALAXIES	 1
1.1 The Conceptualization of the Gaseous Halos	1
1.2 Early Evidence of Gaseous Halos	2
1.3 Early Models of Galaxy Formation and Gaseous Halos	5
1.4 Galactic Winds	6
1.5 Thermal Instabilities	11
1.6 A Modern Renaissance: the Circumgalactic Medium	13
1.7 The Roadmap of the Dissertation	15
 2 MINING THE CIRCUMGALACTIC BARYONS IN THE LOW REDSHIFT UNIVERSE	 18
2.1 Introduction	18
2.2 A Public Sample of Galaxy–QSO Pairs	23
2.3 QSO UV Spectroscopy	30
2.4 Analysis	34
2.4.1 Absorption Constraints of Circumgalactic Baryons	34
2.4.2 Absorption Properties of Individual Galactic Halos	38
2.4.3 Mass Scaling Of Extended Gaseous Halos	43
2.4.4 Mean Absorption Profiles from Stacked CGM Spectra	48
2.4.5 Gas Covering Fraction	51
2.5 Discussion	53
2.5.1 Observed Absence of Heavy Elements at $d \gtrsim 0.7 R_h$ from Star-forming Regions	54
2.5.2 The CGM at Low and High Redshifts	58
2.5.3 Uncertainties due to Pseudo-evolution of Halo Radius R_h	60
2.6 Summary	63
 3 COLUMN DENSITY PROFILES OF THE MULTIPHASE GASEOUS HALOS .	 74
3.1 Introduction	74
3.2 Galaxy Formation Simulations	76
3.2.1 Star Formation	77
3.2.2 Stellar Feedback	78
3.2.3 Cosmic Ray Feedback	79

3.3	CGM Analysis	81
3.3.1	Lines of sight and ion column densities	81
3.3.2	Curve of Growth Analysis	82
3.4	CGM in Simulations	84
3.4.1	Main features of the CGM in simulations	88
3.4.2	Cosmic Ray Feedback on the CGM	88
3.4.3	Evolution of the CGM in runs with different feedback models	90
3.4.4	Evolution of the radial distribution of the CGM column density and rescaling with halo scale radius	92
3.4.5	Exponential form of the column density profile and scaling with ionization energy of ions	95
3.5	Radial column density distributions: simulations vs. observations	100
3.5.1	Rescaling observations and simulations for fair comparison	103
3.5.2	Comparison of the column density profiles in simulations and observations	106
3.5.3	Comparing the CGM Covering Fractions	111
3.6	Discussion	114
3.6.1	Column density profiles of the CGM absorbers: is there an ion boundary of the CGM?	115
3.6.2	Column density profiles of the CGM absorbers: scaling with halo mass and redshift	118
3.6.3	The Effects of Self-Shielding, Non-Equilibrium and Local Starburst Radiation	120
3.6.4	Comparisons with Previous Studies	122
3.7	Conclusions	125
4	OBSERVING THE CIRCUMGALACTIC MEDIUM OF SIMULATED GALAXIES THROUGH SYNTHETIC ABSORPTION SPECTRA	128
4.1	Introduction	128
4.2	Simulations and Synthetic Absorption Spectrum	130
4.2.1	Simulations, star formation, and feedback models	130
4.2.2	Synthetic absorption spectra of the simulated CGM	132
4.2.3	Muti-phase gas viewed from simulations and synthetic spectra	134
4.2.4	Relative contribution of photoionization and collisional ionization	138
4.3	A Bayesian Approach for Extracting Physical Properties of the CGM	142
4.3.1	Bayesian Voigt Profile Fitting	149
4.3.2	Bayesian Ionization Modelling	151
4.4	Conclusions	156
5	ON THE MODEL OF CIRCUMGALACTIC MIST: THE IMPLICATIONS OF CLOUD SIZES IN GALACTIC WINDS AND HALOS	159
5.1	Introduction	159
5.2	Semi-Analytic Consideration	163
5.2.1	Fragmentation in thermal instability	164
5.2.2	Isobaric condensation mode in thermal instability	170

5.3	Simulations	172
5.3.1	Physics and initial conditions in simulations	174
5.3.2	The evolution of cold gas in simulations	177
5.3.3	Column density of cloudlet comparisons in simulations	183
5.4	Implications of cloud sizes in the Circumgalactic context	187
5.4.1	The circumgalactic mist in galactic halos	188
5.4.2	Spatial extents of low, intermediate and high ions from the galactic profile of cloud sizes	191
5.5	Discussions and Conclusions	194
6	SUMMARY AND CONCLUDING REMARKS	198
6.1	Summary of results	198
6.2	Future outlook	199
A	APPENDIX	200
A.1	Introduction	200
A.2	The Package: bayesvp	201
A.2.1	Posterior distribution and sampling	201
A.2.2	Voigt profile model	203
A.2.3	Continuum model	204
A.2.4	Convergence criteria	206
A.2.5	Model comparison	206
A.2.6	Uniform treatment for upper limits, lower limits and detections . . .	210
A.3	Summary	211

LIST OF FIGURES

1.1	This figures show an early example of the association of cold gas and galaxies using background QSOs and galaxies at intermediate redshifts (Steidel, 1993). The upper left panel shows Mg II absorption doublet at redshift similar to the galaxies at $z \approx 0.66$. The right panel is the spectrum of the host galaxies with the 4000Å break near 6640Å, giving a systemic redshift of 0.66. The bottom panels show images of the QSO-galaxy field in two different filter bands (R and K).	3
1.2	A map of the EAGLE simulations (Schaye et al., 2015) in a $100 \times 100 \times 20$ co-moving Mpc box. The pullout images shows a fantastically large dynamical range based using modern computers and improved algorithms. The intensity represents the gas density while the color represents the gas temperature (blue through green through red indicate cooler to hotter). The zoom-in shows a galaxy with stellar mass $3 \times 10^{10} M_{\odot}$ showing the optical band stellar light.	7
1.3	This is a composite image of the nearby galaxy M82, showing the starburst-driven multiphase galactic outflows. The blue color shows X-ray observations by the Chandra Space Telescope; the green and orange show optical emission by 10^4 K gas; the red shows infrared data from the Spitzer Space Telescope. The pullout image shows the central region with extended X-ray emissions, in addition to the two labeled X-ray point sources.	9
1.4	This figure shows the locations of ions in the phase diagram where its ionization fraction are maximized in a combination of collisional ionization (temperature dependence) and photoionization (density dependence; Tumlinson et al., 2017). The square labels of ions are colored according to the ionization states. The lines on the top right panel indicate the redshifts at which the lines are observable in either ground-based or space-based telescopes.	13
2.1	Distribution of projected distance versus redshift for the 195 galaxies (§ 2). The top (right) panel displays the redshift (impact parameter) histograms.	25
2.2	Observed correlation between stellar mass M_{star} and rest-frame r -band absolute magnitude, $M_{\text{SDSS},r}$ for 111 overlapping galaxies that are both in our pair sample and in the NASA–Sloan Atlas (dark solid points). The solid line indicates the best-fit model with $\log M_{\text{star}} = 0.14 - 0.49 M_r$ and a $1\text{-}\sigma$ scatter of $\Delta \log M_{\text{star}} = 0.21$ (error bar in the lower-right corner). The full NASA–Sloan catalog, which has $> 100\text{k}$ galaxies (grey points), is also included to contrast these 111 overlapping galaxies. The full sample displays a consistent mean relation (dashed line) and scatter between M_{star} and M_r . The consistent distribution of the two samples strongly supports that the mean relation derived using the subsample of 111 galaxies is representative of the general galaxy population at low redshifts.	27

2.3	Stellar mass distribution of our galaxy sample (blue shaded histogram). The median mass is $\langle M_{\text{star}} \rangle = 5 \times 10^9 M_{\odot}$, in comparison to the characteristic mass $M_{\text{star}}^* = 5 \times 10^{10} M_{\odot}$ for the field galaxies (Baldry et al., 2012; Muzzin et al., 2013). Our galaxy sample has a mass distribution that encompasses previous CGM studies at both low and high redshifts. Specifically, the COS-Halos sample at $z = 0.1 - 0.4$ (Werk et al., 2013, red open histogram) has a median stellar mass of $4 \times 10^{10} M_{\odot}$ and the starburst sample at $z = 2.2$ from Reddy et al. (2012; green open histogram) has a median stellar mass of $7.9 \times 10^9 M_{\odot}$. The number counts of the Reddy et al. sample has been reduced by a factor of 25 for presentation purpose.	29
2.4	Distribution of specific star formation rate (sSFR) and stellar mass based on a subsample of our galaxies with available UV photometric from the NASA-Sloan Atlas (solid points). SFR is derived based on the observed NUV flux from GALEX and therefore represents a lower limit to the intrinsic value. The blue (red) line marks the star-forming main-sequence (passive, red) galaxies from Schiminovich et al. (2010). The tilt in the distribution of our galaxies relative to the mean relation of star-forming main-sequence galaxies can be explained by more massive galaxies showing higher degrees of dust extinction (e.g., Zahid et al., 2013). Nevertheless, the majority of our galaxies are consistent with being star-forming main-sequence galaxies, and a small fraction are found in the red-sequence regime.	31
2.5	Examples of galaxies with associated CIV absorbers from our search. For each field, we present the QSO spectrum in the left panel and the corresponding SDSS image centered at the galaxy in the right panel. In each panel, the direction to the QSO sightline is indicated by the arrow, and the horizontal bar in the upper-left corner indicates an angular scale of $5''$. The projected separation between the QSO sightline and the galaxy is shown in the lower-right corner. Absorption features due to Si III $\lambda 1206$, Ly α $\lambda 1215$, and CIV $\lambda\lambda 1548, 1550$ at the redshifts of the galaxies are indicated by blue, dotted lines in the left panels. The absorber redshift is shown in the lower right corner of each spectrum.	32
2.6	Relative velocity distributions of Ly α (<i>left</i>) and CIV (<i>right</i>) absorbers with respect to the systemic redshift of their associated galaxies in our pair sample. We detect associated Ly α absorbers in 125 of the 195 galaxies searched, and associated CIV absorbing gas in nine galaxies. The left panel shows that the velocity distributions of all Ly α absorbers detected around galaxies at $d < 500$ kpc (shaded histogram), around galaxies at $d < 250$ kpc (green open histogram), and galaxies at $d > 250$ kpc. Considering the full sample requires a double Gaussian profile to characterize the velocity distribution with a narrow component centered at $\langle v(\text{Ly}\alpha) \rangle = -7 \text{ km s}^{-1}$ and dispersion of $\sigma_v(\text{Ly}\alpha) = 32 \text{ km s}^{-1}$ and a broad component centered at $\langle v(\text{Ly}\alpha) \rangle = -8 \text{ km s}^{-1}$ and dispersion of $\sigma_v(\text{Ly}\alpha) = 188 \text{ km s}^{-1}$ (the solid curve). In contrast, the velocity distribution of CIV absorbing gas around galaxies can be characterized by a single Gaussian distribution of $\langle v(\text{CIV}) \rangle = 9 \text{ km s}^{-1}$ and $\sigma_v(\text{CIV}) = 58 \text{ km s}^{-1}$, excluding the outlier at $v(\text{CIV}) \approx +260 \text{ km s}^{-1}$	35

2.7	Observed spatial distribution of rest-frame absorption equivalent widths of Ly α (top), CII (middle), and CIV (bottom) in the circumgalactic space. The left panels show the variation of absorber strengths versus projected distance d , while the right panels show the spatial variation as a function of halo-radius (R_h) normalized projected distance, d/R_h . Galaxies with no detectable absorption are shown as 2- σ upper limits (points with downward arrows) that are determined based on the noise characteristics (see § 4.1). Measurements from our study are shown in solid points. Measurements for $\sim L_*$ galaxies at $z = 0.14 - 0.36$ from the COS-Halos program (Werk et al., 2013) are shown in open squares. For starburst galaxies at $z = 2.2$ (Steidel et al., 2010) constraints on Ly α and CII absorption are shown in star symbols. But because CIV doublets were not resolved in this study, a range of $W_r(1548)$ is shown in vertical bars based on assumed doublet ratios of 2:1 and 1:1. The measurements at $z = 2.2$ are from stacked spectra of galaxies covering a range in stellar mass (Reddy et al., 2012). Previous measurements of CIV absorption in the low-redshift CGM from Chen et al. (2001b) are shown in triangles. The best-fit power-law model of Chen et al. (1998, 2001a) to describe the observed $W_r(1215)$ versus d anti-correlation is shown in solid line. Finally, mean absorption strengths determined from stacked CGM spectra of the full sample (§ 4.4) are shown in (blue) pentagons with the horizontal bars indicating the bin size.	41
2.8	Similar to Figure 6 but for the observed spatial distribution of rest-frame absorption equivalent widths of Si II (top), Si III (middle), and Si IV (bottom) in the circumgalactic space. Note that the constraints for non-detections in the COS-Halos sample appear to be worse than our typical upper limits, because the COS-Halos team adopted a much larger velocity window for measuring their upper limits (i.e. $\approx 300 \text{ km s}^{-1}$ adopted for the COS-Halos galaxies versus $\approx 50 \text{ km s}^{-1}$ adopted in our study. The velocity window adopted in our study is empirically determined based on detected components, and is therefore more appropriate for constraining the presence/absence of cool, photo-ionized clouds in the circumgalactic space.	42
2.9	Stacked QSOs spectra in different R_h -normalized projected distance intervals from host galaxies. Galaxies at $d < 1.7 R_h$ are grouped into three bins following the binning of the high-redshift starburst sample. Galaxies at $d > 1.7 R_h$ are divided into two d/R_h bins with roughly equal number of galaxies. The median R_h -normalized projected distance and the number of galaxies included in each stack are indicated in each panel. The stacked spectra clearly show declining Ly α absorption strengths with increasing projected distance, and no ionic transitions are detected with $d/R_h > 0.7$	45
2.10	Mean absorption profiles of Ly α (left) and CIV (right) in stacked absorption spectra of galaxies in different d/R_h intervals.	48

- 2.11 Mean gas covering fraction $\langle \kappa \rangle$ measured in each d/R_h bin for different species. H I is shown in the left panel, Si II, Si III, and Si IV are shown in the middle panel, and C II and C IV are shown in the right panel. Gas covering fraction is computed for a detection threshold of $W_0 = 0.05 \text{ \AA}$. Error bars represent the 68% confidence interval. The covering fraction of Ly α absorbing gas remains high at $\approx 60\%$ outside of dark matter halos at $d > R_h$, while the covering fraction of heavy ions is restricted to $\langle \kappa \rangle < 10\%$ in individual d/R_h bins. Considering *all* galaxies at $d > R_h$ further improves the limit on the covering fraction of heavy ions to $\langle \kappa \rangle < 2\%$ at a 95% confidence level. In inner halos ($d \lesssim 0.7 R_h$), we also find a declining mean covering fraction from high- to low-ionization species. . . . 51
- 2.12 R_h -normalized radial profiles of C IV absorbing gas for low-mass galaxies of $M_{\text{star}} < 0.1 M_{\text{star}}^*$ (filled and open squares) and high-mass galaxies of $M_{\text{star}} > 0.1 M_{\text{star}}^*$ (filled and open circles), where $M_{\text{star}}^* = 5 \times 10^{10} M_\odot$ (Baldry et al., 2012; Muzzin et al., 2013). Constraints from stacked spectra of the full sample are included for reference (pentagon points). Both low- and high-mass galaxies exhibit a modest mean covering fraction of strong ($W_r(1548) > 0.1 \text{ \AA}$) C IV absorbing gas $\langle \kappa \rangle \approx 30 - 50\%$ at $d < 0.7 R_h$, and no associated C IV absorbers at $d > R_h$. No significant distinction, beyond R_h -normalization, can be made in the extent of chemical enrichment around low- and high-mass galaxies. In contrast, six of the 17 galaxies in the Borthakur et al. (2013) sample exhibit strong C IV absorption out to $\sim R_h$ (star symbols). While the constraints for the 11 non-detections are weak, the detected C IV absorbers around six galaxies (four starburst and two passive galaxies) are all significantly stronger than the mean values observed in our sample. 56
- 2.13 C IV radial absorption profile normalized by the scale radius r_s of the host dark matter halos (*left*) or by M_{star} of the absorbing galaxies (*right*). The scale radius r_s represents a characteristic radius where the dark matter density profile resembles an r^{-2} profile (Navarro et al., 1997). We also include in the left panel measurements from stacked spectra using the full galaxy sample (blue pentagons). In addition, the scaling coefficient of M_{star} is not a best fit to the C IV data, but adopted from the best-fit coefficient of Mg II absorbers from Chen et al. (2010b). Nonetheless, the result already reveals a clear C IV enriched halo out to M_{star} -normalized projected distance of $\approx 160 \text{ kpc}$ (consistent with the finding of L_B -normalized C IV profile from Chen 2012). In both panels, the C IV radial absorption profiles at $\langle z \rangle = 0.041$ (solid points) and at $\langle z \rangle = 2.2$ (green crosses) appear to be better aligned than what is seen in the bottom right panel of Figure 7 & 8. 61
- 3.1 Star formation histories of the main galaxy in our simulations compared to the average histories for observed galaxies in halos of $M_{\text{vir}}(z=0) = 10^{12} M_\odot$ derived by Behroozi et al. (2013, with dark and light gray shaded bands showing the one- two-sigma confidence regions of their constraints, respectively). The SFHs in simulations are averaged over time bins of $\Delta t_{\text{SF}} = 100 \text{ Myr}$ 83

3.2	Maps of the total gas column density, metallicity, and mass-weighted temperature of the fiducial and $5E_{\text{SNII}}$ models at $z = 0$. The maps show the outflow streams aided by the boosted supernova events while the fiducial run maintains a normal star-forming disk. As a consequence, gas metallicity is well-mixed in $5E_{\text{SNII}}$ model out to large radii at 10% solar to solar value. In contrast, despite the disk in the fiducial run having solar metallicity, the gas hovers at a few percent solar in the circumgalactic space. Note that the color of the star-forming disk in the fiducial run has been clipped in order to ensure the dynamic range of the color map in the halo.	85
3.3	The spatial column density distribution (500 kpc across; $R_{\text{vir}} \approx 250$ kpc) of the chemically enriched CGM as probed by HI and MgII at $z = 0.0$. Solid circle represents R_{vir} , dashed circle is $4r_s$ and dotted circle is r_s . Our fiducial model with star-formation efficiency $\epsilon_{\text{ff}} = 10\%$ produces a stellar component consistent with observations (Agertz & Kravtsov, 2015, 2016) but fails to produce an extended multi-phase gaseous halo. On the other hand, our strong feedback model (with five times supernova energy, $5 \times E_{\text{SNII}}$) matches the CGM profiles despite destroying the star-forming disk. This shows that the CGM provides sensitive orthogonal constraints on galaxy formation, especially feedback recipes.	86
3.4	Similar to Figure 3.3, but for intermediate ion CIV and high ion OVI at $z = 0.0$. The multi-phase gaseous halos are progressively more extended from low and intermediate ions (e.g., MgII and CIV) to high ion (OVI).	87
3.5	Model predictions of MgII column density maps for different redshifts. The region is 250 kpc (physical) on a side. Solid circles represents R_{vir} (it is outside the region at $z = 0$) and dashed circle represents $4r_s$. At high redshifts, $4r_s$ tracks the growth of R_{vir} , but at $z < 2$ R_{vir} becomes larger and grows faster than $4r_s$. For models that match observations of the CGM (ALL_Efb_e001_5ESN and ALL_e010_CR) the physical extent of the CGM is approximately a constant fraction of r_s at all z	89
3.6	Column density profiles for four representative low to high ions for our models and their evolution from $z = 3.0$ to $z = 1.0$. At high redshift ($z = 2 - 3$), the CGM seems to be insensitive to the feedback recipes except for the very weak star formation efficiency $\epsilon_{\text{ff}} = 1\%$. At $z = 1$, the fiducial model starts to deviate from $5E_{\text{SNII}}$ and cosmic ray models. The extent of the profiles in units of R_{vir} shrinks with redshifts mainly due to the fast-evolving R_{vir} relative to the physical extent.	91

3.7	Column density profiles from low (H I and Mg II) to high (C IV and O VI) ions in the ALL_Efb_e001_5ESN run. Various lines are medians of the profiles at different redshifts (blue to red as low z to high z). Dark and light blue shaded regions represent 68% and 95% intervals for the profiles at $z = 0$, respectively. The left column shows the column density profiles in the inner regions remained relatively unchanged over 11 billion years at $z \leq 3$. This is why scaling with slower evolving scale radius, r_s (right panel) brings profiles at different z into agreement better than scaling with R_{vir} (middle column). At the same time, scaling with r_s also accounts for the different masses of the galaxies, which is necessary to put a population of observed galaxies in an equal footing. Note that higher ionization ions have progressively more extended profiles.	93
3.8	Evolution of the virial radius and halo scale radius, r_s . Virial radius is measured from the dark matter halo at each snapshot of the simulation ALL_Efb_e001_5ESN. The corresponding scale radius $r_s \equiv R_{\text{vir}}/c_{\text{vir}}$ is calculated via the $M_{\text{vir}} - c_{\text{vir}}$ model in Diemer & Kravtsov (2015). At high redshifts, the inner halo characterized by $4r_s$ tracks the growth of the halo virial radius, R_{vir} , until $z \sim 2$. Afterwards, R_{vir} continues to grow, while the inner region of the dark matter halo approaches a constant value.	96
3.9	Scale height of the exponential profile of the column density as a function of the impact parameter predicted in our run ALL_Efb_e001_5ESN at $z = 0$ as a function of ionization energy of the corresponding ion. The left axis shows the scale height normalized to the halo scale radius, while the right axis shows the scale height normalized to the halo virial radius. Color points show results of the fits of the column density profile of a particular ion, while the dashed lines show the best fit power law relations (see eqs. 3.8-3.9).	98
3.10	Distribution of halo mass M_{vir} probed by various ions in the compiled low redshift data set. See references in Table 3.3.	101
3.11	Equivalent width distribution with impact parameter for the low- z galaxies observed by the COS instrument onboard <i>HST</i> (see Table 3.3 for description of observational samples and references) and $z \approx 2$ stacked measurements by Steidel et al. (2010). Blue pentagons are detections, while gray pentagons with arrows are upper limits. The color 2D histogram shows distribution of equivalent widths for the lines of sight through the halo of the MW-sized galaxy in the ALL_Efb_e001_5ESN run, assuming turbulent broadening of $b_{\text{NT}} = 20 \text{ km s}^{-1}$. Black dashed lines represent W_r profiles converted from the best fit exponential profiles to the simulated column density profiles for corresponding ions. The left panel shows profiles against the impact parameter d in physical kpc, while in the middle and right panels d is rescaled by the virial radius, R_{vir} , and halo scale radius, r_s , respectively.	104
3.12	Equivalent width profiles for three representative ions in the run with cosmic ray feedback at $z = 1$ as a function of impact parameter scaled by the scale radius, d/r_s (2D color histogram) compared to observations (points; the point types are the same as in Figure 3.11).	105

3.13	Comparison between predicted radial column density distributions of various ion absorbers (2D histograms) with observations at $z \approx 0 - 0.3$ (pentagon points; see Table 3.3). Blue points show detections, while gray points show non-detections with downward arrows denoting 2σ upper limits. The ion for which the profile is constructed is indicated in the upper right. The left column shows profiles predicted by the fiducial model with all feedback + $E_{\text{fb}}, \epsilon_{\text{ff}} = 10\%$. The right column shows the predicted profiles from the model with all feedback + $E_{\text{fb}}, \epsilon_{\text{ff}} = 1\%, 5E_{\text{SNII}}$. The 2D histograms show that the majority of the points in ALL_Efb_e010 are many orders of magnitude below observations, while profiles in the ALL_Efb_e001_5ESN run are closer to observations.	107
3.14	Similar to Figure 3.13, but for Si II, Si III and Si IV.	108
3.15	Similar to Figure 3.13, but for Mg II, O VI and Ne VIII.	109
3.16	Comparison of the predicted covering fraction profiles with measurements of Mg II $\lambda 2796$ (Chen et al., 2010a), C IV $\lambda 1548$ (Liang & Chen, 2014; Bordoloi et al., 2014) and O VI $\lambda 1031$ (Tumlinson et al., 2011; Johnson et al., 2015b) transitions. Detection threshold $W_0(\text{CIV}\lambda 1548) = 0.05 \text{ \AA}$ is adopted from Liang & Chen (2014). We convert the threshold to column density using a curve-of-growth analysis adopting the doppler parameter $b = \sqrt{2}\sigma = 29 \text{ km s}^{-1}$. Detection threshold $\log N_0(\text{O VI } \lambda 1031) = 13.5 \text{ cm}^{-2}$ is adopted from Johnson et al. (2015b). $\log N_0(\text{Mg II } \lambda 2796) = 13 \text{ cm}^{-2}$ is chosen based on the level upper limits in Chen et al. (2010a). Error band represents 68% confidence level bracketed by two limiting cases of consideration of upper limits above the detection threshold (e.g., as detection and as non-detection). In all three ions, ALL_Efb_e001_5ESN and ALL_e010_CR predictions agree with the inner region ($< 2 - 4r_s$) of the data while under-predicting the covering fraction of the outskirts of the halo ($\sim 10r_s$). Note that ALL_e010_CR is at $z = 1$ while the other runs are at $z = 0$. It is conceivable that halo in the CR run will continue to grow at lower redshift, as suggested by Fig. 3.5	113
3.17	Column density profiles of Mg II, C IV and O VI in the <i>outer region</i> of the hot gaseous halo share the same underlying exponential profile traced by the total gas column density profile. All other ions (not shown) also show similar exponential profiles. Column densities for ions have been offset by $\log N_0 = 6$	117
3.18	Comparisons of predicted column density profiles from low to high ions between our simulations with the profiles from simulations of Hummels et al. (2013, denoted as H+13). The profiles in our fiducial run are closest to the medium feedback run of H+13.	121
4.1	Example of a spectrum along a line of sight (LOS) sampling the properties of the CGM around the simulated galaxy. The LOS (white line) through the density (color map) and velocity (vector) field are shown in the right panel, while the left panels show the one-dimensional profiles of the density, temperature, metallicity, and velocity (relative to the galaxy systemic velocity) of the gas the LOS passes through.	132

- 4.2 One-dimensional profiles of column density, temperature, and velocity as a function of distance from the point of closest approach (minimum distance) to the galaxy for two types of lines of sight (LOS). Note that the top plots of both panels show column density contributed by each pixel i at the location dS_i , i.e., $N_i = n_i dS_i$. The total column density is integrated along the LOS, $N = \sum N_i$. For clarity, various lines representing different ions have been artificially shifted by the value shown in the legends. The left column shows the multiphase structure of the absorbing gas in the LOS shown in Figure 4.1. This LOS passes through a density peak with a large gradient in gas density and temperature. High ions (e.g., NV, O VI, O VII and Ne VIII) preferentially reside on the boundary of the density peak while low ions reside in the core where the temperature is low ($\sim 10^4$ K). The right panel shows the LOS that passes through the hot and diffuse coronal halo gas near the virial temperature of the halo ($\sim 10^{5.5}$ K) without any dense gas. The abundance of low ions is therefore small. The temperature and velocity profiles (middle and lower columns) determine the line width of absorbers in the absorption spectra shown in Figure 4.3. 135
- 4.3 The synthetic spectra generated from the two LOS shown in Figure 4.2. Left panel: the LOS that passes through the density peak exhibits strong absorption lines of low ions, as well as high ions (e.g., OVI), reflecting the multiphase structure of the gas along this LOS. Note that a single density peak along this LOS generates three (or four) absorption components due to the strong velocity gradients. However, the low and high ions are approximately kinematically aligned per absorption component. However, the large column density of HI and oscillator strength of Ly α lead to a blending of all three absorption components, making the determination of physical parameters difficult, although such blending does not always occur. Higher order Lyman series lines are needed to separate the absorption components. Right panel: the LOS that passes through the diffuse hot halo exhibits only HI and OVI absorption since low ions cannot survive in the physical condition near the virial temperature of the halo ($\sim 10^{5.5}$ K). We also find detectable O VII lines ($\lambda_{\text{rest}} = 21.6\text{\AA}$) in both cases in our simulations with $\log N \sim 14 \text{ cm}^{-2}$ 136

- 4.4 This figure shows the same LOS as the left panel in Figure 4.3 but with variations of ionization models. This shows how ionization models affect the total column densities of low, intermediate and high ions and their spectral features. The blue spectrum is computed using an ionization model with photoionization turned off (i.e collisional ionization only). The green spectrum assumes the gas is optically thin with 100% of the UVB contribution. Lastly, the black curve takes into account self-absorption of the CGM gas using an approximation for the fraction f_{UVB} that reaches the gas as a function of the local density from Rahmati et al. (2013). Comparing the model with collisional ionization only ($f_{\text{UVB}} = 0$), the optically thin photoionization model ($f_{\text{UVB}} = 1$) extinguishes a large fraction of low ions (e.g., H I, O I, Mg II and Fe II; blue to green), which results in the change of shape of the absorption profile. Absorption lines of O VI, however, change relatively little (green vs blue profiles), which shows that they are primarily collisionally ionized. Taking into account CGM self-absorption (black) generally increases the abundance of low ions compared to the optically thin model (green). 139
- 4.5 Column density profiles of H I, Mg II, C IV and O VI computed for the different treatment of ionization radiation (see legend and caption of Figure 4.4). The comparison of different ionization models show that photo-ionization from the UVB extinguishes an abundant of low and intermediate ions at distances beyond $\approx 0.2R_{\text{vir}} \approx 65$ kpc. On the other hand, photo-ionization does not have a strong effect on the column density of O VI, showing that it is primarily collisionally ionized. The model that takes into account CGM self-absorption generally increases the abundance of low ions. 140
- 4.6 Each row shows synthetic spectrum (with continuum $C(\Delta v) = 1$) around a particular absorption line derived from an LOS through simulation along with the best fit model Voigt profile resulting from our Bayesian fitting (left panel) and the derived posterior probability distribution of the column density of the corresponding. The true column densities are indicated in the right panel by the green vertical dashed lines. We have selected five representative examples of absorption lines that show different degrees to which the lines were detected: from a non-detection (top row) to high signal-to-noise detection (two bottom rows). The same saturated H I line is repeated three times in the middle panels with varying S/N ($= 5, 10, 20$) showing that the resulting fit can be non-Gaussian in low S/N spectra due to the degeneracy between N and b -parameter. For the non-detection/upper limit constraints in C II $\lambda 1334$, the lack of decrement in the flux implies that column density of C II greater than $\sim 10^{13} \text{ cm}^{-2}$ is disfavored by the data. In contrast, the Si IV is well measured with a narrow Gaussian column density PDF. These probability distributions can then be used for ionization modelling even though C II line is not formally “detected” or that the H I line is saturated. 143

- 4.7 These four corner plots show the joint posterior distributions corresponding to absorption lines in Figure 4.6. The top left panel shows the posteriors of the non-detected C II line. The joint probability distributions show noisy fluctuation without a preferred location in the parameter space. The marginalized probability of column density drops precipitously for $\log N/\text{cm}^{-2} > 12$, where high column density is strongly disfavored by the data. The other three panels show the posteriors of the saturated H I line with corresponding S/N of 5, 10 and 20. The elongated shape in the joint probability shows the degeneracy between b and N in a saturated regime. As expected, the constraints improve as S/N of the spectra increases. The corner plots are made using publicly available code `corner.py` (Foreman-Mackey, 2016). 144
- 4.8 Comparison between derived properties and the “true” H I mass-weighted averages of each quantity along corresponding LOS in simulations. Experiment A (green points) represent values derived using the integrated column density along the entire LOS, which may include absorption from multiple absorbers. The column densities in experiment A have uncertainties of 0.1 dex. Generally, the derived values are close to the H I-weighted properties along the LOS. Experiment B (not shown) shows similar results as A . The gray bands show the typical intrinsic scatter along these LOS. The pink points in experiment C show how well one can retrieve the physical properties of absorbers when column densities constraints are poor (see the upper limit constraint of C II in Fig 4.6 for an example). The pink error bars show the ranges that enclose 68% and the gray error bars enclose 99% of probability of the posterior distributions. Naturally, the constraints of the derived parameters are poorer due to the lack of constraining power from upper limits PDF. Nevertheless, this demonstrates a bias-free, formally correct way to retrieve physical properties of the CGM when there’s limited information available. 145
- 4.9 Experiment B: comparison of how well ionization modelling can retrieve the H I density weighted physical properties by including a set of ions with increasing ionization energy. Each line connected by color points represent a single LOS. The color points at a given E_{ion} means that the ionization modelling includes ions *up to* the specified E_{ion} . For example, points at $E_{\text{ion}} = 16.345$ eV means H I ($E_{\text{ion}} = 13.598$ eV), O I (13.618), Mg II (15.035) and Si II (16.345) are included in the ionization modelling. The typical error bars are larger with a low number of ions (not shown), and all derived properties are roughly consistent with the input values. Surprisingly, the derived physical properties approach the H I mass-weighted values by including higher ions. 146
- 4.10 An example of posterior PDFs of the parameters in experiment B, in which all low to high ions with $E_{\text{ion}} \leq 140$ eV (up to O VI, with $\log N_{\text{OVI}}/\text{cm}^{-2} = 14.06$) are included. As shown, the derived properties are fairly consistent with the underlying H I-weighted values. 147

4.11	MCMC sampled posterior distributions of best-fit parameters of an ionization model (including photo- and collisional ionization) derived from constraints of column density in the form of non-Gaussian probability distributions including “non-detected” metal lines. In this example, column density constraints come from H I, O VI, N V and Ne VIII, where Ne VIII is a “non-detection” similar to C II constraints in Figure 4.6. The corresponding input mock spectrum and derived column density PDF is shown top right. Although with the poor column density constraints with Ne VIII, this figure shows one can formally recover the true parameters (shown by blue crosses) in this mock observation to the degree that the data allow. If N V becomes an upper limit when S/N of the spectrum is poorer, the corresponding constraints of temperature and metallicity will be poorer. Explicit priors can also be set to limit the wide range of allowed density based on our physical understanding of O VI-producing gas. Note that the probability distribution of H I column density from Voigt profile fit acts as a prior for the ionization modelling. In other words, the posterior of H I shown here is a re-sampled version of the posterior sampled from Voigt profile fitting.	148
5.1	Cooling (solid) and heating (dashed) curves used throughout this work at $Z/Z_{\odot} = 0.3$ (Gnedin & Hollon, 2012).	165
5.2	The size of cloudlets under the ansatz, $r_{\text{cl}} = c_s(T)t_{\text{cool}}(n, T, Z)$, as a function of density and temperature. The contours are constant cloudlet sizes in parsecs at $Z = 0.3Z_{\odot}$. The contours of different sizes converge to the lower boundary, which is set by the balance of heating and cooling. In other words, the constraint of the cloudlet size disappears as $t_{\text{cool}} \rightarrow \infty$. At this point, the problem becomes scale-free. Note that at the typical condition of observed warm gas ($T < 10^{4.5}\text{K}$ and $n_{\text{H}} \sim 10^{-3}\text{cm}^{-3}$), the expected cloudlet size is of the order of 1 pc or less.	166
5.3	The cloudlet sizes as function of temperature for different ambient pressures. The gradient of blue lines shows higher background pressure is able to compress cloudlets more, leading to smaller sizes. The corresponding column densities of cloudlets are represented by the series of magenta lines of different opacities. As shown, the column densities are insensitive to the change of background pressure.	168
5.4	The evolution of the cloudlet properties given the background pressure with $P/k_{\text{B}} = 100$ (left panel) and $P/k_{\text{B}} = 1000$ (right panel). As shown, cloudlet quickly cools and condenses within a cooling time of the initial state $t_{\text{cool},0}$. Upon reaching the minimum of the cooling time, the radius of the cloudlet follows $r_{\text{cl}} = c_s t_{\text{cross},\text{min}}$ as t_{cool} departs and grows as the net cooling rate approaches zero. The final size of the cloudlet depends on the external pressure as shown, where $r_{\text{cl}} \approx 0.2$ pc for $P/k_{\text{B}} = 1000$, and $r_{\text{cl}} \approx 3$ pc for $P/k_{\text{B}} = 100$	169
5.5	Initial condition of the density region with isobaric density perturbation with radius $\sim 150\text{pc}$. The region has a median temperature of $\sim 10^5\text{K}$ and is tapered by a gaussian on the edges that smoothly transitions into the background temperature at $T_h = 10^7\text{K}$ and density $n_h = 10^{-4}\text{cm}^{-3}$	173

5.6	The evolution of the density (left column) and the temperature (right column) in the fiducial simulation with cooling only ($\beta_{\text{fid}} = \infty$). As shown, the initial warm gas cools quickly to $T \sim 10^4\text{K}$. The cold gas then fragments into pc-size cloudlets through internal dynamics due to cooling and the interaction with the hot winds.	178
5.7	Temperature maps of simulations with the same initial conditions except the strength of the initial magnetic fields ranging from $\beta = 10 - \infty$. These simulations do not include thermal conduction. The left panel shows the adiabatic run where no cold clumps are present. This is in stark contrast to those with radiative cooling, where they produce similar cold clumps with $N_{\text{H}} \sim 10^{17.5}\text{cm}^{-2}$. Larger magnetic field strengths produce more filamentary structure (e.g., $\beta = 10$ in the right panel).	179
5.8	Simulations with thermal conduction at two representative time steps. The left panel ($\beta_c = \infty$) shows a run with cooling and isotropic thermal conduction. The middle and right columns show the evolution with an increased magnetic field strengths, which prolong the survival of cold gas.	180
5.9	Temperature and density distribution of all the computational cells of the cooling only simulation. The dark green line shows the initial distribution where the simulation is initialized in pressure equilibrium. The broader distribution shows the gas maintains approximate pressure equilibrium with the ambient hot gas.	184
5.10	The top panel shows column densities distribution for simulations without thermal conduction. Simulations with conduction are better shown without binning in the cumulative distribution in the bottom panel, since there is only a small number of cloudlets. As shown, the medians of these distributions are all located near $N_{\text{H}} \approx 10^{17.5}\text{cm}^{-2}$, as predicted by the model. Simulations with large initial magnetic field strengths (e.g., $\beta_c = 1 - 10$) show the cloudlet column densities are slightly higher.	185
5.11	Left panel: simulation median pressure profile of L_* like galaxy at redshift $z = 0$ with $M_{\text{vir}} = 10^{12}M_{\odot}$ (Agertz & Kravtsov, 2015; Liang et al., 2016). Right panel: corresponding allowed cloudlet sizes, $r_{\text{cl,min}} = \min(c_{\text{st}}t_{\text{cool}})$, as a function of radius based on the pressure profile. Different color curves show the dependence of metallicity. The higher the metallicity, the more efficient the gas can cool, which leads to smaller cloudlet sizes. Note that the sharp boundary of detections of low and intermediate ions at $0.5 - 0.7R_{\text{vir}}$ (Liang & Chen, 2014) corresponds to similar radii where small cloud sizes are small, in the order of 100 pc or less. In addition, the pressure profile on the left can be approximated by an exponential profile, $P/k_B \propto \exp[-(r/H_P)]$, where $r = R/R_{\text{vir}}$ and $H_P \approx 50$ kpc is pressure scaleheight. We expect the assumption of pressure equilibrium to derive the cloudlet size breaks down roughly at the size of the pressure scaleheight.	187
A.1	Marginalized posterior distributions for parameters in a single component Voigt profile fit with a linear continuum model fit to a synthetic spectrum. The top right panel shows the best fit model relative to the input spectrum over the fitted wavelength interval.	205

A.2	Evolution of the Gelman-Rubin (GR) indicator for model parameters as a function of MCMC steps. A perfect convergence of a parameter corresponds a value of one. Users may choose a level of convergence based on a threshold of GR indicator to decide the number of MCMC steps necessary to run. Note also that parameters converge at different rates.	207
A.3	This figure shows the posterior distributions Voigt profile fit of a synthetic spectrum. This shows that a 2-component model is a good fit to the data. Note also the parameters of the two Voigt profiles are correlated.	208
A.4	Model comparison based on BIC as a function of the number of Voigt components (N_{Voigt}) for the spectrum shown in Figure A.3. Note that there are three free parameters in each component. The model of 2 components with the minimum value is preferred.	209

LIST OF TABLES

2.1	Summary of Galaxy Properties ^a	66
2.2	Summary of Observations of QSOs ^a	67
2.3	Available Absorption Constraints of Different Galaxy Samples	68
2.4	Summary of Individual Halo Absorption Properties. ^{a,b}	69
2.5	Absorption Equivalent Widths Observed in Stacked CGM Spectra in Different Projected Distance Range ^a	70
2.6	Absorption Equivalent Width in Stacked CGM Spectra in different R_h -Normalized Projected Distance Range ^a	71
2.7	Mean Gas Covering Fraction, $\langle \kappa \rangle$, of Different Species ^a	72
2.8	Absorption Equivalent Width in Stacked Spectra of Ly α Absorbers at $d > 0.7 R_h$ ^a	73
3.1	Simulation suite	77
3.2	Scale Heights From Simulations	100
3.3	Observation data sets of CGM	102
5.1	Simulation suite	176

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Andrey Kravtsov, for all the wisdom, knowledge, and guidance that he provided throughout the past few years. I am especially grateful that he took me on as a student nearly halfway through graduate school. I am thankful that Andrey has been incredibly patient with me as I struggle over the years. I also want to thank Andrey for supporting me to dream of my own projects and how he has been incredibly helpful in guiding me at the same time. I consider myself extremely lucky to have such a wonderful advisor. Thank you for all the conversations about science and life in general. I have always come away with excitement after each conversation, and eager to work on whatever we talked about. Thank you!

I would also like to thank Nick Gnedin, Dan Fabrycky, and Josh Frieman, who served on my thesis committee. I am in debt for your guidance and advice. In particular, I want to thank Nick for his guidance throughout the years in group meetings and other venues. I have learned a lot from you.

I also want to thank Oscar Agertz, who I have a chance to collaborate. Thank you for this collaboration opportunity and allow me to use your simulations in two chapters of the dissertation. Without your work and guidance, this dissertation would not have been possible.

During my first three years in graduate school, I had the pleasure of working with Hsiao-Wen on observational topics of the CGM. I am grateful to Hsiao-Wen for shaping me into an expert on quasar absorption lines.

John Carlstrom, thank you for your help, guidance, and the opportunity for me to work with the South Pole Telescope team during my first year. I am grateful to the entire SPT team and the chance to participate in the exciting CMB science, in particular, the first lensing B -mode detection!

I am in debt to other faculty members in the department, especially Rich Kron and Don York. Rich, thank you for your support and guidance. I want to thank Don for questioning

me during my first candidacy exam on showing a picture of the gaseous halo. This has provided a general guidance over the years on thinking about how the circumgalactic medium may actually look like.

David Tytler, thank you for the research opportunity when I was doing my undergraduate degree. It was a pleasure to be part of the team and working with you. Thank you for fostering my interest in astronomy and science research in general. Thank you for supporting me during those years and opening doors to my next chapter.

I would like to express my gratitude to my colleagues in the astronomy and physics department. I would not have made it without all your support and friendship. Sean Johnson, Fakhri Zahedy, Yun-Hsin Huang, and Vadim Semenov have been particularly influential in shaping the work presented in this thesis. Thank you for all the discussion and sharing your wisdom and knowledge with me.

Sacha Favelevic and Ale Argerich, thank you for your hospitality. Tatiana, thank you for all the support, patience, laughter, and all the memories. I truly would not have made it without your support.

Finally, I am in debt to my family who has always trusted me and supported me unconditionally. Grandpa and grandma, thank you for taking care of me when I was younger. My aunt, Angie, thank you for bringing me to the United States. Without you, this Ph.D. and life in the U.S would not have been possible. Dad and mom, thank you for your patience and unconditional love.

The work presented in this thesis was supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program - Grant NNX15AR86H, the Brinson Foundation Predoctoral Fellowship, the McCormick Fellowship, and the William Harper Dissertation Fellowship.

PREFACE

The chapters within this thesis are based on the work presented in the following publications in the Monthly Notices of the Royal Astronomical Society (MNRAS) with the copyright agreement and permission from the Oxford University Press.

1. “Mining circumgalactic baryons in the low-redshift universe”, **Liang C. J.**, Chen H.-W., 2014, MNRAS, 445, 2061
2. “Column density profiles of multiphase gaseous haloes”, **Liang C. J.**, Kravtsov A. V., Agertz O., 2016, MNRAS, 458, 1164
3. “bayesvp: a Bayesian Voigt profile fitting package”, **Liang C. J.**, Kravtsov A. V., arXiv:1710.09852
4. “Observing the circumgalactic medium of simulated galaxies through synthetic absorption spectra”, **Liang, C. J.**, Kravtsov, A. V., Agertz, O., 2018, MNRAS, 479, 1822
5. “On the model of the circumgalactic mist: the implications of cloud sizes in galactic winds and halos”, **Liang C. J.**, Remming, I., 2018, arXiv:1806.10688, submitted for publication to MNRAS in June, 2018.

ABSTRACT

There is no doubt that our atmosphere is an integral part of the ecosystem of the Earth. Everyday weather and long-term climate of the atmosphere are directly linked to activities on the surface of the Earth and vice versa. Gaseous halos, known as the circumgalactic medium (CGM), are the equivalent atmospheres of galaxies. The cycles of birth and death of stars are intricately tied to the cycle of gaseous structures in the galactic sky. This dissertation provides a major step towards characterizing and modeling the physical properties of the CGM and the co-evolution with their host galaxies. First, we statistically map the spatial extents of the CGM in ~ 200 galaxies that span a large range of stellar mass. This mapping shows a continuous detection of neutral hydrogen out to 10 virial radii of the host galaxies while heavy elements are detected only up to $50\% - 70\% R_{\text{vir}}$. We explore the effects on the CGM profiles from different modeling of star formation and feedback processes using a set of high-resolution cosmological zoom-in simulations of a Milk-Way progenitor. This comparison illustrates that the properties of galaxies and the properties of their CGM provide strong complementary constraints on the processes governing galaxy formation. We show that the spatial profiles are self-similar across halo mass and cosmic time once it is re-normalized by the scaled radius r_s of the host dark matter halo. To connect more closely with observations, we develop a synthetic absorption pipeline, as a virtual telescope, to observe the simulated galaxy and their CGM. Finally, using semi-analytic calculations, and a set of magnetohydrodynamic simulations, we present a multiphase model of the gaseous halos around galaxies, the circumgalactic mist. The model implies a large number of cold clumps in the inner halo with a small volume filling factor but a large covering fraction. The model naturally gives rise to spatial extents and differential covering fractions of cold, warm and hot gas. To self-consistently model the co-evolution of the CGM and star formation within galaxies, future simulations must address the mismatch of the spatial resolution and characteristic scale of cold gas.

CHAPTER 1

A BRIEF HISTORY OF THE GASEOUS HALOS AROUND GALAXIES

To understand the formation and evolution of galaxies, in the broadest stroke, is to understand the processes of converting baryonic matter from one phase to another. This dissertation focuses on a critical regime, the gaseous halo around the star-forming sites, which harbors the largest fraction of baryonic matter in the Universe.

1.1 The Conceptualization of the Gaseous Halos

The story of gaseous halos around galaxies started nearly seven decades ago in the 1950s. During this time, observations of interstellar gas within the spiral arms of the Milky Way were common and well established. Gas clouds outside of the galaxy at larger distances were not. The story began when Guido Munch and Harold Zirin were conducting spectroscopic and radio observations of stars far away from the galactic plane. They found evidence of cold gas from the detections of neutral sodium (Na I) and singly ionized calcium (Ca II). Based on the kinematics of these absorption lines detected in the spectra of stars within and above the spiral arms, they concluded that these must be intervening absorbing gas originated from at least above $\sim 500 - 1000$ pc from the galactic plane¹. The naive expectation is that these dense clouds would expand in a low-pressure environment; the associated absorption will thus be slight. To explain the puzzling existence of these dense gas at great distances, Spitzer (1956) proposed that they must be pressure-confined by an interstellar galactic corona”. This corona of gas must be hot with temperature $T \approx 10^6$ K and diffuse with density $n \approx 10^{-4} \text{ cm}^{-3}$ to match the internal pressure of these cold dense clouds ($T \approx 100$ K, $n \approx 10 \text{ cm}^{-3}$). Thus, the concept of a (hot) gaseous halo was born.

1. We have now come to know these as High-Velocity Clouds (HVC) in the modern times.

1.2 Early Evidence of Gaseous Halos

Just two years after the discovery of Munch & Zirin, Schmidt (1963) and Oke (1963) presented optical observations of the radio star-like object 3C 273. The spectrum of the object reveals its peculiar nature with a dominant blue power-law continuum with $F \propto \nu^{0.3}$. In addition, a family of absorption lines of 3C 273 appears to have very large peculiar velocities relative to the Milky-Way. Schmidt proposed his revelation of this unique object being extragalactic in nature, if one assumes the large velocities are due to a cosmological redshift at $z \approx 0.158$.

This exciting discovery of an extragalactic source led to a continued effort for searching this new class of objects throughout the years. These star-like objects were later coined the “quasi-stellar radio sources” (Sandage & Wyndham 1965). It was renamed again to quasi-stellar object (QSO) or quasar after their spectra were more thoroughly studied. It became an exciting observational probe for the studies of the intergalactic medium (IGM). This is because of their bright blue continuum, that allows themselves to be found out to great cosmological distances, with redshift up to $z \approx 6$. Very quickly, we have collected a large sample of QSOs revealing a variety of absorption line systems, including neutral hydrogen and heavy elements in various ionization states. In particular, a class of absorption lines known as the Lyman alpha forest was thought to be intergalactic in nature (e.g., Davé et al., 1997). The column density of these forest of absorption lines were measured and generally fall between $10^{12} - 10^{15} \text{ cm}^{-2}$. Generally, these systems do not have associated absorption lines due to heavy elements. The intergalactic nature of Lyman alpha forest was later confirmed by cosmological simulations of structure formation, which showed that they indeed arise from pristine gaseous filaments that permeate between galaxies and serve as fuel for their continual star formation across cosmic time (e.g., Davé et al., 1997).

Systems with column density of neutral hydrogen great than 10^{15} cm^{-2} were also discovered. In contrast to $\text{Ly}\alpha$ forest, associated heavy elements were detected. In particular, triply ionized carbon (CIV) were found nearly 100% with systems of high H I column (Sar-

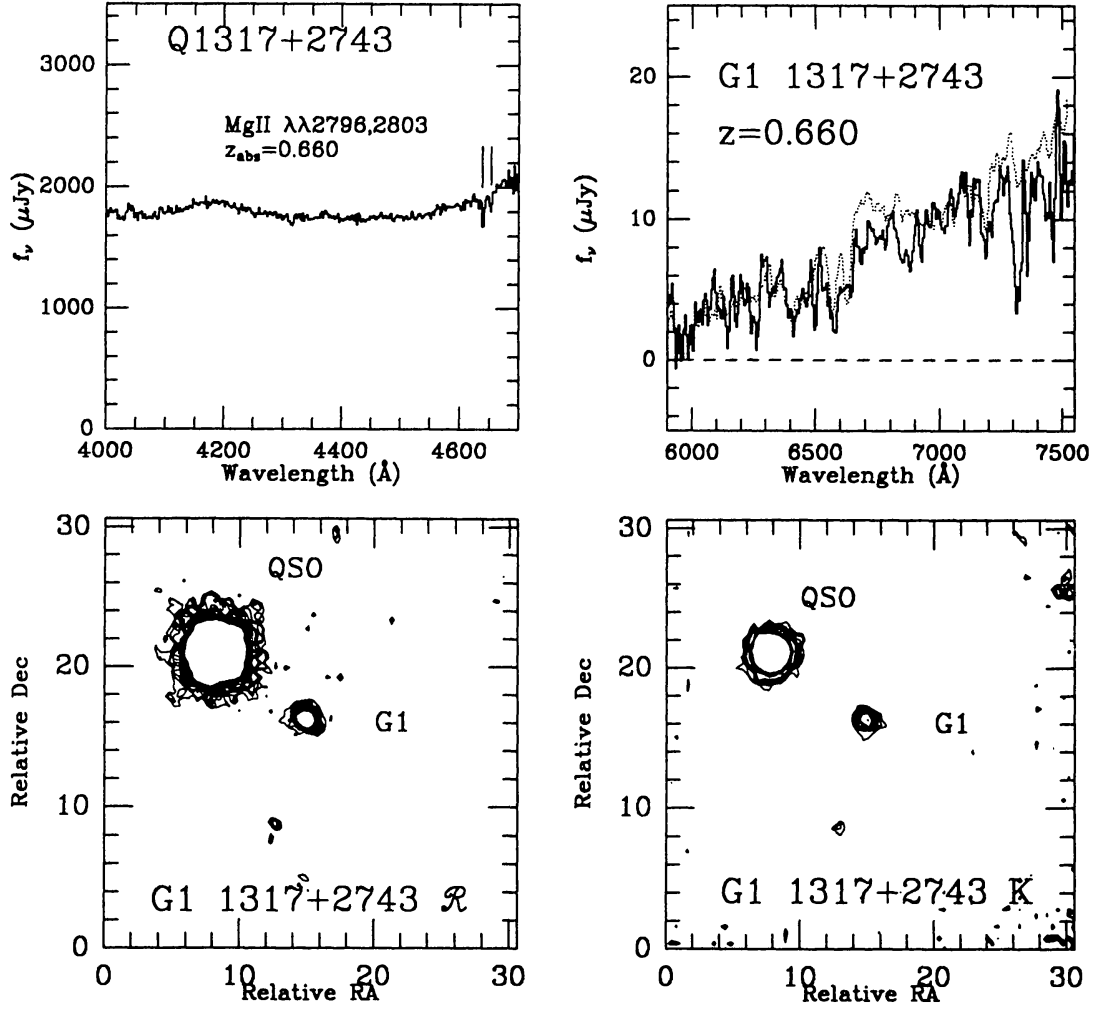


Figure 1.1: These figures show an early example of the association of cold gas and galaxies using background QSOs and galaxies at intermediate redshifts (Steidel, 1993). The upper left panel shows Mg II absorption doublet at redshift similar to the galaxies at $z \approx 0.66$. The right panel is the spectrum of the host galaxies with the 4000Å break near 6640Å, giving a systemic redshift of 0.66. The bottom panels show images of the QSO-galaxy field in two different filter bands (R and K).

gent & Steidel 1992). Singly ionized magnesium (Mg II) were detected in H I systems with even higher column ($> 10^{17} \text{ cm}^{-2}$) at a rate of $\sim 25\%$. This special yet ubiquitous high column-density H I was coined as Lyman Limit systems (LLS) since the entire Lyman series can be detected (Lanzetta 1987; Bergeron & Stasinska 1986; Sargent & Steidel 1992). The high detection rate of heavy elements signals that the gas has been processed by star formation, providing tantalizing hints that they do not live far from star-forming sites of external galaxies.

It was recognized early that there were “too many” metal lines in spectra of high- z QSOs to account for the distribution galaxies. Given the typical size of the luminous component of galaxies ($\sim 10\text{-}20 \text{ kpc}$), it would imply that the number density of galaxies at high redshift would be too great, compared to the observations of the current population of galaxies in the local Universe. Before it was well understood, Bahcall & Spitzer (1969) has in fact proposed that most of the intervening metal absorption lines originate from extended halos of normal galaxies. If this is true, one is tempted to ask the question of how large galaxies would have to be to account for the large detection rates of heavy elements? Bahcall & Spitzer (1969) suggested that this extended gaseous halos must be in the order of $\sim 50h^{-1} \text{ kpc}$ for Mg II and $100h^{-1} \text{ kpc}$ for C IV to match to their detection rates². While the origins of these absorbers were debated, this provides a working hypothesis that pushes the field forward.

The direct evidence of the connection between metal absorption lines and galaxies has remained elusive for many years. Even so, it is becoming more convincing as the circumstantial evidence accumulates and corroborates the hypothesis of gaseous halos from multiple angles. For example, when high-resolution spectrograph became available, it was observed that the velocity variation of metal lines is similar to those within galaxies at $dv \sim 200 \text{ km/s}$. They also exhibit clustering in velocity in the scale of $200 < dv \leq 1000 \text{ km/s}$ (e.g, Sargent et al., 1988; Steidel, 1990).

The difficulty of directly connecting metal line absorber and their host galaxies is that

2. This is close to the current statistical mapping of the CGM, the subject of chapter 1.

they are too faint to measure their redshift. Thus, early studies have focused on Ca II H&K lines $\lambda\lambda 3934, 3969$ and galaxies at the local Universe with $z < 0.03$ (e.g., Boksenberg & Sargent, 1978; Boksenberg et al., 1980). The first direct evidence of an association between metal absorption line and a galaxy is by Bergeron & Boisse (1986) where he searched for galaxies near the redshift of the known absorbers in the QSO spectra.

1.3 Early Models of Galaxy Formation and Gaseous Halos

The emergence of cold dark matter paradigm of galaxy formation came around the 1970-1980s. At this time, it was understood that hierarchical structure formation originates from early small density fluctuations due to gravitational instability in an otherwise homogenous and isotropic density field. Given a cosmological framework, the goal of galaxy formation is to understand how the density perturbations grow over time.

It was recognized early that the mass scale of galaxies is not set by gravity alone. In addition to the nonlinear gravitational collapse of density peaks, the formation of galaxies depends on the physics of hydrodynamics and thermodynamics. The limiting masses of galaxies are thought to depend on how quickly and efficiently the hot halo can cool compared to the dynamical timescales (i.e., free-fall time). In fact, it was first proposed in 1962 that galaxies form out of giant gas clouds via radiative cooling (Eggen et al., 1962). In addition, it is known that there are two major types of galaxies at this point: late-type, star-forming disk and early-type elliptical quiescent galaxies. Gott & Thuan (1976) attempted to explain this bimodality of galaxy populations with the amplitudes of the density peaks and efficiency of cooling of hot halo based on the idea of Eggen et al. (1962).

White & Rees (1978) proposed a two-stage galaxy formation within the CDM framework, which has become the foundation of modern theory of galaxy formation. First, dark matter halos form from hierarchical clustering due to gravitational instabilities in the cosmic density field. The IGM follows the clustering of dark matter and is heating via accretion shocks to form hot gaseous halos with a temperature equal to the virial temperature of the dark matter

halo, which depends on the initial amplitude of the density peaks. Second, the shock-heated hot gas establishes hydrostatic equilibrium and can later cool within the potential wells of the dark matter halo (Silk, 1977; Rees & Ostriker, 1977; Binney, 1977). Figure 1.2 is a map of modern cosmological simulation, which nicely illustrates the basic idea of galaxy formation in the CDM paradigm.

The condensation of cold gas from the hot halo provides star formation fuel across cosmic time. It was recognized early that the amount of cold gas condensing out of the hot halo would result in an overproduction of stars in galaxies. This “overcooling” problem was recognized early. Larson (1974), in particular, invoked that supernovae upon the death of massive stars can provide the necessary heating. White & Rees (1978) followed this idea that some form of feedback mechanisms are needed to provide heating in order to prevent the overcooling in the CDM paradigm of galaxy formation.

The two-stage model by White & Rees (1978) laid out the foundation of the modern theory of galaxy formation. Many details and improvements of the model are worked out in the decades that follow. For example, the gas accreting onto a dark matter halo may not be shock-heated to the virial temperature (Birnboim & Dekel, 2003; Kereš et al., 2009; Faucher-Giguère et al., 2011). These authors predict that galaxies acquire their gas through two modes. A majority of the baryonic content of a galaxy is accreted from filamentary cold gas if the halo mass is less than $10^{11} M_{\odot}$. Above this critical mass scale, galaxies acquire their gas through a hot gaseous halo as prescribed by White & Rees (1978).

1.4 Galactic Winds

The necessary heating invoked by Larson (1974) and White & Rees (1978) provide a natural channel to enrich the CGM (e.g., Muratov et al., 2015). In this picture, enriched cold gas is launched by supernova feedback in the form of hot galactic winds into the hot gaseous halo. Similarly, active galactic nuclei (AGN) are thought to be the main source of feedback that suppresses star formations in the more massive and quiescent galaxies.

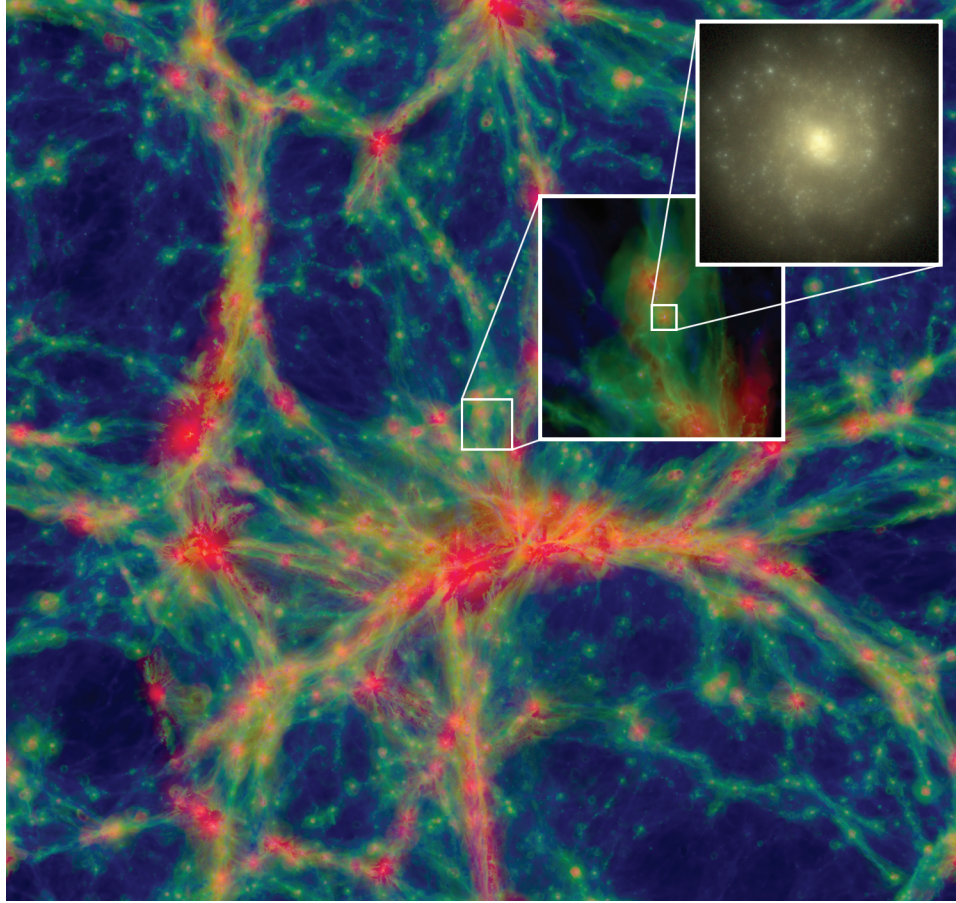


Figure 1.2: A map of the EAGLE simulations (Schaye et al., 2015) in a $100 \times 100 \times 20$ co-moving Mpc box. The pullout images shows a fantastically large dynamical range based using modern computers and improved algorithms. The intensity represents the gas density while the color represents the gas temperature (blue through green through red indicate cooler to hotter). The zoom-in shows a galaxy with stellar mass $3 \times 10^{10} M_{\odot}$ showing the optical band stellar light.

Galactic outflows are commonly observed in both the local and high-redshift Universe (e.g., Shapley et al., 2003; Veilleux et al., 2005). Figure 1.3 shows the composite image of multiple observations of a nearby starburst galaxy M82, which is one of the best examples of multiphase galactic winds. In addition to the hot phase observed through X-ray emitting gas (e.g., Strickland & Heckman, 2009), the cold outflows can be observed via self-absorption of cold gas by the Na I λ 5890, 5896 and Mg II doublets (e.g., Heckman et al., 2000; Weiner et al., 2009) and warm ionized gas (e.g., Lehnert et al., 1999; Heckman et al., 2017). In some cases, even molecular gas can also be detected (e.g., Veilleux et al., 2009; Beirão et al., 2015; Leroy et al., 2015). Galactic outflows are truly multiphase that spans a temperature range from 100K and 10^8 K. A successful model for galactic winds must ambitiously explain all of the existing observations.

In this picture, it is natural to consider how cold clouds can be accelerated by ram pressure of hot galactic winds. Although it is not clear if cold gas is indeed co-moving with the hot wind, a small fraction of cold gas is observed to be moving in the order of 500 km/s or more (see Heckman et al., 2017). This is one of the major challenges in the theoretical community, known as the entrainment of cold gas. Many have reduced the complex cloud-winds interaction down to hot winds striking a single (or multiple) spherical clouds in the early days. In pure hydrodynamics, Klein et al. (1994) shows that the velocity evolution of adiabatic clouds of this wind-cloud interaction is self-similar. In addition, they show that the characteristic cloud crushing timescale is,

$$t_{cc} = \chi^{1/2} R / v_w \quad (1.1)$$

where $\chi = n_{cl}/n_w$ is the density contrast, R is the initial radius of the cloud and v_w is the velocity of the hot winds. Many papers with increasing sophistication ensue. Direct numerical simulations of cold gas embedded in hot winds in two and three dimensions (e.g., Klein et al., 1994; Stone & Norman, 1992). A common finding from these work is that the clouds are destructed before gaining a significant fraction of the hot winds speed, as t_{cc} is

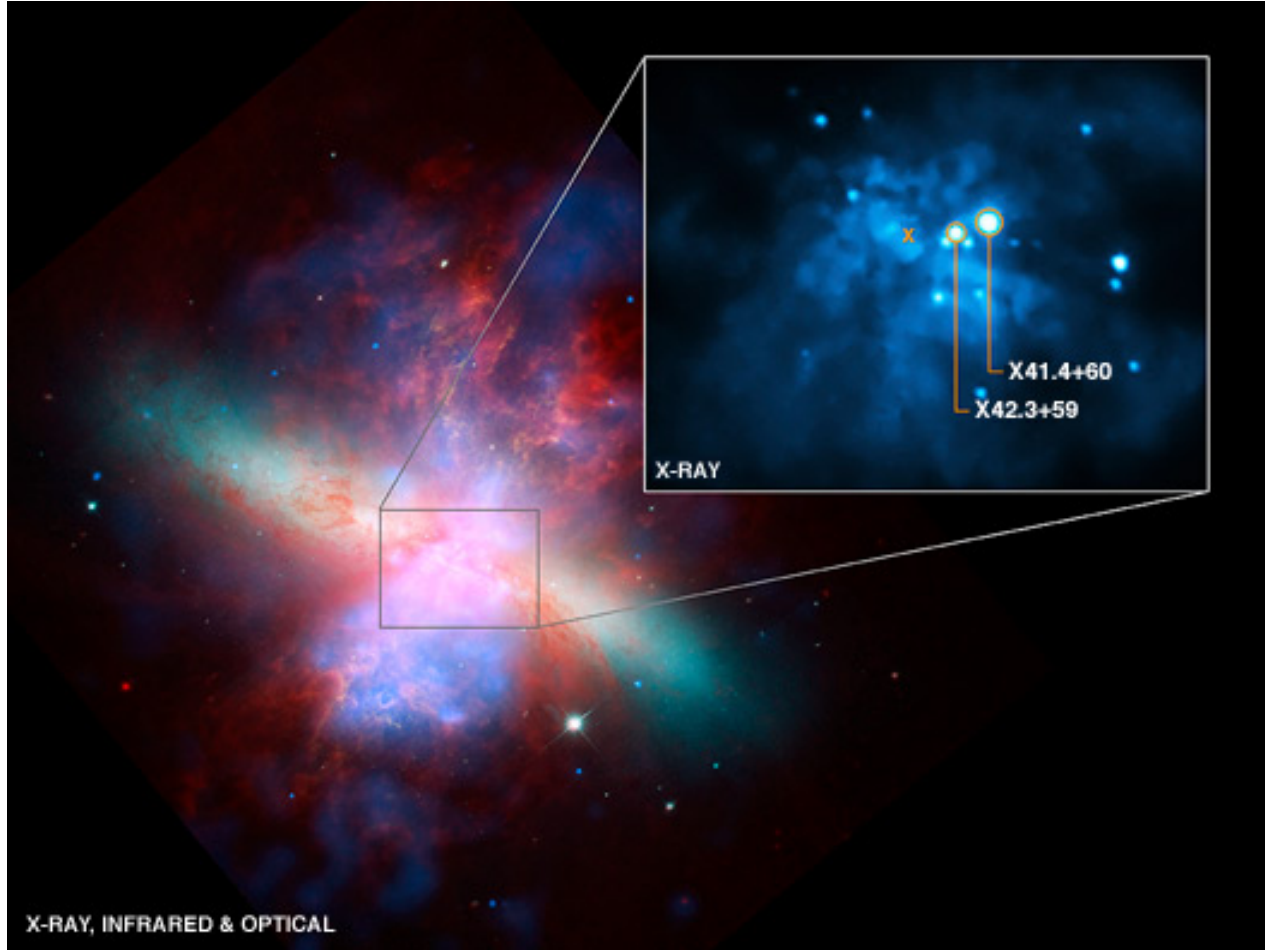


Figure 1.3: This is a composite image of the nearby galaxy M82, showing the starburst-driven multiphase galactic outflows. The blue color shows X-ray observations by the Chandra Space Telescope; the green and orange show optical emission by 10^4 K gas; the red shows infrared data from the Spitzer Space Telescope. The pullout image shows the central region with extended X-ray emissions, in addition to the two labeled X-ray point sources.

very short for the typical parameter values (e.g., Scannapieco & Brüggen, 2015; Zhang et al., 2017). This open problem has led to many more sophisticated attempts. For example, it was shown that radiative cooling can lead to the cloud fragmenting into small and long-lived fragments (Cooper et al., 2009; Mellema et al., 2002). In addition, Brüggen & Scannapieco (2016) showed that thermal conduction causes cold clouds to fully evaporate though it can extend their lifetimes by compressing them into filaments. This compression by evaporation leads to a small cross-section that faces the hot winds, the resulting high column density filaments, therefore, are harder to accelerate and do not move with the velocity of the hot wind. To suppress thermal conduction and dynamical instabilities, magnetic fields may significantly prolong the lifetimes of cold gas so that they can be accelerated over a longer period of time (McCourt et al., 2015).

Overall, simulations thus far have not convincingly demonstrated that cold clouds can be accelerated to be co-moving with the hot winds as observed. Additional processes may be needed, such as additional acceleration mechanisms, as suggested by observations. In addition to ram pressure by hot winds, radiation pressure from young massive stars may also play a role (e.g., Murray et al., 2005; Hopkins et al., 2012; Thompson et al., 2015). In fact, many modern galaxy simulations implement the net effects of radiation pressure through their absorption and scattering by dust grains (e.g., Thompson et al., 2015). The momentum inject rate from radiation pressure in simulations can be modeled as $\dot{p} = (1 + \tau_{\text{IR}})L(t)/c$ (Agertz et al., 2013), where τ_{IR} is the optical depth in the infrared wavelength regime. The first term describes the radiation absorption and scattering in the UV regime and the second term describes the additional boost when re-emitted photons scatter with dust grains.

Yet another possibility is cosmic rays driven winds. It is known that cosmic rays are prevalent in galaxies; there is roughly equipartition of cosmic ray and magnetic energy densities (Beck & Krause, 2005). This suggests that CRs may play a significant role in the ISM and galactic winds. Cosmic rays driven winds is an attractive solution for launching cold gas into the halo. First, cosmic rays are thought to be the product of supernova explosions

where it retains roughly 10% of the nominal energy 10^{51} ergs. Second, the diffusive nature of cosmic rays leads to large scattering optical depth. This provides a steady and gradual pressure gradients that can smoothly drive cold gas out of larger distances in the halo (e.g., Booth et al., 2013; Salem & Bryan, 2014). The effects of cosmic rays driven winds have not yet been fully explored. In particular, it is not clear how one should model the diffusion of cosmic rays since the diffusion coefficient (κ_{CR}) depends on the magnetic fields in galaxies. Chapter 3 shows one of the first attempts of including cosmic rays in cosmological zoom-in simulations of galaxies. Although the implemented isotropic diffusion cosmic rays help drive cold gas into the CGM, it still cannot recover fully the observed spatial extents and covering fractions of cold gas.

1.5 Thermal Instabilities

A competing scenario of launching cold gas out to the hot halo is in situ formation within the halo through thermal instability. In a seminal paper, Field (1965) laid the foundations and basic property of astrophysical thermal instability. To understand the basic idea, let us consider density perturbations in an otherwise uniform medium in pressure equilibrium. Since the rate of energy loss due to cooling depends on the local density of the gas, slightly over-dense regions will cool more. Thermal pressure support is lost in these denser regions; and then the surrounding hot gas will compress the gas further, which results in even denser gas that can cool even more effectively. Therefore, thermal instability describes a runaway cooling process that leads to a multiphase medium, one hot and one cold in pressure equilibrium. The criterion for thermal instability depends sensitively on the net cooling function $\mathcal{L}$. In particular, the gas is unstable if

$$\left(\frac{\partial \mathcal{L}}{\partial T}\right)_P < 0 \quad (1.2)$$

where the subscript P indicates the derivative is taken under constant pressure. This states that runaway cooling ensues if the net cooling rate increase when temperature decreases.

It was recognized early that this criterion holds true for a broad range of temperatures. It was, therefore, applied to a wide range of astrophysical problems, including the famous two- and three-phase models of the ISM (Field et al., 1969; McKee & Ostriker, 1977). Many have also explored the details of the interplay between thermal instability in hot halo and galaxy formation for many years to come. Semi-analytic models suggest that the large fraction of the hot halo can be multiphase and dynamic (e.g., Mo & Miralda-Escude, 1996; Maller & Bullock, 2004). These authors have in fact argued that the cold gas observed by QSOs absorption lines studies can be explained by the same two-phase structure of gaseous halos as postulated by Spitzer (1956): cold gas at $\sim 10^4$ K in pressure equilibrium with the hot corona at the virial temperature of the dark matter halo.

Many have also explored direct hydrodynamical simulations of thermal instabilities in various setting. For example, Joung et al. (2012) ran a series of small box simulations to examine the conditions under which cold clouds can condense and fall out of the hot halo. They find that the ratio of cooling time should be short compared to the dynamical time; otherwise, cold clouds is accelerated under gravity but disrupted by dynamical instabilities. Similar to findings by Kereš & Hernquist (2009), inflowing cold gas with over-densities of 10-20 is expected to be marginally unstable and can seed the onset of thermal instabilities. While Joung et al. (2012) focuses on the dynamics of a single cloud falling under gravity, a different set of simulations in recent years have explored the formation of a multiphase structure in the halo by considering the ratio of cooling time to free-fall time ($t_{\text{cool}}/t_{\text{ff}}$; McCourt et al., 2012; Li et al., 2015). Observations of intracluster medium indicate the H_α luminosity depends sensitively on this ratio, with a stark multiphase transition at $t_{\text{cool}}/t_{\text{ff}} \sim 10$ (e.g., Voit & Donahue, 2015). However, the evidence for similar criterion and the applicability of thermal instability are not yet conclusive in halos of L_* galaxies. Chapter 5 in this dissertation explores the multiphase structure based on thermal instabilities in the regime of L_* galaxies.

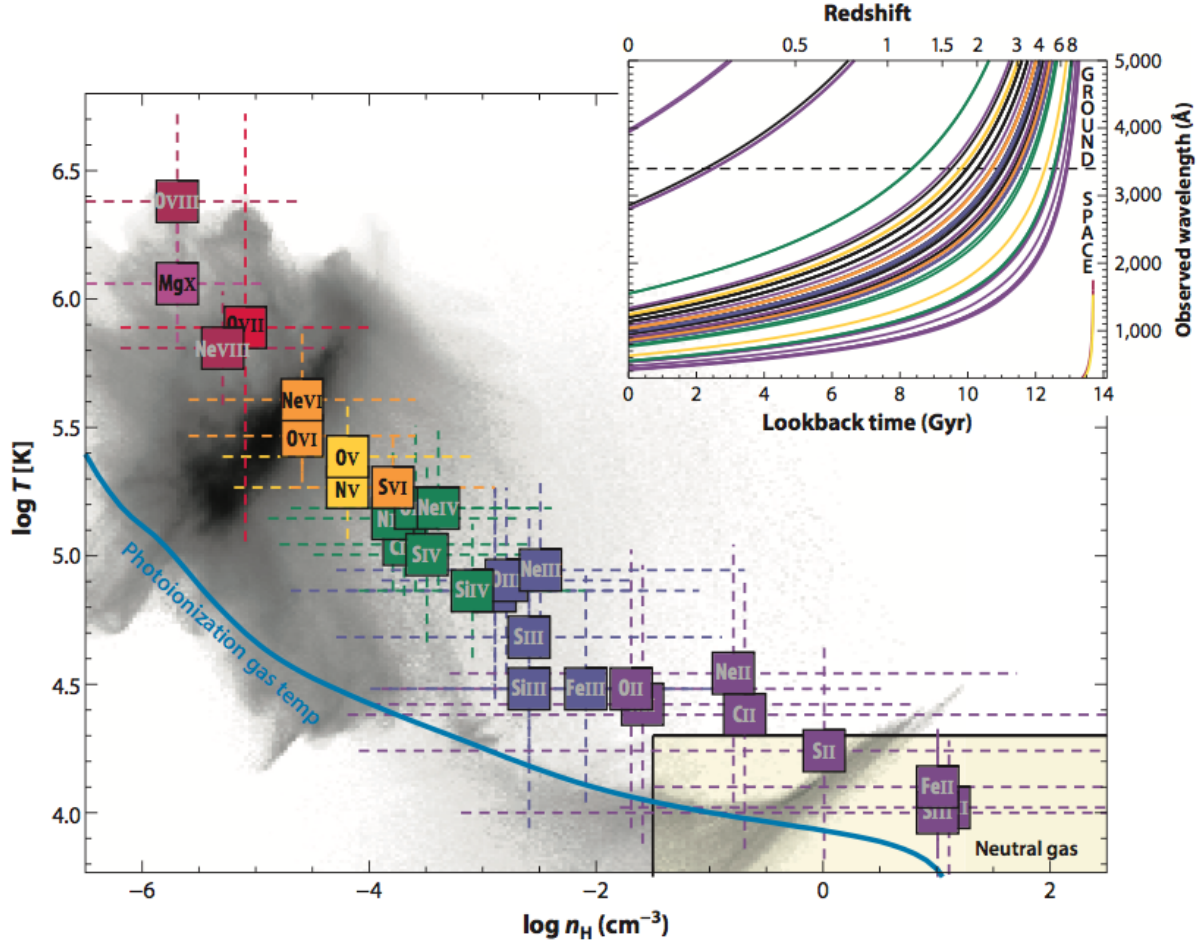


Figure 1.4: This figure shows the locations of ions in the phase diagram where its ionization fraction are maximized in a combination of collisional ionization (temperature dependence) and photoionization (density dependence; Tumlinson et al., 2017). The square labels of ions are colored according to the ionization states. The lines on the top right panel indicate the redshifts at which the lines are observable in either ground-based or space-based telescopes.

1.6 A Modern Renaissance: the Circumgalactic Medium

Despite our improved understanding of the gaseous halos around galaxies since the 1960s, the empirical constraints mainly come from a few dozen of QSOs-galaxy pairs. In addition, this was limited to the Mg II doublet for galaxies at intermediate redshifts. The launch of the *Hubble Space Telescope* (HST) and the Cosmic Origin Spectrograph (COS; Green et al., 2012) brought a renaissance to the galaxy-gaseous halo connection and characterizations. The ultraviolet (UV) wavelength coverage and the high sensitivity of COS allow one to measure

absorption lines of a broad range of ions, probing a wide range of physical conditions in the multiphase CGM. Figure 1.4 shows the location of ions in a temperature-density diagram, indicating what physical condition an ion would probe. The early studies include probing neutral hydrogen using Ly α λ 1215 (Lanzetta et al., 1995; Chen et al., 1998, 2001b), and heavier elements with C IV $\lambda\lambda$ 1548,1550 (Chen et al., 2001a), and highly ionized O VI (Chen & Mulchaey, 2009; Prochaska et al., 2011).

In recent years, there has also been an accelerated progress thanks to a combination of large HST and ground-based observing programs that probe a broad range of ions at once (e.g., H I, Si II, Si III, Si IV, C II, C IV, O VI). For example, the COS-Halo program (Tumlinson et al., 2011; Werk et al., 2013) probes a set of star-forming and quiescent galaxies within 160 kpc. The authors show the CGM harbors a significant fraction of the baryons (Werk et al., 2014). The COS-GASS program explores the connection between the interstellar medium (ISM) and the CGM by combining with H I 21 cm line observations (Borthakur et al., 2015). While these studies focus on the inner regime of the CGM of some specific types of galaxies, Liang & Chen (2014), in contrast, used 195 galaxy-QSO pairs to probe CGM with a broad range of stellar mass of the galaxies ($10^7 - 10^{12} M_{\odot}$) and out to large distances (500 kpc). With a stringent limit, they show that detection rates of low ions (e.g., Si II) drop quickly beyond a small fraction of the virial radius ($0.2 - 0.3 R_{\text{vir}}$), while intermediate ions (e.g., Si IV, C IV) are detected up to $\sim 0.7 R_{\text{vir}}$. Johnson et al. (2015b) extended this study to highly ionized gas where they found continuous detection of O VI up to $\sim R_{\text{vir}}$.

These analyses have provided a comprehensive characterization of multiphase gas in a broad range of galaxies. The task from the theoretical side has been to understand the implications of these observations and incorporate the governing physical processes into our modeling of galaxy formation. Galaxy simulations have previously focused on forming a realistic stellar component of galaxies and ignored the gaseous halos. The current generation of galaxy formation models appears to capture many of the relevant physical processes. State-of-the-art simulations can now produce galaxies with realistic properties (Feldmann et al.,

2011; Sales et al., 2012; Aumer et al., 2013; Stinson et al., 2013; Hopkins et al., 2014; Agertz & Kravtsov, 2015; Schaye et al., 2015). This success in the theoretical community can be mainly attributed to our improved understanding of star formation and feedback processes. Galactic winds are believed to play many crucial roles in shaping the galaxy as a whole. It has shown that star formation feedback can drive large-scale galactic winds to regulate star formation. It can provide the necessary heating to balance the global cooling of gaseous halos.

The increased empirical measurements now place strong constraints on even the modern state-of-the-art galaxy simulations. Despite the success of recent simulations in producing realistic star-forming sites of galaxies, matching the CGM properties proves to be difficult. Simulations can explain some CGM properties, such as the bimodality of O VI and specific star formation rate (sSFR) of the host galaxies (Tumlinson et al., 2011; Oppenheimer et al., 2016). However, most simulations have been unable to produce enough cold/warm gas in the CGM, despite a diversity of feedback processes implemented in them (e.g., Stinson et al., 2012; Ford et al., 2013; Hummels et al., 2013; Shen et al., 2013; Suresh et al., 2015). On the other hand, a variation of large feedback energy in Liang et al. (2016)³ leads to a more realistic CGM by launching cold gas into the halo; however it does so by destroying their central galaxy. As mentioned above, it is also thought that cosmic rays feedback may help launch the cold gas into large distances (e.g., Socrates et al., 2008; Booth et al., 2013; Salem et al., 2014), although it has been shown that these initial attempts were not yet enough (e.g., Liang et al., 2016; Salem et al., 2016).

1.7 The Roadmap of the Dissertation

The following chapters explore the observational and theoretical aspects of the spatial distribution of the multiphase CGM. Chapter 2 focuses on empirical constraints of CGM in cold/warm phase. The spatial profiles of neutral hydrogen ($\text{Ly}\alpha$) and heavy elements (Si II,

3. This is the subject matter in Chapter 3.

Si III, Si IV, C II, C IV) are mapped statistically using an ensemble of 195 galaxy-QSO pairs. The spatial profiles mapped an uncharted regime of the gaseous halos out to 500 kpc from the central galaxies. The differential covering fractions of ions, and their spatial extents themselves provide the empirical constraints on how feedback processes enrich the CGM.

Using a set of cosmological “zoom in” simulations of galaxies, chapter 3 focuses on the spatial distribution of the CGM and conduct systematic comparisons with the observations. We also investigate the spatial distribution of the CGM over cosmic times and its relationship with the host dark matter halos. Simulations predict the CGM column density profiles of ions are exponential in form, $N \propto \exp(-d/h_s)$, where h_s is the scale height of the profiles. In addition, we find that the CGM is self-similar in mass when the profiles are scaled with the virial radius (R_{vir}); the CGM is self-similar in both mass and time when the profiles are presented in units of the scale radius (r_s).

To connect more closely with observations, chapter 4 presents a synthetic absorption pipeline, as a virtual telescope, to observe the simulated galaxy and their CGM. We also develop a Bayesian method for modeling a combination of absorption lines to derive physical properties of absorbers with a formal treatment of detections, including saturated systems, and non-detections in a uniform manner. A fully Bayesian Monte-Carlo Markov Chain (MCMC) Voigt profile fitting `Python` package `bayesvp` is developed along this work. The details of the package are described in the Appendix.

In chapter 5, we present a multiphase model of the CGM based on the idea that the observed cold halo gas may be a composite of cold, dense and small cloudlets. Using semi-analytic calculations and a set of magnetohydrodynamic (MHD) simulations, we show that the resulting cold gas from thermal instabilities conforms to a characteristic column density of $N_{\text{H}} \approx 10^{17} \text{cm}^{-2}$ as predicted by the $c_s t_{\text{cool}}$ ansatz. The model implies a large number of cold clumps in the inner galactic halo with a small volume filling factor but large covering fraction. The model also naturally give rise to spatial extents and differential covering fractions of cold, warm, and hot gas. This implies that future galaxy formation simulations

must address the mismatch of the spatial resolution and relevant scales of the cold gas.

CHAPTER 2

MINING THE CIRCUMGALACTIC BARYONS IN THE LOW REDSHIFT UNIVERSE

2.1 Introduction

Circumgalactic space is thought to contain a vast amount of baryons (e.g, Spitzer, 1956; Fukugita, 2004) that are both multiphase and dynamic (e.g., Mo & Miralda-Escude, 1996; Maller & Bullock, 2004). In principle, this circumgalactic gas provides the fuel necessary for sustaining star formation in galaxies, while at the same time it is also being replenished with both newly accreted (and presumably chemically pristine) intergalactic gas and chemically enriched materials either through gas stripping due to satellite interaction or through starburst driven winds. Therefore, circumgalactic space provides a critical laboratory for studying the baryon cycle that regulates star formation and galaxy growth.

Over the past two decades, absorption spectroscopy of background QSOs have provided an effective tool for probing the otherwise unseen, diffuse gas in halos around distant galaxies. While each QSO provides a one-dimensional mapping through each single halo, observing an ensemble of close galaxy and QSO pairs leads to statistical measurements of mean halo profiles (i.e. absorption strength as a function of projected distance) averaged over the entire galaxy sample. Commonly seen transitions in circumgalactic space include hydrogen $\text{Ly}\alpha \lambda 1215$ (e.g., Lanzetta et al., 1995; Chen et al., 1998; Tripp et al., 1998; Wakker & Savage, 2009; Steidel et al., 2010; Prochaska et al., 2011; Thom et al., 2012; Rudie et al., 2013), the C IV $\lambda\lambda 1548, 1550$ doublet (e.g., Chen et al., 2001a; Adelberger et al., 2005; Steidel et al., 2010)

the Mg II $\lambda\lambda 2796, 2803$ doublet (e.g., Bowen et al., 1995; Chen & Tinker, 2008; Kacprzak et al., 2008; Barton & Cooke, 2009; Chen et al., 2010a; Gauthier et al., 2010; Bordoloi et al., 2011; Werk et al., 2013) and the O VI $\lambda\lambda 1031, 1036$ doublet (Chen & Mulchaey, 2009; Wakker & Savage, 2009; Prochaska et al., 2011; Tumlinson et al., 2011; Johnson et al., 2013)

A general finding from different studies is that most galaxies are surrounded by chemically enriched gas out to $\approx 100\text{--}300$ kpc in projected distance. Competing scenarios for explaining the presence of heavy elements at > 50 kpc from star-forming regions include super galactic winds (e.g., Murray et al., 2011; Booth et al., 2013) and tidally disrupted satellites (e.g., Wang, 1993; Gauthier, 2013).

A particularly interesting absorption feature to adopt for probing the baryon content of galactic halos is the CIV doublet. These doublet transitions are strong and their rest-frame wavelengths, 1548.20 and 1550.77 Å, enable uniform surveys of chemically enriched gas in the optical window over a broad redshift range from redshift $z \approx 1$ to $z \gtrsim 5$. CIV absorbers are found to originate primarily in photoionized gas of temperature $T \approx 4 \times 10^4$ K with some contribution from shock heated gas in galaxies and galactic halos (e.g., Rauch et al., 1997; Boksenberg et al., 2003; Simcoe et al., 2004). Therefore, CIV absorption transitions provide an effective tracer of chemically enriched warm gas in and around star-forming regions.

A number of surveys have been carried out to characterize the statistical properties of CIV absorbers found along random QSO sightlines (e.g., Songaila, 2001; Boksenberg et al., 2003; Simcoe, 2011; Cooksey et al., 2013). While the shape of the CIV equivalent width frequency distribution function appears to remain the same, the total cosmic mass density in C^{3+} ions is found to increase from $z = 4$ to $z \approx 1.5$ by a factor of 2 (e.g., Cooksey et al., 2013). The increasing CIV mass density together with an increasing background radiation intensity with decreasing redshift (e.g., Haardt & Madau, 2012) indicate an increasing chemical enrichment level with time in the gas traced by the CIV absorption transitions (e.g., Oppenheimer & Davé, 2008).

At the same time, only three studies have been carried out to characterize extended CIV gaseous halos around galaxies, two at $z \lesssim 0.5^1$ and one at $z \approx 2.2$. A clear understanding has yet to be established. At $z \approx 0.4$, Chen et al. (2001a) studied the incidence of CIV

1. Though we have recently learned that Bordoloi et al. (2014) has also been conducting a survey of CIV absorbers in halos around low-redshift galaxies.

absorbing gas at projected distances $d \lesssim 300$ kpc from a sample of 50 galaxies. These authors reported the presence of a distinct boundary at a B -band luminosity-normalized projected distance of $\hat{d} \equiv d \times (L_B/L_B^*)^{-0.5} \approx 160$ kpc, beyond which no C IV absorbers are detected. The lack of C IV absorption is based on a subsample of 35 galaxies that occur at $\hat{d} > 160$ kpc from their background QSO sightlines. In contrast, 14 of the 15 galaxies at $\hat{d} < 160$ kpc have an associated C IV absorber, although the rest-frame absorption equivalent width, $W_r(1548)$, exhibits a large scatter between individual galaxies. At $z \approx 2.2$, Steidel et al. (2010) examined the mean spatial absorption profiles of 512 galaxies based on stacked spectra of background galaxies that occur at $d \lesssim 125$ kpc. These authors found that the mean C IV absorption strength declines rapidly at projected distances of $\approx 50 - 100$ kpc. Comparing the observations of Chen et al. (2001a) and Steidel et al. (2010) yielded little distinction in the C IV-traced circumgalactic medium (CGM) at low and high redshifts (Chen, 2012). Both the luminosity-normalized spatial extent and mean absorption equivalent width of the CGM around galaxies of comparable mass (but with very different on-going star formation rate) have changed little over the redshift interval $z = 0.4 - 2.2$, although there exists a large scatter in $W_r(1548)$ in the low-redshift study. Nevertheless, a lack of variation in the spatial profile of the chemically enriched CGM between two distinct epochs poses a serious challenge to the theoretical models of gas flows around galaxies (e.g., Hummels et al., 2013; Ford et al., 2013).

On the other hand, Borthakur et al. (2013) targeted a sample of 20 galaxies at $z < 0.2$ and were able to constrain the incidence of C IV absorbing gas for 17 of these galaxies. Of the 17 galaxies studied, eight occur at $\hat{d} < 160$ kpc and nine at larger distances. These authors detected C IV absorbers for three of the eight galaxies at $\hat{d} < 160$ kpc and three of the nine galaxies at $\hat{d} > 160$ kpc in their sample. The finding of a flat rate of C IV incidence with increasing projected distance is intriguing and appears to be in stark contrast to the earlier finding of (Chen et al., 2001a). While Borthakur et al. (2013) attributed the detected C IV absorbers to starburst activity in the associated galaxies, it is not straightforward to

reconcile the discrepant trend found for extended CIV halos between the Chen et al. (2001a) and Borthakur et al. (2013) samples.

To address the discrepant trend found for extended CIV halos around galaxies at low redshifts and to improve the empirical understanding of how the chemically enriched CGM evolves with time, we have searched public archives to find spectroscopically identified galaxies at small projected distance to a UV bright, background QSO for which high-quality echelle/echellette spectra are already available in the *Hubble Space Telescope* (HST) data archive. Our search has yielded the first large sample of ~ 300 close QSO and galaxy pairs that enables a systematic study of extended CIV halos over the projected distance interval of $d = 0 - 500$ kpc. As described below, this unique sample allows us to place unprecedented limits on the extent of chemical enrichment around low-redshift galaxies based on observations of not only the CIV absorption doublets but also a whole host of ionic transitions from low- and intermediate-ionization states.

Recently, Tumlinson et al. (2013) carried out a large program to use the Cosmic Origins Spectrograph (COS Green et al., 2012) on board HST to study the CGM at $d < 150$ kpc of 44 L_* galaxies at $z = 0.15 - 0.35$. This program, known as the “COS-Halos” survey, was designed to map the multi-phase CGM using O VI and other metal-line diagnostics (Tumlinson et al., 2011; Werk et al., 2013). The low-redshift cut at $z = 0.15$ was dictated by the requirement of detecting the O VI doublet using COS and the FUV channel. However, the spectral coverage of the COS FUV gratings misses the CIV doublet at $z \gtrsim 0.15$, and therefore CIV transitions are not included in the COS-Halos analysis (e.g., Werk et al., 2013).

Our study differs from the COS-Halos program in two fundamental aspects. First, our search is designed to probe the CIV absorption in galactic halos for comparisons with high-redshift studies. Second, because fainter galaxies are also more numerous, it is more likely to find a faint foreground galaxy near a QSO sightline than a luminous galaxy discussion is § 2. As a result, low-luminosity ($< 0.1 L_*$) and low-mass ($< 0.1 M_{\text{star}}^*$) dwarf galaxies contribute

to a large fraction ($\approx 50\%$) of our random QSO and galaxy pair sample. Therefore, *our study complements the COS-Halos effort both in terms of the mass regime of the galaxies and in terms of the ionization state of the CGM.*

Here we report initial findings for the low-redshift CGM based on a sample of 195 galaxies for which constraints on the stellar mass are available. The available stellar mass estimate of each galaxy allows us to infer the total mass of the dark matter in which the galaxy resides (e.g., Behroozi et al., 2013; Kravtsov et al., 2014). This in turn allows us to assess the intrinsic size differences between different galaxies based on the estimated dark matter halo radius from a standard halo model. While the CGM is expected to be regulated by various complicated physical processes, including gravitational motion and star formation feedback, addressing the effects of all these processes at once is challenging. As a first step toward establishing a clear understanding of the processes that drive the observed CGM properties, we have designed our study here to account for the intrinsic size differences (driven by the gravitational potential) of individual galactic halos by normalizing the observed projected distances with the underlying halo radius. We expect that the results will facilitate future studies that consider the effect of additional processes such as star formation feedback.

This paper is organized as follows. In Section 2, we describe the selection criteria for establishing the close galaxy–QSO pair sample and summarize the properties of the galaxy members in the pair sample. In Section 3, we describe the data reduction and analysis procedures of the absorption spectra of the QSOs. In Section 4, we examine the correlation between the strengths of different ionic transitions and galaxy properties, and compare the results of our study with previous findings. Finally, we discuss the implication of our results in Section 5. Throughout the paper, we adopt the standard Λ cosmology, $\Omega_\Lambda = 0.7$, $\Omega_M = 0.3$ with a Hubble constant $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

2.2 A Public Sample of Galaxy–QSO Pairs

We have assembled a large sample ($n_{\text{gal}} \sim 300$) of spectroscopically-identified galaxies that occur at small projected distances $\lesssim 500$ kpc from the sightline of a background UV bright QSO. This galaxy sample is ideal for characterizing extended gaseous halos around galaxies based on the absorption features imprinted in the spectra of the background QSOs.

We first searched the HST archive for QSOs that have been observed with either the Space Telescope Imaging Spectrograph (STIS; Woodgate et al. 1998) or the Cosmic Origins Spectrograph (COS Green et al., 2012) as of Cycle 20. This search yielded a sample of 150 QSOs at $z_{\text{QSO}} = 0.058 - 1.476$. Next, we searched for spectroscopically identified galaxies in public archives, including the Nearby General Catalog (NGC; Sinnott, 1988), the Sloan Digital Sky Survey (SDSS; York et al., 2000) spectroscopic galaxy sample, the Two Micron All Sky Survey (2MASS) galaxies (Huchra et al., 2012), and the Two Degree Field Galaxy Redshift Survey (2dFGRS; Colless et al., 2001). A galaxy and a background QSO were considered a suitable pair if they satisfied the following criteria. (1) At the redshift of the galaxy, the physical projected distance d between the galaxy and the QSO sightline is $d \leq 500$ kpc. (2) The galaxy must be at $|\Delta v| > 10000$ km s $^{-1}$ below the redshift of the QSO to avoid ambiguity between foreground absorbers and those originating in QSO outflows (e.g., Wild et al., 2008). The maximum separation limit of 500 kpc is sufficiently large to encompass the gaseous halos of L_* galaxies known empirically (e.g., Chen et al., 1998, 2001b; Gauthier & Chen, 2011; Prochaska et al., 2011; Rudie et al., 2012). It is also motivated by theoretical expectations of extended halo gas out to ~ 300 kpc in radius from star-forming regions (e.g., Ford et al., 2013). Therefore, including pairs separated by greater than 300 kpc allows us to explore the presence of chemically enriched gas beyond the fiducial virial radius of dark matter halos.

Finally, we formed an “isolated” galaxy and QSO pair sample by considering only galaxies without close neighbors, in order to minimize the ambiguity in associating an absorber with a galaxy. Previous studies have also shown that galaxies in group environments tend to

have more extended halo gas than isolated galaxies that results in a larger scatter in the observed spatial distribution of absorption strength with projected distance (Chen et al., 2010a; Bordoloi et al., 2011; Gauthier & Chen, 2011). In our study, a galaxy was considered an “isolated” one, if no other galaxies have been spectroscopically identified within projected separation of $d = 500$ kpc and velocity separation of $|\Delta v| = 500 \text{ km s}^{-1}$ were present. These criteria are driven by the model expectation that an L_* galaxy at $z = 0$ has a virial radius of ≈ 300 kpc and maximum circular velocity of $\approx 200 \text{ km s}^{-1}$. We chosen a more conservative limit to allow for uncertainties in galaxy redshifts.

Our search criteria of suitable galaxy and QSO pairs yielded a final sample of 213 “isolated” galaxies that occur at $d \lesssim 500$ kpc from the sightline of a background QSO for studies of the low-redshift CGM. Excluding 18 galaxies that are in the COS-Dwarfs program (Bordoloi et al., 2014) led to a total of 195 galaxies in our sample. The redshifts of the galaxies range from $z = 0.002$ to $z = 0.176$ with a median of $\langle z \rangle = 0.041$, and the projected distances of the QSOs range from $d \approx 32$ kpc to $d \approx 500$ kpc with a median of $\langle d \rangle = 362$ kpc. The projected distance versus redshift distribution of the galaxies sample is shown in Figure 1, in which the top and right panels display the histograms in redshift and in projected distance, respectively.

To facilitate a detailed investigation of how the observed CGM absorption properties depend on galaxy properties, we also made use of available measurements of rest-frame UV luminosities and stellar mass from the NASA–Sloan Atlas² (Blanton et al., 2005, 2011). This public atlas contains 145,155 galaxies with known redshifts at $z \lesssim 0.05$. For each galaxy, it includes optical and UV photometric measurements from the SDSS Data Release 8 (Blanton et al., 2011) and GALEX Data Release 6 (Schiminovich et al., 2007), as well as derived quantities such as stellar mass (M_{star}) and rest-frame UV and optical absolute magnitudes computed by the K-correct code (Blanton & Roweis, 2007). The rest-frame UV absolute magnitudes allows us to estimate an unobscured star formation rate (SFR) for each galaxy

2. <http://nsatlas.org/>

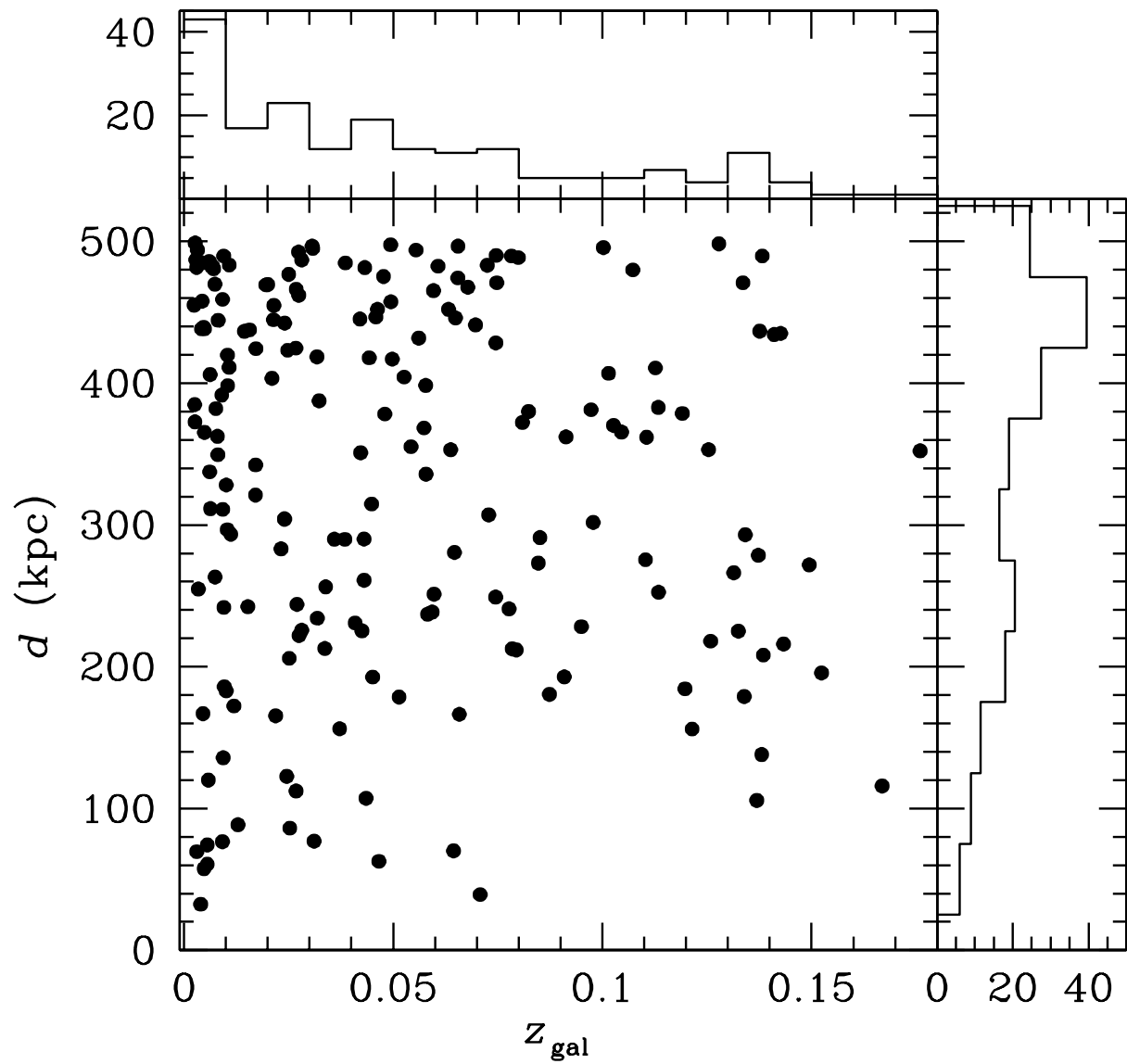


Figure 2.1: Distribution of projected distance versus redshift for the 195 galaxies (§ 2). The top (right) panel displays the redshift (impact parameter) histograms.

based on the calibration coefficient of Kennicutt & Evans (2012).

The available stellar mass estimate of each galaxy allows us to infer the total mass of the dark matter halo (M_h) in which the galaxy resides (Behroozi et al., 2013; Kravtsov et al., 2014). This in turn allows us to assess the intrinsic size differences between different galaxies based on the estimated dark matter halo radius R_h from a standard halo model (see § 4 for a more detailed discussion).

Cross-matching the “isolated” galaxy sample with the NASA–Sloan Atlas resulted in 116 overlapping galaxies. We visually inspected the optical images of each galaxy and removed five objects due to erroneous photometry. This exercise resulted in a sample of 111 galaxies with accurate M_{star} and star formation rate (SFR) measurements available from the NASA–Sloan Atlas. However, to utilize the full “isolated” galaxy sample for investigating how the observed CGM absorption properties depend on galaxy properties, we needed to estimate the stellar mass for each of the remaining 100 galaxies based on available photometric data. We used the subsample of 111 galaxies as a guide. Comparing the stellar mass provided by the NASA–Sloan Atlas with intrinsic luminosities in different bandpasses revealed a tight correlation between M_{star} and rest-frame absolute r -band magnitude, $M_{\text{SDSS},r}$, over four orders of magnitude in M_{star} (Figure 2). Based a linear regression analysis, we find that the M_{star} vs. $M_{\text{SDSS},r}$ correlation is best characterized by

$$\log M_{\text{star}} = 0.14 - 0.49 M_{\text{SDSS},r} \quad (2.1)$$

with a $1\text{-}\sigma$ scatter of $\Delta \log M_{\text{star}} = 0.21$. Equation (1) enables us to obtain a stellar mass estimate and associated uncertainty for galaxies based on $M_{\text{SDSS},r}$. In Figure 2, we also include the entire NASA–Sloan catalog, which has $> 100\text{k}$ galaxies (grey points), to contrast the 111 galaxies adopted for deriving Equation (1). The comparison shows that the 111 overlapping galaxies and the full NASA–Sloan catalog share a consistent mean relation and scatter between M_{star} and $M_{\text{SDSS},r}$, particularly at the high-mass end with

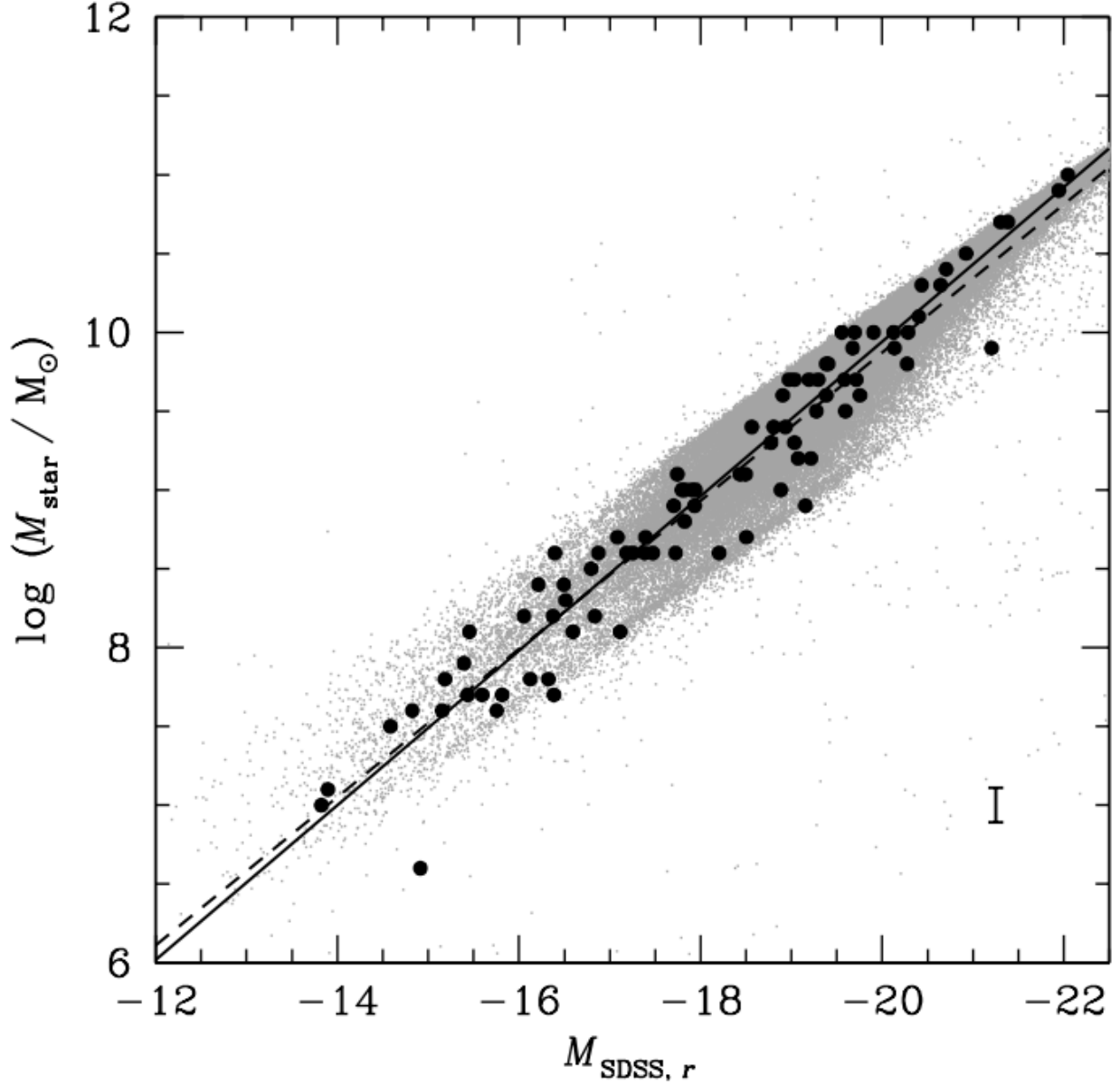


Figure 2.2: Observed correlation between stellar mass M_{star} and rest-frame r -band absolute magnitude, $M_{\text{SDSS}, r}$ for 111 overlapping galaxies that are both in our pair sample and in the NASA-Sloan Atlas (dark solid points). The solid line indicates the best-fit model with $\log M_{\text{star}} = 0.14 - 0.49 M_r$ and a 1- σ scatter of $\Delta \log M_{\text{star}} = 0.21$ (error bar in the lower-right corner). The full NASA-Sloan catalog, which has $> 100\text{k}$ galaxies (grey points), is also included to contrast these 111 overlapping galaxies. The full sample displays a consistent mean relation (dashed line) and scatter between M_{star} and M_r . The consistent distribution of the two samples strongly supports that the mean relation derived using the subsample of 111 galaxies is representative of the general galaxy population at low redshifts.

$M_{\text{star}} > 10^{10} M_{\odot}$. Equation (1) is therefore representative of the general galaxy population at low redshifts.

We present the distribution of stellar masses of our galaxies in Figure 3. The blue shaded histogram shows the stellar mass distribution for the full sample of 195 galaxies, while the black open histogram shows the stellar mass distribution for the subsample of 102 galaxies with M_{star} inferred from M_r following Equation (1). Figure 3 shows that the galaxies in the full sample span a wide range in stellar mass, from $M_{\text{star}} = 1.5 \times 10^5 M_{\odot}$ to $M_{\text{star}} = 1.4 \times 10^{11} M_{\odot}$, with a median $\langle M_{\text{star}} \rangle = 5.0 \times 10^9 M_{\odot}$, which is $\approx 0.1 M_{\text{star}}^*$ (Baldry et al. 2012; Muzzin et al. 2013). Comparing the blue shaded and black open histograms shows that including galaxies with M_{star} inferred from Equation (1) allows us to expand our CGM study to include higher-mass galaxies, covering a mass range that overlaps with the COS-Halos sample at $z = 0.1 - 0.4$ (Werk et al., 2013, red open histogram) and the starburst sample at $z = 2.2$ from (Reddy et al., 2012, green open histogram)³.

In addition, examining the SFR versus M_{star} correlation helps to characterize the galaxies in our pair sample in the context of the general galaxy population. Considering the 111 galaxies with both M_{star} and rest-frame UV photometric measurements from the NASA-Sloan Atlas, we present in Figure 4 the specific Star Formation Rate (sSFR; SFR per unit stellar mass) and stellar mass of these galaxies (solid points) in comparison to the mean relations of star-forming main sequence galaxies (blue line) and passive, red galaxies (red) from Schiminovich et al. (2010). Recall that we measured an unobscured SFR for each galaxy based on the observed NUV flux without correcting for possible dust extinction. Therefore, our measurements likely represent lower limits to the intrinsic SFR. The tilt in the distribution of our galaxies relative to the mean relation of star-forming main-sequence galaxies can be explained by the observed correlation between dust extinction and stellar

3. We note that while different galaxy samples overlap in the stellar mass range, there are strong distinctions in SFR and in the knowledge of environments and redshift precisions. Specifically, the starburst sample at $z = 2.2$ have a mean SFR of $\sim 30 M_{\odot} \text{ yr}^{-1}$ (e.g., Erb et al., 2006) and low-redshift galaxies have on average unobscured SFR of $\sim 0.1 M_{\odot} \text{ yr}^{-1}$. In addition, while our sample only includes galaxies in an isolated environment, the environments of the high-redshift starburst sample are unknown.

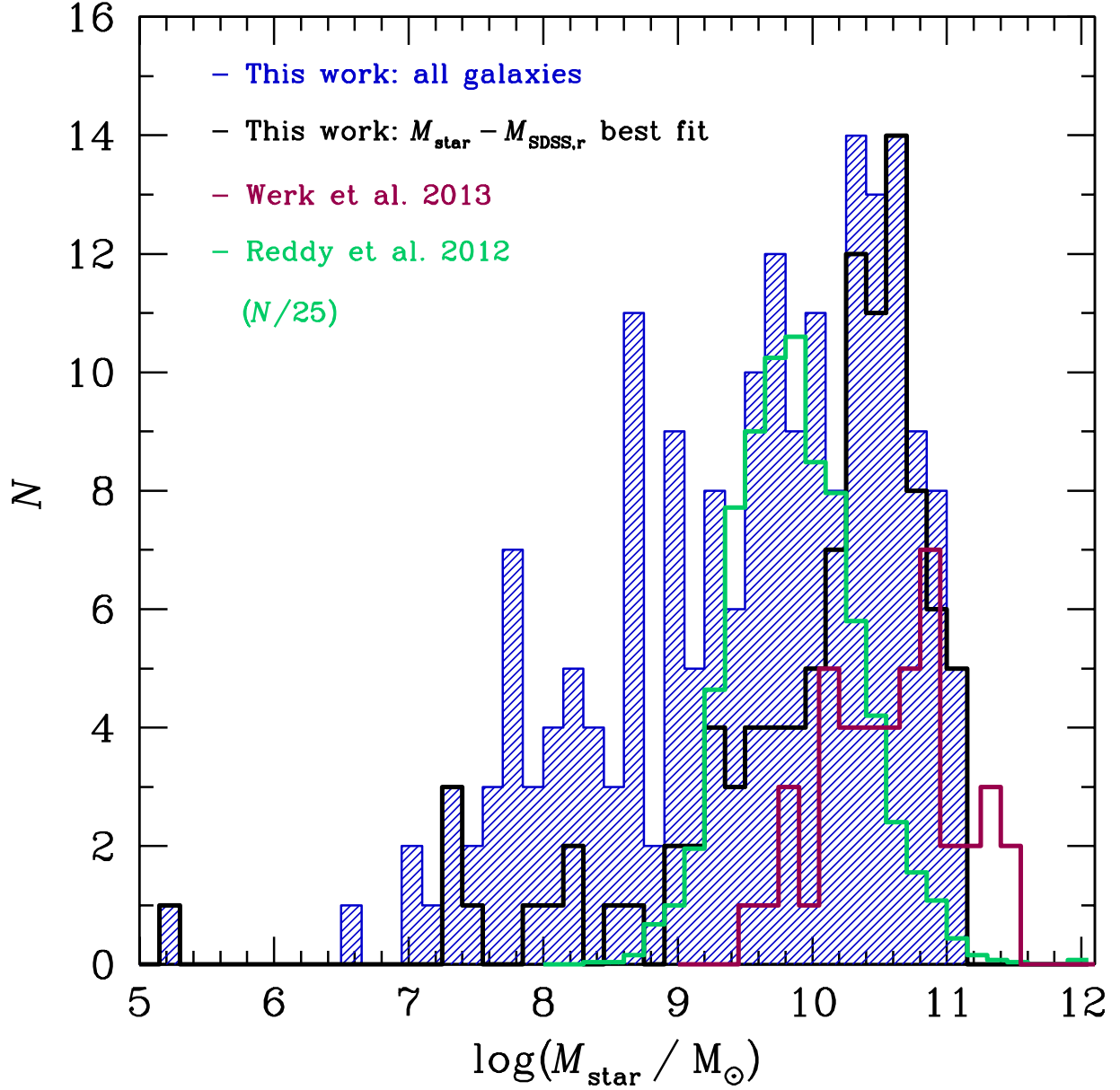


Figure 2.3: Stellar mass distribution of our galaxy sample (blue shaded histogram). The median mass is $\langle M_{\text{star}} \rangle = 5 \times 10^9 M_{\odot}$, in comparison to the characteristic mass $M_{\text{star}}^* = 5 \times 10^{10} M_{\odot}$ for the field galaxies (Baldry et al., 2012; Muzzin et al., 2013). Our galaxy sample has a mass distribution that encompasses previous CGM studies at both low and high redshifts. Specifically, the COS-Halos sample at $z = 0.1 - 0.4$ (Werk et al., 2013, red open histogram) has a median stellar mass of $4 \times 10^{10} M_{\odot}$ and the starburst sample at $z = 2.2$ from Reddy et al. (2012; green open histogram) has a median stellar mass of $7.9 \times 10^9 M_{\odot}$. The number counts of the Reddy et al. sample has been reduced by a factor of 25 for presentation purpose.

mass, namely more massive galaxies on average exhibit higher degrees of dust extinction (e.g., Zahid et al., 2013). It is clear that the majority of the galaxies in our sample are consistent with being star-forming main-sequence galaxies with a small fraction found in the red-sequence regime.

We note that while the stellar mass derived using the K-correct code may be systematically overestimated by 0.1 dex in the low mass regime of $M_{\text{star}} \lesssim 10^{10} M_{\odot}$ (e.g., Moustakas et al., 2013), such error translates to a systematic uncertainty of ≈ 0.06 dex in M_h , well within the scatter in the stellar-mass-to-halo-mass relation of Behroozi et al. (2013). The resulting uncertainty in R_h is then negligible.

Of the 195 galaxies in our pair sample, 12 occur within 100 kpc in projected distance from a background QSO sightline. We summarize the properties of the galaxy sample in Table 1, which lists for each galaxy the identification, right ascension (RA) and declination (Dec), spectroscopic redshift (z_{gal}), unobscured SFR (when available), stellar mass (M_{star}), halo mass (M_h), halo radius (R_h), and absolute r -band magnitude $M_{\text{SDSS},r}$.

2.3 QSO UV Spectroscopy

Our galaxy sample contains 195 “isolated” galaxies that occur near the sightlines of 96 independent background QSOs with high-quality, far-ultraviolet echellette spectra available in the HST data archive. Of the 96 QSOs, 13 were observed using STIS and the E140M grating and 83 were observed using COS and the G130M/G160M gratings. STIS and the E140M grating offers a contiguous spectral coverage from $\approx 1140 \text{ \AA}$ to $\approx 1740 \text{ \AA}$ with a full-width-at-half-maximum resolution of $\text{FWHM} \approx 7 \text{ km s}^{-1}$, while COS with G130M and G160M at different central wavelengths offers a nearly contiguous spectral coverage from $\approx 1150 \text{ \AA}$ to $\lambda \approx 1770 \text{ \AA}$ with a spectral resolution of $\text{FWHM} \approx 16 \text{ km s}^{-1}$.

Individual flux-calibrated spectra from STIS were retrieved from the MAST server. For each QSO, individual echelle orders were continuum-normalized and co-added to form a single, stacked spectrum using our own software. The continuum was determined using a

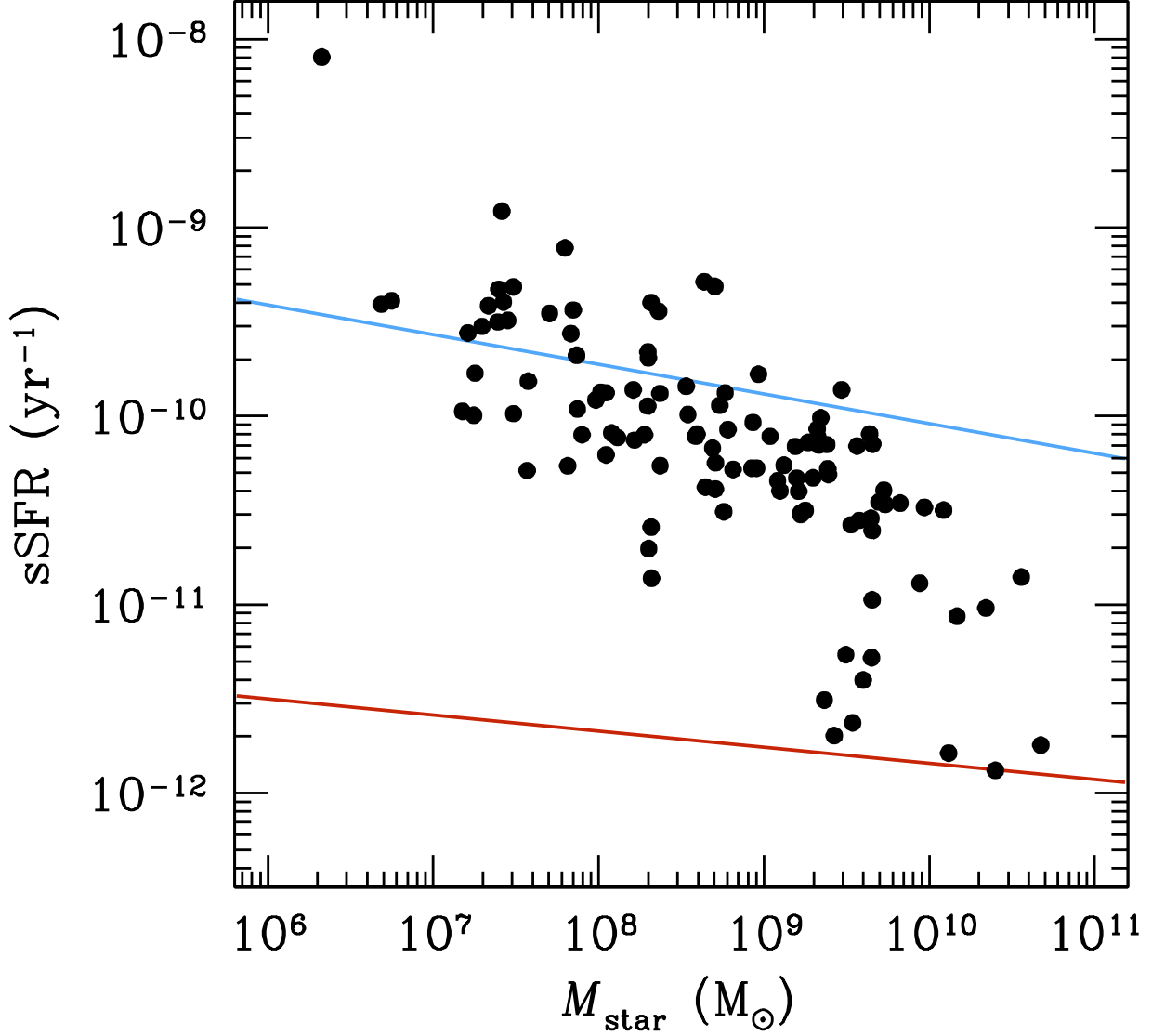


Figure 2.4: Distribution of specific star formation rate (sSFR) and stellar mass based on a subsample of our galaxies with available UV photometric from the NASA-Sloan Atlas (solid points). SFR is derived based on the observed NUV flux from GALEX and therefore represents a lower limit to the intrinsic value. The blue (red) line marks the star-forming main-sequence (passive, red) galaxies from Schiminovich et al. (2010). The tilt in the distribution of our galaxies relative to the mean relation of star-forming main-sequence galaxies can be explained by more massive galaxies showing higher degrees of dust extinction (e.g., Zahid et al., 2013). Nevertheless, the majority of our galaxies are consistent with being star-forming main-sequence galaxies, and a small fraction are found in the red-sequence regime.

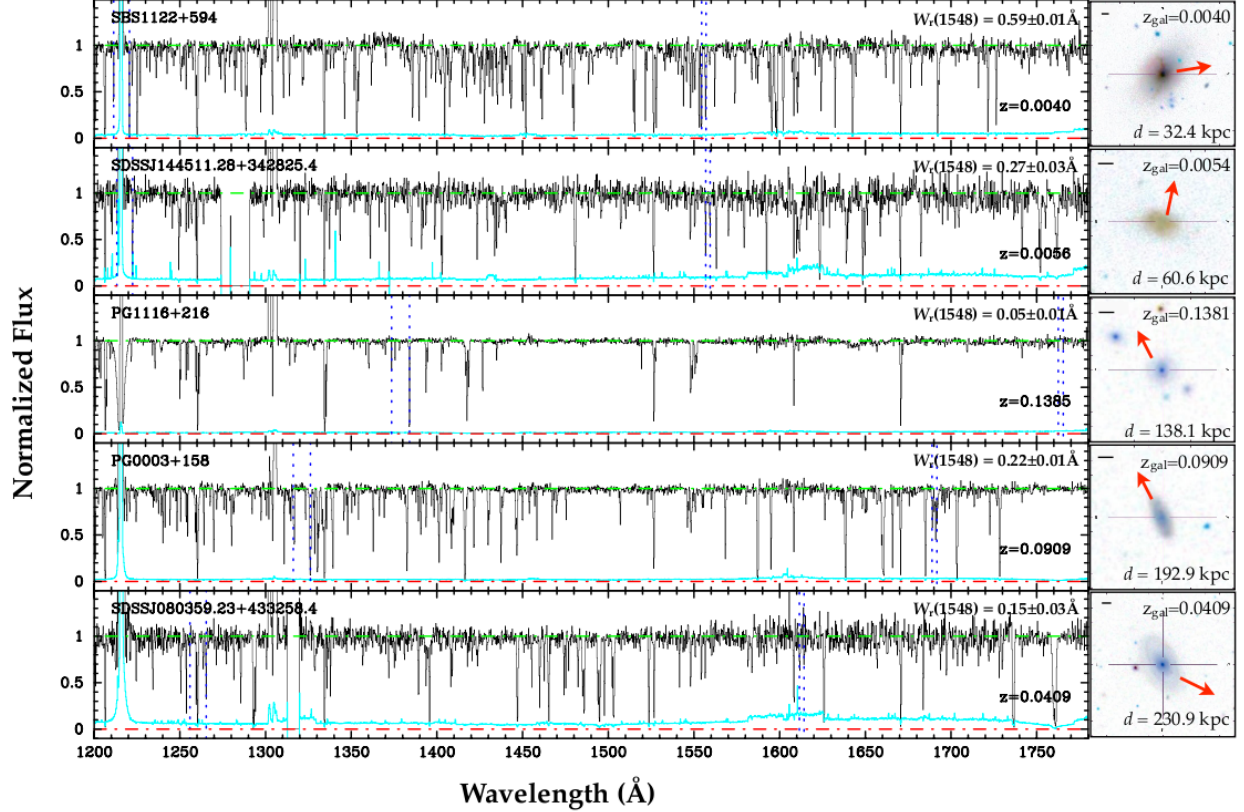


Figure 2.5: Examples of galaxies with associated CIV absorbers from our search. For each field, we present the QSO spectrum in the left panel and the corresponding SDSS image centered at the galaxy in the right panel. In each panel, the direction to the QSO sightline is indicated by the arrow, and the horizontal bar in the upper-left corner indicates an angular scale of $5''$. The projected separation between the QSO sightline and the galaxy is shown in the lower-right corner. Absorption features due to Si III λ 1206, Ly α λ 1215, and CIV $\lambda\lambda$ 1548, 1550 at the redshifts of the galaxies are indicated by blue, dotted lines in the left panels. The absorber redshift is shown in the lower right corner of each spectrum.

low-order polynomial fit to spectral regions that are free of strong absorption features. The mean signal-to-noise (S/N) of the combined echelle spectra of 13 QSOs have on average $S/N > 7$ over the majority of the spectral regions covered by the data.

Individual one-dimensional COS spectra were retrieved from the HST archive and combined using our own software. In summary, individual spectra were first aligned using common absorption lines along each sightline, including both intervening absorbers at $z > 0$ and known Milky Way features, such as Si III λ 1206 and C II λ 1334 in the G130M data and Si II λ 1526 and Al II λ 1670 in the G160M data. In a few sightlines, the pipeline generated individual spectra exhibit wavelength calibration errors that vary with wavelength, from $\Delta\lambda \approx -0.06 \text{ \AA}$ at $\lambda = 1160 \text{ \AA}$ to $\Delta\lambda \approx +0.06 \text{ \AA}$ at $\lambda \approx 1600 \text{ \AA}$. We correct for the wavelength calibration errors by employing a low-order polynomial fit to the observed wavelength-dependent shift in individual spectral segment. Then the individual wavelength-corrected spectra were re-binned to a pixel resolution of $\delta v = 7.5 \text{ km s}^{-1}$ and coadded into a single combined spectrum using our own co-addition code.

The final processed and combined spectra have a median signal-to-noise of $S/N \approx 2 - 26$. A summary of the STIS and COS observations for each QSO is presented in Table 2, which lists from columns (1) through (5) the name, RA, Dec, emission redshift of the QSO, and the program ID under which the data were taken. We also list in columns (6) and (7) of Table 2 the median S/N of the G130M and G160M spectra from COS and in column (8) the median S/N of the STIS/E140M spectra.

To facilitate absorption-line measurements, we also continuum normalized each QSO spectrum using a continuum model determined from a low-order polynomial fit to spectral regions that were free of strong narrow line features. The continuum normalization included strong damping wings in each QSO spectrum due to either the Milky Way Ly α absorption or intervening damped Ly α absorbers.

We did not attempt to fit strong emission features such as the geocoronal Ly α and O I $\lambda\lambda$ 1302, 1304 lines. These narrow emission lines were instead manually masked and were

not included in spectral regions for identifying absorption features.

2.4 Analysis

The process described in § 3 yielded high-quality, continuum-normalized UV echellette spectra of 96 QSOs. These UV spectra allow us to investigate the baryon content of the circumgalactic space within 500 kpc in projected distance of 195 independent, foreground galaxies based on the presence/absence of the corresponding absorption-line features, in particular the CIV doublet. In this section, we describe absorption-line constraints of individual galactic halos. In addition, we present stacked QSO spectra at the rest-frame of the galaxies, and improved absorption-line constraints afforded by the stacked spectra. We show that with much improved S/N , these stacked spectra allow us to place unprecedented limits for the absorption properties of gaseous halos around galaxies.

2.4.1 Absorption Constraints of Circumgalactic Baryons

In order to obtain robust constraints for the baryon content in halos around our sample galaxies, we first identify and mask contaminating features associated with other strong Ly α absorbers at $z \leq z_{\text{QSO}}$ along each QSO sightline, including higher-order Lyman series and ionic absorption lines. Then for each of the 195 galaxies, we search in the associated QSO spectra for the corresponding Ly α absorption line within the velocity range of $|v| \leq 500$ km s $^{-1}$ from the systemic redshifts of the galaxies. The allowed large velocity interval is guided by previous observations (e.g., Lanzetta et al., 1995) which show a velocity dispersion of ~ 190 km s $^{-1}$ between galaxies and their associated Ly α absorbers. For each detected Ly α absorber, we further examine whether associated ionic transitions, such as CIV, CII, SiIV, SIII, and SiII, are also present in the spectral range covered by the STIS/COS data. Absorption transitions considered in our study are summarized in Table 3, together with those considered by Werk et al. (2013) and Steidel et al. (2010) for comparison. We also

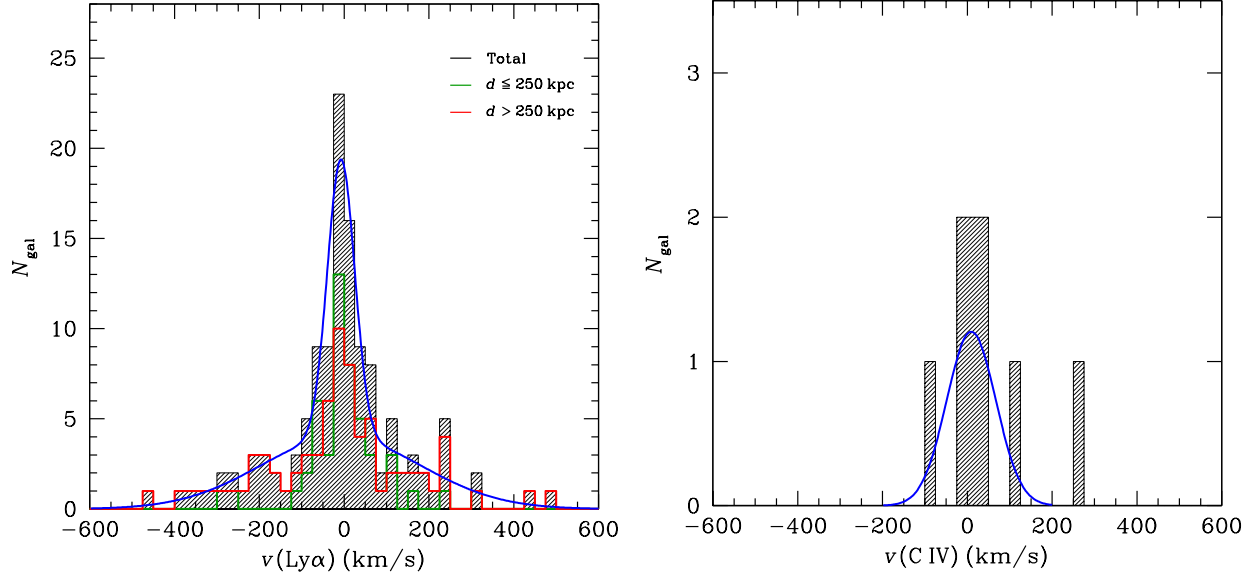


Figure 2.6: Relative velocity distributions of Ly α (*left*) and CIV (*right*) absorbers with respect to the systemic redshift of their associated galaxies in our pair sample. We detect associated Ly α absorbers in 125 of the 195 galaxies searched, and associated CIV absorbing gas in nine galaxies. The left panel shows that the velocity distributions of all Ly α absorbers detected around galaxies at $d < 500$ kpc (shaded histogram), around galaxies at $d < 250$ kpc (green open histogram), and galaxies at $d > 250$ kpc. Considering the full sample requires a double Gaussian profile to characterize the velocity distribution with a narrow component centered at $\langle v(\text{Ly}\alpha) \rangle = -7 \text{ km s}^{-1}$ and dispersion of $\sigma_v(\text{Ly}\alpha) = 32 \text{ km s}^{-1}$ and a broad component centered at $\langle v(\text{Ly}\alpha) \rangle = -8 \text{ km s}^{-1}$ and dispersion of $\sigma_v(\text{Ly}\alpha) = 188 \text{ km s}^{-1}$ (the solid curve). In contrast, the velocity distribution of CIV absorbing gas around galaxies can be characterized by a single Gaussian distribution of $\langle v(\text{CIV}) \rangle = 9 \text{ km s}^{-1}$ and $\sigma_v(\text{CIV}) = 58 \text{ km s}^{-1}$, excluding the outlier at $v(\text{CIV}) \approx +260 \text{ km s}^{-1}$.

include in Table 3 the sample size, redshift range, and stellar mass range of each galaxy sample for reference.

We detect associated $\text{Ly}\alpha$ absorbers around 125 of the 195 galaxies in our survey sample, and associated C IV absorbing gas in nine galaxies. Roughly 30% of the detected $\text{Ly}\alpha$ absorbers show multiple components within a velocity interval of 500 km s^{-1} and only one of the nine detected C IV displays an additional secondary component at $\Delta v = -87 \text{ km s}^{-1}$. No $\text{Ly}\alpha$ (C IV) measurements can be made for 27 (35) galaxies due to the presence of contaminating features. This leaves 43 (151) galaxies that show no trace of $\text{Ly}\alpha$ (C IV) absorbing gas at $d \leq 500 \text{ kpc}$ and $|v| \leq 500 \text{ km s}^{-1}$. Examples of the galaxies in our sample that display associated C IV absorbers are presented in Figure 5, which also highlights the presence of associated $\text{Si III } \lambda 1206$ and $\text{Ly}\alpha \lambda 1215$ in blue dotted lines.

To characterize the absorber properties, we measure the rest-frame absorption equivalent widths for all the observed transitions. We focus our current analysis on absorption equivalent width measurements in order to make direct comparisons with previous studies. A more detailed profile analysis is deferred to a future paper. In cases where an absorption transition is not detected, we measure a $2\text{-}\sigma$ upper limit to the rest-frame absorption equivalent width over a spectral window that corresponds to the median line width (FWHM) of individual, observed components of the targeted transition from the full sample. For galaxies without associated $\text{Ly}\alpha$ absorption, the upper limits of the underlying $\text{Ly}\alpha$ and C IV absorption strengths are estimated at the systemic redshifts of the galaxies. For galaxies with associated $\text{Ly}\alpha$ absorption but no detectable C IV features, the upper limits of the underlying C IV absorption strengths are estimated at the systemic velocities of the $\text{Ly}\alpha$ absorbers which are the velocity centroids of the strongest components determined by the fitting routine described below.

To measure the line widths and velocity centroids of each detected absorber, we model the observed absorption profile as a collection of individual Gaussian components and consider only the minimum number of components necessary to fully describe the observed absorption

features. The fitting routine returns for each component the velocity centroid, the $1\text{-}\sigma$ line width, and the maximum absorption depth. Considering all the individual $\text{Ly}\alpha$ components observed in 125 galaxies, we measure a median component line width of $\text{FWHM}(\text{Ly}\alpha) = 90 \text{ km s}^{-1}$. Considering all the individual C IV components observed in nine galaxies, we measure a median component line width of $\text{FWHM}(\text{C IV}) = 49 \text{ km s}^{-1}$. We note that the observed median line widths are significantly larger than the line width expected of thermal broadening in a warm gas of $T \sim 4 \times 10^4 \text{ K}$, which is $\approx 7.5 \text{ km s}^{-1}$ for C IV absorbing gas. The observed line widths are therefore likely driven by the gas dynamics in galactic halos.

We present in Figure 6 the relative velocity distribution of $\text{Ly}\alpha$ and C IV absorbers with respect to the systemic redshifts of the galaxies. Here we also adopt the velocity centroid of the strongest component found in each absorber as the systemic velocity of the absorber. The left panel of Figure 6 shows that the velocity distribution of $\text{Ly}\alpha$ absorbing gas around galaxies is best represented by a double Gaussian profile with a narrow component centered at $\langle v(\text{Ly}\alpha) \rangle = -7 \text{ km s}^{-1}$ and dispersion of $\sigma_v(\text{Ly}\alpha) = 32 \text{ km s}^{-1}$ and a broad component centered at $\langle v(\text{Ly}\alpha) \rangle = -8 \text{ km s}^{-1}$ and dispersion of $\sigma_v(\text{Ly}\alpha) = 188 \text{ km s}^{-1}$ (the solid curve). Excluding the outlier at $v(\text{C IV}) \approx +260 \text{ km s}^{-1}$, the velocity distribution of C IV absorbing gas around galaxies can be characterized by a single Gaussian distribution of $\langle v(\text{C IV}) \rangle = 9 \text{ km s}^{-1}$ and $\sigma_v(\text{C IV}) = 58 \text{ km s}^{-1}$ (right panel of Figure 6).

However, we note that the sample size of C IV absorbing galaxies is small. It is possible that the difference seen between $\text{Ly}\alpha$ and C IV absorbing gas in Figure 6 is driven by sampling uncertainties. To test this possibility, we randomly generated a large number of subsamples of nine $\text{Ly}\alpha$ absorbers from the parent sample of 125 $\text{Ly}\alpha$ absorbers, and performed a Kolmogorov–Smirnov (KS) test to compare each subsample of $\text{Ly}\alpha$ absorbers with the full sample of C IV absorbers. Roughly 1/3 of the time, the KS test gives $> 95\%$ chance that the two samples are drawn from the same underlying parent population. Whether or not $\text{Ly}\alpha$ and C IV absorbers share the same origin remains to be tested with a larger sample of C IV absorbing galaxies.

Considering the Ly α distribution alone, we find that the broad component is dominated by galaxies at $d > 250$ kpc (red open histogram in the left panel Figure 6). We therefore hypothesize two different origins of the detected Ly α absorbers: (1) halo gas associated with the galaxies contributing primarily to the narrow Gaussian distribution component and (2) intergalactic gas associated with correlated large-scale filaments that dominate the broad wings. We suggest that if larger samples of CIV absorbers were found to lack a broad velocity component, that would further support this scenario.

We also repeat the search for C II λ 1334, Si II λ 1260, Si III λ 1206, and Si IV λ 1393 absorption features associated with these galaxies. The observed median line width for C II absorbing gas is FWHM ≈ 67 km s $^{-1}$, while the observed median line widths for silicon ions are FWHM $\approx 27, 36$, and 47 km s $^{-1}$ for Si II, Si III, and Si IV, respectively. We therefore adopt these velocity intervals for measuring 2- σ upper limits for respective ionic transitions⁴. The observed absorption properties of the gaseous halos around individual galaxies are summarized in Table 4, which we list for each galaxy the angular separation θ from the QSO sightline, the projected proper distance d , redshift z_{gal} , the redshift of the associated Ly α absorber if detected, the measured rest-frame absorption equivalent widths or 2- σ upper limits of different transitions.

2.4.2 Absorption Properties of Individual Galactic Halos

To examine the relation between the incidence and strength of CGM absorption features and galaxy properties, we present in Figure 7 the observed distribution of rest-frame absorption equivalent widths of Ly α , C II, and C IV versus projected distance of the associated galaxy for all galaxies in our galaxy sample. The results can be summarized as follows.

First, robust constraints in Ly α absorption can be determined for 168 galaxies (solid

4. The velocity window adopted here for measuring upper limits of non-detections is empirically determined based on detected components, and is on average 4–6 times smaller than the velocity window adopted by Werk et al. (2013) for determining their upper limits. As a result, the reported upper limits in Werk et al. (2013) are roughly 3 times higher than what we determine here.

points in the upper-left panel of Figure 7), 43 of which do not have a detectable Ly α absorption feature to a sensitive upper limit (grey, solid points with arrows). All 12 galaxies at $d < 100$ kpc exhibit a corresponding Ly α absorber, leading to a 100% mean covering fraction of H I absorbing gas within 100 kpc of star-forming regions. Beyond $d = 100$ kpc, the incidence of Ly α absorption remains high with only 43 of 156 galaxies showing no corresponding Ly α absorption features of $W_r(1215) \gtrsim 0.1 \text{ \AA}$, leading to a mean gas covering fraction of $> 70\%$ over the projected distance interval of $100 - 500$ kpc from galaxies. In addition, the observed mean Ly α absorption strength gradually declines with increasing d , consistent with the mean relation reported by Chen et al. (1998, 2001a; solid line). For comparisons, we include measurements and upper limits for $\sim L_*$ galaxies at $z = 0.14 - 0.36$ from the COS-Halos program Werk et al. (2013) and for starburst galaxies at $z = 2.2$ from Steidel et al. (2010).

Second, robust constraints in C II absorption can be determined for 141 galaxies (solid points in the middle-left panel of Figure 7), 134 of which do not have a detectable C II absorption feature to a sensitive upper limit (grey, solid points with arrows). Three of the four galaxies at $d \lesssim 60$ kpc exhibit associated C II absorption and none of the eight galaxies at $d = 60 - 100$ kpc show a corresponding C II absorber of $W_r(1334) > 0.04 \text{ \AA}$. This leads to a mean covering fraction of $\approx 67\%$ in C II absorbing gas at $d \lesssim 60$ kpc of star-forming regions and 25% at $d < 100$ kpc. Beyond $d = 100$ kpc, we detect C II absorption near four additional galaxies at $d = 106, 116, 138$, and 225 kpc, and the remaining galaxies do not show associated C II absorber to $2\text{-}\sigma$ upper limits of better than $W_r(1334) > 0.1 \text{ \AA}$. The lack of observed C II absorption leads to a strong constraint on the mean C II absorbing-gas covering fraction of $\approx 3\%$ over the projected distance interval of $100 - 500$ kpc from galaxies. Likewise, we include measurements and upper limits for $\sim L_*$ galaxies at $z = 0.14 - 0.36$ from the COS-Halos program Werk et al. (2013) and for starburst galaxies at $z = 2.2$ from Steidel et al. (2010) for comparisons.

A possible explanation for a substantially lower incidence of C II absorption at $d > 100$

kpc is an increased ionizing radiation field that offsets a larger fraction of carbon atoms into higher ionization states. We examine the incidence of C IV absorption features in the bottom-left panel of Figure 7, where we show that robust constraints in C IV absorption can be determined for 160 galaxies (solid points) and 151 of these do not have a detectable C IV absorption feature to a sensitive upper limit (grey, solid points with arrows). Of the 12 galaxies at $d < 100$ kpc, four show a corresponding C IV absorber, leading to a mean covering fraction of $\approx 33\%$ in C IV absorbing gas within 100 kpc of star-forming regions. Beyond $d = 100$ kpc, five galaxies show associated C IV absorption and the rest do not show corresponding C IV absorption of $W_r(1548) \gtrsim 0.1 \text{ \AA}$, leading to a mean C IV absorbing-gas covering fraction of $\approx 3.4\%$ over the projected distance interval of 100 – 500 kpc from galaxies. For comparisons, we include measurements and upper limits for $\sim L_*$ galaxies at $z \approx 0.4$ from (Chen et al. 2001b; open triangles) and for starburst galaxies at $z = 2.2$ from Steidel et al. (2010; pluses). Note that the C IV doublets were not resolved in the stacked spectra of Steidel et al. (2010). We therefore display a possible range of $W_r(1548)$ based on two extremes in the assumed line ratio (e.g. 2:1 or 1:1) between the C IV $\lambda 1548$ and C IV $\lambda 1550$ members.

The QSO spectra also allow us to examine the spatial distribution of silicon ions in the CGM. Figure 8 displays the observed absorption strength of Si II $\lambda 1260$, Si III $\lambda 1206$, and Si IV $\lambda 1393$ versus projected distance of the associated galaxy for all galaxies in our sample. We find that the Si II $\lambda 1260$ transition appears to be contaminated for 37 galaxies, and that only three of the remaining 158 galaxies show associated Si II absorption at $d = 39, 115$, and 138 kpc. (upper-left panel of Figure 8). With the exception of one galaxy at $d = 497$ kpc, for which no sensitive limit can be placed due to the poor quality of the data, none of the other 154 galaxies show associated Si II absorption to sensitive limits better than $0.06 \text{ \AA}$. In addition, the Si III $\lambda 1206$ transition appears to be contaminated for 62 galaxies, and eight of the remaining 135 galaxies show associated Si III absorption. Of the 10 galaxies at $d < 100$ kpc for which constraints for Si III absorption can be obtained, four exhibit associated Si III

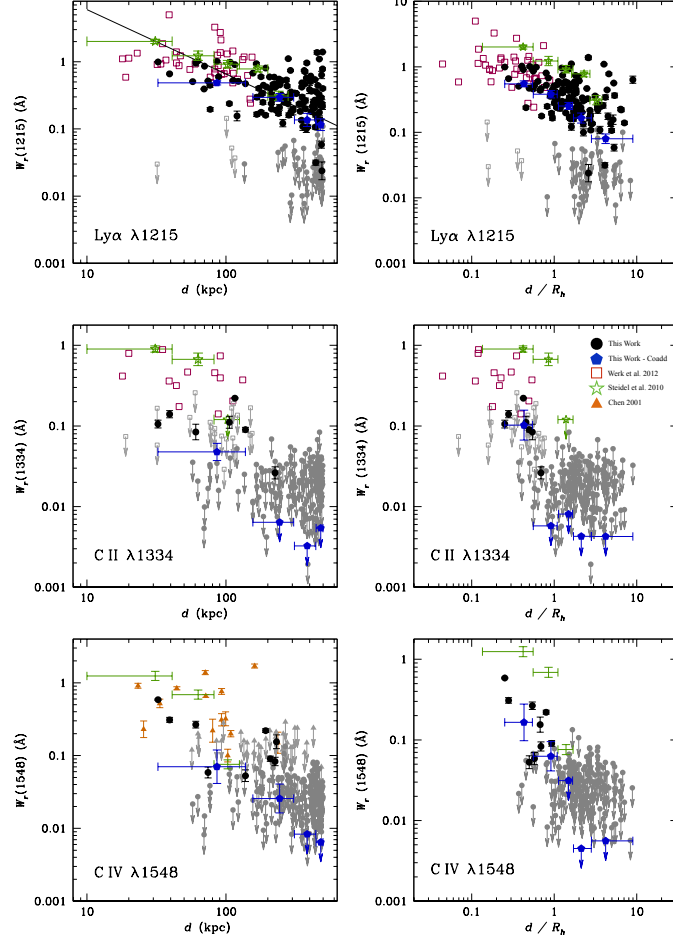


Figure 2.7: Observed spatial distribution of rest-frame absorption equivalent widths of Ly α (top), C II (middle), and C IV (bottom) in the circumgalactic space. The left panels show the variation of absorber strengths versus projected distance d , while the right panels show the spatial variation as a function of halo-radius (R_h) normalized projected distance, d/R_h . Galaxies with no detectable absorption are shown as $2\text{-}\sigma$ upper limits (points with downward arrows) that are determined based on the noise characteristics (see § 4.1). Measurements from our study are shown in solid points. Measurements for $\sim L_*$ galaxies at $z = 0.14 - 0.36$ from the COS-Halos program (Werk et al., 2013) are shown in open squares. For starburst galaxies at $z = 2.2$ (Steidel et al., 2010) constraints on Ly α and C II absorption are shown in star symbols. But because C IV doublets were not resolved in this study, a range of $W_r(1548)$ is shown in vertical bars based on assumed doublet ratios of 2:1 and 1:1. The measurements at $z = 2.2$ are from stacked spectra of galaxies covering a range in stellar mass (Reddy et al., 2012). Previous measurements of C IV absorption in the low-redshift CGM from Chen et al. (2001b) are shown in triangles. The best-fit power-law model of Chen et al. (1998, 2001a) to describe the observed $W_r(1215)$ versus d anti-correlation is shown in solid line. Finally, mean absorption strengths determined from stacked CGM spectra of the full sample (§ 4.4) are shown in (blue) pentagons with the horizontal bars indicating the bin size.

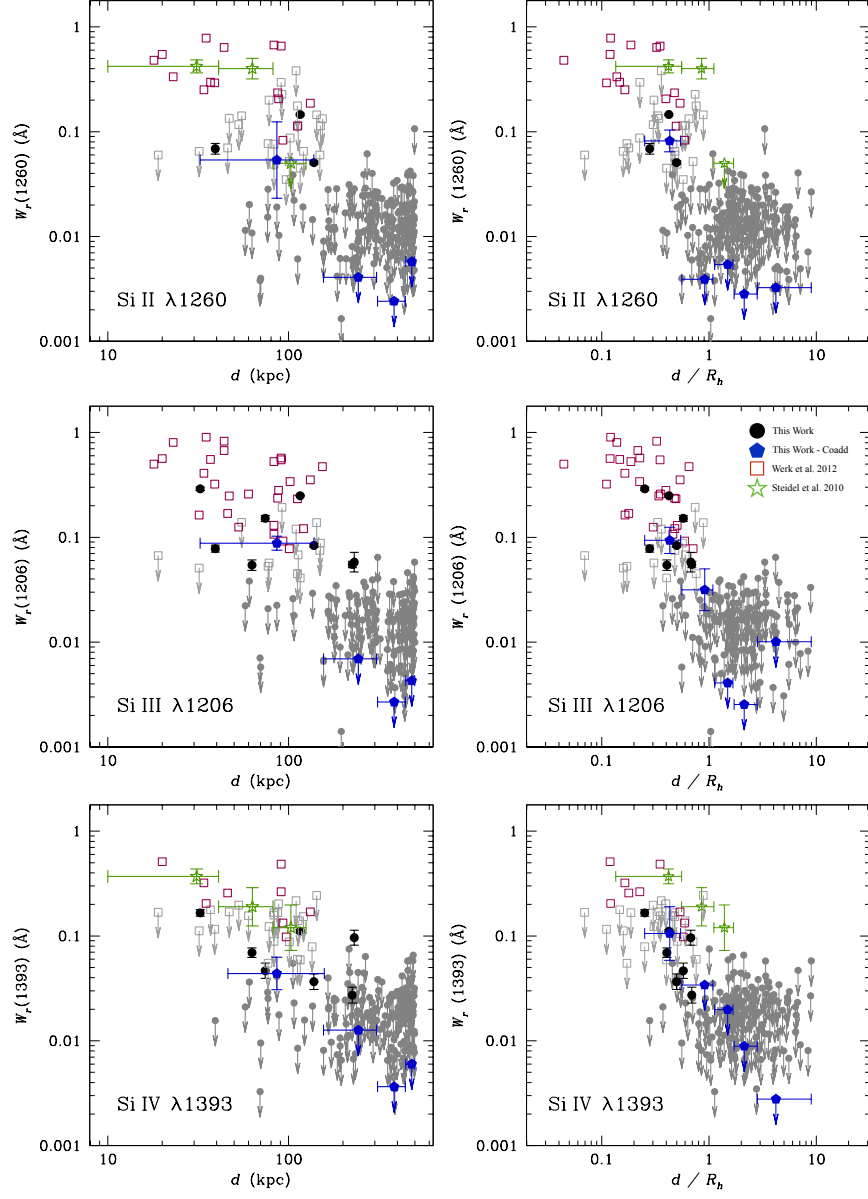


Figure 2.8: Similar to Figure 6 but for the observed spatial distribution of rest-frame absorption equivalent widths of Si II (top), Si III (middle), and Si IV (bottom) in the circumgalactic space. Note that the constraints for non-detections in the COS-Halos sample appear to be worse than our typical upper limits, because the COS-Halos team adopted a much larger velocity window for measuring their upper limits (i.e. $\approx 300 \text{ km s}^{-1}$ adopted for the COS-Halos galaxies versus $\approx 50 \text{ km s}^{-1}$ adopted in our study). The velocity window adopted in our study is empirically determined based on detected components, and is therefore more appropriate for constraining the presence/absence of cool, photo-ionized clouds in the circumgalactic space.

absorption, leading to a mean Si III covering fraction of $\approx 40\%$ (middle-left panel of Figure 8). At $d > 100$ kpc, four of the 125 galaxies exhibit associated Si III, leading to a mean covering fraction of $\approx 3\%$. Finally, robust constraints in the Si IV $\lambda 1393$ absorption can be determined for 150 galaxies, seven of these exhibit associated Si IV absorption with four of the detections found at $d > 100$ kpc (bottom-left panel of Figure 8). None of the other 143 galaxies show associated Si IV absorption to sensitive limits better than $0.07 \text{ \AA}$.

The left panels of Figures 7 & 8 display a stark contrast between the spatial distributions of hydrogen atoms and heavy elements. While hydrogen gas is observed all the way out to 500 kpc in projected distance, the same galaxies that display moderately strong H I absorption do not exhibit associated heavy elements at $d > 200$ kpc. The lack of heavy elements at large distances persists through all ionization states examined in our study, from low-ionization species such as C^+ and Si^+ ions to high-ionization species such as C^{3+} and Si^{3+} ions. It indicates that if heavy ions are present in the outer halos, then they remain in a tenuous, hot phase in which gaseous clumps giving rise to the observed absorption lines cannot survive. Similar to Figure 7, we include previous measurements and upper limits for $\sim L_*$ galaxies at $z = 0.14 - 0.36$ from the COS-Halos program (Werk et al., 2013, open squares) and for starburst galaxies at $z = 2.2$ from (Steidel et al., 2010, star symbols) for comparisons.

2.4.3 Mass Scaling Of Extended Gaseous Halos

In addition to the distinct spatial extent between hydrogen gas and heavy ions, the left panels of Figure 7 & 8 also display significant scatter in the W_r versus d space. In particular, different galaxy samples appear to segregate in different parts of the W_r - d parameter space. Both the $z = 2.2$ starburst galaxy sample and the COS-Halos sample contain primarily galaxies at $d < 100$ kpc, which exhibit on average stronger absorption in $\text{Ly}\alpha$ and in ionic transitions. Such distinction appears particularly strong in C II and Si II absorption strengths.

Recall that our galaxies span a broad range of stellar mass with 50% of the sample characterized as low-mass dwarf galaxies of $M_{\text{star}} < 0.2 M_{\text{star}}^*$, while the COS-Halos galaxies

are primarily M_{star}^* galaxies (Werk et al., 2013) and the $z = 2.2$ starburst sample from Steidel et al. (2010) has a median stellar mass that lies between the COS-Halos and our samples (e.g., Reddy et al., 2012, see also Figure 3). It has been shown empirically at low redshifts that more massive galaxies on average possess more extended gaseous halos (e.g., Chen et al., 2010b). This is also expected by theoretical models (Mo & Miralda-Escude, 1996; Maller & Bullock, 2004; Ford et al., 2013). As a first step toward establishing a clear understanding of the processes that drive the observed CGM properties, it is therefore necessary to investigate whether the apparent distinction in the observed gaseous extent of different galaxy samples can be explained by the intrinsic size differences of individual galactic halos. The results will facilitate future studies that consider the effect of additional processes such as star formation feedback.

To address the possible mass-scaling of gaseous radius, we first estimate the dark matter halo mass M_h for each galaxy based on its known stellar mass M_{star} from the NASA-Sloan Atlas and the stellar mass-halo mass relation of Kravtsov et al. (2014) at $z \lesssim 0.1$. We then compute the halo radius R_h , following the prescription of Bryan & Norman (1998).

$$M_h = \frac{4\pi}{3} \Delta_h \langle \rho_m \rangle(z) R_h^3, \quad (2.2)$$

where $\langle \rho_m \rangle(z)$ is the mean matter density of the Universe at z , Δ_h represents the overdensity over which a halo is defined, and $\Delta_h = 18\pi^2 + 82x - 39x^2$ with $x \equiv \Omega_M (1+z)^3 [H_0/H(z)]^2 - 1$.

1. The halo radius is then computed according to

$$R_h = \frac{261.3}{(1+z)} \left(\frac{\Delta_h(z) \Omega_M}{97.2} \right)^{-1/3} \left(\frac{M_h}{10^{12} \text{ M}_\odot} \right)^{1/3} \text{ kpc}. \quad (2.3)$$

Lastly, we divide the projected distance of each galaxy by its halo radius and examine how the observed absorption strengths of different transitions vary with R_h -normalized projected distance, d/R_h . The results are shown in the right panels of Figures 7 & 8.

For comparison, we also calculate the halo radius for each COS-Halos galaxy using the

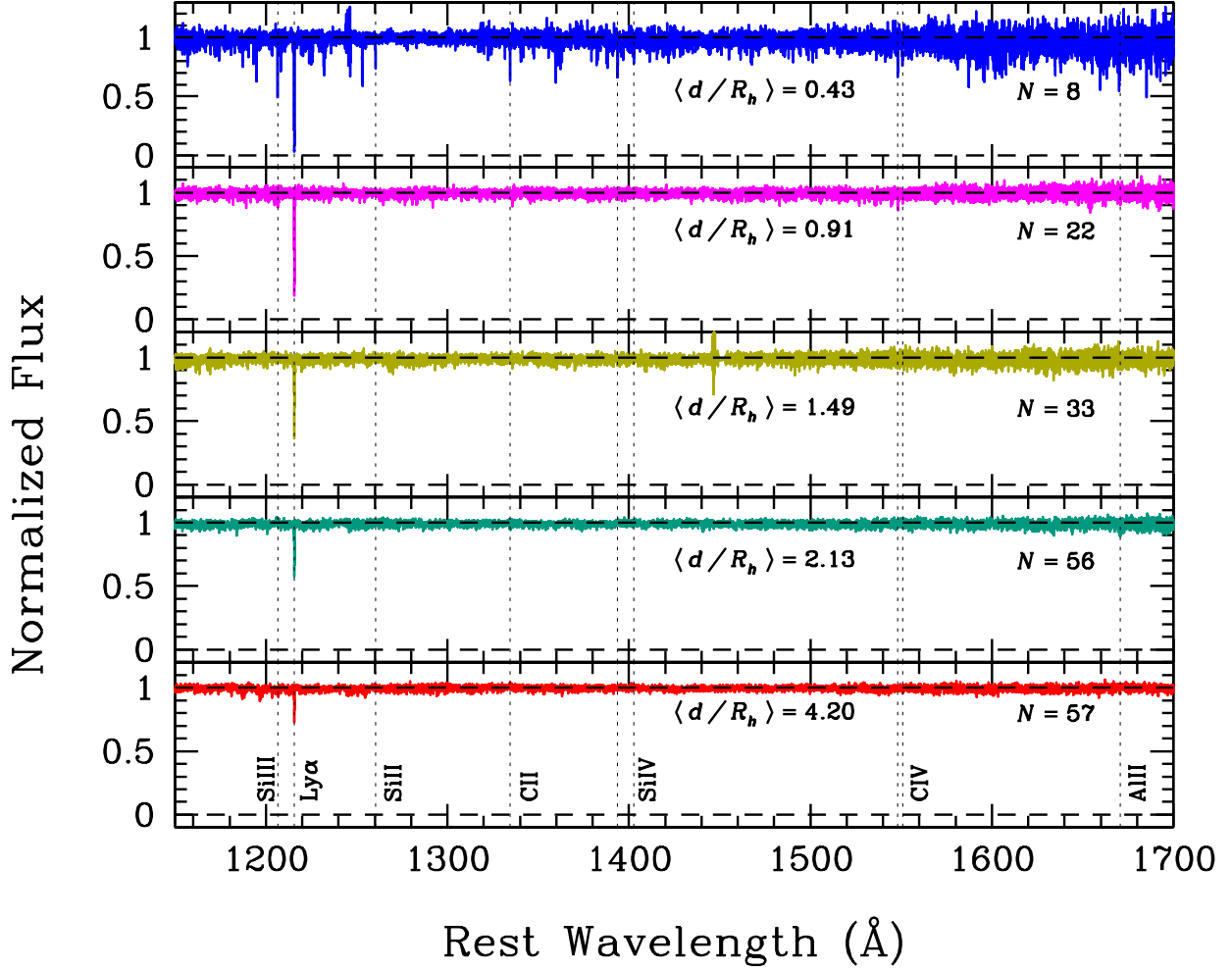


Figure 2.9: Stacked QSOs spectra in different R_h -normalized projected distance intervals from host galaxies. Galaxies at $d < 1.7 R_h$ are grouped into three bins following the binning of the high-redshift starburst sample. Galaxies at $d > 1.7 R_h$ are divided into two d/R_h bins with roughly equal number of galaxies. The median R_h -normalized projected distance and the number of galaxies included in each stack are indicated in each panel. The stacked spectra clearly show declining Ly α absorption strengths with increasing projected distance, and no ionic transitions are detected with $d/R_h > 0.7$.

published M_{star} from Werk et al. (2013) normalize the galaxy projected distance by the corresponding R_h , and include their absorption measurements in the R_h -normalized panels. In addition, we adopt the median halo mass of $\langle \log (M_{\text{star}} / M_{\odot}) \rangle = 9.9 \pm 0.5$ for the starburst galaxies at $z = 2.2$ from Reddy et al. (2012). Using a stellar mass-halo mass relation appropriate for high-redshift galaxies from Behroozi et al. (2013), we estimate a mean halo mass of $\log (M_h / M_{\odot}) = 11.6 \pm 0.2$ and a mean halo radius of $R_h \approx 74 \pm 13$ kpc⁵. Because the dispersion in R_h as a result of dispersion in M_{star} of the starburst sample is much less than the bin size adopted for d , we normalize their mean projected distance by a single mean halo radius and include their measurements in the right panels of Figures 7 & 8. We cannot include the earlier CIV measurements of Chen et al. (2001b) because the mass of each galaxy is unknown.

After accounting for a possible mass scaling of gaseous radius, we note three important features in the right panels of Figures 7 & 8. First, the QSOs in our pair sample probe regions as close as $0.2 R_h$ in galactic halos to $\approx 10 R_h$ outside of galactic halos. While Ly α absorbers are frequently observed at projected distances much beyond R_h , no ionic transitions are found at these large radii. The few detections at $d > 100$ kpc turn out to be associated with most massive galaxies in the sample and the QSO sightlines appear to still probe the inner halos of these galaxies. The large sample allows us for the first time to place stringent limits on the possible presence of heavy elements beyond galactic halos. We place a strict *upper limit* of 3% (at a 95% confidence level) on the incidence of $W_r > 0.05$ Å metal-line absorbers over the range of projected distances from R_h to $9 R_h$ (see § 4.5 and Figure 11 below).

Second, the COS-Halos sample probes the metal-enriched CGM over the projected distance range from $\approx 0.05 R_h$ to $\approx 0.5 R_h$. Recall that the median mass of the COS-Halos sample at $z = 0.1 - 0.4$ is $4 \times 10^{10} M_{\odot}$ (Werk et al., 2013), while our galaxy sample spans a

5. The inferred halo mass is a factor of two less than what is estimated by Steidel et al. (2010) or by Trainor & Steidel (2012) based on the observed galaxy clustering amplitude. For consistency, we adopt the halo mass estimate based on the stellar mass to halo mass relation.

broad range in stellar mass with a median $\langle M_{\text{star}} \rangle = 5 \times 10^9 M_{\odot}$. Therefore, the COS-Halos sample probes primarily the inner halos of massive galaxies and our galaxy sample extends to outer halos of lower- and high-mass galaxies. The observed absorption strengths versus R_h -normalized projected distance from the COS-Halos observations now connect smoothly with the measurements for our sample. These include H I, C II, Si II, Si III, and Si IV. The on average stronger absorbers found around COS-Halos galaxies can therefore be explained by higher gas density in the inner halos, at least around massive galaxies. Considering our sample with the COS-Halos sample together shows that both the absorption strengths and incidence (covering fraction) of heavy elements decline steeply beyond $0.3 R_h$ and no heavy ions are detected beyond $\approx 0.7 R_h$.

Finally, the starburst galaxy sample probes the metal-enriched CGM at $z = 2.2$ over the projected distance range from $\approx 0.1 R_h$ to $\approx R_h$. The mean absorption strength of the metal-enriched CGM around starburst galaxies at $z = 2.2$ is found to decline rapidly at beyond R_h . In addition, the observed mean absorber strengths at a fixed d/R_h remain to be stronger around these starburst galaxies than what is observed around either low-mass dwarf or high-mass galaxies at $z \sim 0$. The differential observed absorption strength versus R_h -normalized projected distance between low- and high-redshift galaxy samples suggests for the first time that the spatial absorption profile of the CGM may have evolved since $z = 2.2$ (cf., Chen, 2012). However, a direct comparison between the low- and high-redshift measurements is difficult because the high-redshift measurements were made in stacked spectra which presumably includes detections and non-detections among galaxies with a broad mass range (Table 3), and there was no distinction made between independent and group galaxies. An additional uncertainty also arises due to uncertainties in halo models (Diemer et al., 2013, see further discussions in § 5.3).

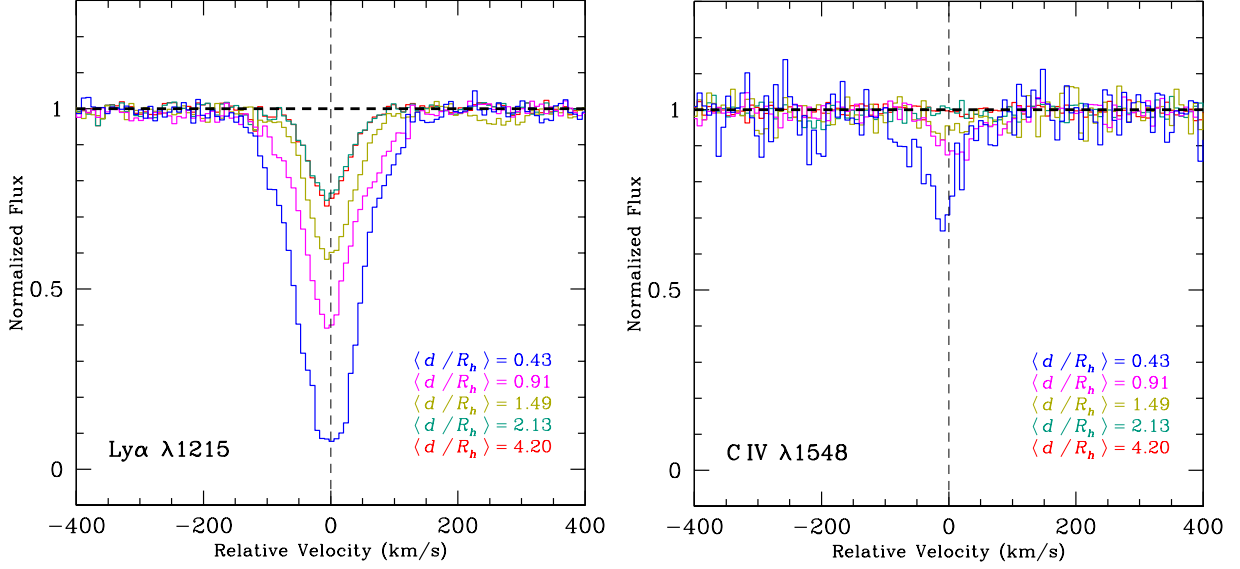


Figure 2.10: Mean absorption profiles of Ly α (left) and C IV (right) in stacked absorption spectra of galaxies in different d/R_h intervals.

2.4.4 Mean Absorption Profiles from Stacked CGM Spectra

To further improve upon the sensitivity in searching for weak absorption features and to facilitate a more direct comparison between our study and that of Steidel et al. (2010), we experiment with co-adding the absorption spectra of individual galaxies in the rest frame. We note, however, that interpreting the comparisons between low- and high-redshift measurements is not trivial (see § 5.3 for discussion).

For each input spectrum, we first mask contaminating features due to either the Milky Way interstellar medium or other strong absorbers at redshifts different from the foreground galaxy. Then each continuum normalized QSO spectrum and its associated error spectrum are shifted to the rest frame of the foreground galaxy. To account for the velocity offset between a galaxy and its associated absorbing gas (Figure 6), the rest frame of each galaxy is set to be the redshift of the observed Ly α absorber⁶. For galaxies with no detected Ly α absorbers, the rest frame is set to be the systemic redshift of the galaxy. Individual rest-

6. We have also experimented with adopting the systemic redshift of the galaxies as the rest frame. No significant changes are seen.

frame spectra are then weighted and stacked to form a final mean spectrum. The weighting factor of each spectrum is defined as the inverse median variance which is determined using the associated error spectrum over the wavelength range from 1100 Å to 1450 Å in G130M data and from 1390 Å to 1750 Å in G160M data (e.g. columns 5 and 6 in Table 2).

We form different stacked CGM spectra for galaxies grouped in different projected distance intervals, as well as in different R_h -normalized projected distance intervals. Given that a large fraction of detected heavy ions occur at $d < 150$ kpc (left panels of Figures 7 & 8), we have adopted the first bin in projected distances to include galaxies at $d < 150$ kpc and divided the remaining galaxies into three projected distance intervals with roughly equal number ($N \approx 50$) of galaxies. For R_h -normalized measurements (right panels of Figures 7 & 8), we have adopted the same binning as the high-redshift starburst galaxy sample at $d < 1.7 R_h$ and divided the rest of the galaxy sample into two d/R_h bins with roughly equal number ($N \approx 56$) of galaxies. The stacked CGM spectra for different R_h -normalized distance intervals are presented in Figure 9.

In each stacked CGM spectrum, we measure mean absorption strengths of different transitions. However, uncertainty estimates of the observed mean absorption strengths in stacked spectra are complicated for two main reasons. First, Figures 7 and 8 show that in each d and d/R_h bin there is a combination of detections and non-detections. The mean properties of each stacked spectrum are therefore likely dominated by sampling errors. In addition, we have adopted an inverse-variance weighted mean to form each stacked spectrum. The final stacks may be dominated by one or two high S/N sightlines. To assess these sampling errors, we employ a bootstrap resampling technique.

For each subsample, we randomly sample the galaxies with replacement to construct a new subsample that contains the same number of input galaxies for stacking. We then measure the mean absorption equivalent width or a $2\text{-}\sigma$ upper limit if no absorption features are found in the new stack. We repeat this exercise N_{boot} times and examine the distribution of the resulting mean absorption measurements. We find that a convergence is reached in the

distribution when $N_{\text{boot}} \gtrsim 350$. We adopt the $1\text{-}\sigma$ dispersion from the bootstrap resampling distribution as the uncertainties in the observed mean absorption in stacked spectra. In the case of non-detections, we quote the 95% upper bound of the distribution as the $2\text{-}\sigma$ upper limit to the underlying absorption strength. The measurements and associated uncertainty estimates, as well as upper limits for non-detections are included in Figures 7 and 8 as solid blue points. The results are also summarized in Tables 5 and 6, which record for each distance interval the range of projected distance d (or d/R_h), median projected distance $\langle d \rangle$ (or $\langle d/R_h \rangle$), the number of galaxies included in the stack, and the mean absorption strengths in $\text{Ly}\alpha$, Si III, Si II, Si IV, C II, and C IV. Both Tables 5 & 6 and Figure 9 confirm that with higher S/N in the stacked spectra the mean $\text{Ly}\alpha$ absorption strength declines steadily with increasing projected distance, while no heavy ions are detected at 95% upper limits of $W_0 \lesssim 0.03 \text{ \AA}$ at $d \gtrsim R_h$.

The strong upper limits presented in Tables 5 and 6 for non-detections demonstrate that using stacked spectra we are indeed able to improve the constraints for the presence/absence of absorbing gas by $\sim \sqrt{N_{\text{gal}}}$. No ionic transitions are found outside of galactic halos to unprecedentedly sensitive limits of better than $0.03 \text{ \AA}$ at the $2\text{-}\sigma$ level of significance. In addition, comparing the absorption-line measurements made in stacked CGM spectra confirms that massive starburst galaxies at $z = 2.2$ (green symbols in the right panels of Figures 7 & 8) exhibit significantly stronger mean CGM absorption than the present-day galaxies (solid blue points). The differences are most pronounced in low-ionization species traced by C II and Si II absorption features. This is in stark contrast to the comparable mean absorption strengths of Mg II absorbers found at fixed luminosity-normalized projected distances (Chen, 2012). The discrepancy is less for higher ionization species such as Si IV and C IV. Most interestingly, while we observe strong differential mean absorption strengths in different ionization states, such as between Si II and Si IV and between C II and C IV, the mean absorption strengths of different ions appear to be comparable at large distances from starburst galaxies at $z = 2.2$. The comparable strengths between low- and high-ionization

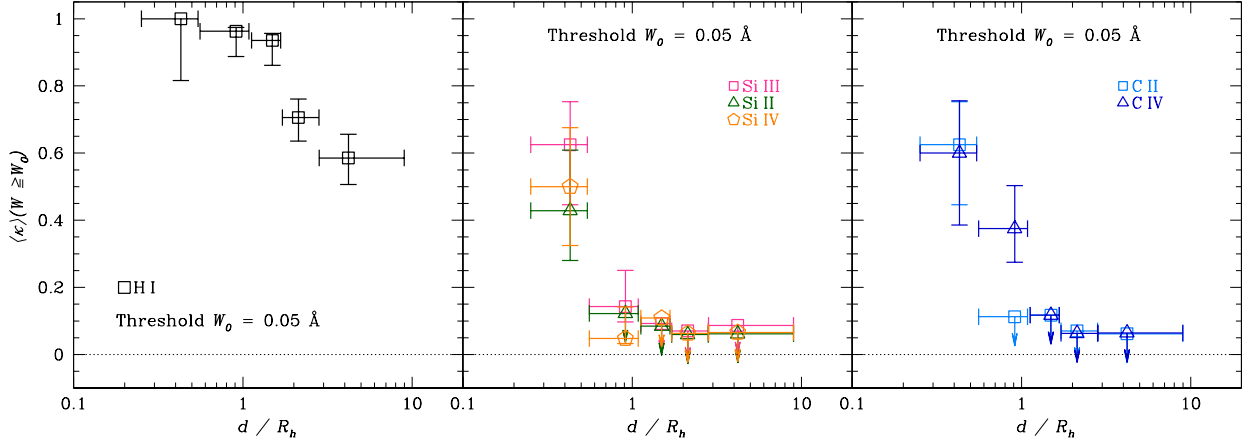


Figure 2.11: Mean gas covering fraction $\langle \kappa \rangle$ measured in each d/R_h bin for different species. H I is shown in the left panel, Si II, Si III, and Si IV are shown in the middle panel, and C II and C IV are shown in the right panel. Gas covering fraction is computed for a detection threshold of $W_0 = 0.05 \text{ \AA}$. Error bars represent the 68% confidence interval. The covering fraction of Ly α absorbing gas remains high at $\approx 60\%$ outside of dark matter halos at $d > R_h$, while the covering fraction of heavy ions is restricted to $\langle \kappa \rangle < 10\%$ in individual d/R_h bins. Considering *all* galaxies at $d > R_h$ further improves the limit on the covering fraction of heavy ions to $\langle \kappa \rangle < 2\%$ at a 95% confidence level. In inner halos ($d \lesssim 0.7 R_h$), we also find a declining mean covering fraction from high- to low-ionization species.

transitions suggests that either the ionization condition in the high-redshift CGM is very different or the lines are saturated (in contrast to the non-saturated lines observed in our low-redshift sample) and therefore the line widths are driven by the underlying velocity field of the gas.

2.4.5 Gas Covering Fraction

We have shown in § 4.4 that stacked CGM spectra enable strong limits on the possible presence of heavy elements outside of galactic halos, namely at projected distances $d > R_h$. Within galactic halos, however, the observed mean absorption strength in a finite distance interval is the product of the spatial absorption profile and partial gas covering. Figures 7 and 8 show that not only does the absorption equivalent widths of individual absorbers decrease with increasing distance, but the fraction of detections also decreases with increasing distance. An independent measure of the gas covering fraction is therefore necessary for an

accurate characterization of the spatial absorption profile of the CGM.

The large number of galaxies in our sample with sensitive limits available on the presence/absence of extended gas in individual galactic halos allows us to quantify how the covering fraction κ of absorbing gas varies as a function of projected distance based on an ensemble average. Here we consider galaxy subsamples defined in fixed d/R_h intervals in order to obtain a global average of how the gas covering fraction depends on halo radius-normalized distance.

Following the prescription described in Chen et al. (2010a), we employ a maximum likelihood analysis to estimate κ for different species. First, the probability that a galaxy inside a dark matter halo of radius R_h exhibits an absorber of rest-frame absorption equivalent width $W_r \geq W_0$ at projected distance d is written as

$$P(\kappa|W_0) = \kappa(d/R_h) H[r_2 - d/R_h] H[d/R_h - r_1] \quad (2.4)$$

where $H(x)$ is the Heaviside step function with $H = 1$ if $x > 0$ and $H = 0$ otherwise, and r_1 and r_2 together define the distance interval over which κ is calculated. The likelihood function of observing an ensemble of n galaxies with associated absorbers of $W_r \geq W_0$ and m galaxies with no absorbers detected to sensitive limits of $W_r < W_0$ can then be written as

$$\begin{aligned} \mathcal{L}(\kappa|W_0) &= \prod_{i=1}^n \kappa(d_i/R_h^i) H[r_2 - d_i/R_h^i] H[d_i/R_h^i - r_1] \times \\ &\quad \prod_{i=1}^m \left(1 - \kappa(d_i/R_h^i) H[r_2 - d_i/R_h^i] H[d_i/R_h^i - r_1] \right) \\ &= \langle \kappa \rangle^n (1 - \langle \kappa \rangle)^m. \end{aligned} \quad (2.5)$$

Equation (5) is simply a binomial distribution function of a sample of $m + n$ objects, in which n are detections, m are non-detections, and $\langle \kappa \rangle$ represents the mean probability of detecting an absorber that maximizes the likelihood $\mathcal{L}$. In the form of a likelihood function, we are able to determine the confidence level for the best fit $\langle \kappa \rangle$. We evaluate $\langle \kappa \rangle$ for

$W_0 = 0.05 \text{ \AA}$ by maximizing the likelihood function $\mathcal{L}$ for different absorption transitions. Figures 7 and 8 show that the majority of the STIS and COS spectra included in our analysis all have sufficient S/N to allow strong constraints for ruling out the presence of an absorber of $W_r > 0.05 \text{ \AA}$ for all transitions considered in our study. In addition, at this sensitive limit of $W_0 = 0.05 \text{ \AA}$, we are already probing tenuous gas of H I column density $\log N(\text{H I}) \lesssim 13$ or chemically enriched CGM of C IV column density $\log N(\text{C IV}) \lesssim 13.5$, assuming photo-ionized gas of $T \sim 4 \times 10^4 \text{ K}$ (e.g., Boksenberg et al., 2003). Figure 11 shows the estimated $\langle \kappa \rangle$ as a function of d/R_h for different transitions. Error bars in $\langle \kappa \rangle$ represent the 68% confidence interval. For non-detections, we present the 95% single-sided upper limit. The results are also summarized in Table 7.

Figure 11 clearly shows that the covering fraction of Ly α remains high ($\approx 60\%$) outside of dark matter halos at $d > R_h$, while the covering fraction of heavy ions is restricted to $\lesssim 10\%$ in individual d/R_h bins. Considering *all* galaxies at $d > R_h$ further improves the limit on the covering fraction of heavy ions to $\langle \kappa \rangle_{\text{ions}} < 3\%$. At $d \approx 0.6 R_h$, low-ionization transitions display a rapidly declining covering fraction from $\langle \kappa \rangle \gtrsim 50\%$ at $d = 0.2 - 0.6 R_h$ to $\langle \kappa \rangle = 6\%$ and $< 12\%$ at $d = 0.6 - 1.1 R_h$ for C II and Si II, respectively. The high gas covering fraction at $d \lesssim 0.6 R_h$ is consistent with measurements of $\langle \kappa \rangle \approx 58 - 70\%$ for C II and Si II from the COS-Halos sample which probes inner galactic halos at $d < 0.2 R_h$, and with Mg II observations at $d \lesssim 0.3 R_h$ by Chen et al. (2010a). In contrast, the covering fraction of C IV declines from $\langle \kappa \rangle = 60\%$ at $d < 0.6 R_h$ to $\langle \kappa \rangle = 38\%$ at $d = 0.6 - 1.1 R_h$. *Combining previous results with our findings suggests that halo gas becomes progressively more ionized from $d \lesssim 0.2 R_h$ to larger distances.*

2.5 Discussion

Based on absorption spectroscopy carried out in the vicinities of 195 galaxies at $z \lesssim 0.17$, we have obtained strong constraints for the absorption properties of the CGM at low redshift. We observe distinct spatial absorption profiles between H I and heavy elements, with

extended Ly α absorbing gas found at distances far beyond the halo radius R_h but no heavy ions detected at $d \gtrsim 0.7 R_h$. Here we discuss the implications of our study.

2.5.1 Observed Absence of Heavy Elements at $d \gtrsim 0.7 R_h$ from Star-forming Regions

A particularly interesting finding from our absorption-line search is that while between 60% and 90% of the galaxies at $d \gtrsim 0.7 R_h$ show associated Ly α absorbers in the QSO spectra, all but two of these Ly α absorbers exhibit associated heavy ions to sensitive upper limits. The lack of detected heavy ions applies to both low-ionization species probed by the C II and Si II absorption transitions and intermediate ionization species probed by Si III, Si IV, and C IV. It implies that either the chemical enrichment in galactic halos has a finite extent at $d \approx 0.7 R_h$ or cold/warm gas giving rise to the observed absorption lines cannot survive at these large distances ⁷.

In addition, the observed absence of heavy ions at $d \gtrsim 0.7 R_h$ is found around galaxies that span over four orders of magnitudes in M_{star} (Figure 3). While the R_h -normalized radial profile in principle accounts for the size-scaling between galaxies of different mass, it does not include possible variations of halo gas properties as a result of the star formation history of the galaxies. Specifically, star formation rate is known to correlate strongly with galaxy mass (e.g. Figure 4) and star formation feedback is expected to substantially alter the physical properties of halo gas (e.g., Agertz & Kravtsov, 2015). At the same time, high-mass galaxies of $M_{\text{star}} \sim 10^{11} M_{\odot}$ are often found in quiescent mode with little on-going star formation (e.g., Muzzin et al., 2013), which would naturally result in a reduced absorption strength in halos around these massive galaxies (e.g., Gauthier et al., 2010).

To investigate possible variation in the R_h -normalized radial absorption profile of the CGM, we divide the galaxy sample into low- and high-mass subsamples. Figure 12 shows

7. We have since learned that the latter is the likely scenario from detection of O VI up to R_h (Johnson et al., 2015b) and simulation results from Liang et al. (2016).

the CIV absorption properties for low-mass galaxies of $M_{\text{star}} \lesssim 0.1 M_{\text{star}}^*$ (filled and open squares) and high-mass galaxies of $M_{\text{star}} \gtrsim 0.1 M_{\text{star}}^*$ (filled and open circles), where $M_{\text{star}}^* = 5 \times 10^{10} M_{\odot}$ (Baldry et al., 2012; Muzzin et al., 2013). Both low- and high-mass galaxies exhibit a modest mean covering fraction of strong ($W_r(1548) > 0.1 \text{ \AA}$) CIV absorbers $\langle \kappa \rangle \approx 30 - 50\%$ at $d < 0.7 R_h$, and no associated CIV absorbers at $d > R_h$. The sample is too small to make a strong distinction, beyond R_h -normalization, in the extent of chemical enrichment between low- and high-mass halos in our sample.

A lack of low ions outside of the halos where dwarf galaxies reside provides important constraints for the amount of heavy elements that can escape these low-mass halos. Starburst driven outflows are thought to be particularly effective in removing baryons from low-mass halos because of a shallow potential well (e.g., Larson, 1974; Martin, 1999). Energetic outflows are also expected to suppress continuing growth of low-mass galaxies due to a reduced accretion rate (e.g., Scannapieco et al., 2001). These effects together can explain known scaling relations such as the mass-metallicity relation found in dwarf galaxies (e.g., Dekel & Woo, 2003). However, model that predict enrichment levels to > 0.1 solar metallicity are inconsistent with the observed limits we obtain at $d > 0.7 R_h$ from low-mass dwarf galaxies at $z \sim 0$. This places an important constraint on feedback models that would enrich the CGM to much higher metallicity at these distances.

Our finding of a lack of detected heavy elements beyond $d \approx 0.7 R_h$ appears to be discrepant from the observations of Borthakur et al. (2013). who reported detections of strong CIV out to $d \sim R_h$ from massive galaxies of $\log(M_{\text{star}}/M_{\odot}) \approx 10.1$ (star symbols in Figure 12). A possible explanation is to attribute the observed CIV absorbers to outflows driven by on-going star-formation activity, which indeed was part of the selection criterion of the Borthakur et al. sample. However, two of the galaxies with associated strong CIV absorbers are classified as quiescent galaxies with little on-going star formation since recent past.

An alternative scenario different from starburst driven outflows in explaining the presence

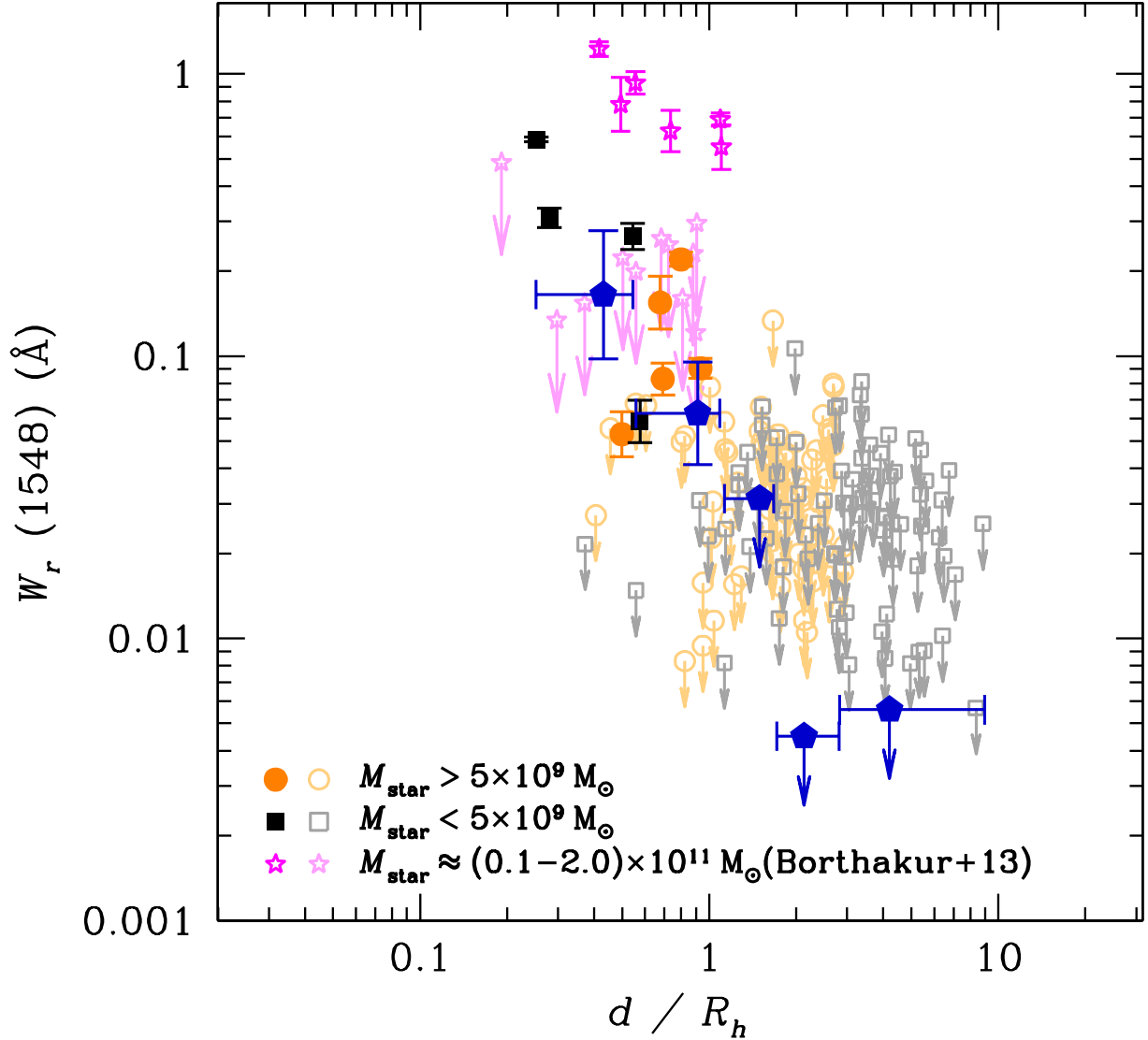


Figure 2.12: R_h -normalized radial profiles of CIV absorbing gas for low-mass galaxies of $M_{\text{star}} < 0.1 M_{\text{star}}^*$ (filled and open squares) and high-mass galaxies of $M_{\text{star}} > 0.1 M_{\text{star}}^*$ (filled and open circles), where $M_{\text{star}}^* = 5 \times 10^{10} M_\odot$ (Baldry et al., 2012; Muzzin et al., 2013). Constraints from stacked spectra of the full sample are included for reference (pentagon points). Both low- and high-mass galaxies exhibit a modest mean covering fraction of strong ($W_r(1548) > 0.1 \text{ \AA}$) CIV absorbing gas $\langle \kappa \rangle \approx 30 - 50\%$ at $d < 0.7 R_h$, and no associated CIV absorbers at $d > R_h$. No significant distinction, beyond R_h -normalization, can be made in the extent of chemical enrichment around low- and high-mass galaxies. In contrast, six of the 17 galaxies in the Borthakur et al. (2013) sample exhibit strong CIV absorption out to $\sim R_h$ (star symbols). While the constraints for the 11 non-detections are weak, the detected CIV absorbers around six galaxies (four starburst and two passive galaxies) are all significantly stronger than the mean values observed in our sample.

of heavy elements at large distances is through gas stripping in an overdense group environment (Kawata & Mulchaey, 2008; McCarthy et al., 2008). Such scenario can be tested with deep galaxy survey data of fields around these strong metal-line absorbers (e.g., Gauthier, 2013).

To improve upon the detection limits of heavy ions in Ly α absorbing gas, we repeat the stacking exercise discussed in § 4.4 but consider only galaxies at $d \gtrsim R_h$ which exhibit associated Ly α absorbers. The goal of this exercise is to investigate whether some constraints can be obtained on the chemical enrichment level in known gaseous clouds revealed by Ly α absorption.

The absorption-line measurements and constraints in stacked spectra of Ly α absorbers are summarized in Table 8. We divide the galaxies at $d \gtrsim R_h$ into three subsamples, each covering different R_h -normalized projected distance intervals. The goal is to maximize the S/N of the stacked spectra, while searching for possible trend of chemical enrichment with distance. In this particular exercise, we also search for the presence of the N V $\lambda\lambda$ 1238, 1242 doublet features. Similar to the O VI absorption doublet, the N V doublet is seen in highly ionized gas and is covered in the STIS and COS FUV spectral range. Observations of the N V absorption doublet together with C IV and other low-ionization transitions, therefore, provides a measure of the ionization state of the gas (see for example Figure 32 of Thom & Chen, 2008). The results of our search remains the same; namely no ionic transitions are detected in Ly α absorbers at $d \gtrsim R_h$ from galaxies in our sample. Adopting the absorption constraints for Ly α and C IV from Table 8 and assuming typical Doppler parameters, $b(\text{HI}) = 30 \text{ km s}^{-1}$ (e.g., Tilton et al., 2012) and $b(\text{C IV}) = 7 \text{ km s}^{-1}$ (e.g., Boksenberg et al., 2003), we infer a column density ratio of $\log N(\text{C IV})/N(\text{HI}) < -1.3$. In contrast, cosmological simulations that incorporate strong outflows to reproduce statistical properties of the general galaxy population in the low-redshift universe predict a column density ratio of $\log N(\text{C IV})/N(\text{HI}) \sim -0.6$ at the virial radius of galactic halos of $10^{11-12} M_\odot$ (Ford et al., 2013). For photo-ionized gas, this column density ratio implies that the chemical enrichment

level of Ly α absorbing gas at $d \gtrsim R_h$ cannot exceed 0.1 solar metallicity.

2.5.2 *The CGM at Low and High Redshifts*

When comparing a luminosity-normalized spatial extent of the CGM at different redshifts, Chen (2012) found that no distinctions can be made between low- and high-redshift galaxies that have dissimilar star formation histories. The adopted luminosity normalization was meant to account for a likely luminosity-size scaling relation commonly seen in optical disks (e.g., Cameron & Driver, 2007). This approach was motivated by the expectation that the CGM is regulated by various complicated physical processes, including gravitational motion and star formation feedback. Normalizing the observed projected distance by the intrinsic luminosity offers a means of accounting for the intrinsic size differences of individual galactic halos, and represents the first step toward establishing a clear understanding of the processes that drive the observed CGM properties. The results will facilitate future studies that consider the effect of additional processes such as star formation feedback.

But while rest-frame B -band luminosity (L_B) is a good proxy of mass as shown in Chen et al. (2001b), it is an indirect measure of galaxy mass. Our study presented here utilizes a sample of galaxies with known stellar masses and normalizes the observed extent of halo gas by the inferred dark matter halo radius under a simple spherical over-density approximation (e.g., Maller & Bullock, 2004).

The finding in § 4.3 is surprisingly very different from the finding of Chen (2012). Specifically, when accounting for a size-scaling based on halo mass (instead of rest-frame B -band luminosity), high-redshift starburst galaxies exhibit both systematically stronger mean absorption strength at a fixed R_h -normalized projected distance at $d/R_h \lesssim 0.7$ and more extended chemically-enriched gas to $d \gtrsim R_h$ than low-redshift galaxies. Furthermore, the differences in the mean absorption strengths of the low- and high-redshift CGM are most pronounced in low-ionization species traced by C II and Si II absorption lines, suggesting distinct ionization conditions between the CGM at $z \approx 0$ and at high redshifts near starburst galaxies.

We note, however, that a fundamental difference between the low- and high-redshift measurements is the treatment of galaxies with close neighbors. In our study presented here, we have considered only “isolated” galaxies with no known close neighbors spectroscopically identified within 500 kpc in projected distance and $|\Delta v| < 500 \text{ km s}^{-1}$ in velocity separation. No such distinction was made in the high-redshift sample, and therefore the high-redshift CGM measurements include both isolated galaxies and those with neighbors. It has been suggested that galaxies in group environments display a more extended, chemically enrichment halo gas than galaxies in the field (Chen et al., 2010a; Bordoloi et al., 2011). At the same time, the overall cool gas content is also found suppressed in galaxy groups (Chynoweth et al., 2011). It is not clear that the more extended, strong CGM absorption around high-redshift starburst galaxies is due to a significant contribution of galaxies in group environments. A detailed study of the CGM in different group environments at low redshifts will be presented in a separate paper.

In addition, Steidel et al. (2010) noted that the Si II absorption lines are saturated at $d \lesssim 40 \text{ kpc}$ in the stacked low-resolution spectra. On the other hand, the ionic transitions observed in individual galaxies at low redshifts are not commonly saturated. Steidel et al. interpreted the observed declining absorption equivalent width (at least from $d = 0$ to $d = 40 \text{ kpc}$) as likely due to a combination of decreasing line-of-sight velocity spread and decreasing covering fraction of the absorbing gas. In contrast, observations individual galaxies at low redshift indicate that the declining mean absorption equivalent width in stacked spectra is best explained by a declining gas covering fraction along with declining optical depth with increasing projected distance. Consequently, interpreting the observed difference between low- and high-redshift observations becomes more complicated.

At $d > 40 \text{ kpc}$, however, Steidel et al. (2010) also pointed out that the ionic absorption lines appeared to become unsaturated around high-redshift starburst galaxies. For unsaturated lines in stacked spectra, the observed declining equivalent width is best explained by a declining gas covering fraction along with a declining optical depth, similar to the

low-redshift observations. Comparisons between our means stacked spectra and those from Steidel et al. could therefore be made for observations at $d > 40$ kpc.

Remaining caveats include redshift uncertainties. Individual spectra of these high-redshift galaxies are of low S/N and low spectral resolution with $\text{FWHM} \approx 350 \text{ km s}^{-1}$. Random motions between galaxies and the absorbing gas (e.g. Figure 6) can smooth out the signal. Chen (2012) performed a simple Monte Carlo simulation to show that these uncertainties can result in an overall reduction of the absorption strength by $\approx 20\%$, increasing the discrepancy between low- and high-redshift samples. At the same time, uncertainties in sky subtraction of low S/N spectra may also bias the continuum estimate that is difficult to assess. It is expected that these uncertainties can be resolved with absorption spectroscopy of background QSOs near individual starburst galaxies (e.g., Rudie et al., 2012; Turner et al., 2014).

2.5.3 *Uncertainties due to Pseudo-evolution of Halo Radius R_h*

We note, however, that a fundamental caveat in comparing R_h -normalized CGM profiles at different redshifts is in how R_h is calculated. Equations (2) and (3) outlines a commonly adopted approach to calculate halo mass M_h and halo radius R_h based on a spherical overdensity model. A dark matter halo is defined as a collapsed and virialized object with a mean enclosed density equal to Δ_h times the mean background density $\langle \rho_m \rangle$ at the redshift of interest. By this definition, however, as the universe expands, $\langle \rho_m \rangle$ decreases. Consequently, both M_h and R_h would increase with time even for a static halo. Such pseudo-evolution of dark matter halos has been investigated by several groups, and a consistent finding in numerical simulations is that most halos of $M_h \lesssim 10^{12} M_\odot$ completed matter accretion by $z \sim 1$ with little/no net growth in mass over the last seven billion years (e.g., Prada et al., 2006; Diemand et al., 2007; Diemer et al., 2013). Therefore, applying an R_h normalization may impose a pseudo-evolution in the properties of galaxies residing in these low-mass halos.

Given the uncertainties in R_h , we explore different scaling relations for an accurate com-

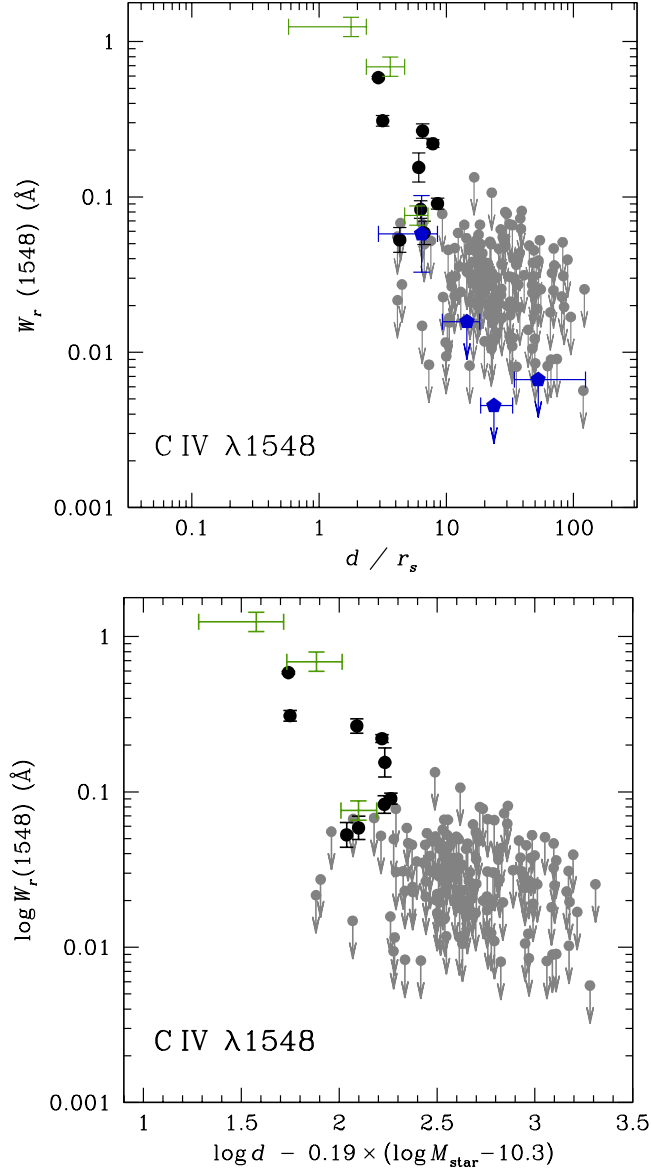


Figure 2.13: C IV radial absorption profile normalized by the scale radius r_s of the host dark matter halos (*left*) or by M_{star} of the absorbing galaxies (*right*). The scale radius r_s represents a characteristic radius where the dark matter density profile resembles an r^{-2} profile (Navarro et al., 1997). We also include in the left panel measurements from stacked spectra using the full galaxy sample (blue pentagons). In addition, the scaling coefficient of M_{star} is not a best fit to the C IV data, but adopted from the best-fit coefficient of Mg II absorbers from Chen et al. (2010b). Nonetheless, the result already reveals a clear C IV enriched halo out to M_{star} -normalized projected distance of ≈ 160 kpc (consistent with the finding of L_B -normalized C IV profile from Chen 2012). In both panels, the C IV radial absorption profiles at $\langle z \rangle = 0.041$ (solid points) and at $\langle z \rangle = 2.2$ (green crosses) appear to be better aligned than what is seen in the bottom right panel of Figure 7 & 8.

parison of the CGM absorption profiles in different epochs. In Figure 13, we present CIV radial absorption profile normalized by the scaled radius r_s of the host dark matter halos (left panel) or by M_{star} of the absorbing galaxies (right panel).

The scale radius r_s represents a characteristic radius where the dark matter density profile resembles an r^{-2} profile (Navarro et al., 1997). The r_s normalization is motivated by the understanding that when a halo stops accreting, the scaled radius r_s remains constant while the halo mass M_h and halo radius R_h continue to increase as the universe expands and $\langle \rho_m \rangle$ decreases. As a result, the halo concentration $c \equiv R_h/r_s$ would also increase. Indeed, halo concentrations are found to increase rapidly for halos of $M_h \lesssim 10^{12} M_\odot$ at $z \lesssim 1$ (e.g. Zhao et al. 2009).

To calculate r_s , we adopt the mean c - M_h relation of Zhao et al. (2009), which parameterizes c as a function of halo mass and redshift, $c(M_h, z)$. The left panel of Figure 13 shows that the r_s -normalized CIV radial profile at $\langle z \rangle = 2.2$ declines rapidly at $d \approx (4 - 6) r_s$. This is consistent with observations of the $\langle z \rangle = 0.041$ sample that display a covering fraction of CIV absorbing gas at $d < 6 r_s$ with $\langle \kappa \rangle_{\text{CIV}} \approx 50\%$ and no detections found beyond $8 r_s$.

The agreement in the r_s -normalized CGM absorption profiles at low and high redshifts is in stark contrast to the distinction seen in the R_h -normalized CGM profiles presented in Figures 7 & 8. We note, however, an inherent uncertainty in r_s as a result of the scatter in the mean c - M_h relation (~ 0.15 dex). Nonetheless, it underscores the fundamental difficulty in making direct comparisons of galaxies and their halo gas properties in two different epochs.

A pseudo-evolution in dark matter halo mass (and therefore halo radius R_h) has important implications beyond a constant CGM radial absorption profile. It would also suggest that chemical enrichment have reached out to the “edge” of the dark matter halo ($\sim 4 r_s$) or likely beyond, although a clear border for the observed metal-line absorbers outside of halo radius still implies a finite extent of chemical enrichment around galaxies. To quantify the extent of chemical enrichment, we will rely on numerical simulations as a guide for deter-

mining the physical size of the underlying dark matter halos of our galaxies. Such study is beyond the scope of this paper. We defer the analysis to a future paper.

Given the uncertainty in R_h and r_s from dark matter halo models, we return to applying empirical properties of the galaxies for scaling halo gas properties. Instead of B -band luminosity, however, we consider normalizing the projected distance of galaxies by stellar mass, M_{star} . The right panel of Figure 13 shows the M_{star} -normalized CIV radial profiles for the high- and low-redshift galaxy samples. We note that the adopted scaling coefficient of M_{star} is not a best fit to the CIV data, but adopted from the best-fit coefficient of Mg II absorbers from Chen et al. (2010a). Nonetheless, the result already reveals a clear CIV enriched halo out to M_{star} -normalized projected distance of ≈ 160 kpc (in an excellent agreement with the finding of L_B -normalized CIV profile from Chen 2012). Similar to the r_s -normalized profiles in the left panel, the M_{star} -normalized CIV radial absorption profiles at $\langle z \rangle = 0.041$ and at $\langle z \rangle = 2.2$ appear to be well aligned. With in mind the uncertainty in R_h and the agreement in M_{star} -normalized CGM radial profile, we therefore conclude that no strong evidence has been found to show a vastly different spatial distribution in the CGM absorption properties, despite a completely disparate star formation properties.

2.6 Summary

We have analyzed archival UV spectra of QSOs located at projected distance $d \leq 500$ kpc from 195 independent galaxies in the foreground. The redshifts of the galaxies range from $z = 0.002$ to $z = 0.176$ with a median of $\langle z \rangle = 0.041$, and the projected distances of the QSOs range from $d \approx 32$ kpc to $d \approx 500$ kpc with a median of $\langle d \rangle = 362$ kpc. The galaxy sample covers a broad range in stellar mass from $M_{\text{star}} = 1.5 \times 10^5 M_{\odot}$ to $M_{\text{star}} = 1.4 \times 10^{11} M_{\odot}$. The median and $1\text{-}\sigma$ dispersion of stellar masses in the sample are $\log(M_{\text{star}}/M_{\odot}) = 9.7 \pm 1.1$. The available UV spectra of the QSOs allow us to study the multiphase CGM based on observations of a suite of absorption features including Ly α , C II, CIV, Si II, Si III, and Si IV. Using this large QSO and galaxy pair sample, we have obtained

strong constraints for the absorption properties of the CGM at $z \lesssim 0.176$. The results of our analysis are summarized as the following:

(1) We observe a stark contrast between the spatial distributions of hydrogen atoms and heavy elements. While hydrogen gas is observed all the way out to 500 kpc in projected distance with a mean covering fraction of $\approx 60\%$, the same galaxies that display moderately strong HI absorption exhibit few associated metal absorption lines at $d > 200$ kpc. The lack of heavy elements at large distances persists through all ionization states included in the study, suggesting that either there exists a chemical enrichment edge at $d \approx 0.7 R_h$ or gaseous clumps giving rise to the observed absorption lines cannot survive at these large distances.

(2) The distinction between chemically-enriched and metal-poor gas is further enhanced after accounting for a mass scaling of gaseous radius in galactic halos. We infer a halo mass M_h and halo radius R_h for each galaxy based on the known stellar mass M_{star} , and find that no ionic transitions are detected at $d \gtrsim 0.7 R_h$ to sensitive upper limits of rest-frame absorption equivalent width $W_r \approx 0.05 \text{ \AA}$ or better. The observed absence of heavy elements at $d \gtrsim 0.7 R_h$ applies to both low-mass dwarfs and high-mass galaxies. Considering all galaxies at $d > R_h$, we place a strict *upper limit* for the covering fraction of heavy elements of 3% (at a 95% confidence level) over the projected distance range from R_h to $9 R_h$.

(3) To further improve upon the sensitivity in searching for weak absorption features, we experiment with co-adding the absorption spectra of individual galaxies in the rest frame. Using stacked spectra we are indeed able to improve the constraints for the presence/absence of absorbing gas by factors of five to six. No ionic transitions are found at $d > R_h$ to unprecedentedly sensitive limits of $\approx 0.03 \text{ \AA}$ or better at the 95% confidence level.

(4) Within the gaseous halo probed by our galaxy-QSO pair sample, we observe differential covering fraction between low- and high-ionization gas. The mean covering fractions of both low- and high-ionization absorbing gas remain high at $\langle \kappa \rangle \gtrsim 50\%$ over the distance range of $d = (0.25 - 0.6) R_h$. At $d = (0.6 - 1) R_h$, the covering fractions of C II and Si III decline

to $\langle\kappa\rangle \lesssim 14\%$, while the covering fraction of C IV remains at $\langle\kappa\rangle_{\text{C IV}} = 38\%$. Combining previously observed high covering fraction of low-ionization gas in inner halos ($d < 0.2 R_h$) with our findings suggests that halo gas becomes progressively more ionized from $d < 0.2 R_h$ to larger distances.

(5) Combining our study of outer gaseous halos of low- and high-mass galaxies with the observations of the inner halos around massive galaxies from the COS-Halos program establishes a consistent R_h -normalized spatial absorption profile of the low-redshift CGM. Both the absorption strengths and incidence (covering fraction) of heavy elements decline steeply beyond $0.3 R_h$ and become vanishingly small beyond $0.7 R_h$. In contrast, massive starburst galaxies at $z = 2.2$ exhibit significantly stronger mean absorption than galaxies at $z \sim 0$. The differences are most pronounced in low-ionization species traced by C II and Si II absorption features, suggesting distinct ionization conditions between the CGM at low and high redshifts.

(6) We caution a potential pseudo-evolution in the CGM radial absorption profiles, when utilizing R_h -normalized projected distances to account for the intrinsic size difference between galaxies of different halo mass. The commonly adopted spherical overdensity calculation for halo mass imposes a pseudo-evolution for static halos as the universe expands and the mean matter density decreases with time. Normalizing the observed CGM radial profile by the scale radius r_s or by the stellar mass M_{star} confirms the previous finding that no significant changes can be found between the CGM at $z \approx 2.2$ and $z \approx 0$.

Table 2.1: Summary of Galaxy Properties^a

Galaxy Name	RA(J2000)	Dec(J2000)	z_{gal}	$\log \frac{\text{SFR}^b}{M_{\odot}/\text{yr}}$	$\log \frac{M_{\text{star}}}{M_{\odot}}$	$\log \frac{M_h^c}{M_{\odot}}$	$R_h(\text{kpc})$	$M_{\text{SDSS},r}^d$	Flag ^e
SDSS J000545.07+160853.3	00:05:45.07	+16:08:53.26	0.0372	-0.76	9.3	11.2	137.0	-18.8	0
SDSS J000548.11+161123.4	00:05:48.12	+16:11:23.40	0.1134	...	10.5	11.9	217.1	-21.2	1
2MASX J000556.17+160804.1	00:05:56.16	+16:08:04.17	0.0909	...	10.7	12.0	241.5	-21.5	1
SDSS J004208.22-102929.1	00:42:08.23	-10:29:29.10	0.0421	-0.50	9.3	11.2	133.6	-19.0	0
SDSS J004214.99-104414.9	00:42:14.99	-10:44:14.90	0.0360	-0.88	9.5	11.3	144.9	-19.6	0
SDSS J022617.97+001203.8	02:26:17.97	+00:12:03.85	0.0913	...	9.8	11.4	153.8	-19.7	1
2MASX J024248.74-082356.3	02:42:48.70	-08:23:56.65	0.0144	-1.83	9.7	11.3	154.7	-19.0	0
SDSS J025405.61-005236.6	02:54:05.64	-00:52:36.35	0.0032	-2.33	7.1	10.0	55.8	-13.9	0
SDSS J025906.86+004306.4	02:59:06.87	+00:43:06.34	0.0432	-1.19	8.9	11.0	118.2	-17.7	0
SDSS J025914.53+003359.6	02:59:14.53	+00:34:00.73	0.0092	-2.44	7.2	10.1	60.6	-15.5	0
SDSS J025938.43+003939.3	02:59:38.44	+00:39:39.24	0.0311	-1.53	8.0	10.5	83.0	-16.8	0
2MASX J025942.96+003907.8	02:59:42.94	+00:39:08.19	0.1342	...	11.0	12.4	327.2	-22.2	1

^a The full table in machine-readable form is available at http://lambda.uchicago.edu/public/local/galaxy_table.dat.

^b Star formation rate derived from GALEX Near UV band with $\lambda_{\text{eff}} \approx 2267 \text{ \AA}$.

^c The halo mass M_h and halo radius R_h are estimated following the prescription described in § 4.3.

^d Absolute r band magnitude estimated from Sergic + Exponential light profile fit (Bernadi et al. 2009).

^e Galaxies with M_{star} from the NASA-Sloan Atlas are noted as “0”, those derived with Eq 1 are noted as “1”.

Table 2.2: Summary of Observations of QSOs^a

QSO Name	RA(J2000)	Dec(J2000)	z_{QSO}	PID	S/N_{G130M}^b	S/N_{G160M}^b	S/N_{STIS}^c
PG 0003+158	00:05:59.24	+16:09:49.01	0.4509	12038	16	13	...
SDSS J004222.29−103743.8	00:42:22.29	−10:37:43.79	0.4240	11598	6	4	...
SDSS J022614.46+001529.7	02:26:14.44	+00:15:30.04	0.6151	11598	8	5	...
SDSS J024250.85−075914.2	02:42:50.86	−07:59:14.25	0.3777	12248	6	5	...
SDSS J025937.46+003736.3	02:59:37.47	+00:37:36.38	0.5342	12248	6	5	...
PKS 0312−770	03:11:55.25	−76:51:50.85	0.2252	8651	...	...	5
SDSS J040148.98−054056.5	04:01:48.98	−05:40:56.58	0.5701	11598	6	5	...
PKS 0405−123	04:07:48.43	−12:11:36.66	0.5726	11541	...	...	48
RX J0439.6−5311	04:39:38.72	−53:11:31.40	0.2430	11520	11	7	...
HE 0439−5254	04:40:12.02	−52:48:17.70	1.0530	11520	11	8	...

^a The full table in machine-readable form is available at http://lambda.uchicago.edu/public/local/qso_table.dat.

^bMedian S/N of COS G130M grating in spectral range of 1150 − 1450Å and G160M grating in 1450 − 1750Å.

^cMedian S/N of STIS spectra in spectral range of 1200 − 1700Å.

Table 2.3: Available Absorption Constraints of Different Galaxy Samples

	# of Galaxies	$\langle z \rangle^a$	d	Range probed (kpc)	$\langle \log M_{\text{star}}/M_{\odot} \rangle^b$	Transitions ^c
This work	195	0.04 ± 0.04		32 – 500	9.7 ± 1.1	Ly α , C II, C IV, Si II, Si III, Si IV
Werk et al. 2013	44	0.22 ± 0.05		18 – 154	10.6 ± 0.5	Ly α , C II, Si II, Si III, Si IV
Steidel et al. 2010	512	2.2 ± 0.3		10 – 125(280) ^d	9.9 ± 0.5^e	Ly α , C II, C IV, Si II, Si IV

^aMedian redshift and dispersion of the galaxy samples.

^bMedian stellar mass and dispersion of the galaxy samples.

^cIncluding only transitions common to this work.

^dMaximum impact parameter is 280 kpc for Ly α and 125 kpc otherwise.

^eMedian and dispersion of stellar mass for Steidel et al. 2010, which is a representative sample of galaxies in Reddy et al. 2012.

Table 2.4: Summary of Individual Halo Absorption Properties.^{a,b}

Galaxy	θ ($''$)	d (kpc)	z_{gal}	$z_{\text{Ly}\alpha}$	HI $W_r(1215)$ (mÅ)	Si III $W_r(1206)$ (mÅ)	Si II $W_r(1260)$ (mÅ)	Si IV $W_r(1393)$ (mÅ)	C II $W_r(1334)$ (mÅ)	C IV $W_r(1548)$ (mÅ)
SDSS J112644.33+590926.0	393.2	32.4	0.0040	0.0040	993 $\pm$ 15	290 $\pm$ 11	...	166 $\pm$ 12	106 $\pm$ 12	586 $\pm$ 12
SDSS J134249.99-005329.0	29.1	39.3	0.0708	0.0717	660 $\pm$ 16	78 $\pm$ 6	69 $\pm$ 8	$\leq$ 16	140 $\pm$ 14	309 $\pm$ 24
NGC 3485	582.5	57.6	0.0048	0.0048	506 $\pm$ 12	$\leq$ 22	$\leq$ 11	$\leq$ 21	$\leq$ 35	$\leq$ 22
SDSS J144520.23+341948.1	529	60.6	0.0056	0.0056	962 $\pm$ 94	$\leq$ 38	$\leq$ 20	$\leq$ 36	84 $\pm$ 19	266 $\pm$ 28
2MASX J08091327+4618424	68.6	62.7	0.0466	0.0464	1026 $\pm$ 10	54 $\pm$ 6	$\leq$ 11	69 $\pm$ 7	$\leq$ 21	$\leq$ 27
SDSS J122815.96+014944.1	1107.1	69.5	0.0030	0.0034	393 $\pm$ 8	$\leq$ 7	$\leq$ 4	$\leq$ 3	$\leq$ 5	$\leq$ 8
SDSS J121413.94+140330.4	56.6	70.1	0.0644	0.0644	859 $\pm$ 9	$\leq$ 6	$\leq$ 4	$\leq$ 10	$\leq$ 10	$\leq$ 15
2MASX J043936.88-530045.5	646.1	74.3	0.0056	0.0055	518 $\pm$ 9	152 $\pm$ 10	...	47 $\pm$ 8	$\leq$ 17	59 $\pm$ 10
SDSS J025914.53+003359.6	406	76.6	0.0092	0.0094	446 $\pm$ 33	...	$\leq$ 15	$\leq$ 29	$\leq$ 33	$\leq$ 39
SDSS J025938.43+003939.3	123.7	77.0	0.0311	0.0311	196 $\pm$ 12	$\leq$ 21	$\leq$ 28	$\leq$ 25	$\leq$ 31	$\leq$ 31

^aThe full table in machine-readable form is available at http://lambda.uchicago.edu/public/local/absorber_table.dat.

^bUpper limits are quoted at 2σ level

Table 2.5: Absorption Equivalent Widths Observed in Stacked CGM Spectra in Different Projected Distance Range^a

d (kpc)	$\langle d \rangle^b$ (kpc)	Number	Ly α $\lambda 1215$ (mÅ)	Si III $\lambda 1206$ (mÅ)	Si II $\lambda 1260$ (mÅ)	Si IV $\lambda 1393$ (mÅ)	C II $\lambda 1334$ (mÅ)	C IV $\lambda 1548$ (mÅ)
32.4 – 138.1	86.1	17	486 $\pm$ 34	88 $\pm$ 13	45 $\pm$ 23	44 $\pm$ 16	48 $\pm$ 12	70 $\pm$ 37
156.2 – 307.1	243.2	52	295 $\pm$ 39	≤ 7	≤ 4	≤ 13	≤ 6	25 $\pm$ 12
311.0 – 442.4	383.9	54	136 $\pm$ 31	≤ 3	≤ 2	≤ 4	≤ 3	≤ 8
444.3 – 498.9	480.7	53	116 $\pm$ 23	≤ 4	≤ 6	≤ 6	≤ 5	≤ 6

^a 1- σ uncertainties and 2- σ upper limits are determined from a bootstrap resampling calculation ($N_{\text{boot}} = 350$)

^b Median impact parameter in each bin.

Table 2.6: Absorption Equivalent Width in Stacked CGM Spectra in different R_h -Normalized Projected Distance Range^a

d/R_h	$\langle d/R_h \rangle^b$	Number	Ly α $\lambda 1215$ (mÅ)	Si III $\lambda 1206$ (mÅ)	Si II $\lambda 1260$ (mÅ)	Si IV $\lambda 1393$ (mÅ)	C II $\lambda 1334$ (mÅ)	C IV $\lambda 1548$ (mÅ)
0.25 – 0.54	0.43	8	550 $\pm$ 70	93 $\pm$ 32	67 $\pm$ 25	204 $\pm$ 94	102 $\pm$ 56	164 $\pm$ 86
0.56 – 1.09	0.91	22	382 $\pm$ 50	32 $\pm$ 15	$\leq$ 4	$\leq$ 34	$\leq$ 6	63 $\pm$ 26
1.13 – 1.67	1.49	33	252 $\pm$ 31	$\leq$ 4	$\leq$ 5	$\leq$ 20	$\leq$ 8	$\leq$ 31
1.71 – 2.82	2.13	56	165 $\pm$ 29	$\leq$ 3	$\leq$ 3	$\leq$ 9	$\leq$ 4	$\leq$ 4
2.82 – 8.97	4.20	57	80 $\pm$ 14	$\leq$ 10	$\leq$ 3	$\leq$ 3	$\leq$ 4	$\leq$ 6

^a 1- σ uncertainties and 2- σ upper limits are determined from a bootstrap resampling calculation ($N_{\text{boot}} = 350$)

^b Median R_h -normalized projected distance in each bin.

Table 2.7: Mean Gas Covering Fraction, $\langle \kappa \rangle$, of Different Species^a

d/R_h	$\langle d/R_h \rangle^b$	Ly α $\lambda 1215$	Si III $\lambda 1206$	Si II $\lambda 1260$	Si IV $\lambda 1393$	C II $\lambda 1334$	C IV $\lambda 1548$
0.25 – 0.54	0.43	1.00 ^{+0.18} _{-0.18}	0.63 ^{+0.13} _{-0.18}	0.43 ^{+0.18} _{-0.15}	0.50 ^{+0.18} _{-0.18}	0.63 ^{+0.13} _{-0.18}	0.60 ^{+0.16} _{-0.23}
0.56 – 1.09	0.91	0.96 ^{+0.01} _{-0.08}	0.14 ^{+0.11} _{-0.05}	≤ 0.12	0.05 ^{+0.09} _{-0.02}	0.06 ^{+0.11} _{-0.02}	0.38 ^{+0.13} _{-0.10}
1.13 – 1.67	1.49	0.94 ^{+0.02} _{-0.07}	≤ 0.09	≤ 0.08	≤ 0.11	≤ 0.12	≤ 0.12
1.71 – 2.82	2.13	0.71 ^{+0.05} _{-0.07}	≤ 0.07	≤ 0.06	≤ 0.06	≤ 0.07	≤ 0.06
2.82 – 8.97	4.20	0.59 ^{+0.07} _{-0.08}	≤ 0.09	≤ 0.06	≤ 0.07	≤ 0.06	≤ 0.06

^a Based on the fraction of absorbers with rest-frame absorption equivalent width greater than $W_0 = 0.05$ Å.

^b Median R_h -normalized projected distance in each bin.

Table 2.8: Absorption Equivalent Width in Stacked Spectra of Ly α Absorbers at $d > 0.7 R_h^a$

d/R_h	$\langle d/R_h \rangle^b$	Number	Ly α 1215 (mÅ)	NV 1238 (mÅ)	SiIII 1206 (mÅ)	SiII 1260 (mÅ)	SiIV 1393 (mÅ)	CII 1334 (mÅ)	CIV 1548 (mÅ)
0.96 – 1.63	1.26	32	308 ± 28	≤ 15	≤ 5	≤ 5	≤ 14	≤ 6	≤ 33
1.64 – 2.59	1.85	31	281 ± 34	≤ 8	≤ 6	≤ 8	≤ 15	≤ 15	≤ 17
2.59 – 8.97	4.06	30	131 ± 19	≤ 7	≤ 9	≤ 9	≤ 8	≤ 4	≤ 6

^a 1- σ uncertainties and 2- σ upper limits are determined from a bootstrap resampling calculation ($N_{\text{boot}} = 350$)

^b Median R_h -normalized projected distance in each bin.

CHAPTER 3

COLUMN DENSITY PROFILES OF THE MULTIPHASE GASEOUS HALOS

3.1 Introduction

A complete picture of galaxy evolution requires an understanding of the interplay between inflow and cooling of gas, star formation, and associated feedback that can drive star-forming gas back into the halo or beyond via large-scale winds. The current generation of galaxy formation models appears to capture many of these processes in a way that leads to galaxies with realistic properties (e.g., Feldmann et al., 2011; Sales et al., 2012; Stinson et al., 2013; Aumer et al., 2013; Hopkins et al., 2014; Agertz & Kravtsov, 2015; Schaye et al., 2015). So far, however, galaxy formation models have been tested almost exclusively against properties of the stellar component of galaxies (with some recent exceptions, e.g., Hummels et al., 2013; Ford et al., 2013, 2015; Schaye et al., 2015).

Yet, we know that stars comprise only a relatively small fraction of the baryon budget. A variety of observational constraints and inferences from the abundance matching technique (e.g., see Fig. 10 and 11 in Kravtsov et al., 2014, for a compilation of recent constraints) shows that even the galaxies that are most efficient in forming stars (forming in halos of $M_h \approx 10^{12} M_\odot$) convert at most $\approx 30 - 40\%$ of the available baryon budget into stars. The bulk of the remainder of the baryon budget is outside of the disk, because cold gas fractions in the evolved galaxies we observe today are also small. The “missing” baryons are thought to reside in the circumgalactic medium (CGM) around galaxies, which can be thought of as a halo of tenuous gas resulting from the complex interactions between static gaseous halo, inflowing gas, and feedback-driven outflows from galaxies. Studying the galactic baryons in the CGM is therefore paramount for the understanding of the feedback processes that regulate inflows and outflows of gas.

Although details of such processes are still debated, efficient feedback recipes have been

implemented in galaxy formation simulations (e.g., Hopkins et al., 2011; Brook et al., 2012; Stinson et al., 2013; Vogelsberger et al., 2013; Agertz et al., 2013; Roškar et al., 2014; Trujillo-Gomez et al., 2015). At the low mass end ($M_h < 10^{12} M_\odot$), stellar feedback in the form of energy and momentum injection from radiation pressure, stellar winds, and supernova explosions is believed to be responsible for the suppression of star formation (Dekel & Silk, 1986; Efsthathiou, 2000). On the other hand, active galactic nuclei (AGN) feedback is thought to be critical in limiting star formation and stellar masses of galaxies at the high mass end ($M_h > 10^{12} M_\odot$; e.g., Silk & Rees, 1998; Benson et al., 2003). More recently, cosmic ray (CR) driven winds have also been shown to be a promising feedback mechanism in the context of galaxy formation simulations (e.g., Socrates et al., 2008; Booth et al., 2013; Salem et al., 2014).

It is quite likely that stellar feedback leaves a specific imprint on the properties of the circumgalactic gaseous halos (e.g., Barai et al., 2013; Hummels et al., 2013; Suresh et al., 2015). In particular, comparisons between simulated CGM and observations have indicated that the observed incidence of absorbers in the CGM favors larger intensity of feedback processes (e.g., Stinson et al., 2012; Hummels et al., 2013; Suresh et al., 2015).

Observations of the CGM have advanced tremendously in recent years. The Cosmic Origins Spectrograph (COS; Green et al., 2012) on board the *Hubble Space Telescope (HST)* has offered a unique opportunity to systematically probe low-redshift ($z \approx 0$) galactic halos in the UV regime. Commonly observed absorption lines arising in the CGM are the Ly α λ 1215, Si II λ 1260, Si III λ 1206, C II λ 1334, C IV λ 1548, 1550 (e.g., Chen, 2012; Borthakur et al., 2013; Liang & Chen, 2014; Bordoloi et al., 2014), and O VI λ 1031, 1036 (e.g., Prochaska et al., 2011; Tumlinson et al., 2011; Johnson et al., 2015b). In the intermediate redshift range, many researchers have probed the gaseous halos using optical ground-based facilities via Mg II (Chen et al., 2010a; Gauthier et al., 2010; Bordoloi et al., 2011). Similar studies for the UV transitions have also been conducted in the high redshift universe ($z \approx 2$; Steidel et al., 2010; Rudie et al., 2012; Turner et al., 2014).

In this paper, we explore how the choices of parameters controlling star formation and stellar feedback in cosmological galaxy formation simulations affect the properties of the CGM around simulated galaxies, such as the profile of the absorber column densities. We also use the simulation in which the CGM properties are closest to observations to explore the physical origin of the CGM absorbers and the processes that shape their properties.

The paper is structured as follows. In section 2, we outline the simulation suite used in our study, including details of star formation and the stellar and cosmic ray feedback models. In section 3, we describe the post-processing of simulations and our CGM analyses. Main qualitative features of the CGM in our simulations are discussed in section 4, where we demonstrate that the radial profiles of the CGM absorber column density have an exponential form and evolve only very weakly with redshift. In addition, we show that profiles of the CGM spanning four decades of stellar mass and 11 billion years in cosmic time can be cast as a single profile using a simple re-scaling of radius ($d \rightarrow d/r_s$). We then present comparisons between our predictions from simulations and observations of the CGM in section 5, where we show that observations also appear to obey scaling with halo scale radius. In section 6 we discuss our results and their implications, as well as compare them to results of other recent studies of the CGM in simulations. Finally, in section 7, we summarize our results and conclusions.

3.2 Galaxy Formation Simulations

In this study, we use a suite of galaxy formation simulations, which consists of “zoom-in” re-simulations of the evolution of a Milky Way-sized halo started from the same initial conditions, but with different parameters of star formation and stellar feedback recipes, as described in Agertz & Kravtsov (2015). In addition, we use a new re-simulation with a cosmic ray feedback model described in Booth et al. (2013).

All simulations are run using the Adaptive Mesh Refinement (AMR) code **RAMSES** (Teyssier, 2002) from the same initial conditions in the WMAP5 Λ CDM cosmology with $\Omega_\Lambda = 0.73$,

Table 3.1: Simulation suite

Simulations	Description	Final Redshift
KMT09 $E_{\text{fb}}, f_{\text{fb}} = 0.5, t_{\text{dis}} = 10\text{Myr}$		
ALL_Efb_e010	All feedback, $\epsilon_{\text{ff}} = 10\%$	$z = 0$
ALL_Efb_e001	All feedback, $\epsilon_{\text{ff}} = 1\%$	$z = 1.5$
ALL_Efb_e001_5ESN	All feedback, $E_{\text{SNII}} = 5 \times 10^{51} \text{ erg}$, $\epsilon_{\text{ff}} = 1\%$	$z = 0$
Cosmic rays , $E_{\text{CR}} = \xi_{\text{CR}} E_{\text{SNII}}$		
ALL_e010_CR	All feedback, $\xi_{\text{CR}} = 10\%$, $\epsilon_{\text{ff}} = 10\%$	$z = 1.0$

$\Omega_{\text{m}} = 0.27$, $\Omega_{\text{b}} = 0.045$, $\sigma_8 = 0.8$ and $H_0 = 70 \text{ km s}^{-1} \text{Mpc}^{-1}$ (Komatsu et al., 2009). At $z = 0$, the halo masses of the simulated $\sim L_*$ galaxies have similar values of $M_{\text{vir}} \approx 10^{12} M_{\odot}$, corresponding to a virial radius of $R_{\text{vir}} \approx 260 \text{ kpc}$. The dark matter particle mass in the highest resolution region is $m_{\text{DM}} = 3.2 \times 10^5 M_{\odot}$. The physical resolution at maximum refinement level reaches $\Delta x \approx 75 \text{ pc}$. A summary of the parameters of our simulations is provided in Table 1.

Unless otherwise stated, all quantities in the paper are given in physical units. We briefly outline the main ingredients of the simulations in the following subsections. A more detailed description can be found in Agertz & Kravtsov (2015), while the specifics on the implementation of stellar and cosmic ray feedback are detailed in Agertz et al. (2013) and Booth et al. (2013), respectively.

3.2.1 Star Formation

In each star forming cell, the number of star particles N is determined from the Poisson distribution $P(N|\lambda_P)$ with the mean $\lambda_P = \dot{\rho}_* \Delta x^3 \Delta t / m_*$, where star formation rate is given by:

$$\dot{\rho}_* = f_{\text{H}_2} \frac{\rho_{\text{g}}}{t_{\text{ff}}} \epsilon_{\text{ff}}. \quad (3.1)$$

Here f_{H_2} is the local mass fraction of molecular hydrogen computed using the model of Krumholz et al. (2009), as described in Agertz & Kravtsov (2015), ρ_{g} is the local gas density of a given cell, $t_{\text{ff}} = \sqrt{3\pi/(32G\rho_{\text{g}})}$ is the free-fall time, and ϵ_{ff} is the star formation efficiency

per free-fall time assumed to be constant in time and space. The simulations analyzed in this paper use two constant values of the efficiency: $\epsilon_{\text{ff}} = 1\%$ and 10% . The parameter m_* , set to $10^4 M_\odot$ in our simulations, is a unit mass of star particles and all star particles are created as a multiple of it (Rasera & Teyssier, 2006; Dubois & Teyssier, 2008).

Parameters of the star formation recipe have a direct impact on the CGM because the feedback strength depends on the mass and spatial distribution of young stars, as we will discuss in section 3.4 (see also Agertz & Kravtsov, 2015; Stinson et al., 2012).

3.2.2 *Stellar Feedback*

Our simulations account for energy, momentum, mass, and the injection of heavy elements (with an effective yield of $1\% - 3\%$) from type Ia supernovae (SNIa), type II supernovae (SNII), stellar winds and radiation pressure from massive stars, as well as for the secular mass loss into the surrounding interstellar medium (ISM) from the AGB stars. Injection of momentum, energy and mass is done continuously during each simulation time step, which is computed according to the Courant-Friedrichs-Lewy (CFL) condition. In the regions of active star formation and feedback the time steps become as short as a few thousand years.

When a supernova explodes, we retain a fraction, f_{fb} , of the explosion energy in a separate energy variable E_{fb} , which is evolved using an equation similar to that of the internal energy of the gas, but with the explicitly specified dissipation time scale, t_{dis} :

$$\frac{\partial}{\partial t}(E_{\text{fb}}) + \nabla \cdot (E_{\text{fb}} \mathbf{v}_{\text{gas}}) = -P_{\text{fb}} \nabla \cdot \mathbf{v}_{\text{gas}} - \frac{E_{\text{fb}}}{t_{\text{dis}}} \quad (3.2)$$

This energy variable is used to define an additional contribution to gas pressure, $P_{\text{fb}} = (\gamma - 1)E_{\text{fb}}$, which can be considered as a crude approximation for effective pressure of the hot gas (Ageritz & Kravtsov, 2015) or subgrid turbulence unresolved in simulations and can allow for more efficient transfer of supernova momentum and energy into the surrounding ISM. The remaining $1 - f_{\text{fb}}$ fraction of energy is added to the thermal energy of the gas in

the cell containing the stellar particles. The dissipation time scale is adopted to be $t_{\text{dis}} = 10$ Myr for all simulations, comparable to a couple crossing times in massive giant molecular clouds.

In addition, we model momentum injection due to radiation pressure, stellar winds, and supernovae by directly adding momentum into surrounding cells using the rate:

$$\dot{p}_{\text{rad}} = (\eta_1 + \eta_2 \tau_{\text{IR}}) \frac{L(t)}{c}, \quad (3.3)$$

where τ_{IR} is the optical depth in the infrared (IR) band, $L(t)$ is the luminosity of the stellar population, taken from the stellar evolution code **STARBURST99** (Leitherer et al., 1999). The first term corresponds to direct momentum injection from UV photons, where η_1 is chosen to be unity. The second term corresponds to momentum transfer by IR photons re-emitted by dust after the UV photons are absorbed, where $\eta_2 = 2$ is adopted in Agertz & Kravtsov (2015).

3.2.3 Cosmic Ray Feedback

Simulations of isolated disk galaxies in Booth et al. (2013) and Salem et al. (2014) show that cosmic ray (CR) driven winds contain more cool/warm gas ($T < 10^5$ K) than the winds driven via direct momentum and thermal energy injection from supernovae. This is because in the CR-driven winds the gas is gradually accelerated by a large-scale pressure gradient established by the cosmic rays diffusing out of the disk, while in the standard winds produced by energy and momentum injection at SN sites the gas is launched at large velocity and is then shock-heated to high temperatures. Therefore, we also consider a new re-simulation of the same halo as in the other simulations, but instead incorporates the CR feedback using the isotropic diffusion model described in Booth et al. (2013).

Briefly, the CRs are assumed to be produced by supernovae with a fraction ξ_{CR} of the SN energy converted into the energy of cosmic rays: $E_{\text{CR}} = \xi_{\text{CR}} E_{\text{SN}}$. The rest of E_{SN} ,

is added to the thermal energy of the gas in the SN parent cell. In the simulation used in this analysis, we adopted $\xi_{\text{CR}} = 10\%$ consistent both with recent empirical evidence for CR acceleration in the SN remnants (Ackermann et al., 2013) and theoretical models (Malkov & Drury, 2001; Caprioli & Spitkovsky, 2014).

The cosmic rays are modelled as an ultra-relativistic ideal fluid with $\gamma_{\text{CR}} = 4/3$, which exerts pressure $P_{\text{CR}} = (\gamma_{\text{CR}} - 1)E_{\text{CR}}$, where E_{CR} is the energy density of the cosmic ray particles followed using a separate variable. The variable is evolved using the advection-diffusion equation, which includes terms corresponding to cooling losses due to decays and Coulomb interactions of CR with gas and magnetic fields, an isotropic diffusion term assuming the constant diffusion coefficient of $\kappa = 3 \times 10^{-27} \text{ cm}^2 \text{ s}^{-1}$, and a source term due to the SN remnant production of cosmic rays. We also include the heating term into the energy equation of baryon gas due to interactions with CRs (see details in Booth et al., 2013).

It is important to keep in mind that properties of galaxy forming in a halo and properties of its CGM may depend on the choices of our model parameters (e.g, the diffusion coefficient). The parameters have sizeable uncertainties and thus, in principle, they need to be varied to explore their full effect on the CGM (e.g Salem et al., 2014). In this pilot study, we choose to investigate results only for the fiducial parameter values, chosen to be consistent with current observational constraints and theoretical expectations.

Note also that CRs can lead to gamma ray emission via pion production. However, in our model, we find that only $< 1\%$ of the total energy would be emitted as gamma rays, consistent with a detailed CR modelling for the Milky Way (Strong et al., 2010). Therefore, the gamma rays should not have a significant effect on the low-density circumgalactic medium, but can be used as interesting observational constraints on the cosmic ray feedback in the future.

3.3 CGM Analysis

3.3.1 Lines of sight and ion column densities

Observed quasar spectra reveal existence of a wide range of ion species tracing different temperatures, which indicates that the CGM is multi-phase. It is thus important to use the entire range of ion species to test whether simulations reproduce the observed CGM structure.

We compute abundances of all the ions for which extensive observational measurements have been reported in the literature. These ions cover a wide range of ionization energies, E_{ion} , as shown in Table 2. We will adopt a convention to name species with $E_{\text{ion}} < 54.4$ eV as *low ions* (e.g., H I, Mg II, Si II, Si III, Si IV and C II), those with $54.4 < E_{\text{ion}} < 100$ eV as *intermediate ions* (C IV), and those with $E_{\text{ion}} > 100$ eV as *high ions* (O VI).

Following Smith et al. (2011) and Hummels et al. (2013), we assume that the number density of ion j of an element X can be computed as:

$$n_{X_j}(n_H, T, Z) = f_{X_j}(T, n_H) f_X(Z) n_H, \quad (3.4)$$

where $f_X(Z) = n_X/n_H$ is the fraction of the element X relative to hydrogen and $f_{X_j}(T, n_H) = n_{X_j}/n_X$ is the ionization fraction of X in state j .

We compute an interpolation table of $f_{X_j}(T, n_H, J_{\text{UV}})$ using the CLOUDY code (version 13.02; last reviewed in Ferland et al., 2013) under assumption of ionization equilibrium for a grid of n_H , T , and redshift (z , corresponding to different cosmic mean $J_{\text{UV}}(z)$). For J_{UV} we adopt the redshift dependent ionization background HM05 (Haardt & Madau, 2012), which includes contribution from QSOs and galaxies and assume that gas is optically thin throughout the CGM. We assume solar pattern of heavy element abundances (Asplund et al., 2009) and solar metallicity of $Z_{\odot} = 0.02$, and approximate the mean molecular weight by a constant $\mu = 0.62$ for ionized gas with $T > 10^4$ K, and $\mu = 1$ otherwise.

The column density of ion X_j along a given line of sight through simulation volume is computed as:

$$N_{X_j} = \int_L n_{X_j}(n_H, T, Z, z) dl, \quad (3.5)$$

where n_{X_j} is 3D number density.

To sample the CGM uniformly under all possible viewing angles, we uniformly sample impact parameter d between the putative QSO and simulated galaxy and solid angle at the point in the line of sight (LOS) closest to the galaxy. To make sure that column densities for the LOSs near the edge of the box (i.e., large d) are not biased low, we have chosen a large box size $L_{\text{box}} = 1$ Mpc (comoving) compared to the virial radius of the simulated dark matter halo. Furthermore, to ensure that column densities are not biased up to $d \sim 2.5R_{\text{vir}}$, we do not include the LOS with path lengths $\lesssim 0.27L_{\text{box}}$ in our analysis and plots.

3.3.2 Curve of Growth Analysis

In observational samples, direct column density measurements are often available only in relatively high signal-to-noise and high resolution spectra, while some studies only provide the equivalent width W_r of absorbers. To combine all the data sets in a uniform manner, we follow Hummels et al. (2013) and apply a curve of growth analysis to convert from column density to equivalent width and vice versa using an approximation of equation (9.8) & (9.27) in Draine (2011):

$$W_r \approx \lambda \times \begin{cases} \sqrt{\pi} \frac{b}{c} \frac{\tau_0}{1+\tau_0/(2\sqrt{2})} & \text{for } \tau_0 < 1.254 \\ \left[\left(\frac{2b}{c} \right)^2 \ln \left(\frac{\tau_0}{\ln 2} \right) + \frac{b\gamma_{lu}\lambda_{lu}}{c^2} \frac{(\tau_0 - 1.254)}{\sqrt{\pi}} \right]^{1/2} & \text{for } \tau_0 > 1.254 \end{cases} \quad (3.6)$$

where

$$\tau_0 \approx \sqrt{\pi} \frac{e^2}{m_e c} \frac{N_l f_{lu} \lambda_{lu}}{b}. \quad (3.7)$$

W_r is the rest frame equivalent width in units of wavelength, τ_0 is the optical depth given the column density N_l and b parameter of the specie in state l , γ_{lu} , f_{lu} and λ_{lu} are the

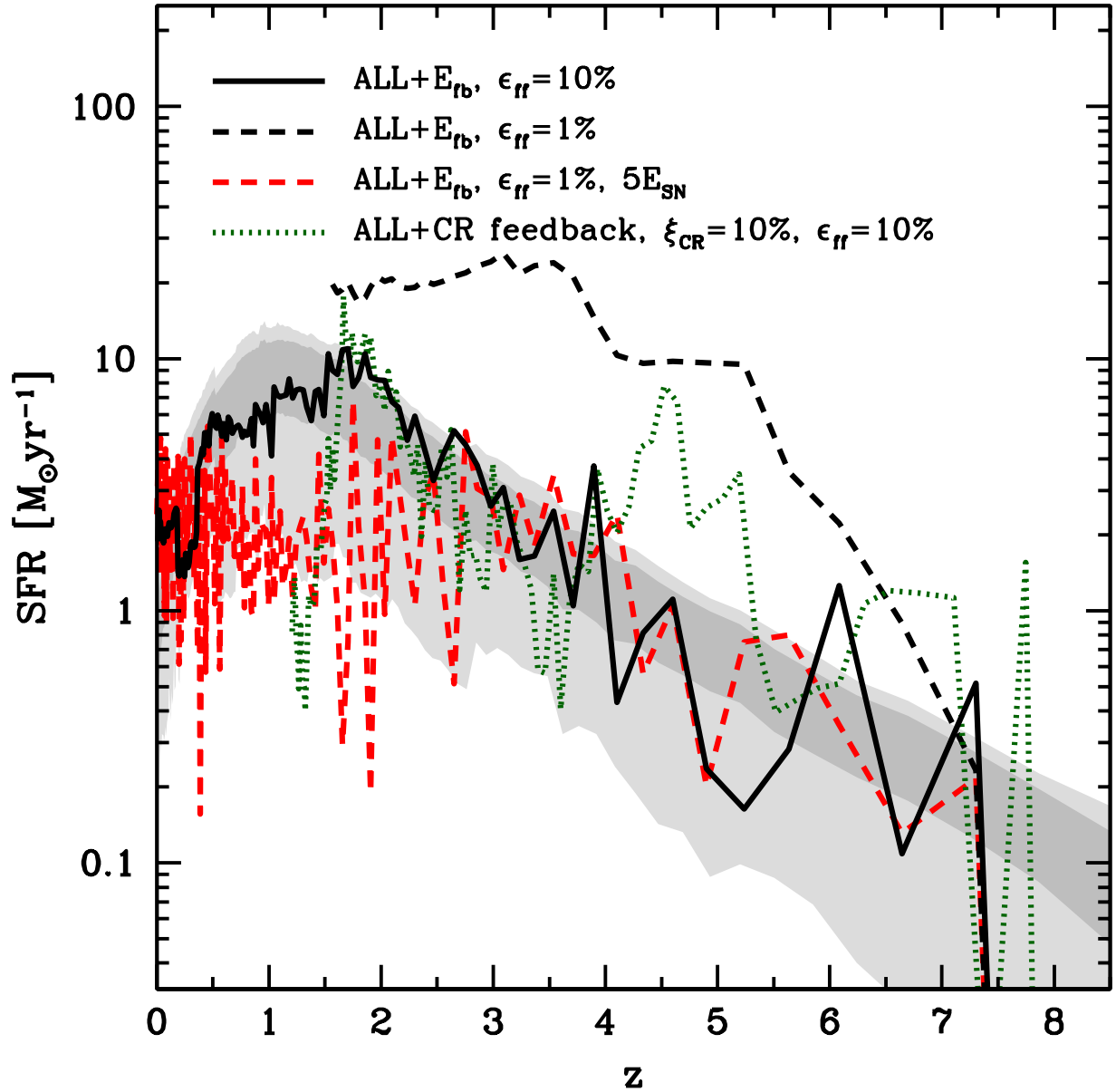


Figure 3.1: Star formation histories of the main galaxy in our simulations compared to the average histories for observed galaxies in halos of $M_{\text{vir}}(z=0) = 10^{12} M_{\odot}$ derived by Behroozi et al. (2013, with dark and light gray shaded bands showing the one- two-sigma confidence regions of their constraints, respectively). The SFHs in simulations are averaged over time bins of $\Delta t_{\text{SF}} = 100$ Myr.

intrinsic width, oscillator strength and wavelength of the specie transition from state l to u . The atomic data for all transitions are taken from a compilation study by Morton (2003).

To compute N from W_r , we invert the function $W_r(N, b)$ using its bivariate spline approximation. When only W_r measurements are available, the set of b values are determined by a Gaussian width σ , as $b \equiv \sqrt{2}\sigma$, which are taken from the literature (e.g., see Hummels et al., 2013; Liang & Chen, 2014). To estimate W_r from N in simulations, we treat b as a free parameter, because simulations do not model non-thermal components of the gas. We discuss this in detail in section 3.5.

3.4 CGM in Simulations

In this section we discuss the main qualitative features of the CGM in our simulated galaxies. We discuss their dependence on the feedback models (different implemented physics), as well as varying parameters of star formation and feedback recipes.

Figure 3.1 shows star formation histories of all the simulations used in this study compared to the range of star formation histories derived for observed galaxies hosted by halos of similar mass Behroozi et al. (2013). We can see that the main progenitor in the ALL_Efb_e001 run has star formation rates at $z > 2$ that are up to an order of magnitude higher than the rate inferred by Behroozi et al. (2013). This leads to an overestimate of stellar mass of the final object by a factor of two and creation of a dense central spheroid in the stellar distribution of the galaxy at low z . In other runs, feedback is sufficiently efficient to suppress star formation to the level consistent with the inference of Behroozi et al. (2013). At lower z the star formation rate in the fiducial run ALL_Efb_e010 continues to track the empirical SFH, while the rates in the ALL_Efb_e001_5ESN and ALL_e010_CR runs fall below the gray band. As we discuss below, the bursty nature of star formation in the ALL_Efb_e001_5ESN simulation, however, allows it to drive significant outflows out of the central disk even at low redshifts, as can be seen in Figure 3.2. The ALL_e010_CR run tracks the SFH derived from observations quite well down to $z = 1.5$, although the SFH starts to decrease sharply after

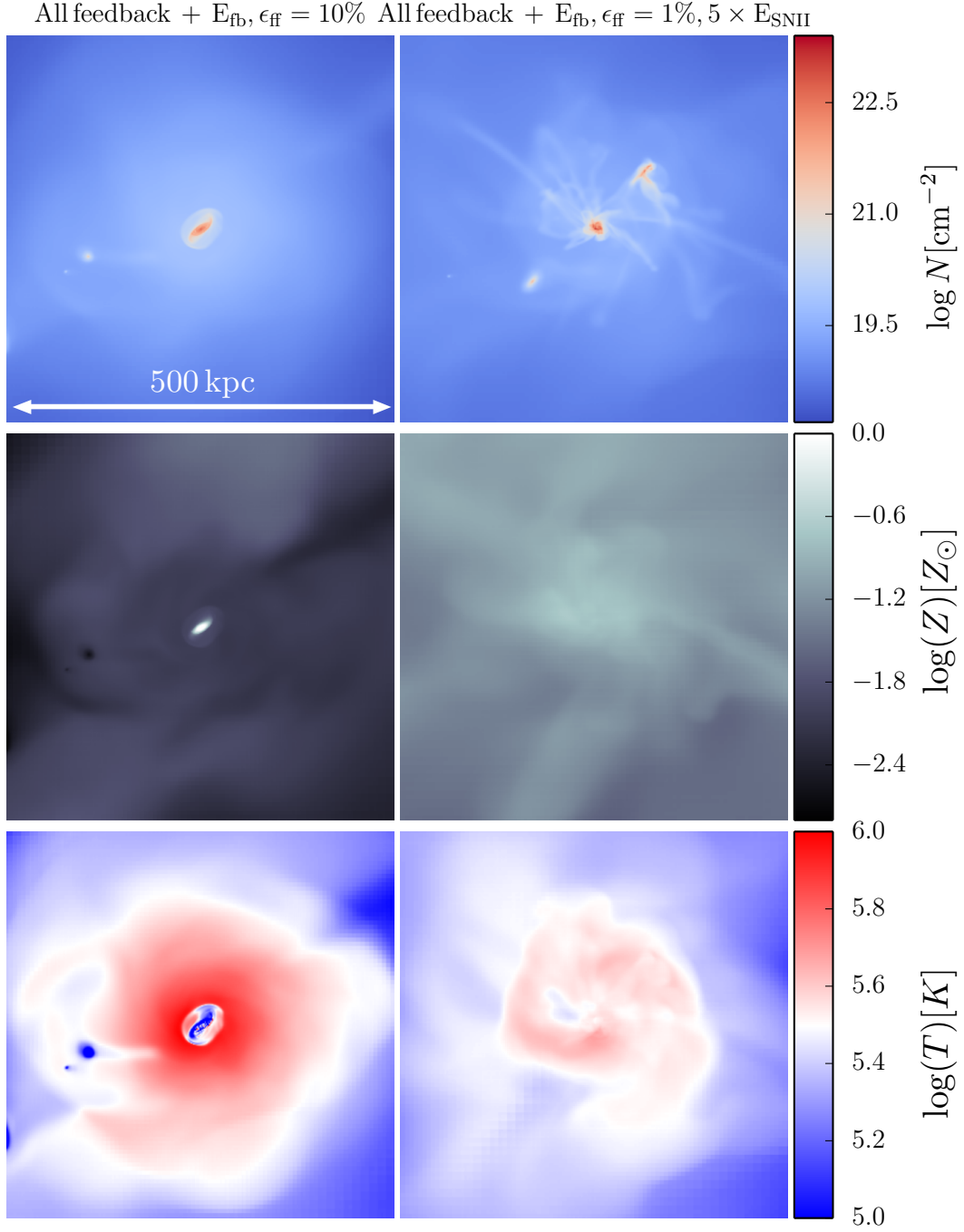


Figure 3.2: Maps of the total gas column density, metallicity, and mass-weighted temperature of the fiducial and $5E_{\text{SNII}}$ models at $z = 0$. The maps show the outflow streams aided by the boosted supernova events while the fiducial run maintains a normal star-forming disk. As a consequence, gas metallicity is well-mixed in $5E_{\text{SNII}}$ model out to large radii at 10% solar to solar value. In contrast, despite the disk in the fiducial run having solar metallicity, the gas hovers at a few percent solar in the circumgalactic space. Note that the color of the star-forming disk in the fiducial run has been clipped in order to ensure the dynamic range of the color map in the halo.

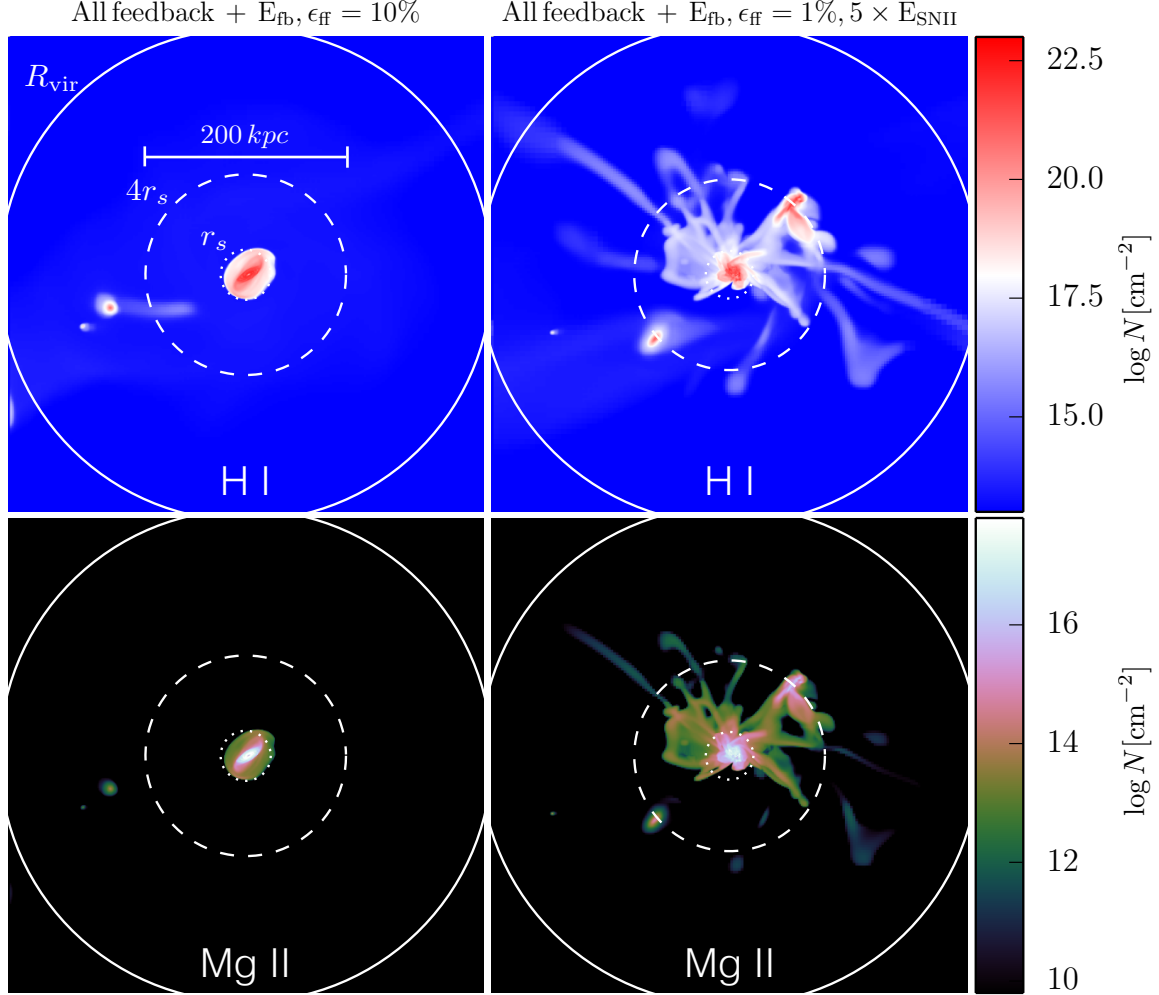


Figure 3.3: The spatial column density distribution (500 kpc across; $R_{\text{vir}} \approx 250$ kpc) of the chemically enriched CGM as probed by HI and Mg II at $z = 0.0$. Solid circle represents R_{vir} , dashed circle is $4r_s$ and dotted circle is r_s . Our fiducial model with star-formation efficiency $\epsilon_{\text{ff}} = 10\%$ produces a stellar component consistent with observations (Agertz & Kravtsov, 2015, 2016) but fails to produce an extended multi-phase gaseous halo. On the other hand, our strong feedback model (with five times supernova energy, $5 \times E_{\text{SNII}}$) matches the CGM profiles despite destroying the star-forming disk. This shows that the CGM provides sensitive orthogonal constraints on galaxy formation, especially feedback recipes.

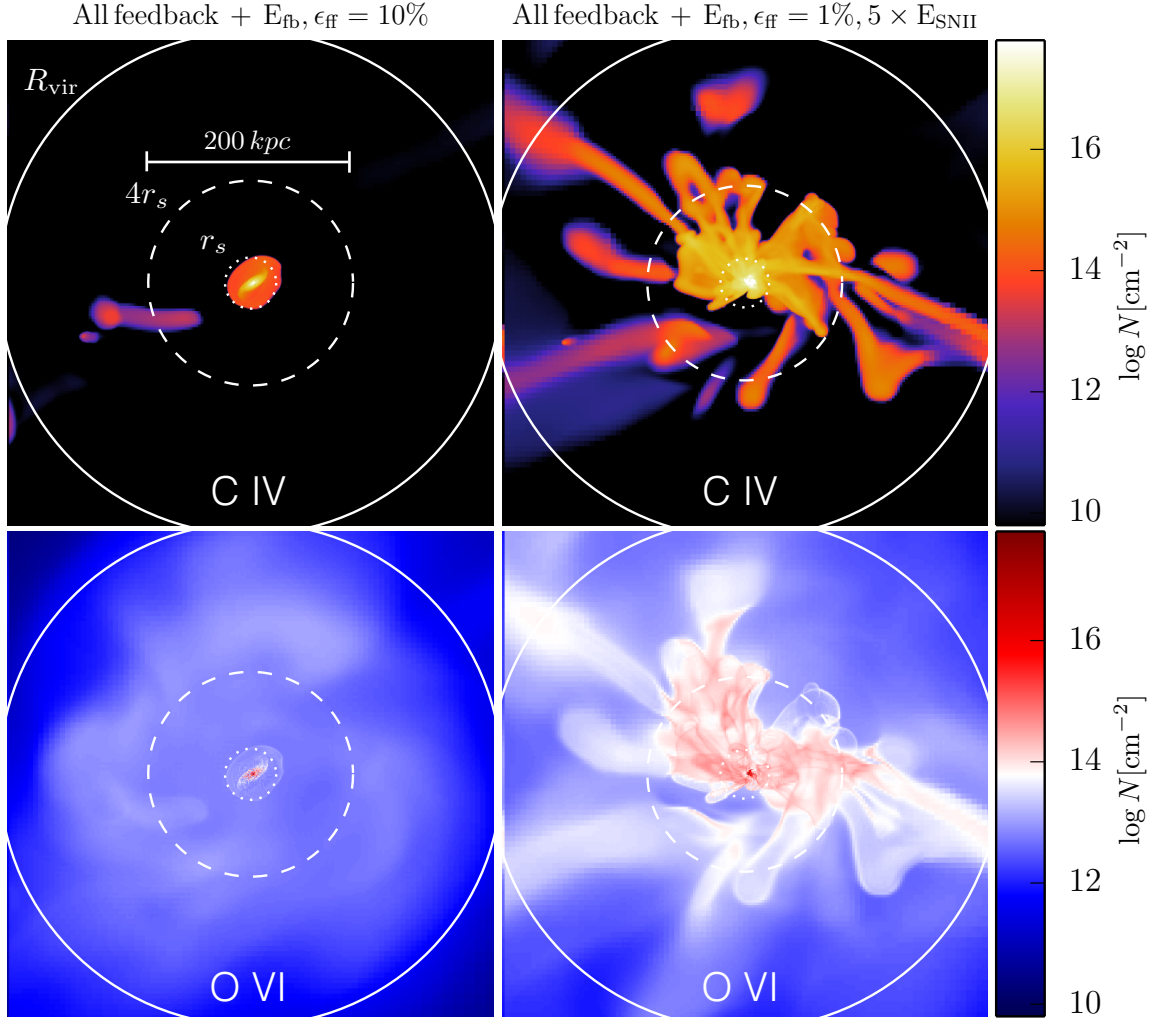


Figure 3.4: Similar to Figure 3.3, but for intermediate ion C IV and high ion O VI at $z = 0.0$. The multi-phase gaseous halos are progressively more extended from low and intermediate ions (e.g., Mg II and C IV) to high ion (O VI).

this epoch.

As shown by Agertz & Kravtsov (2015, 2016), the run `ALL_Efb_e010` not only reproduces the observed star formation history (SFH), but produces a galaxy with realistic stellar mass-halo mass ratio, bulge-to-disk ratio, metallicity, and the Kennicutt-Schmidt relation. On the other hand, the `ALL_Efb_e001_5ESN` run, with less efficient star formation but more energetic SNe events, under-produces stellar mass at low z , and fails to produce a disk-dominated galaxy.

3.4.1 *Main features of the CGM in simulations*

Figures 3.2, 3.3, and 3.4 show the overall thermodynamic properties of the CGM and column density maps of specific ions at $z = 0$. Here we focus on the runs `ALL_Efb_e010` and `ALL_Efb_e001_5ESN` and will compare results with other runs in the next subsection. The figures clearly show that these two runs produce not only very different central galaxies, but very different CGMs. The CGM of the `ALL_Efb_e010` run is hot, low-metallicity, and is almost devoid of gas with $T \sim \text{few} \times 10^5$ K in its inner regions. In contrast, the CGM in the `ALL_Efb_e001_5ESN` simulation does have such warm gas due to gas lifted out of the disk by recent outflows. These can be seen as “tails” and “streaks” in the column density map.

Figures 3.3 and 3.4 also show that the radial distribution of the high ions is more extended compared to the low ions. Overall, analyses of Ford et al. (2014) and Ford et al. (2015) show that higher ionization energy ions like, O VI, originate from the gas that was ejected from the disk with outflows at higher redshifts, while low-ionization energy ions originate in gas associated with recent outflows.

3.4.2 *Cosmic Ray Feedback on the CGM*

Recent studies (e.g., Booth et al., 2013; Salem & Bryan, 2014) have shown that CR-driven winds can create large outflow mass loading factors, especially in dwarf galaxies. Figures 3.5 and 3.6 show that both boosted SN-driven feedback in the run `ALL_Efb_e001_5ESN` and

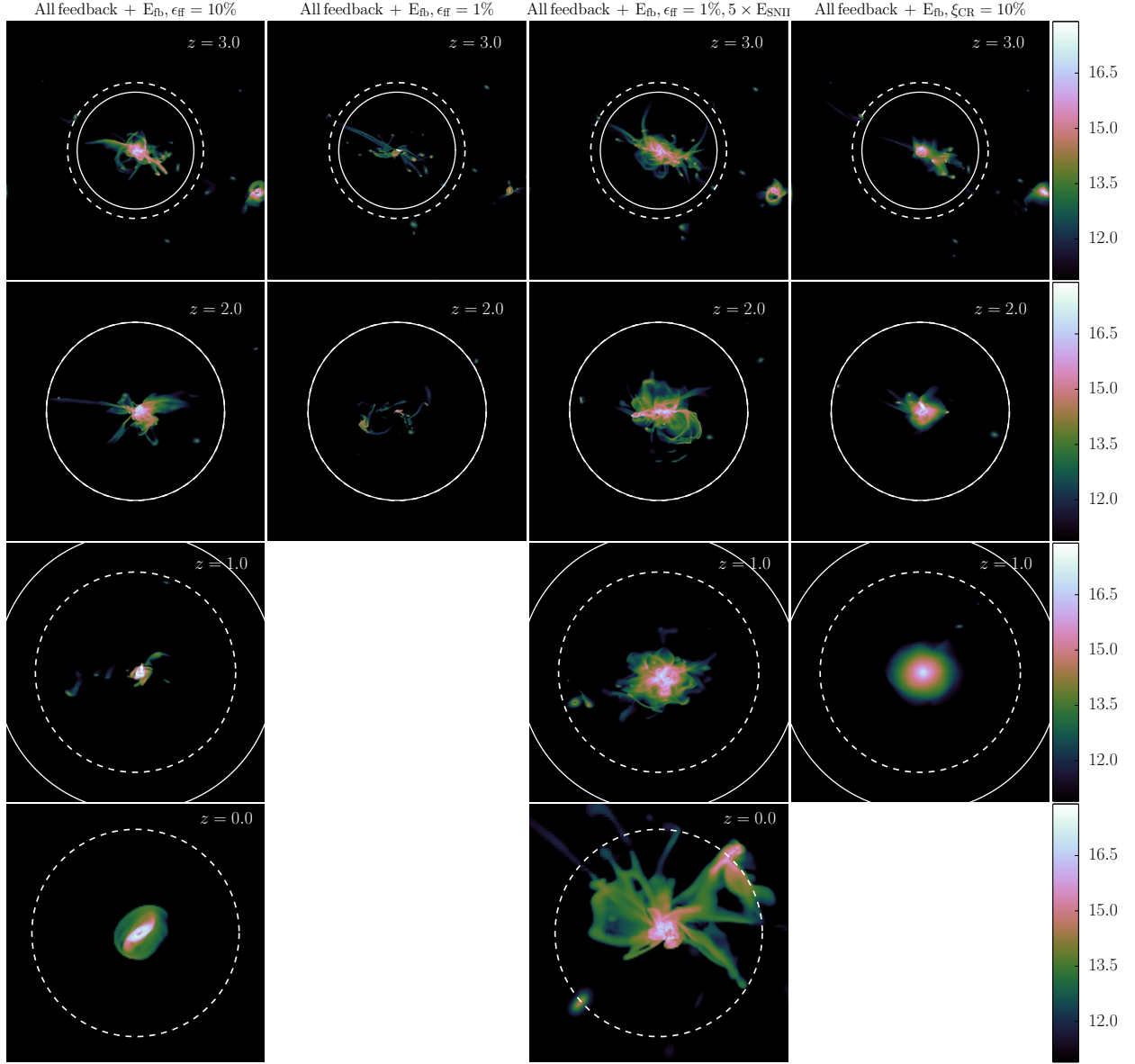


Figure 3.5: Model predictions of MgII column density maps for different redshifts. The region is 250 kpc (physical) on a side. Solid circles represents R_{vir} (it is outside the region at $z = 0$) and dashed circle represents $4r_s$. At high redshifts, $4r_s$ tracks the growth of R_{vir} , but at $z < 2$ R_{vir} becomes larger and grows faster than $4r_s$. For models that match observations of the CGM (ALL_Efb_e001_5ESN and ALL_e010_CR) the physical extent of the CGM is approximately a constant fraction of r_s at all z .

CR-driven outflows in the run `ALLe010_CR` produce column density profiles of similar shape and extent.

However, the processes that launch gas out into the CGM in these simulations are quite different. In the `ALL_Efb_e001_5ESN` run outflows are driven by energy and momentum injection at the sites of the SNe explosions in the disk and are quite bursty. In the CR run, on the other hand, the outflows are driven by a large-scale pressure gradient established by the cosmic rays diffusing out of the disk. The acceleration to this gradient is more gradual, which results in lower outflow velocities and resulting lower gas temperature (see Booth et al., 2013; Salem & Bryan, 2014; Salem et al., 2014, for more details on the wind properties in the CR feedback model). The resulting distribution of the CGM in the CR run is much smoother than in the `ALL_Efb_e001_5ESN` run, as can be seen in the right column of Figure 3.5.

3.4.3 *Evolution of the CGM in runs with different feedback models*

Figure 3.5 also shows evolution of the CGM in our different runs at different z . We choose to show Mg II $\lambda\lambda 2796, 2803$ lines as an illustrative example, but similar behavior is seen for other ions that trace the cold/warm gas (e.g., C IV). Figure 3.5 shows that the extended gaseous halo is established by $z = 3$ in all runs except `ALL_Efb_e001`. In the case of `ALL_Efb_e001`, the gas in the halo collapses onto the central disk where it is converted into stars, which results in high star formation rates (see Fig. 3.1). The ubiquity of the extended CGM in all runs at high z is due to the shallow gravitational potential at early epochs, which allows even moderate feedback to eject gas from the disk and launch it far into the halo (and even beyond). It is thus more difficult to use CGM for differentiating feedback recipes at high redshift ($z \geq 2 - 3$).

At redshifts $z \leq 2$, the galaxy in the simulation `ALL_Efb_e010` continues to form stars with the rate that matches the empirical average SFH of Behroozi et al. (2013) for galaxies of this mass. However, the feedback implementation fails to drive winds far into the halo

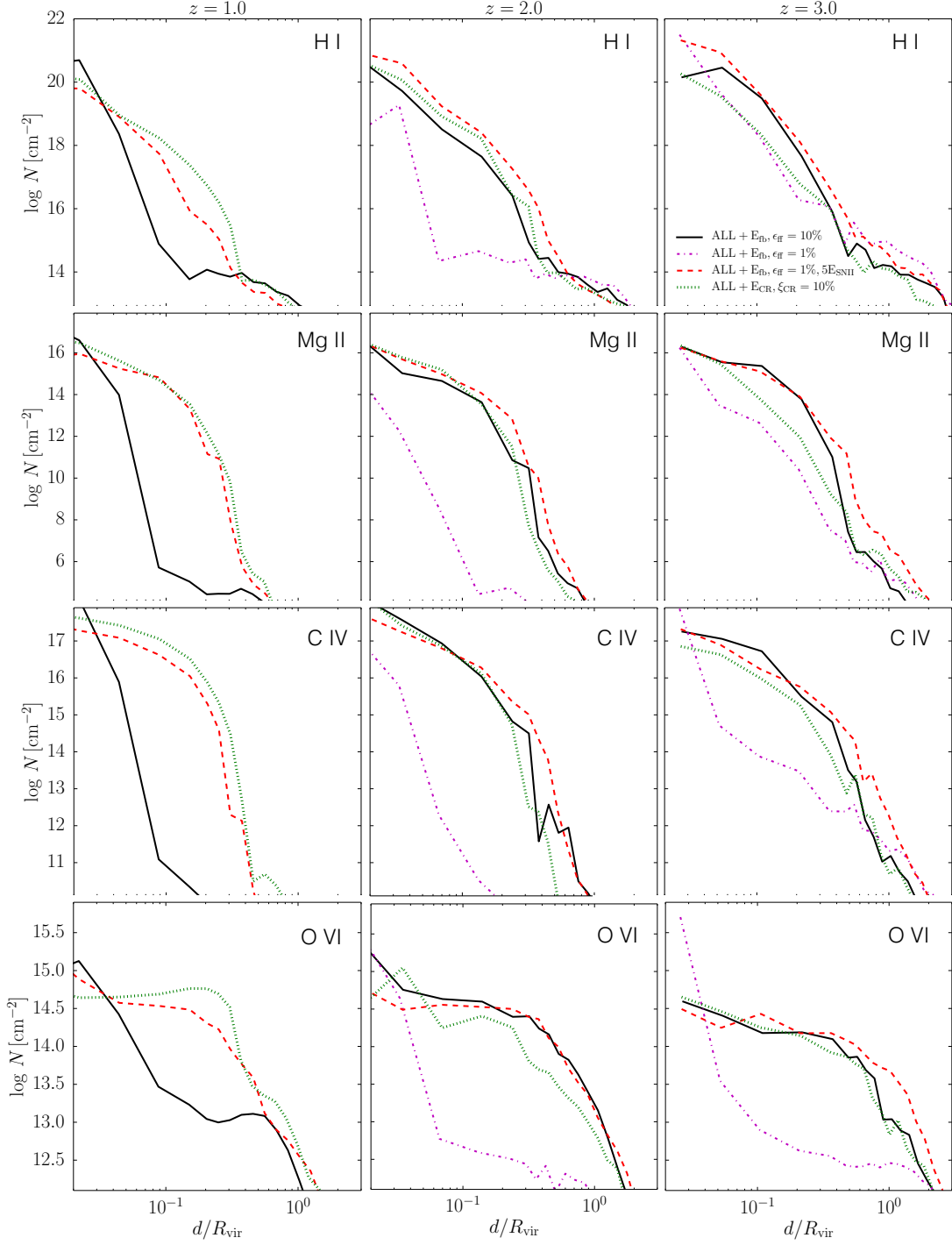


Figure 3.6: Column density profiles for four representative low to high ions for our models and their evolution from $z = 3.0$ to $z = 1.0$. At high redshift ($z = 2 - 3$), the CGM seems to be insensitive to the feedback recipes except for the very weak star formation efficiency $\epsilon_{\text{ff}} = 1\%$. At $z = 1$, the fiducial model starts to deviate from $5E_{\text{SNI}}$ and cosmic ray models. The extent of the profiles in units of R_{vir} shrinks with redshifts mainly due to the fast-evolving R_{vir} relative to the physical extent.

at $z < 1$; the SN action at these redshifts is limited to small fountains and stirring of the gas within the disk. On the other hand, the runs `ALL_Efb_e001_5ESN` and `ALL_e010_CR` drive significant outflows down to the smallest redshifts to which they were run. The winds lift cool gas from the disk into the halo in extended plumes, which are prominent in Figures 3.3–3.5.

Another interesting feature of the evolution shown in Figure 3.5 is that the extent of the high-column density area increases visibly slower than the virial radius (shown by the solid lines). The same can be seen in the column density profiles at different z in Figure 3.6, in which profiles are rescaled by host halo $R_{\text{vir}}(z)$ at the corresponding redshift. For example, although visible physical extent of the Mg II distribution increases from $z = 3$ to $z = 1$ in runs `ALL_Efb_e001_5ESN` and `ALL_e010_CR` in Figure 3.5, the extent in units of virial radius in Figure 3.6 decreases over the same redshift interval.

The virial radius of host halos has now been used in a number of studies to rescale the input parameter of galaxies at different redshifts in order to compare their profiles (Churchill et al., 2013; Werk et al., 2014; Liang & Chen, 2014; Bordoloi et al., 2014). However, the evolution of the profile in simulations indicates that the actual scaling may be different. This motivates a closer look at the evolution of the column density distribution and corresponding scaling across redshifts.

3.4.4 Evolution of the radial distribution of the CGM column density and rescaling with halo scale radius

The boundary of dark matter halos is usually defined as the radius enclosing a given density contrast Δ (e.g., $\Delta = 200$) relative to a reference density, such as the critical, $\rho_{\text{crit}}(z)$, or mean, $\rho_{\text{m}}(z)$, density of the universe. Common choices are R_{vir} (adopted in this paper), $R_{200\text{m}}$ and $R_{200\text{c}}$. These definitions of the radii are loosely motivated by simple models of spherical top-hat collapse. For real halos the increase of the virial radius roughly tracks the actual physical splashback radius R_{sp} (More et al., 2015), which delineates the outer region

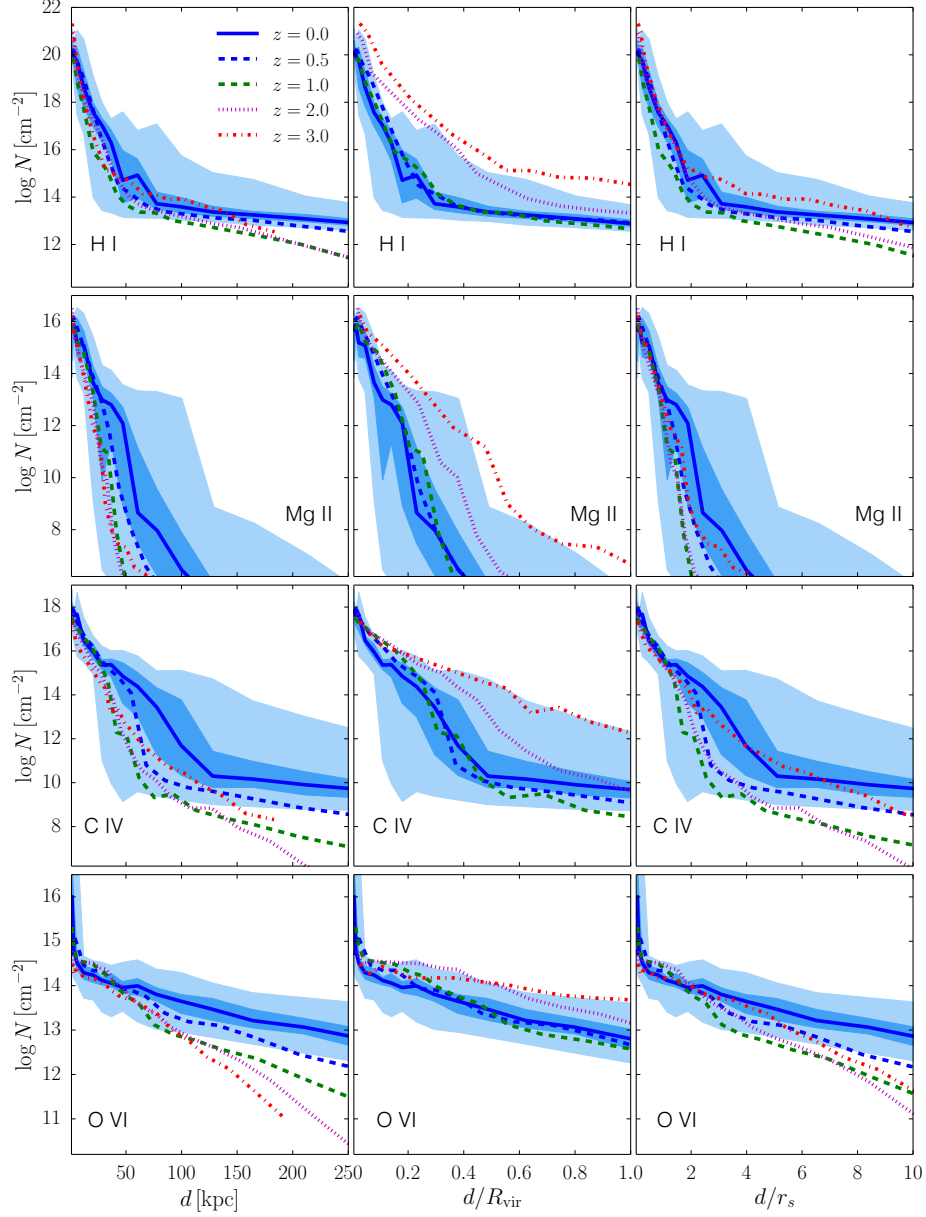


Figure 3.7: Column density profiles from low (H I and Mg II) to high (C IV and O VI) ions in the ALL_Efb_e001_5ESN run. Various lines are medians of the profiles at different redshifts (blue to red as low z to high z). Dark and light blue shaded regions represent 68% and 95% intervals for the profiles at $z = 0$, respectively. The left column shows the column density profiles in the inner regions remained relatively unchanged over 11 billion years at $z \leq 3$. This is why scaling with slower evolving scale radius, r_s (right panel) brings profiles at different z into agreement better than scaling with R_{vir} (middle column). At the same time, scaling with r_s also accounts for the different masses of the galaxies, which is necessary to put a population of observed galaxies in an equal footing. Note that higher ionization ions have progressively more extended profiles.

enclosing recently accreted matter. During epochs when halos accrete quickly, the entire halo profile scales well with the virial and splashback radii. However, when the accretion rate slows down the inner density profile remains almost static in physical units (e.g., Cuesta et al., 2008) and thus no longer scales well with the outer virial or splashback radius. Galaxies at $z \leq 2$ are in the latter regime, which means that scaling with the virial radius may not be optimal, especially since observations probe the CGM mostly in the inner halo ($\leq 4 - 6r_s$).

Evolution of the CGM for run `ALL_Efb_e001_5ESN` is illustrated in Figure 3.7, which shows evolution of the column density profiles of different ions from $z = 3$ to $z = 0$ in physical kpc and scaled with R_{vir} in the left and middle columns. Over the range of redshifts shown the dark matter halo mass grows from $M_{\text{vir}} = 3 \times 10^{11} M_{\odot}$ to $M_{\text{vir}} = 10^{12} M_{\odot}$, while R_{vir} grows by a factor of 4.6. The figure shows that the profile in physical kpc evolves only very weakly over 11 billion years since $z = 3$ and that scaling with virial radius introduces spurious evolution.

Recently, More et al. (2015) advocated a mass and radius definition to characterize evolution of the inner regions of halo profiles based on halo scale radius r_s , defined as the radius where logarithmic slope of the dark matter density profile is -2 . In particular, they argued that the virial radius in the fast accretion regime is approximately equal to $4r_s$, while in the slow accretion regime the r_s and $4r_s$ cease evolution (Bullock et al., 2001) while R_{vir} continues to grow due to pseudo-evolution (Diemer et al., 2013), as can be seen in Figure 3.8.

If the gravitational potential well in the inner region plays a major role in controlling the extent of the CGM, it is natural to assume that the CGM profiles should scale with r_s . Indeed, the right column of Figure 3.7 shows that the column density profiles scale with r_s much better than with R_{vir} .¹

Note that the scaling of profiles around galaxies of different stellar masses at a given z

1. To estimate the scale radius $r_s \equiv R_{\text{vir}}/c_{\text{vir}}$, we use the halo-concentration model of Diemer & Kravtsov (2015), implemented in the publicly available python code `Colossus`: <http://bdiemer.bitbucket.org>

with r_s (or any multiple of it) is very similar to scaling with R_{vir} at the same z because $r_s = R_{\text{vir}}/c_{\text{vir}}$ and $c_{\text{vir}} \propto M_{\text{vir}}^{-0.1}$ is a weak function of mass. However, both R_{vir} and c_{vir} evolve fast with redshift in a way that leaves r_s almost constant at $z < 2$ for galaxy sized halos. Thus, re-scaling of profiles of objects at different z with r_s and R_{vir} is quite different.

Although concentration at a fixed halo mass exhibits scatter of ≈ 0.14 dex, which will be added in quadrature to the scatter in the M_* at fixed M_{vir} during conversion from M_* to M_{vir} , Figure 3.7 shows that such additional scatter may be worth the reduced bias in the redshift rescaling. As we will show in the next section, observed column density profiles also appear to favor rescaling with r_s .

3.4.5 *Exponential form of the column density profile and scaling with ionization energy of ions*

The column density profiles plotted on the log-log scale in Figure 3.6 exhibit a sharp turnover at a particular impact parameter. Although often interpreted as indication of an “edge” to the ion distribution, such sharp turnover can also simply reflect the smooth exponential decrease of column density. Indeed, the same profiles of the ALL_Efb_e001.5ESN run shown on the log-linear scale in Figure 3.7 can clearly be well approximated by an exponential profile (i.e., a straight line in the log-linear plot). Some ions exhibit two exponential components, but the outer component is at the column densities that are well below the sensitivity of current observations.

The exponential profile $N(d) = N_0 \exp(-d/h_s)$, or equivalently $\log_{10} N(d) = \log_{10} N_0 - \frac{d}{h_s \ln(10)}$, is characterized by a single scale, which we will call the *scale height*, h_s , to differentiate from the exponential scale *lengths* of galaxy light distribution.

The scale height of the profile provides a simple estimate of the extent of the CGM. Nielsen et al. (2013) and Borthakur et al. (2015) show that the observed column density profiles of some of the ions can indeed be described by an exponential profile. Note, however, that in both studies the exponential distributions are fitted to the equivalent width profiles instead

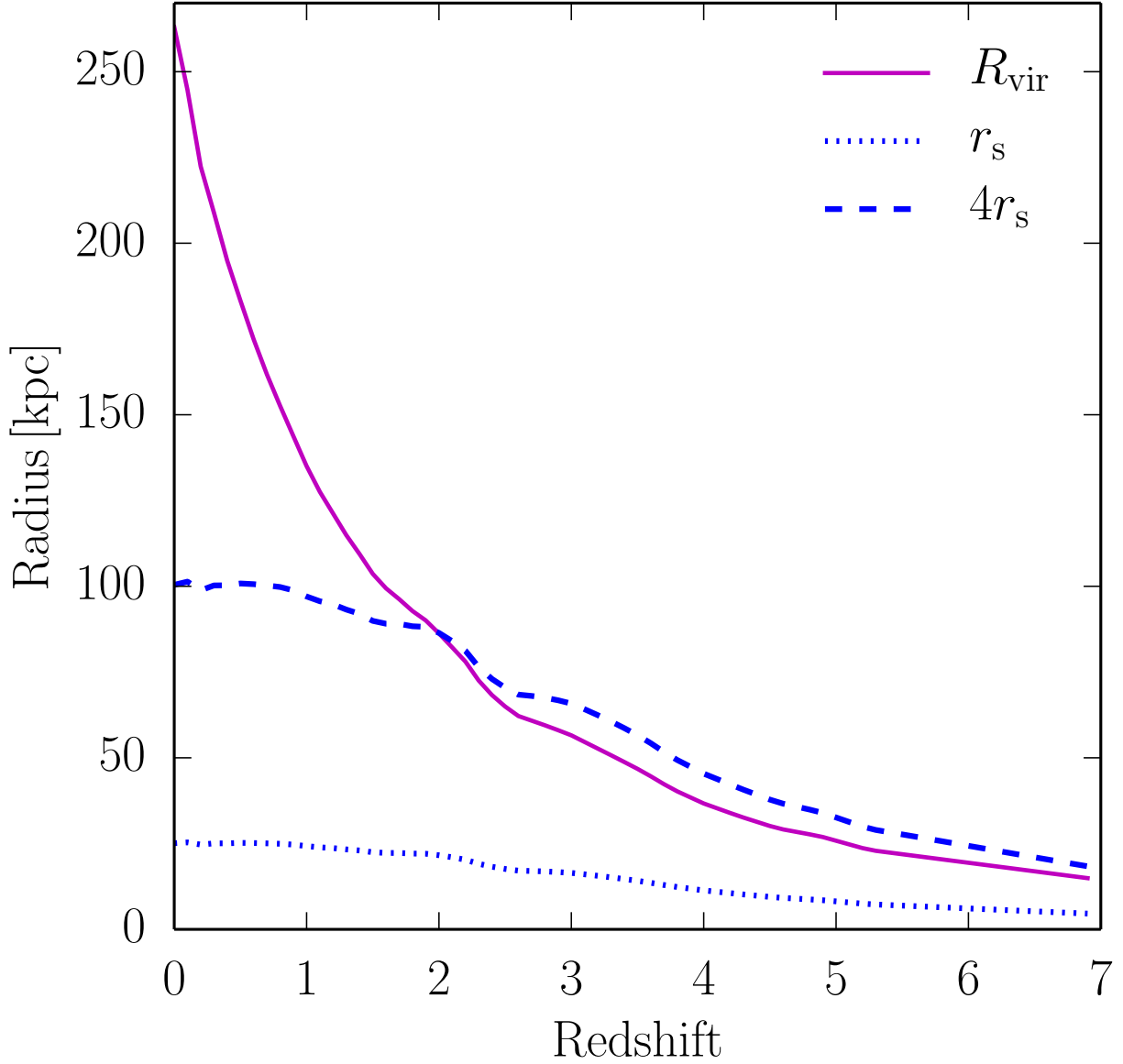


Figure 3.8: Evolution of the virial radius and halo scale radius, r_s . Virial radius is measured from the dark matter halo at each snapshot of the simulation `ALL_Efb_e001_5ESN`. The corresponding scale radius $r_s \equiv R_{\text{vir}}/c_{\text{vir}}$ is calculated via the $M_{\text{vir}} - c_{\text{vir}}$ model in Diemer & Kravtsov (2015). At high redshifts, the inner halo characterized by $4r_s$ tracks the growth of the halo virial radius, R_{vir} , until $z \sim 2$. Afterwards, R_{vir} continues to grow, while the inner region of the dark matter halo approaches a constant value.

of the column density. Due to non-linearity of the conversion between N and W_r , scale heights from our fits should not be *directly* compared with ones from W_r profiles. Instead, a conversion from N to W_r is needed before comparisons (see Figure 3.11).

The best fit scale heights of the run `ALL_Efb_e001_5ESN` for the inner exponential profiles of different ions are shown in Fig. 3.9 and presented in Table 3.2. Interestingly, Figure 3.9 shows that the scale height exhibits a tight power law scaling with ionization energy: $h_s \propto E_{\text{ion}}^m$. Previous theoretical (e.g., Hummels et al., 2013; Ford et al., 2015) and observational (Liang & Chen, 2014; Johnson et al., 2015b) studies have also found that higher ions have more extended distributions compared to the low ions. Here we present the first quantitative characterization of this trend and demonstrate that it is a remarkably tight function of the ionization energy of ions.

The best fit power law relation between scale height normalization to the virial radius, R_{vir} , or scale radius, r_s , is characterized by the slope m and offset b : $\log_{10}(x) = m \log_{10} E_{\text{ion}} + b$, where

$$\text{for } x = h_s/r_s : \quad m = 0.742 \pm 0.005, \quad b_{r_s} = -1.29 \pm 0.02. \quad (3.8)$$

$$\text{for } x = h_s/R_{\text{vir}} : \quad m = 0.742 \pm 0.005, \quad b_{R_{\text{vir}}} = b_{r_s} - \log c_{\text{vir}} \quad (3.9)$$

where $c_{\text{vir}} \approx 10.5$ is the concentration parameter of the halo at $z = 0$, computed from the $M_{\text{vir}} - c_{\text{vir}}$ relation (Diemer & Kravtsov, 2015). We also find a similar power law slope $m = 0.72$ for the `ALL_e10_CR` at $z = 1$. We estimated the best-fit scale heights h_s for each ion by a simple χ^2 minimization fit to the simulated scaled profiles $\log N_{\text{ion}} - d/r_s$. The fit is done to the range of radii, $0.1 < d/r_s < 4$, and column densities $\log N_c < \log N < 17.2 \text{ cm}^{-2}$ (21 for HI), where $\log N_c$ is the column density where we can clearly see the transition between the inner and outer exponential distributions. The limits are chosen to exclude the ISM of the galaxy and limit the fit only to the inner exponential component.

The errors on the parameters are estimated by bootstrapping the column density mea-

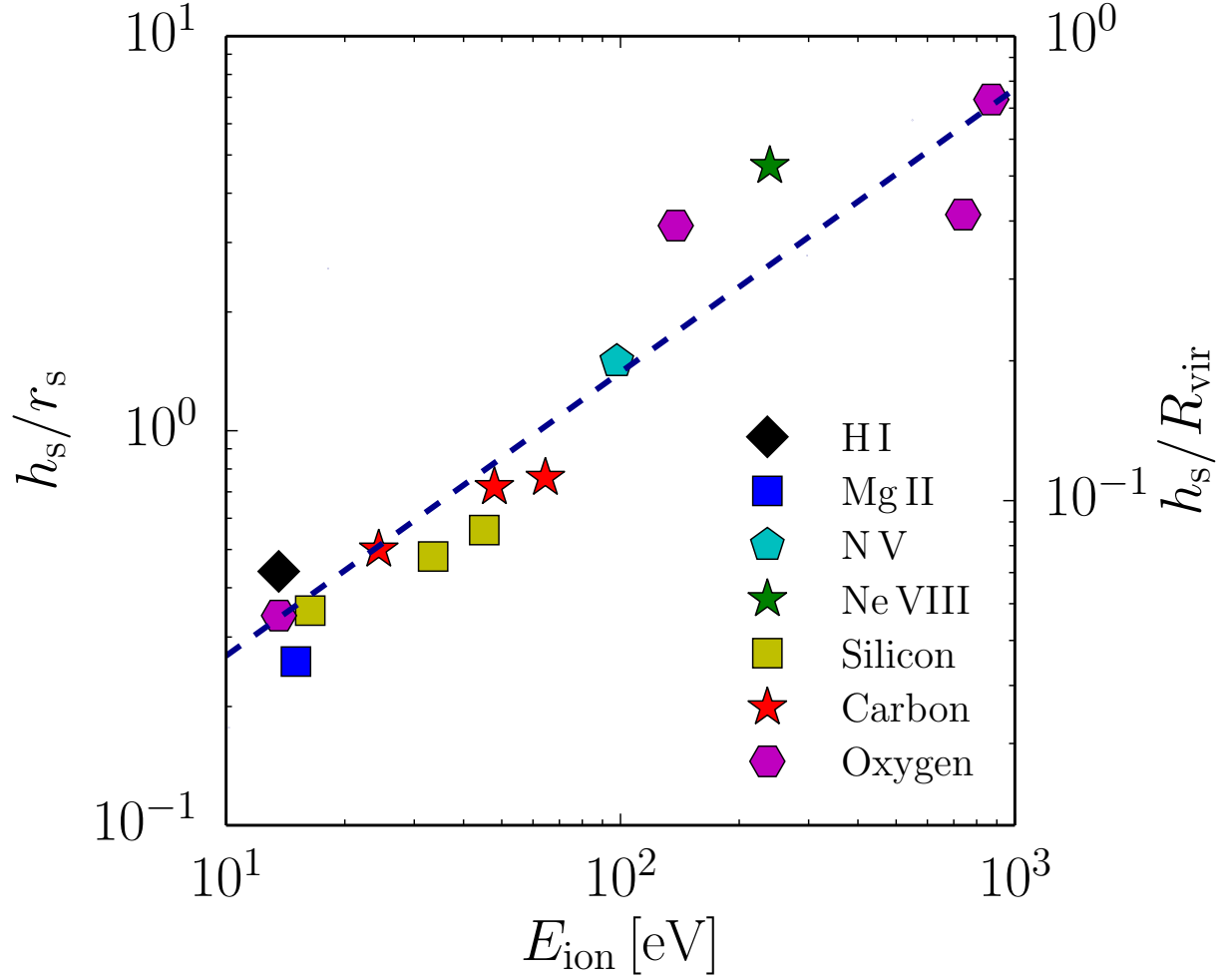


Figure 3.9: Scale height of the exponential profile of the column density as a function of the impact parameter predicted in our run ALL_Efb_e001_5ESN at $z = 0$ as a function of ionization energy of the corresponding ion. The left axis shows the scale height normalized to the halo scale radius, while the right axis shows the scale height normalized to the halo virial radius. Color points show results of the fits of the column density profile of a particular ion, while the dashed lines show the best fit power law relations (see eqs. 3.8-3.9).

surements of many lines of sight and thus account for the scatter in the profiles. We also varied the selected ranges of the data for the fit and find that h_s can shift systematically by $\sim 5 - 10\%$. Nevertheless, the power law dependence of h_s on E_{ion} remains the same. We therefore caution readers that when comparing scale heights with observations, the data should be fitted in the same range.

The trend of the scale height with ionization energy is likely due to two separate reasons. First, abundance of low ions is determined by photoionization equilibrium and their distribution traces a particular range of the ionization factor, U , and thus a particular range of gas density profiles (see, e.g., §16.9 in Mo et al., 2010). Ions with higher ionization energies, such as CIV and OVI, can be created both by photo- and collisional ionization. Overall, they favor regions with higher U and thus lower gas densities, which can be partly responsible for the trend we observe. The dependence of scale-height on the ionization energy of ions may thus be related to the fact outflowing plumes of gas expand as they move out to regions of lower density and pressure.

To test this idea, we carried out an analysis with and without an ionizing background to compare the relative contribution from photo- and collisional ionization. Indeed, with the addition of photo-ionization, low ions are lifted into higher states in the low density regime (at large distances). This effectively boosts the scale heights of high ions, while decreasing the scale height of the low ions. Thus, the ionization factor dependence on density is at least partially responsible for the $h_s - E_{\text{ion}}$ scaling.

The fact that higher ions arise from shocked gas that was carried out by older winds (e.g., Ford et al., 2015) may also contribute to the correlation. An NFW-like potential is quite shallow in terms of its logarithmic slope out to $r \approx 1r_s$ and steepens at large radii. It is thus tempting to associate the characteristic scale height of the low ions ($\approx 2r_s$) with the region of the shallow potential out to which recent winds propagate easily and then stop when the derivative of the potential steepens. However, we have checked the evolution of gas outflows in our simulations and found that even the recent outflows in the ALL_Efb_e001_5ESN run

Table 3.2: Scale Heights From Simulations

Ionization State	$E_{\text{Ion}} [\text{eV}]^a$	h_s/R_{vir}	h_s/r_s
HI	13.60	0.042 ± 0.004	0.44 ± 0.04
Mg II	15.04	0.025 ± 0.001	0.26 ± 0.01
N V	97.89	0.142 ± 0.012	1.50 ± 0.13
Ne VIII	239.09	0.447 ± 0.006	4.69 ± 0.07
C II	24.38	0.047 ± 0.003	0.50 ± 0.03
C III	47.89	0.068 ± 0.003	0.72 ± 0.04
C IV	64.49	0.073 ± 0.004	0.76 ± 0.04
Si II	16.35	0.033 ± 0.002	0.35 ± 0.02
Si III	33.49	0.046 ± 0.002	0.48 ± 0.02
Si IV	45.14	0.053 ± 0.003	0.56 ± 0.03
O I	13.62	0.032 ± 0.002	0.34 ± 0.02
O VI	138.12	0.315 ± 0.005	3.31 ± 0.05
O VII	739.31	0.336 ± 0.004	3.53 ± 0.04
O VIII	871.38	0.659 ± 0.012	6.91 ± 0.13

^aIonization energy from a compilation of atomic data in Morton (2003). Scale heights h_s from exponential fits, $\log_{10} N_{\text{ion}} = \log_{10} N_0 - d/(h_s \ln(10))$, of the first component of the column density profiles for ALL_Efb_e001_5ESN at $z = 0$. We see a power law dependence $h_s \propto E_{\text{ion}}^{0.74}$. Note that $h_s/r_s = c_{\text{vir}} \times h_s/R_{\text{vir}}$, where $c_{\text{vir}} \approx 10.5$ is the concentration parameter, computed via $M_{\text{vir}} - c_{\text{vir}}$ model in Diemer & Kravtsov (2015). Uncertainties on h_s are computed via bootstrap sampling.

propagate to considerably larger radii than r_s . The scale height of low ions is thus likely determined by the overall shape of the gaseous halo density profile and the corresponding profile of the ionization factor.

The *prediction* of the exponential profile and the tight correlation of the scale height with E_{ion} can be tested against observations. The correct treatment of the profile fitting in the observational data requires proper consideration of the upper limits and is out of the scope of our paper, but we plan to return to it in a future study.

3.5 Radial column density distributions: simulations vs. observations

To compare our simulations with observations, we have compiled a set of existing measurements of the CGM absorbers probed with different ions around galaxies with stellar mass

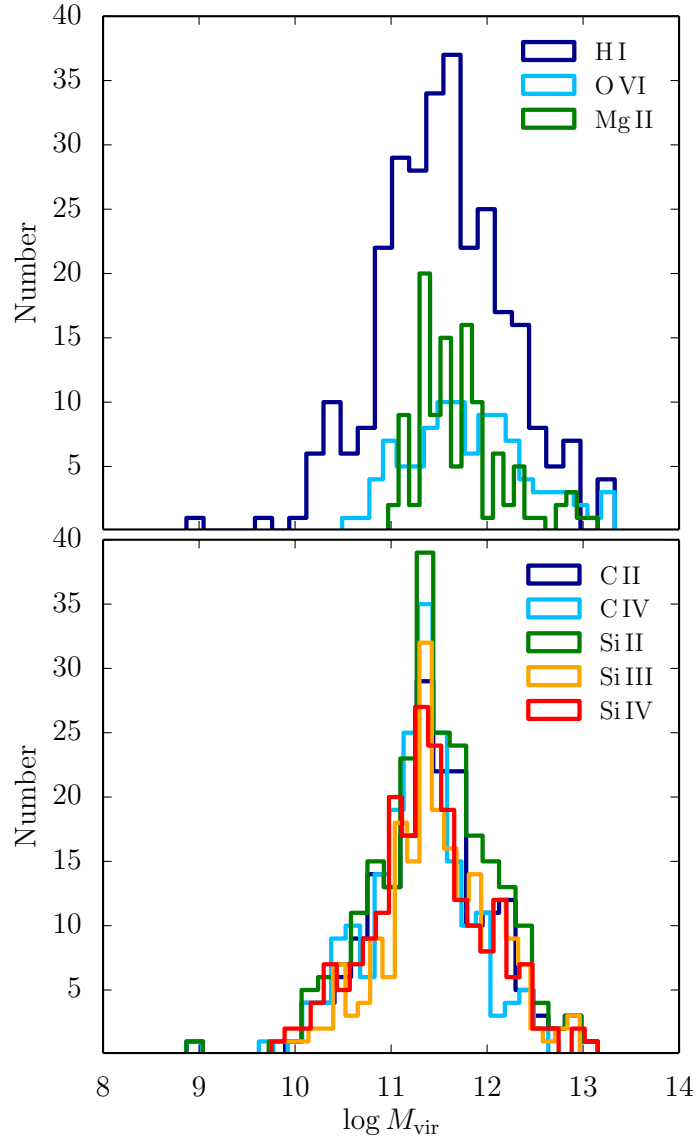


Figure 3.10: Distribution of halo mass M_{vir} probed by various ions in the compiled low redshift data set. See references in Table 3.3.

Table 3.3: Observation data sets of CGM

	# of Galaxies	$\langle z \rangle^a$	Impact Parameter Range (kpc)	$\log M_*/M_\odot^b$	Transitions
Johnson et al. 2015	71	0.271 ± 0.088	63 – 991	10.14 ± 0.92	Ly α , O VI
Liang & Chen 2014	195	0.041 ± 0.044	32 – 499	9.92 ± 1.20	Ly α , C II, C IV, Si II, Si III, Si IV
Bordoloi et al. 2014	43	0.027 ± 0.023	14 – 135	9.5 ± 0.48	C IV
Werk et al. 2013	44	0.221 ± 0.053	18 – 154	10.61 ± 0.50	Ly α , C II, Si II, Si III, Si IV
Tumlinson et al. 2011					O VI
Chen et al. 2010	75	0.239 ± 0.094	8.5 – 119	9.9 ± 0.58	Mg II
Steidel et al. 2010	512 (stacked) ^c	2.2 ± 0.3	10 – 125(280) ^d	9.85 ± 0.46^e	Ly α , C II, C IV, Si II, Si IV

^aMedian redshift and dispersion of the galaxy samples.

^bMedian stellar mass and dispersion of the galaxy samples.

^cMean equivalent width W_r measured from stacked background galaxy spectra.

^dMaximum impact parameter is 280 kpc for Ly α and 125 kpc otherwise.

^eMedian and dispersion of stellar mass for Steidel et al. (2010), a representative sample of galaxies in Reddy et al. (2012).

estimates. The data set is summarized in Table 3.3, which specifies the number of galaxies, their mean redshifts and stellar masses, and the range of impact parameters probed by a particular sample. The absorber galaxy samples span a wide range of stellar masses from dwarf galaxies with $M_* \approx 10^7 M_\odot$ to L_* galaxies with $M_* \approx 10^{11} M_\odot$ at $z \approx 0$, and Lyman-break galaxies with $\langle M_* \rangle \approx 10^{11} M_\odot$ at $z \approx 2$. Furthermore, Figure 3.10 shows the ranges of virial halo masses for the galaxies which generated a particular ion absorber, where virial masses were estimated from galaxy stellar masses using our adopted stellar mass–halo mass relation (see below). Figure 3.10 shows that although the overall range of halo masses is broad, individual ions may be probed by a narrower range of halo masses.

In this section, we first discuss details of how the comparison between observations and simulations is made. We then compare model predictions with observed column density, equivalent width and covering fraction profiles around galaxies and discuss the implications for galaxy formation and feedback.

3.5.1 *Rescaling observations and simulations for fair comparison*

Observations of low- z absorbers in our data set span more than four decades of stellar mass (Chen et al., 2010a; Tumlinson et al., 2011; Liang & Chen, 2014; Werk et al., 2014; Bordoloi et al., 2014; Johnson et al., 2015b), while our simulations model a single progenitor of mass close to L_* galaxies. We thus need to rescale both data and simulations in a way that makes the comparison sensible.

In the previous section, we showed that the column density profiles in simulations scale most optimally across redshifts not with the virial radius of their parent halo, but with the scale radius of their halo density profile, defined as the radius where the profile has logarithmic slope of -2 . To test whether a similar scaling applies to observed profiles in Figure 3.11 we show the equivalent width profiles of absorbers in the ALL_Efb_e001_5ESN simulations along with observational data at redshifts $z \in 0 - 0.3$ and $z \approx 2$ with radii in physical kpc, and rescaled by the virial radius of host halo $R_{\text{vir}}(z)$ and $r_s(z)$. We show

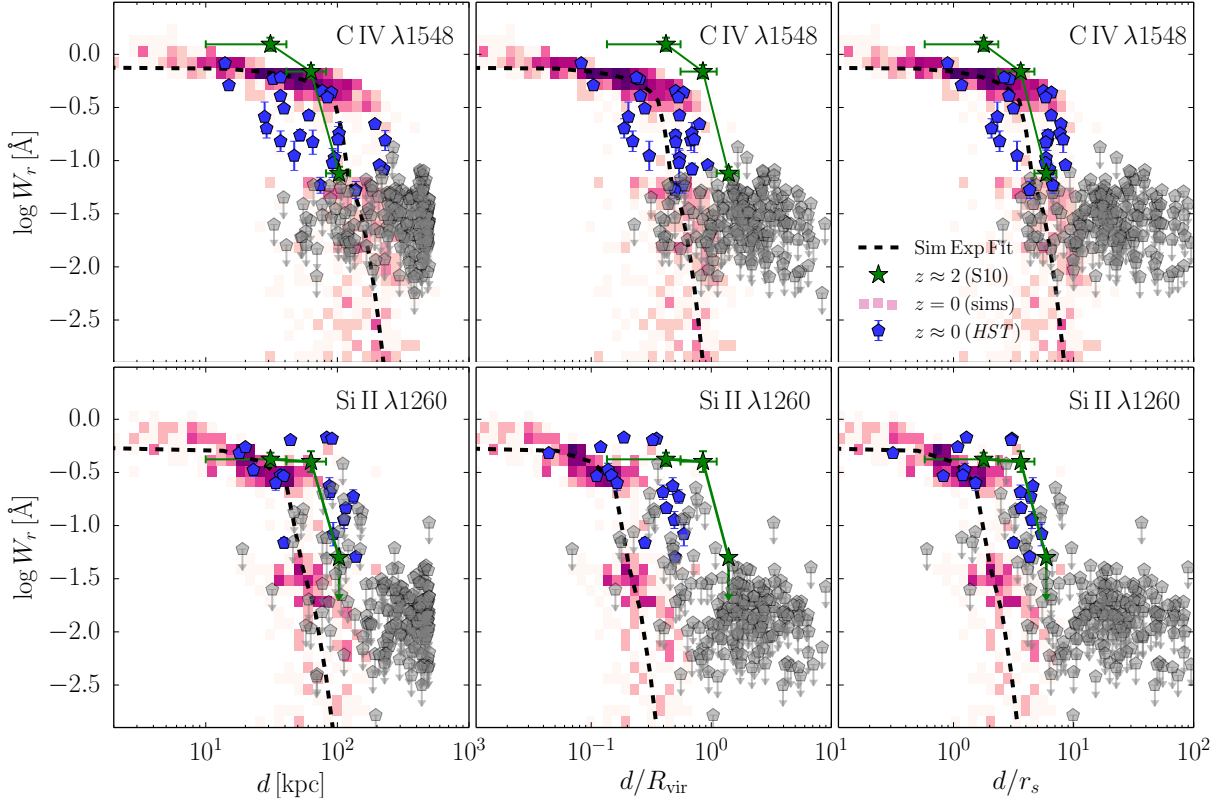


Figure 3.11: Equivalent width distribution with impact parameter for the low- z galaxies observed by the COS instrument onboard *HST* (see Table 3.3 for description of observational samples and references) and $z \approx 2$ stacked measurements by Steidel et al. (2010). Blue pentagons are detections, while gray pentagons with arrows are upper limits. The color 2D histogram shows distribution of equivalent widths for the lines of sight through the halo of the MW-sized galaxy in the ALL_Efb_e001_5ESN run, assuming turbulent broadening of $b_{\text{NT}} = 20 \text{ km s}^{-1}$. Black dashed lines represent W_r profiles converted from the best fit exponential profiles to the simulated column density profiles for corresponding ions. The left panel shows profiles against the impact parameter d in physical kpc, while in the middle and right panels d is rescaled by the virial radius, R_{vir} , and halo scale radius, r_s , respectively.

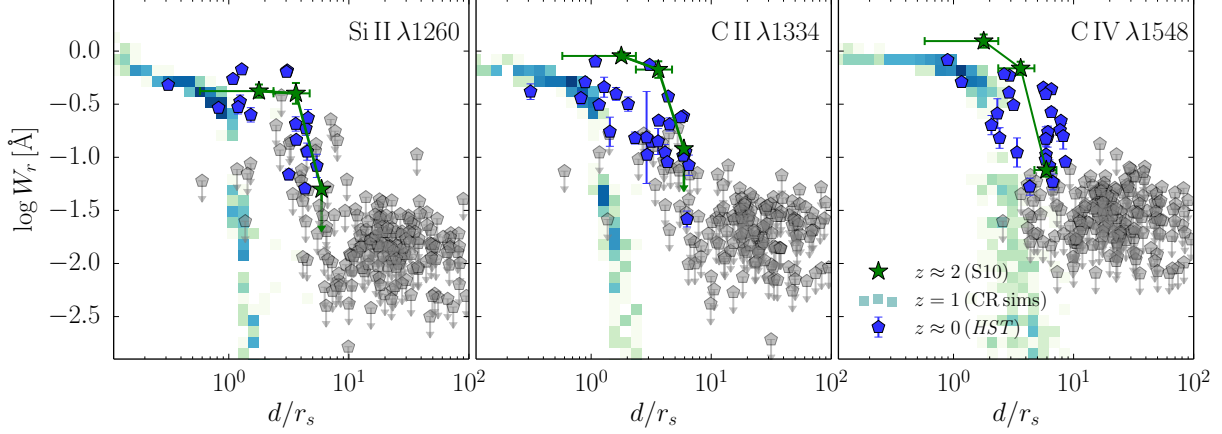


Figure 3.12: Equivalent width profiles for three representative ions in the run with cosmic ray feedback at $z = 1$ as a function of impact parameter scaled by the scale radius, d/r_s (2D color histogram) compared to observations (points; the point types are the same as in Figure 3.11).

evolution of the W_r profiles for C IV and Si II only, because the profiles for these ions exhibit some of the sharpest “edges” and because results for other ions are similar.

We convert our predicted column density into rest frame equivalent width W_r using the curve of growth analysis discussed in section 3.3. To account for turbulence not fully resolved in simulations in the relatively low resolution regions of gaseous halos, we use the b parameter as $b^2 = b_T^2 + b_{NT}^2$, and adopt $b_{NT} \approx 20 \text{ km s}^{-1}$ (see also Oppenheimer et al., 2012), although we checked that using b_{NT} also provides a reasonable match to observations.

In the simulations the virial radius of the galaxy is known from the density profile of the matter distribution. In observations, however, we need to estimate it statistically using stellar masses. We do this using the average stellar mass-halo mass relation derived by Kravtsov et al. (2014) using the abundance matching. This relation uses the measurement of galaxy stellar mass function by Bernardi et al. (2013), which is based on updated photometry of the SDSS main galaxy sample that corrects for background subtraction errors in the previous SDSS data releases.

We assume a non-evolving $M_* - M_{\text{vir}}(z)$ relation (e.g., Behroozi et al., 2013) to convert M_* to M_{vir} for each observed galaxy in the dataset. We then use M_{vir} to compute virial radius

using its definition $R_{\text{vir}}(z) = [3M_{\text{vir}}(z)/(4\pi\Delta_{\text{vir}}\bar{\rho}(z))]^{1/3}$, where $\Delta_{\text{vir}}(z)$ is the redshift-dependent “virial” density contrast computed using approximation to the spherical collapse overdensity given by Bryan & Norman (1998).² The scale radius is obtained from the virial radius assuming the median concentration of halos of mass M_{vir} at redshift z , predicted by the model of Diemer & Kravtsov (2015): $r_s = R_{\text{vir}}/c_{\text{vir}}(M_{\text{vir}}, z)$.

Figure 3.11 shows that evolution in the physical radius from $z \approx 2$ to the present epoch in the observational sample is quite mild (as found previously by Chen, 2012; Liang & Chen, 2014). The figure also shows that the drop in the W_r profiles is sharper when r_s rescaling is used instead of R_{vir} . Furthermore, the profile rescaled by r_s does not evolve significantly between $z \approx 0$ and $z \approx 2$.

This alignment of the profiles at different z is remarkable because properties of the low- z and $z \approx 2$ populations of galaxies are drastically different. For example, typical star formation rates differ by over an order of magnitude. Given the theoretical arguments for r_s scaling and the fact that it appears to work for real CGM, we will scale impact parameters in the simulations and observational sample by the radius of $4r_s$ in the comparisons presented below. We choose the multiple of four because at $z \approx 2$ this radius is actually close to the virial radius of galaxy halos (see More et al., 2015, for a detailed discussion).

3.5.2 *Comparison of the column density profiles in simulations and observations*

Figures 3.13, 3.14 and 3.15 compare column density distributions with $d/4r_s$ in the two simulations that were run to $z = 0$, ALL_Efb_e010 and ALL_Efb_e001.5ESN, and observations for all commonly observed transitions. It is immediately apparent that the absorber column densities in the fiducial run ALL_Efb_e010 with $\epsilon_{\text{ff}} = 10\%$ for all ions are consistently low by up to several orders of magnitude compared to observations. Interestingly, Agertz &

2. Adopting other commonly used definitions of the “virial radius”, such as $R_{200\text{m}}$ or $R_{200\text{c}}$, would not qualitatively affect our conclusions.

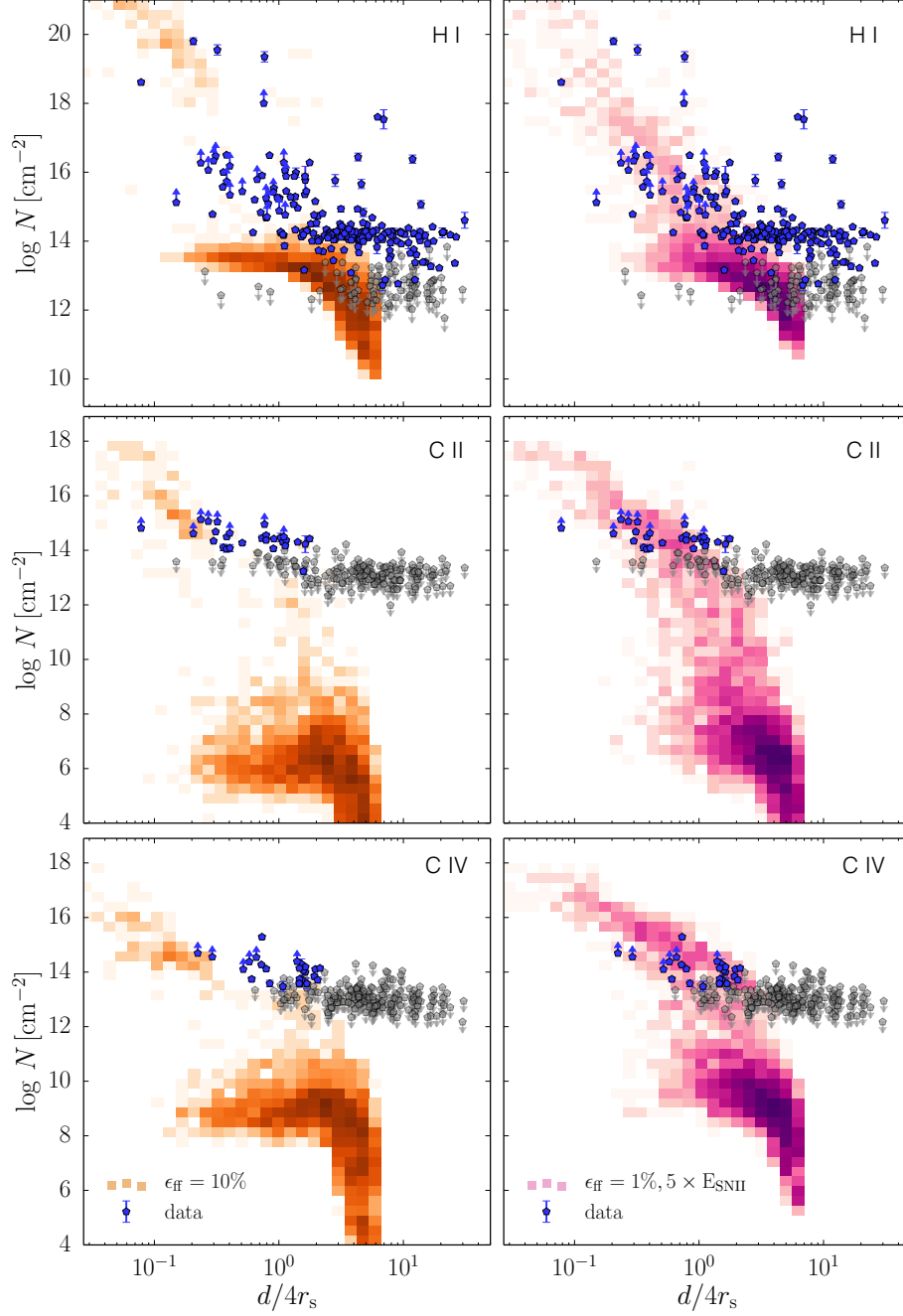


Figure 3.13: Comparison between predicted radial column density distributions of various ion absorbers (2D histograms) with observations at $z \approx 0 - 0.3$ (pentagon points; see Table 3.3). Blue points show detections, while gray points show non-detections with downward arrows denoting 2σ upper limits. The ion for which the profile is constructed is indicated in the upper right. The left column shows profiles predicted by the fiducial model with all feedback + E_{fb} , $\epsilon_{\text{ff}} = 10\%$. The right column shows the predicted profiles from the model with all feedback + E_{fb} , $\epsilon_{\text{ff}} = 1\%, 5 \times E_{\text{SNI}}$. The 2D histograms show that the majority of the points in ALL_Efb_e010 are many orders of magnitude below observations, while profiles in the ALL_Efb_e001_5ESN run are closer to observations.

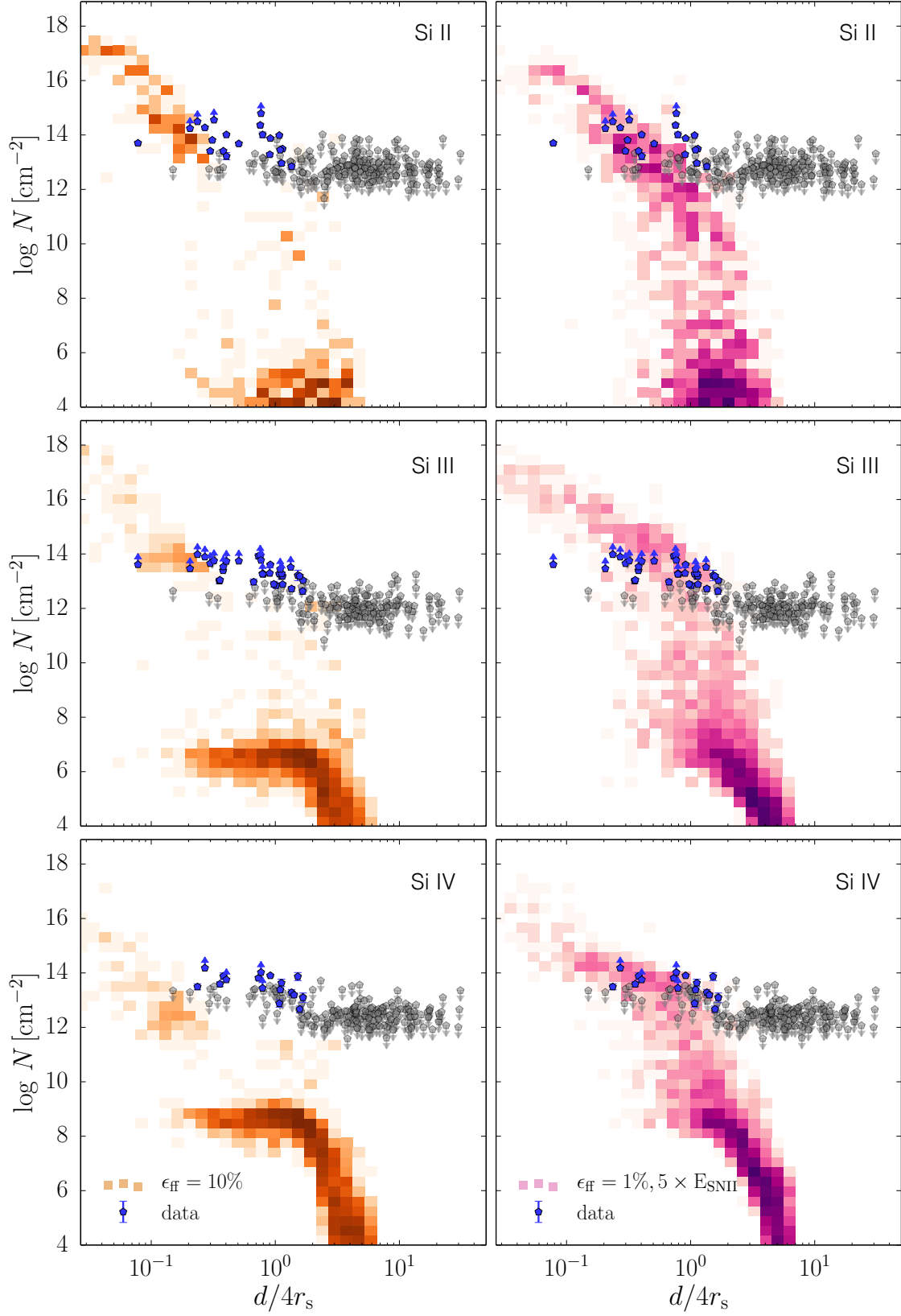


Figure 3.14: Similar to Figure 3.13, but for Si II, Si III and Si IV.

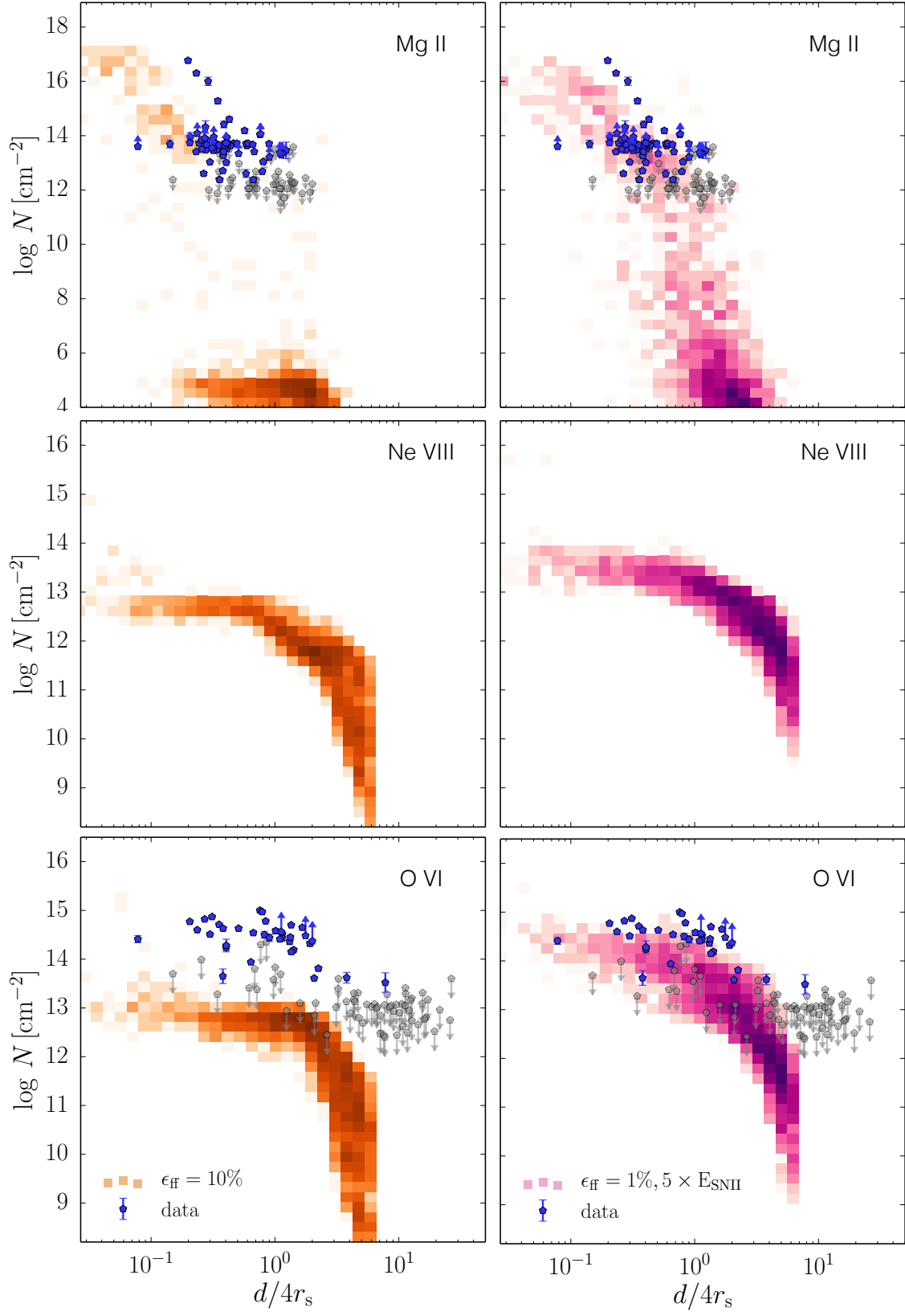


Figure 3.15: Similar to Figure 3.13, but for Mg II, O VI and Ne VIII.

Kravtsov (2015, 2016) show that this run predicts properties of the stellar and cold HI gas component of the galaxy in very good agreement with properties of a typical late type galaxy of this mass. Thus, there is a striking difference in how successful this run is in producing realistic stellar components of galaxies and the discrepant gaseous CGM it predicts.

In contrast, the `ALL_Efb_e001.5ESN` run matches the CGM column density distribution of all ion species in the $\log N - d/4r_s$ plane reasonably well, especially given that we are comparing a single object to galaxies of a range of stellar masses. However, this run produced stellar component of the galaxy with a spheroidal morphology, exponential stellar surface density profile and unrealistically large half-mass radius. *This shows that the CGM provides a stringent and orthogonal constraint on the galaxy formation simulations relative to the optical observations of galaxies.*

It is interesting that the CGM in the `ALL_Efb_e001.5ESN` run reproduces the sharp drop in the column density of absorbers apparent for all transitions, which coincides with the location where a large fraction of measurements are upper limits. Liang & Chen (2014) have estimated this drop to be at $d/R_{\text{vir}} \approx 0.7$ for Si II, Si III Si IV, C II and C IV. A similar drop of the C IV column densities was reported by Bordoloi et al. (2014) at $d/R_{200\text{m}} \approx 0.5$, and for O VI by Johnson et al. (2015b) at $d/R_{\text{vir}} \approx 1$.

The difference between $d/R_{200\text{m}} = 0.5$ and $d/R_{\text{vir}} = 0.7$ found for low and intermediate ions is simply due to the difference between R_{vir} ($\Delta \approx 360$) and $R_{200\text{m}}$ definition and the specific adopted $M_* - M_h$ relation. Once consistent definitions are adopted, the two samples of data are in agreement. Indeed, if we re-compute R_{vir} using the Bryan & Norman approximation and Kravtsov et al. (2014) $M_* - M_h$ for all galaxies compiled in this study, the low-ionization metal boundary is consistent with $d/R_{\text{vir}} \approx 0.7$.

Liang & Chen (2014) found a hint that the C IV extends beyond $d/R_{\text{vir}} = 0.7$, but based only on absorbers around two galaxies. Here, combining the data sets of Liang & Chen (2014) and Bordoloi et al. (2014), we see a stronger indication that C IV indeed extends to $d \approx 6 - 7r_s \approx 0.9R_{\text{vir}}$ and its distribution is thus more extended compared to the low-ions

(e.g., Si II and Mg II). Taking into account the result of Johnson et al. (2015b) that the O VI distribution extends to $d \approx R_{\text{vir}}$, observations indicate that the extent of the absorbing gas increases with increasing ionization energy of the ions. This is qualitatively consistent with the tight trend between the scale height of the exponential column density profile and ionization energy of the corresponding ion that we found in simulations (see Figure 3.9 and Table 3.2). It remains to be seen whether observations will confirm the tight correlation of the profile scale height with ionization energy predicted by our simulations.

3.5.3 Comparing the CGM Covering Fractions

The scatter of the column density in the region of sharp drop of ion column densities shown in Figures 3.13, 3.14 and 3.15 is large, and at some radii may span orders of magnitude. The large scatter means that it may be misleading to compare median or average profiles and one has to compare distributions of column densities directly, as we did above. Another way to characterize the column density distribution is via the covering fraction κ , which measures a fraction of area covered by absorbers with column density or equivalent width larger than a given threshold.

Different detection thresholds reported for different observations present a challenge for a uniform comparison. Some studies adopt rest frame equivalent width of $W_0 = 0.05 - 0.1$ Å and $\log N_0 = 13.5 \text{ cm}^{-2}$ as detection thresholds (Liang & Chen, 2014; Johnson et al., 2015b). We adopt these values here for comparison with the derived covering fraction of data from Liang & Chen (2014) and Johnson et al. (2015b) for O VI. For observations where only equivalent width measurements are available, Liang & Chen (2014) adopted the threshold $W_r = 0.05$ Å. We convert this value from W_0 to $\log N_0 = 13.75$ for the calculation of the covering fraction, using their estimated b parameter of 29 km s^{-1} .

In principle, one can use a different way of treating the upper limits in which the upper limit is considered either as detection with the corresponding column density or equivalent width, or as effectively zero column density. The former choice gives the largest possible

covering fraction, while the latter gives the smallest possible covering fraction. These two limits are shown as blue bands in the panels of Figure 3.16. Overall, the two methods give estimates consistent with each other, but the second approach in some cases results in a wider range of allowed covering fractions. We also perform a bootstrap re-sampling of the simulations to estimate the uncertainty of the covering fraction profile in simulated galaxies.

Figure 3.16 shows the covering fraction profile predictions for Mg II, C IV and O VI for runs ALL_Efb_e010, ALL_e010_CR and ALL_Efb_e001_5ESN. As discussed above, the column densities of most lines of sights in the ALL_Efb_e010 run is orders of magnitude too low compared to observations. These low column densities are below the detection limit of the observations, resulting in low covering fraction. This is apparent for Mg II, C IV and even O VI. The only region with non-zero κ is with distance $d < r_s$ (or $\approx 0.1R_{\text{vir}}$ for the simulated halos), at the extended disk where density is high and temperature is low.

On the other hand, the ALL_Efb_e001_5ESN run produces cold/warm outflows that reach to larger distances, producing a higher covering fraction. The covering fraction profile predicted in this run is much closer to the profile derived from observations. Although the match is not perfect, we note that we are comparing a single object in simulations to a sample of galaxies spanning a wide range of stellar masses. Thus, meaningful conclusions can be drawn only from larger samples of simulated galaxies. Nevertheless, the fact that simulations reproduce high covering fraction extent to large radii in rough agreement with observations is encouraging.

The run ALL_e010_CR, which includes cosmic ray feedback, predicts profiles at $z = 1$ (the lowest z to which the simulation was run) that are intermediate between those of ALL_Efb_e010 and ALL_Efb_e001_5ESN. For O VI it is considerably closer to ALL_Efb_e001_5ESN. This indicates that cosmic ray-aided winds can help to bring predictions in better agreement with observations.

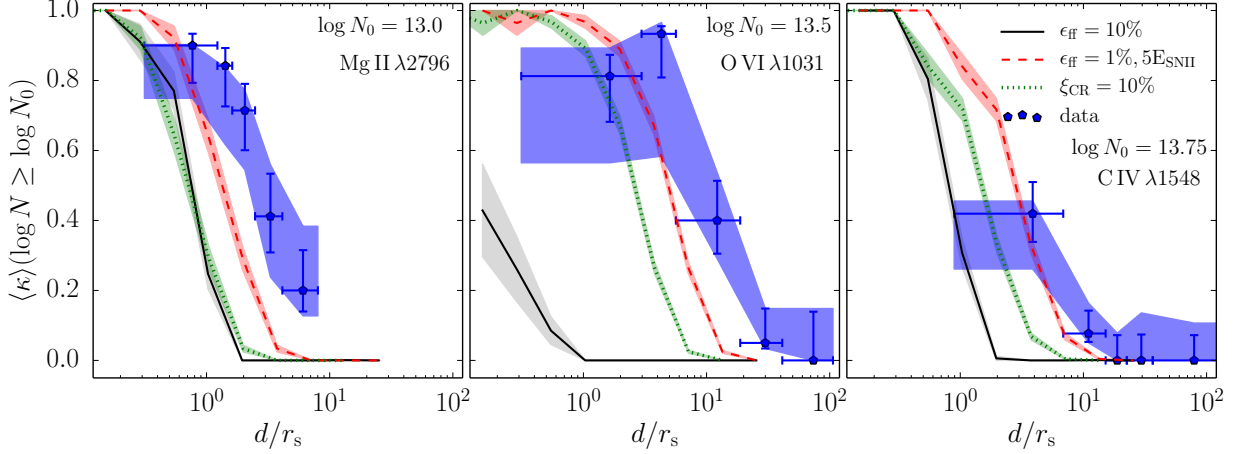


Figure 3.16: Comparison of the predicted covering fraction profiles with measurements of Mg II $\lambda 2796$ (Chen et al., 2010a), C IV $\lambda 1548$ (Liang & Chen, 2014; Bordoloi et al., 2014) and O VI $\lambda 1031$ (Tumlinson et al., 2011; Johnson et al., 2015b) transitions. Detection threshold $W_0(\text{C IV } \lambda 1548) = 0.05 \text{ \AA}$ is adopted from Liang & Chen (2014). We convert the threshold to column density using a curve-of-growth analysis adopting the doppler parameter $b = \sqrt{2}\sigma = 29 \text{ km s}^{-1}$. Detection threshold $\log N_0(\text{O VI } \lambda 1031) = 13.5 \text{ cm}^{-2}$ is adopted from Johnson et al. (2015b). $\log N_0(\text{Mg II } \lambda 2796) = 13 \text{ cm}^{-2}$ is chosen based on the level upper limits in Chen et al. (2010a). Error band represents 68% confidence level bracketed by two limiting cases of consideration of upper limits above the detection threshold (e.g., as detection and as non-detection). In all three ions, ALL_Efb_e001_5ESN and ALL_e010_CR predictions agree with the inner region ($< 2 - 4r_s$) of the data while under-predicting the covering fraction of the outskirts of the halo ($\sim 10r_s$). Note that ALL_e010_CR is at $z = 1$ while the other runs are at $z = 0$. It is conceivable that halo in the CR run will continue to grow at lower redshift, as suggested by Fig. 3.5

3.6 Discussion

In the preceding section we have presented comparisons of properties of gaseous halos forming around a progenitor of a $\approx L_*$ galaxy in simulations with different parameters and implementations of the star formation – feedback loop. We find that properties of the gaseous CGM, in particular, spatial distribution of various ions, are highly sensitive to the details of these implementations and parameter choices. For example, the fiducial simulation of Agertz & Kravtsov (2015, 2016), which produces a very realistic central galaxy with the correct stellar mass, size, angular momentum, rotation curve, bulge-to-disk ratio, stellar and gas surface density profiles does not produce extended CGM and is in striking discrepancy with observations. This illustrates that properties of galaxies and properties of their CGM provide strong *complementary* constraints on the processes governing galaxy formation.

Variations of the stellar feedback model, such as adding feedback modelling due to cosmic rays or simply changing parameters governing star formation and stellar feedback, affect properties of gaseous halos around simulated galaxies and produce a more extended CGM in better agreement with observations (see, e.g., Figs. 3.13-3.15 above). This is particularly true for ions with higher ionization energies, such as C VI and O VI, indicating that these simulations correctly capture the thermal and density structure and metallicity profile of hotter gas.

There are indications that the distribution of low-ions in all of our simulations is somewhat less extended than indicated by observations. In the fiducial run, this is due to the lack of low temperature gas at large radii at low z (see also Hummels et al., 2013). This is mainly because in the fiducial run there are no global winds at $z < 0.5 - 1$ and thus there is no cold/warm gas lifted from the disk of the galaxy into the halo. In contrast, in the $5E_{\text{SN}}$ run, the winds continue to $z = 0$ and therefore the CGM is filled with cold/warm gas at low redshifts.

The low metallicity of the halo in that run may also play a role in the low column density of absorbers. For example, the total column density of all carbon is still lower than the

observed column density of CIV alone by a factor of ~ 10 . We think this is because the winds at high z in the fiducial run are so strong that they disperse metals over a large volume, lowering the overall metallicity.

Some recent observations indicate the presence of low-ion absorbing clouds with sizes below the resolution of the simulations (Muzahid, 2014; Crighton et al., 2015), which is particularly poor in the gaseous halo far away from the dense regions of the disk. It is possible that a sub-grid model of small clouds created by nonlinear thermal instabilities (Joung et al., 2012) can ease the disagreement. Gas inflow along filaments and feedback-driven outflows could plausibly excite such instabilities, which then could create clouds that would boost the amount of cold/warm gas present. Some low-redshift observations, on the other hand, indicate large physical sizes (Davis et al., 2015) and non-hydrostatic states of many absorbers (Werk et al., 2014) consistent with the properties of recent outflows of the kind we observe in our ALL_Efb_e001.5ESN and ALL_e010_CR runs. Low-ion absorbers in galaxies thus may originate both from large-scale outflows and from small clouds forming by thermal instability that outflows help to excite.

Overall, the comparison of simulation results with different star formation and feedback models has produced valuable insights about the radial distribution of absorbers, which we discuss below. Below we will also discuss how the results of our simulations compare to other recent theoretical studies on this subject.

3.6.1 Column density profiles of the CGM absorbers: is there an ion boundary of the CGM?

As we discussed in section 3.5.1, observations of absorption lines of a variety of ions, such as Si II, Si III, Si IV, C II, C IV, and O VI, exhibit a sharp drop in the incidence of absorber detections and their covering fraction beyond a certain radius (Liang & Chen, 2014; Bordoloi et al., 2014; Johnson et al., 2015b). Some researchers have interpreted this drop-off as an “ion boundary.”

Our simulations of a $\sim L_*$ galaxy progenitor at different redshifts predict approximately exponential column density profiles for all ions. When plotted on a log-log scale, the exponential profiles exhibit a sharp turnover at approximately the “half-mass radius” or $\approx 1.68h_s$, where h_s is the scale height of the exponential profile. Thus, our simulations indicate that the perceived sharp turnover in the incidence of high column density absorbers is simply a manifestation of a continuous underlying exponential profile.

It is worth noting that de-projection of the projected exponential column density profile via the Abel integral gives the modified Bessel function of the second kind, $K_0(x)$ for the radial profile of 3D number density (see, e.g., Appendix in Pitts & Tayler, 1997). The $K_0(r)$ profile has a shallow slope at small r , which steepens at larger r and we find that the $K_0(x)$ function does approximate the gas density profile in our simulations at $r \sim 20 - 200$ kpc quite well. The origin of the approximately exponential total column density profile, $N_{\text{H}}(d)$, is thus in the shape of the 3D gas density profile.

The column density profiles of specific ions are additionally shaped by the metallicity profile and by the ionization factor and thermal structure of the gas that affect their photo- and collisional ionization rates. These factors are ion-specific and give rise to profiles that vary systematically with the ionization energy of ions. The metallicity profile in the central regions varies in different simulations. In the ALL_Efb_e001_5ESN run the metallicity rapidly increases with decreasing radius at $r \lesssim 50$ kpc, while in the fiducial ALL_Efb_e010 run the metallicity remains low and its profile remains flat down to small radii. The difference is that the former simulation drives outflows at low redshifts that enrich the central region (~ 50 kpc) and provides a steady supply of fresh warm gas that gives rise to low-ion absorbers.

We find that column density profiles of all ions trace the overall shape of $N_{\text{H}}(d)$ quite closely at $d \gtrsim 50$ kpc, as shown in Figure 3.17, because the metallicity profile at these radii in our simulations is quite flat. The low-ions at the large radii have column densities much below the sensitivity limits of observations. However, higher ionization ions, such as O VI, do produce detectable absorption. Thus, the exponential column density profile of ions like

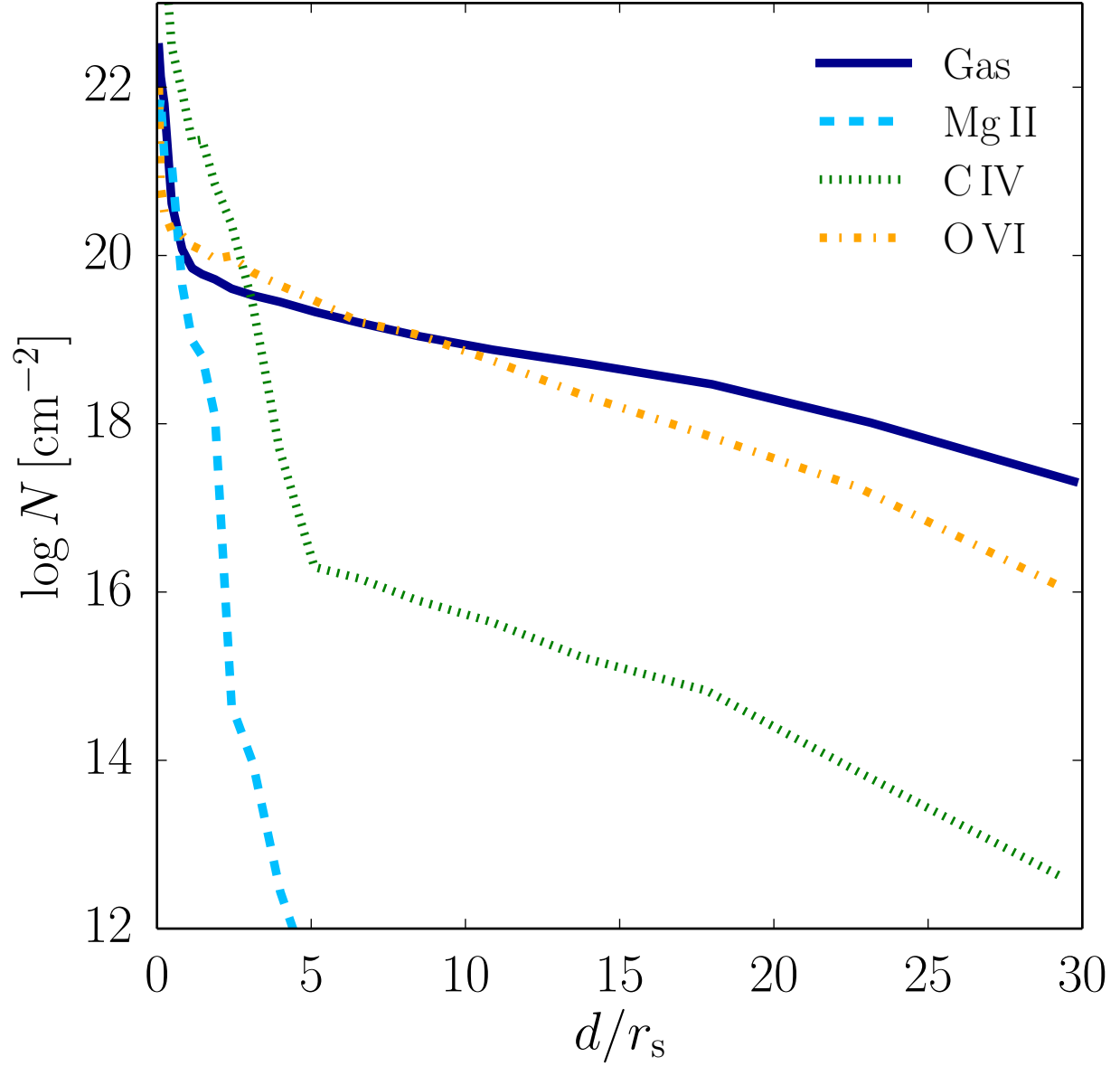


Figure 3.17: Column density profiles of Mg II, CIV and OVI in the *outer region* of the hot gaseous halo share the same underlying exponential profile traced by the total gas column density profile. All other ions (not shown) also show similar exponential profiles. Column densities for ions have been offset by $\log N_0 = 6$.

O VI simply reflects the underlying total gas profile. At $r \lesssim 2r_s \approx 50$ kpc the profiles are more difficult to interpret, but it appears that the cycling outflows at these radii do result in a metallicity profile that is also approximately exponential in projected radius. Other factors, such as dependence of the U -factor on radius, probably also contribute to the shape of the column density profiles of specific ions. In particular, the U -factor almost certainly plays a role in setting the tight correlation of scale height with ionization energy we find for the low ions.

Encouragingly, there have been a few recent studies that show exponential profiles are a good fit to the observed data. Study by Nielsen et al. (2013) finds that an exponential profile provides a better fit to the Mg II equivalent width profile, as compared to a power law. Borthakur et al. (2015) finds an approximately exponential form for the Ly α equivalent width profile. Bordoloi et al. (2014) have also used an exponential function to describe the sharp drop in the equivalent width of CIV absorber column density in their sample, although overall they used a more complicated profile given by a product of a power law and an exponential function.

3.6.2 Column density profiles of the CGM absorbers: scaling with halo mass and redshift

Interestingly, as shown in section 3.5 above, the collection of observations considered in our study indicates that the scale height of the column density profiles for observed absorbers around galaxies of a wide range of stellar masses occurs at approximately the same multiple of the halo scale radius, r_s at different redshifts. The transition in galaxies at $z \approx 2$ in the sample of Steidel et al. (2010) also occurs at the same multiple of r_s . Regardless of the precise form of the profile, this scaling is consistent with the scaling predicted in our simulations for the inner regions of the CGM profiles.

The scaling of the overall column density profile with r_s is due to the slow evolution of the inner gaseous profiles. Simulations of halo evolution in CDM models show that the inner

profiles of galaxy-sized halos are in equilibrium and evolve little after $z \approx 1 - 2$ (Prada et al., 2006; Diemand et al., 2007; Cuesta et al., 2008; Diemer et al., 2013; Zemp, 2014; Diemer & Kravtsov, 2014; Correa et al., 2015), while the profiles in the outskirts ($r \gtrsim R_{200c}$) continue to accrete mass and evolve (Diemer & Kravtsov, 2014; More et al., 2015, see also discussion in Section 3.4.4 above). The virial radius of the halo tracks the evolution of the outer profile and the outer splashback boundary of the halo fairly closely and thus evolves considerably faster than the profile in the inner regions. The evolution of the latter can be better characterized by the halo scale radius (More et al., 2015), which also exhibits slow evolution at low redshifts (Bullock et al., 2001). If the overall gaseous profile is in approximate equilibrium, this can explain the apparent scaling of the column density profile evolution with r_s .

Note that *at a fixed redshift*, the scalings with R_{vir} and r_s are nearly equivalent because $r_s \equiv R_{\text{vir}}/c_{\text{vir}}$ and concentration is a very weak function of halo mass: $c_{\text{vir}} \propto M_{\text{vir}}^{-s} \propto R_{\text{vir}}^{-0.3s}$, where $s \approx 0.08 - 0.1$ at $z = 0$ and is shallower at higher redshifts (e.g., Bullock et al., 2001; Neto et al., 2007; Diemer & Kravtsov, 2015). Thus scaling with r_s at a fixed z is equivalent to scaling with $R_{\text{vir}}^{\approx 1-1.25}$. At the same time, concentration exhibits a rather strong evolution with redshift due to fast evolution of R_{vir} and thus redshift evolution of r_s and R_{vir} and corresponding scaling as a function of redshift will be very different. Therefore, our results indicate that for samples spanning a wide range of redshift, scale radius r_s should be used to rescale radial scales, such as the impact parameter.

The remaining question is why the column density profiles shaped by stellar feedback driven outflows in simulated galaxies scale with r_s . We believe this can be understood from the following considerations. The specific energy of the wind when it is ejected is $E = v_w^2/2 + \phi(r_w)$, where v_w is initial wind velocity at the launch radius, r_w and ϕ is gravitational potential at r_w . If the outflowing gas is gravitationally bound ($E < 0$), the wind will stop at some radius $r_{\text{out}} < R_{\text{vir}}$. If we assume that energy losses due to ram pressure force from the tenuous halo gas on the wind can be neglected, then the wind stops at $E \approx \phi(r_{\text{out}})$. If r_{out} is sufficiently large we can approximate the potential by the form

expected from the NFW profile:

$$\phi(r_{\text{out}}) \approx -4.63 V_{\text{max}}^2 \frac{\ln(1 + x_{\text{out}})}{x_{\text{out}}}, \quad (3.10)$$

where V_{max} is the maximum circular velocity of the halo, $x_{\text{out}} \equiv r_{\text{out}}/r_s$ and r_s is the scale radius of the NFW profile. Although the potential in the region where stars launch winds is likely not described by the NFW form due to contributions of the baryons in the galaxy, its overall amplitude should still scale as $\phi(r_w) \propto -V_{\text{max}}^2$. At the same time, for both momentum- and energy-driven winds we expect $v_w \propto V_{\text{max}}$ (Murray et al., 2005; Booth et al., 2013; Muratov et al., 2015), the scaling that is also indicated by observations (Schwartz & Martin, 2004; Rupke et al., 2005; Schwartz et al., 2006, although see Heckman et al. 2015, submitted). Therefore, the above equations suggest that

$$\frac{\ln(1 + x_{\text{out}})}{x_{\text{out}}} \approx c_1 + c_2 \Psi(r_{\text{out}}), \quad (3.11)$$

where c_1 , c_2 are constants and $\Psi(r_{\text{out}})$ does not depend explicitly on V_{max} . This equation means that x_{out} is independent of V_{max} and hence halo mass and thus $r_{\text{out}} \propto r_s$. Although individual fountains may have a range of initial velocities and launch radii, from the above considerations we can see that the characteristic extent of the profile they shape may naturally scale with r_s .

3.6.3 The Effects of Self-Shielding, Non-Equilibrium and Local Starburst Radiation

In this work, we use three simplifying assumptions in the calculation of ionization abundance: we assume that gas is optically thin, that it is in ionization equilibrium and ionized by the cosmic UVB only (i.e., not accounting for the local radiation from the stars in the simulated galaxy). We qualitatively discuss their effects and implications on the CGM here.

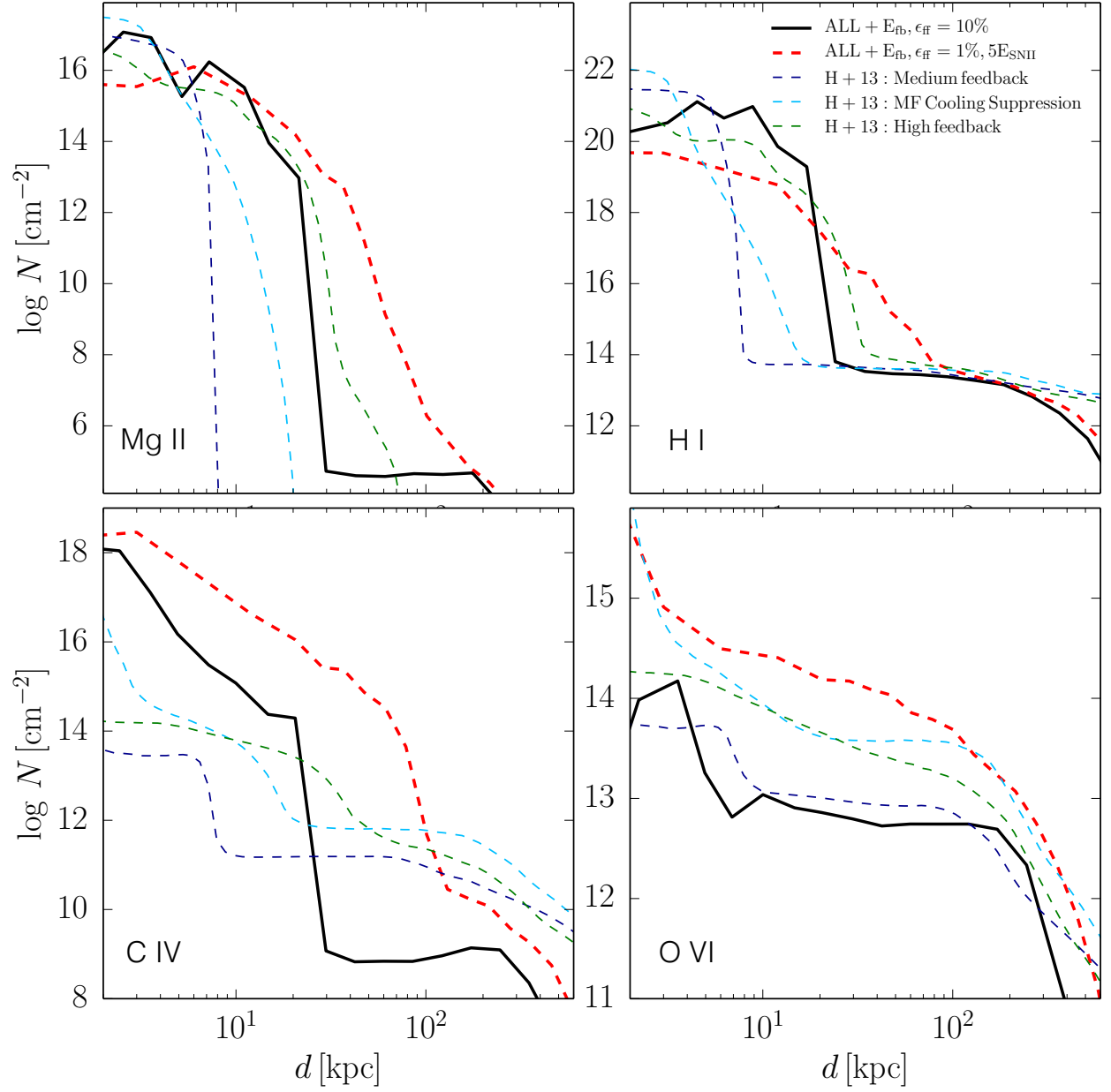


Figure 3.18: Comparisons of predicted column density profiles from low to high ions between our simulations with the profiles from simulations of Hummels et al. (2013, denoted as H+13). The profiles in our fiducial run are closest to the medium feedback run of H+13.

First, self-shielding may increase the column density of low ions compared to our estimate. We find, however, that self-shielding can only change the column density significantly near the star-forming disk where density is high. For example, Rahmati et al. (2013) shows that self-shielding affects gas at column densities greater than $\log N \approx 17 \text{ cm}^{-2}$ and our Figure 3.7 shows that this corresponds to $r \lesssim 25 \text{ kpc}$.

Second, the ionizing radiation from local star-forming regions will likely increase the ratio of low to high ions due to a softer stellar spectrum compared to QSOs. Assuming escape fraction of $f_{\text{esc}} = 1 - 2\%$, the contribution of local radiation is negligible at radii beyond $\approx 30\text{-}50 \text{ kpc}$ (Shen et al., 2012). Note that adding local radiation would affect our result in the opposite direction from self-shielding because it would increase ionizing radiation in the regions of high gas column density.

Finally, deviations from photo-ionization equilibrium may boost ionization at lower temperatures. However, this effect becomes less important at lower metallicity, especially in the presence of photo-ionization by a radiation background (Oppenheimer & Schaye, 2013), the regime of our analysis.

Given the current small observational statistics, and lack of strong arguments that these effects are significant at large radii, neglecting them should not affect our conclusions significantly.

3.6.4 *Comparisons with Previous Studies*

A number of recent studies have explored predictions for the observable properties of the CGM in cosmological simulations of galaxy formation. Some studies have focused exclusively on predictions at high redshifts ($z \approx 2 - 3$ Barnes et al., 2011; van de Voort et al., 2012; Shen et al., 2012, 2013; Pallottini et al., 2014; Suresh et al., 2015; Faucher-Giguère et al., 2015). Much of the high- z work has focused on exploring predictions for the distribution and observable properties of neutral hydrogen and dependence of predictions on the implementation of feedback and associated winds (e.g., Barnes et al., 2011; van de Voort et al.,

2012; van de Voort & Schaye, 2012; Pallottini et al., 2014; Faucher-Giguère et al., 2015).

Studies that systematically analyzed properties of the CGM at low redshifts used two kinds of simulations. Ford et al. (2013, 2014, 2015) used a statistical sample of galaxies formed in a simulation of $32h^{-1}$ Mpc and $48h^{-1}$ Mpc boxes, but which modelled the wind launching process phenomenologically by imposing a particular specified wind scaling with halo properties (e.g., wind velocity and loading factor) as a function of halo mass. In these studies the authors have placed particular emphasis on the origin and dynamical cycles of the CGM. Ford et al. (2013, 2014) showed that low ions, such as Mg II, trace dense gas close to galaxies that were part of recent outflows and which will re-accrete onto a disk on a $\sim$ Gyr time scale, while high ions, such as O VI, have more extended distributions and originate from ancient outflows (see also Shen et al., 2012). Our results are in qualitative agreement with the results of these studies. Moreover, the column density profiles presented in Ford et al. (2014) and Ford et al. (2015) have a form in qualitative agreement with results of our simulations. In particular, their profiles are approximately exponential in the inner regions, with some ions exhibiting two exponential component profiles (see, e.g., Fig. 10 in Ford et al., 2013, and Fig. 7 in Ford et al. 2014).

In the second type of simulations, which includes simulations presented in this work, formation of individual galaxies was modelled using different numerical codes (**Enzo**, **RAMSES**, **Gasoline**) in the “zoom-in” simulation where all mass and spatial resolution is focused on a single object (Stinson et al., 2012; Shen et al., 2013; Hummels et al., 2013; Roškar et al., 2014; Marasco et al., 2015). The higher resolution in the disk region in such simulations allows a more sophisticated treatment of stellar feedback, which allows treating the wind launching process self-consistently without imposing particular wind scaling properties. Not surprisingly, the CGM properties in such simulations were shown to be quite sensitive to the details of star formation and stellar feedback implementation (Hummels et al., 2013; Roškar et al., 2014; Suresh et al., 2015; Marasco et al., 2015). Again, this is in agreement with our results, which indicate strong dependence of the properties of the CGM on the

feedback parameters and implementation. We should note that the high sensitivity of the CGM properties to wind modelling is specific to the zoom-in simulations with self-consistent wind launching. Ford et al. (2014, 2015) did not find significant differences in the CGM produced by the different phenomenological wind models used in their simulations.

In Figure 3.18, we compare the CGM profiles from two of our simulations with SN feedback with the profiles of Hummels et al. (2013).³ The figure shows that the profiles from our `ALL_Efb_e010` are in good agreement with the “Medium Feedback” model of Hummels et al. (2013), which produce a CGM inconsistent with observations. Hummels et al. (2013) found that artificially delaying cooling after gas is heated by SNe makes wind launching more efficient and produces a more extended CGM halo, although the extent of the ion distribution is still somewhat smaller than indicated by observations. In all transitions, our model `ALL_Efb_e001_5ESN` can be viewed as a similar variation of the feedback parameters, although instead of suppressing cooling, it assumes larger energy release per supernova. Physically, this can correspond to a larger fraction of thermal energy injected by a supernova retained by the gas or to top-heavy IMF. Our profiles for this simulation are indeed closest to the “high feedback” run of Hummels et al. (2013), although our profiles are somewhat more extended and are thus in better agreement with observations.

One may question the physical validity of the cooling delay or increasing energy per supernova. However, we think that these choices may approximate some other physical processes that result in more efficient wind driving. For example, when we include cosmic rays with isotropic treatment of their diffusion, simulations produce CGM profiles quite close to those of the `ALL_Efb_e001_5ESN` run, albeit without assuming larger energy per supernova. Overall, these results (and previous studies that produce realistic CGM) show that feedback processes that contribute to wind launching should be quite efficient.

As we discussed in the previous section, the column density profiles of the CGM ions are predicted to scale linearly with halo scale radius, r_s , although we could naively expect that

3. The profiles are taken from the website of the authors: <http://chummels.org/CGM.html>

strong feedback could break such self-similar scaling with halo parameters. Other recent studies have reported similar findings. For example, van de Voort & Schaye (2012) show that outskirts of gaseous halos of galaxies in the OWLS simulation evolve as expected from the “virial” scaling relations. Pallottini et al. (2014) find that CGM profiles around high-redshift galaxies ($z = 4$) are self-similar once scaled with R_{vir} . They have found that the HI column density profiles as a function of normalized impact parameter, d/R_{vir} , have power law form, while we find exponential form (or double exponential) for all ions and neutral hydrogen absorbers. Our results indicate that column density profiles of the CGM ions do exhibit self-similarity, but they scale better with halo scale radius, r_s , rather than with the virial radius.

3.7 Conclusions

We presented a detailed analysis of the observable properties of the circumgalactic medium in a suite of high-resolution cosmological galaxy re-simulations of a Milky-Way sized halo with a variety of star-formation and feedback models. We have also contrasted these predictions with a large set of existing observations of absorbers around galaxies of different mass from $z \approx 0$ to $z \approx 2$. Our specific findings can be summarized as follows.

1. Our simulations indicate that the CGM probed by absorbers of different ions arise in gas that was ejected from the disk by stellar feedback. At low redshifts, the outflows in the MW-sized halo modelled in our simulations are not sufficiently energetic to leave the halo and instead produce plumes that reach a finite radius and then turn around to infall back onto the disk. We show that low ionization absorbers originate in such fountain outflows (Fig. 3.3 and Fig. 3.4).
2. We find that the column density profiles of various ions and of neutral hydrogen in our simulations are an exponential function of the impact parameter (see Section 3.4.5). Ions with higher ionization energy have more extended profiles with the scale height

of the exponential distribution tightly correlated with the ion ionization energy (see Fig. 3.9, eqs. 3.8-3.9): $h_s \propto E_{\text{ion}}^{0.74}$. At $z \approx 0$, the scale height of warm gas traced by low-ionization species, such as Mg II and C IV, range in $\approx 0.03 - 0.07 R_{\text{vir}}$, while higher ionization species, such as O VI and Ne VIII, have scale heights of $\approx 0.32 - 0.45 R_{\text{vir}}$ (Table 3.2 and Fig. 3.9).

3. The predicted correlation of scale height with ionization energy of ions is in qualitative and quantitative agreement with observations. In particular, one of our simulations reproduces the radii where observations show a sharp turnover in the column density distribution for different ions (Fig. 3.13, 3.14 and 3.15). We emphasize, however, that this turnover is not due to a sharp boundary, but is a manifestation of the continuous exponential column density profile. (Fig. 3.11).
4. We find that the total gas column density profile can be approximated by a double exponential profile. The outer exponential distribution is well traced by the distribution of all ions, although column densities of low ions at these radii are predicted to be much below current sensitivity limits. However, the model predicts that the observed column densities of higher ions track the shape of the total gas column density profile at large r quite closely.
5. We find that the physical extent of the CGM in our simulations evolves slower than the virial radius at $z \leq 2$. We show that column density profiles, in the simulations that match the observed CGM, evolve similarly to the halo scale radius, r_s . Thus, column density profiles of galaxies at different redshifts can be rescaled using r_s corresponding to their halo mass. This reveals a *self-similarity* in the simulated and observed CGM across four decades of stellar mass *and* 11 billion years in cosmic time (Fig 3.7 and Fig. 3.11).
6. We show that the addition of supernova-produced cosmic ray fluid and associated pressure on the gas produces a CGM with profiles close to low- z observations. The

CGM in this case is much more uniform and less patchy compared to the simulations without cosmic rays. These differences can potentially be tested by future observational measurements of the covering fraction profiles. Overall, we find that CR-driven winds in the galactic halo contain cooler gas ($T < 10^5$ K) compared to winds driven by supernova feedback without cosmic rays in agreement with previous studies (Fig. 3.5, 3.6 and 3.16).

All but one simulation presented here reproduce the star formation history expected for a typical $\approx L_*$ galaxy reasonably well. In agreement with other recent studies, we show that properties of the CGM are quite sensitive to the details of the star formation–feedback loop. For example, the fiducial simulation of Agertz & Kravtsov (2015, 2016), `ALL_Efb_e010`, which produces a very realistic central galaxy with the correct stellar mass, size, angular momentum, rotation curve, bulge-to-disk ratio, stellar and gas surface density profiles fails to reproduce existing observations of the CGM. At the same time, variations of star formation efficiency, energy per supernova, or introducing cosmic ray feedback appear to bring the predicted CGM in better agreement with observations. This illustrates that the properties of galaxies and the properties of their CGM provide strong *complementary* constraints on the processes governing galaxy formation. Our results clearly show that future tests of galaxy formation physics should make use of the growing data set of the CGM measurements. Potentially, information on the hot gas component via the highest ions or direct X-ray imaging could also provide valuable constraints on the models of star formation and feedback.

CHAPTER 4

OBSERVING THE CIRCUMGALACTIC MEDIUM OF SIMULATED GALAXIES THROUGH SYNTHETIC ABSORPTION SPECTRA

4.1 Introduction

Due to the low surface brightness of the circumgalactic medium (CGM), nearly all of the observational studies (with few exceptions, e.g., Martin et al., 2014; Cantalupo et al., 2014) rely on analyses of absorption spectra from background light sources (e.g., quasars). We derive various empirical constraints from mapping the CGM using background quasars absorption spectroscopy. For example, the abundance of cold/warm absorbers exponentially drops off with increasing distance away from the host galaxies (Nielsen et al., 2013; Liang & Chen, 2014; Bordoloi et al., 2014; Borthakur et al., 2016). From the spatial profiles and ionization modelling of absorption lines, some researchers derive the total mass and metallicity of the cold phase of the CGM (Lehner et al., 2013; Werk et al., 2014; Wotta et al., 2016; Stern et al., 2016; Prochaska et al., 2017).

At the same time, a large set of simulation studies is compared with the constraints of the CGM properties from observations to improve our theoretical understanding. These include the CGM mass budget in different phases (Suresh et al., 2015), the sensitivity of column density profiles and covering fraction of ions with models of star formation and feedback processes (Shen et al., 2012; Stinson et al., 2012; Hummels et al., 2013; Faucher-Giguère et al., 2015; Liang et al., 2016), and also the metallicity distribution in the CGM (Hafen et al., 2016).

Just as observers model absorption lines for column density and subsequent ionization analysis for the absorbing gas, simulators should create synthetic absorption spectra to make comparisons more direct (e.g., Oppenheimer & Davé, 2006; Bird et al., 2015; Churchill

et al., 2015; Hummels et al., 2016). In addition, there is likely more information than we currently extract from the observed spectra. Analyzing synthetic spectra generated from simulated gaseous halos with known properties may reveal fruitful insights into how well this inversion process works. Therefore, we develop a pipeline that realistically ‘observe’ the CGM of simulated galaxies by generating synthetic absorption spectra through random lines of sight (LOS). We study the multiphase structure of the CGM with ionization modelling by combining knowledge about underlying properties from our simulations and the absorption features from the resulting synthetic spectra.

It is worthwhile to note that the spatial resolution of halo gas in the typical cosmological state-of-the-art zoom-in simulations of the kind we use here is far from resolving the parsec scales that may be characteristic of at least some absorbers. There are theoretical arguments that CGM may be subject to thermal instabilities, which can result in multiphase medium with small-scale structure unresolved in such simulations (e.g., Fall & Rees, 1985; Begelman & Fabian, 1990; McCourt et al., 2016). Simulations of small-scale regions focusing on disruption of macroscopic cold/warm regions in a hot halo also show the emergence of very small-scale structures in gas distribution (e.g., McCourt et al., 2016; Schneider & Robertson, 2017, Liang & Remming, in prep). At the same time, some of the current galaxy formation simulations come close to reproducing column density profiles and covering fractions of some of the ionic species (Stinson et al., 2012; Ford et al., 2016; Oppenheimer et al., 2016, 2017; Liang et al., 2016). For example, one of our simulations with boosted feedback ($E_{\text{fb}} = 5E_{\text{SN}}$), the `ALL_Efb_e001_5ESN` run (Agertz & Kravtsov, 2015, 2016), reproduces the observed column density profiles of all low and intermediate ions; but it marginally underproduces O VI abundance. The relative contribution of small-scale ($\sim \text{pc}$) features to the overall column densities of various absorbers is still under debate. However, given the fact that simulations are producing column densities close to observations, it is reasonable to employ them in tests of the kind presented in this paper.

In this paper, we focus on the spatial distribution and physical properties, such as tem-

perature, density, and metallicity of the CGM gas in our simulations and how well these can be derived from modelling their signatures in the absorption spectrum. We go beyond what is usually done by developing a Bayesian approach using full probability distribution of column densities of ions. This allows us to treat detections and non-detections in a uniform manner. The paper is organized as follows. In section 2, we describe the simulations and the methodology of generating synthetic absorption spectra. We discuss the multiphase structure, the origins of absorbers such as cold/warm “clouds” and hot halo gas as seen from synthetic spectra and our simulations. In addition, we explore the effects of various ionization models on the abundances of commonly observed ions. In section 3, we describe a new methodology for deriving physical properties of absorbers, assuming ionization mechanism is known, that include proper treatment of upper and lower limits of column densities. Finally, we discuss and summarize our findings in section 4.

4.2 Simulations and Synthetic Absorption Spectrum

4.2.1 *Simulations, star formation, and feedback models*

In this study we use “zoom-in” simulation of $\sim L_\star$ galaxy presented in Agertz & Kravtsov (2015, 2016). The simulation was run with the Adaptive Mesh Refinement (AMR) code RAMSES (Teyssier, 2002) in the Λ CDM cosmology with $\Omega_\Lambda = 0.73$, $\Omega_m = 0.27$, $\Omega_b = 0.045$, $\sigma_8 = 0.8$ and $H_0 = 70 \text{ km s}^{-1} \text{Mpc}^{-1}$ (Komatsu et al., 2009). The mass of the dark matter particle in the high-resolution “zoom-in” region is $m_{\text{DM}} = 3.2 \times 10^5 M_\odot$, while the peak spatial resolution in the high-density ISM regions is $\Delta x \approx 75 \text{ pc}$. The CGM we analyze here is re-binned from the AMR grid at 683 pc resolution everywhere in the halo, typical of current generation state-of-the-art simulations since refinement criteria focus on the densest regime. At $z = 0$, the galaxy has a stellar mass of $M_\star \approx 2 \times 10^{10} M_\odot$ and is forming stars with a rate $\approx 2 M_\odot \text{yr}^{-1}$ in a dark matter halo $M_{\text{vir}} \approx 10^{12} M_\odot$ ($R_{\text{vir}} \approx 260 \text{ kpc}$).

The specific simulation we use in this study is one of the re-simulations of galaxies

started from the same initial conditions, but varying parameters and implementations of star formation and stellar feedback, details of which are presented in Agertz & Kravtsov (2015, 2016). In Liang et al. (2016), we have compared the CGM around galaxies in these different re-simulations to available observations of low-, intermediate-, and high-ionization absorbers and showed that one simulation, labeled **ALL Efb_e001_5ESN**, has incidence and covering fractions of absorbers closest to observations. We, therefore, choose this simulation (at $z = 0$) as the basis for tests presented in this study.

The star formation in this simulation was parameterized as:

$$\dot{\rho}_\star = f_{\text{H}_2} \frac{\rho_g}{t_{\text{ff}}} \epsilon_{\text{ff}}. \quad (4.1)$$

where the star formation rate $\dot{\rho}_\star$ depends on the molecular hydrogen fraction f_{H_2} (computed using model by Krumholz et al., 2009), the local gas density ρ_g and corresponding free fall time t_{ff} , and finally the star formation efficiency per free fall time ϵ_{ff} . In the simulation we use here $\epsilon_{\text{ff}} = 0.01$ was adopted.

The simulation implements feedback recipe that accounts for radiation pressure, stellar winds, and energy and momentum injection by supernovae. In the simulation we use in this study, each supernova is assumed to release 5×10^{51} ergs of energy, a value that is set higher than the traditionally adopted values to increase the effects of feedback. In addition, a dual-energy scheme is used in which a fraction of the energy released by supernovae is stored and evolved as a separate variable to model the unresolved turbulence and thermal pressure of hot gas generated by supernovae. This energy density, E_{fb} , introduces a new form of pressure, $P_{\text{fb}} = (\gamma - 1)E_{\text{fb}}$, and is evolved with its own equation similar to that of the internal energy, but with a dissipation time scale (t_{dis}) of 10 Myr:

$$\frac{\partial}{\partial t}(E_{\text{fb}}) + \nabla \cdot (E_{\text{fb}} \mathbf{v}_{\text{gas}}) = -P_{\text{fb}} \nabla \cdot \mathbf{v}_{\text{gas}} - \frac{E_{\text{fb}}}{t_{\text{dis}}} \quad (4.2)$$

The dissipation time scale is chosen to match values similar to a few sound crossing times

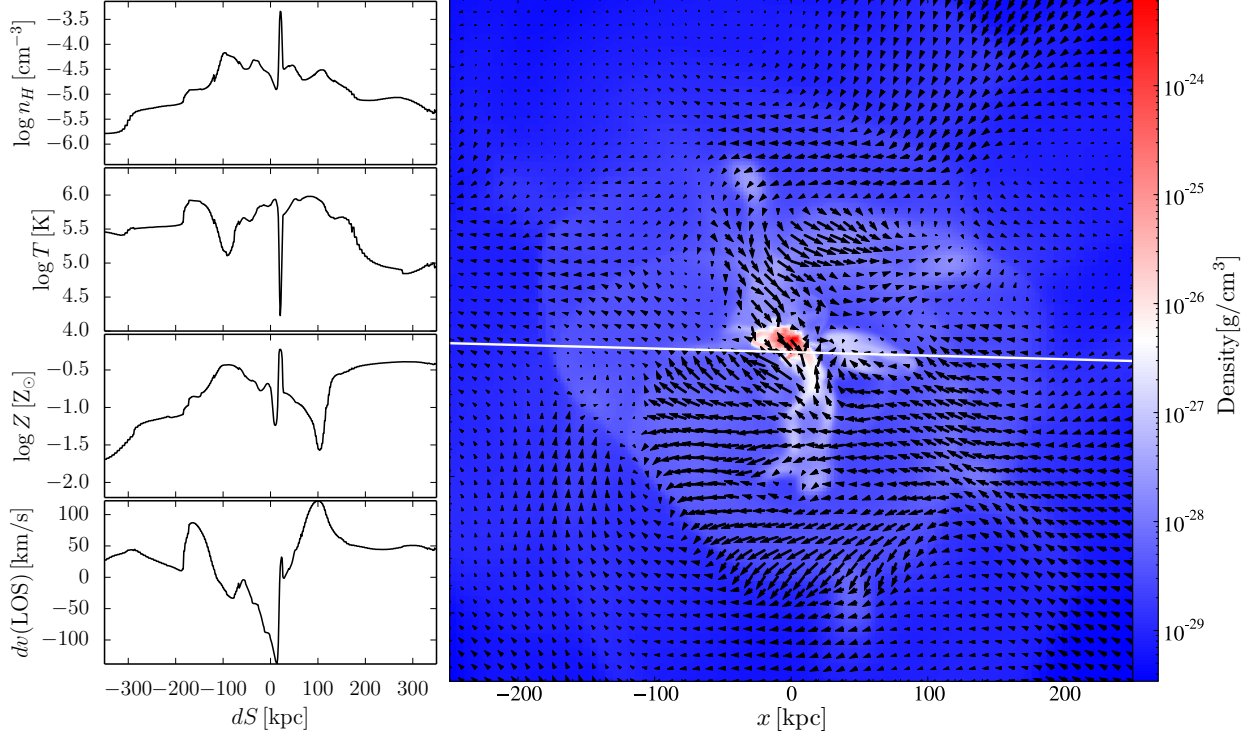


Figure 4.1: Example of a spectrum along a line of sight (LOS) sampling the properties of the CGM around the simulated galaxy. The LOS (white line) through the density (color map) and velocity (vector) field are shown in the right panel, while the left panels show the one-dimensional profiles of the density, temperature, metallicity, and velocity (relative to the galaxy systemic velocity) of the gas the LOS passes through.

of typical giant molecular clouds (GMC). Such additional energy variable regulates star formation by controlling the pressure within the ISM and thickness of star-forming disk (see Agertz et al., 2013, for details).

4.2.2 Synthetic absorption spectra of the simulated CGM

The synthetic spectra are computed with our custom code that converts physical properties of absorbing gas (such as density, temperature, and metallicity) into absorption lines, as described below.

First, we compute the number densities of ions from random LOS through the simulations. Briefly, the number density of an element X (e.g., C, N, O) at state j is computed as:

$$n_{X_j}(n_H, T, Z) = f_{X_j}(T, n_H) f_X(Z) n_H, \quad (4.3)$$

where $f_X(Z) = n_X/n_H$ is the fraction of the element X relative to hydrogen assuming solar pattern. The ionization fraction, $f_{X_j}(T, n_H) = n_{X_j}/n_X$, is computed with both photo- and collisional ionization using the **CLOUDY** code (Ferland et al., 2013). We assume that each simulation cell i represents an isothermal cloud with temperature T and constant hydrogen density n_H .

The integrated normalized spectrum flux at wavelength location λ is thus computed assuming optically thin photoionization from the Haardt & Madau (2012) ultra-violet background (UVB) as

$$F(\lambda) = \prod_i \exp[-\tau_i(\lambda)], \quad (4.4)$$

where i is the index running through computational cells along a given LOS and the optical depth τ_i of each cell i is given by:

$$\tau_i(\lambda|N_i, b_i, z_i) = N_i \frac{\pi e^2}{m_e c} f_{\text{osc}} \Phi(\lambda|b_i, z_i). \quad (4.5)$$

Here N_i is the column density of the corresponding ion $N_{X_j} = n_{X_j} dl$ computing using ion number density in each cell and its size dl , b_i is the Voigt width parameter set by the temperature and velocity distribution of gas along LOS, and z is the redshift of the cell relative to the assumed systemic redshift of the galaxy, which includes both the peculiar gas motion and the Hubble flow contribution, $v_H = H_z r$, where H_z is the Hubble constant at the systemic redshift z . Also, e , m_e , c and f_{osc} are the relevant constants for the ion: the electron charge, electron mass, speed of light and oscillator strength of the transition between energy levels of the atom, respectively. Finally, $\Phi(\lambda|b_i, z_i)$ is the Voigt profile, which we computed using the real part of the Faddeeva function using the **wofz** function of the **SciPy** package.

Figure 4.1 shows an example of a random LOS going through our simulated galaxy in the ALL_Efb_e001_5ESN run. The density map shows an aerial view of the LOS that goes through some inflowing materials close to the star-forming regions. The LOS records the one-dimensional information shown quantitatively in the left panel of Figure 4.1, covering a wide range of density, temperature, metallicities and projected velocities. The corresponding ionic abundance and absorption spectra are shown in Figure 4.2 and 4.3. These indicate the multiphase nature of the gas along this LOS. With the tools to generate synthetic spectra for any ions and at any viewing angle of the simulated CGM, one can then systematically study both the cold and hot gas in the gaseous halos, which we will discuss in the following sections.

4.2.3 Multi-phase gas viewed from simulations and synthetic spectra

Absorption spectroscopy routinely shows that the CGM is multiphase and dynamic; the evidence for multiphase gas is found in a wide range of galaxies (e.g., Hennawi et al., 2006; Steidel et al., 2010; Tumlinson et al., 2011; Werk et al., 2014; Bordoloi et al., 2014; Liang & Chen, 2014; Prochaska et al., 2014; Johnson et al., 2015a; Borthakur et al., 2016; Lau et al., 2016; Heckman et al., 2017). Multiphase gas is generically expected where cold/warm absorbing gas is moving through a hot halo. In our simulations, the cold/warm gas in the hot halo corresponds to outflows of the ISM gas driven by supernovae feedback and, partly, accretion of gas from the IGM. In this section, we show how the underlying physical properties of multiphase gas manifest in spectroscopic observations by generating synthetic spectra through LOS in the simulated CGM.

In Figure 4.2, we show the physical properties of two types of LOS in our simulations. The left panel shows an LOS that goes through a dense absorber with large gradients in density and temperature. It is this large range of density and temperature in a localized region that simultaneously give rise to an abundance of low, intermediate and high ions. Naturally, these ions are not co-spatial. The low ions (H I and Mg II) are produced in the densest core

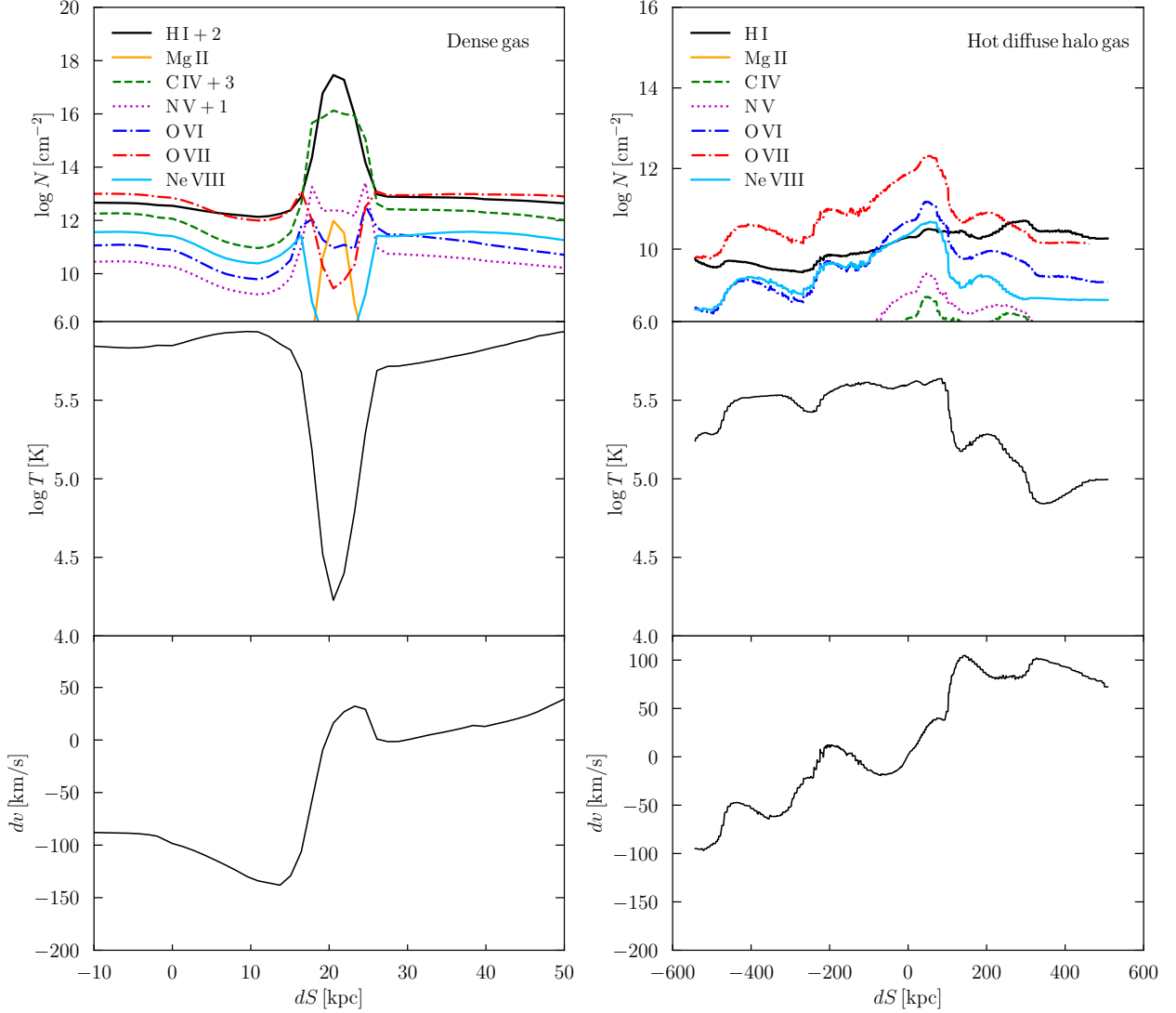


Figure 4.2: One-dimensional profiles of column density, temperature, and velocity as a function of distance from the point of closest approach (minimum distance) to the galaxy for two types of lines of sight (LOS). Note that the top plots of both panels show column density contributed by each pixel i at the location dS_i , i.e., $N_i = n_i dS_i$. The total column density is integrated along the LOS, $N = \sum N_i$. For clarity, various lines representing different ions have been artificially shifted by the value shown in the legends. The left column shows the multiphase structure of the absorbing gas in the LOS shown in Figure 4.1. This LOS passes through a density peak with a large gradient in gas density and temperature. High ions (e.g., NV, O VI, O VII and Ne VIII) preferentially reside on the boundary of the density peak while low ions reside in the core where the temperature is low ($\sim 10^4$ K). The right panel shows the LOS that passes through the hot and diffuse coronal halo gas near the virial temperature of the halo ($\sim 10^{5.5}$ K) without any dense gas. The abundance of low ions is therefore small. The temperature and velocity profiles (middle and lower columns) determine the line width of absorbers in the absorption spectra shown in Figure 4.3.

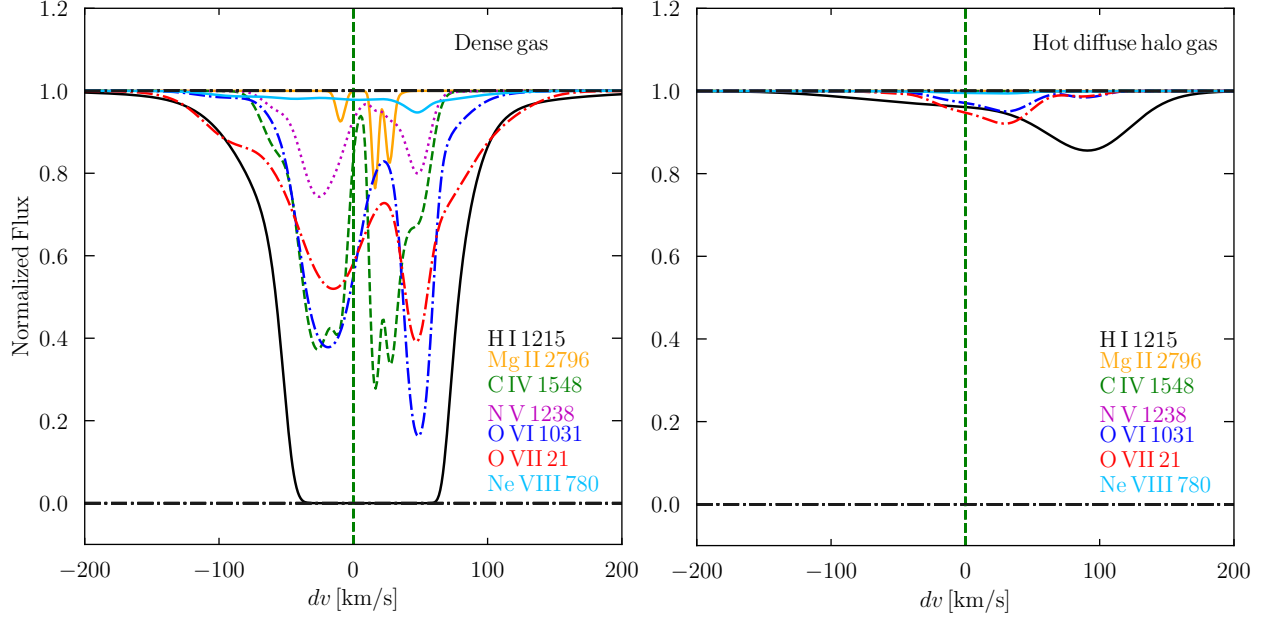


Figure 4.3: The synthetic spectra generated from the two LOS shown in Figure 4.2. Left panel: the LOS that passes through the density peak exhibits strong absorption lines of low ions, as well as high ions (e.g., OVI), reflecting the multiphase structure of the gas along this LOS. Note that a single density peak along this LOS generates three (or four) absorption components due to the strong velocity gradients. However, the low and high ions are approximately kinematically aligned per absorption component. However, the large column density of HI and oscillator strength of Ly α lead to a blending of all three absorption components, making the determination of physical parameters difficult, although such blending does not always occur. Higher order Lyman series lines are needed to separate the absorption components. Right panel: the LOS that passes through the diffuse hot halo exhibits only HI and OVI absorption since low ions cannot survive in the physical condition near the virial temperature of the halo ($\sim 10^{5.5}$ K). We also find detectable O VII lines ($\lambda_{\text{rest}} = 21.6\text{\AA}$) in both cases in our simulations with $\log N \sim 14\text{ cm}^{-2}$.

of the absorber whereas the intermediate and high ions (C IV, N V, O VI and Ne VIII) are spatially extended in the boundaries where the density is low and the temperature is high. By contrast, the right panel of Figure 4.2 shows an LOS that goes through the hot halo itself with characteristic temperature $T_{\text{vir}} \approx 10^{5.5}$ K for an L_* -like galaxy. For this type of LOS, the physical condition leads to highly ionized gas where the abundance of O VI and Ne VIII is orders of magnitude higher than low ions.

Using the formalism outlined in sec 2.2, Figure 4.3 shows the corresponding synthetic spectra for the two LOS in velocity space centered at the systemic redshift of the galaxy. The left panel shows simultaneously the absorption lines from low to high ions, indicating the existence of multiphase gas discussed above. Note that the absorption lines are split into multiple velocity components (roughly three) due to the velocity shear across the absorber. The number of absorption components depends on both the spatial resolution to resolve individual absorbers and velocity/spectral resolution for their motion. The alignment of the strongest absorption component between the low and high ions depends on how quickly the velocity varies across the size of the absorber. The alignment of O VI with low ions seen in observations (e.g., Werk et al., 2016) suggests that the velocity field does not vary significantly over a range of sub-kiloparsec based on cloud size estimates. These are inferred from the required low density assuming photoionization modelling in Stern et al. (2016).

The right panel of Figure 4.3 shows the corresponding absorption spectrum of the hot halo gas. In this particular LOS, we see detectable broad O VI and H I lines but not Ne VIII lines. In this type of LOS, the absorption profile is typically very broad due to the large variation of the velocity field across the entire halo (~ 250 kpc), although rare exceptions do exist in some particular viewing angles. In addition, the dominant ionization state of oxygen is O VI at $T_{\text{vir}} \sim 10^{5.5}$ K. Our simulations suggest the observed LOS with similar characteristics (i.e., broad O VI and no low ions) are likely probing the hot halo gas of galaxies with $T_{\text{vir}} \sim 10^{5.5}$ K (Savage et al., 2011; Werk et al., 2016).

Although the integrated column density for Ne VIII within the virial radius is $\approx 10^{13} \text{cm}^{-2}$

(see Fig. 15 in Liang et al., 2016), the optical depth is spread out in velocity space which makes it difficult to detect observationally because it can be easily hidden in the noise and in the uncertainty of continuum placement. In section 3, we present a Bayesian Voigt profile routine that includes utilities such as a continuum model to account for this uncertainty of continuum placement. We also describe a method of constraining physical properties of the hot halo with a limited number of detections and poor constraints based on the Bayesian approach.

4.2.4 *Relative contribution of photoionization and collisional ionization*

In Figure 4.2 and 4.3, the abundance of ions and their spectral features are computed using an ionization equilibrium model that assumes collisional and photoionization in the optically thin regime. The relative contribution of these ionization processes for different ions in observed absorption systems is not yet completely understood. It is thus interesting to explore the relative contribution of photo- to collisional ionization using our simulations. The interesting questions include, for example, the degree to which the abundance of low ions in the core of absorbers can be boosted in the optically thick regime.

To correctly compute the incident ionizing flux, one must in principle perform radiative transfer calculation. Although no radiative transfer was carried out in the simulation, we compute the optical depth approximately in post-processing using a local approximation for the attenuation of the ionizing photons calibrated by Rahmati et al. (2013). Specifically, we use Eq. (6) to parameterize the ratio of the attenuated photoionization rate (Γ_{Phot}) to the total background rate (Γ_{UVB}) as a function of the local density of a computational cell:

$$f_{\text{UVB}}(n_{\text{H}}) \equiv \frac{\Gamma_{\text{Phot}}}{\Gamma_{\text{UVB}}} = 0.98 \left[1 + \left(\frac{n_{\text{H}}}{n_{\text{H,ssh}}} \right)^{1.64} \right]^{-2.28} + 0.02 \left[1 + \frac{n_{\text{H}}}{n_{\text{H,ssh}}} \right]^{-0.84} \quad (4.6)$$

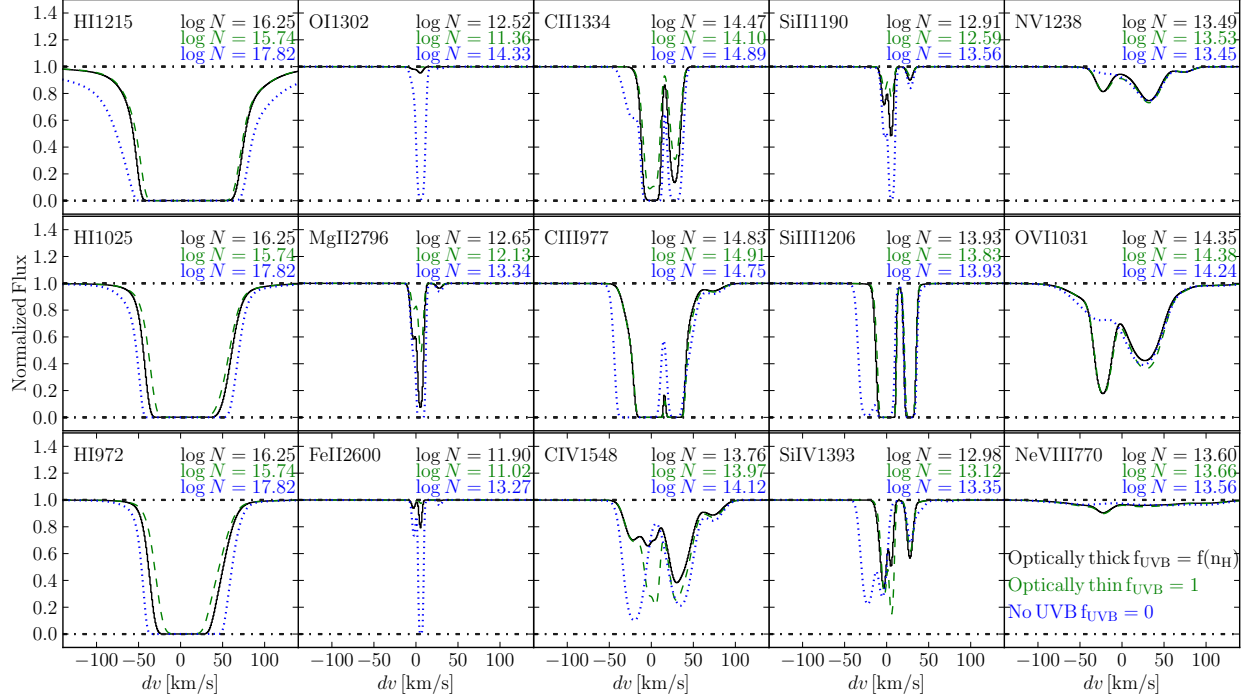


Figure 4.4: This figure shows the same LOS as the left panel in Figure 4.3 but with variations of ionization models. This shows how ionization models affect the total column densities of low, intermediate and high ions and their spectral features. The blue spectrum is computed using an ionization model with photoionization turned off (i.e collisional ionization only). The green spectrum assumes the gas is optically thin with 100% of the UVB contribution. Lastly, the black curve takes into account self-absorption of the CGM gas using an approximation for the fraction f_{UVB} that reaches the gas as a function of the local density from Rahmati et al. (2013). Comparing the model with collisional ionization only ($f_{\text{UVB}} = 0$), the optically thin photoionization model ($f_{\text{UVB}} = 1$) extinguishes a large fraction of low ions (e.g., HI, OI, MgII and FeII; blue to green), which results in the change of shape of the absorption profile. Absorption lines of OVI, however, change relatively little (green vs blue profiles), which shows that they are primarily collisionally ionized. Taking into account CGM self-absorption (black) generally increases the abundance of low ions compared to the optically thin model (green).

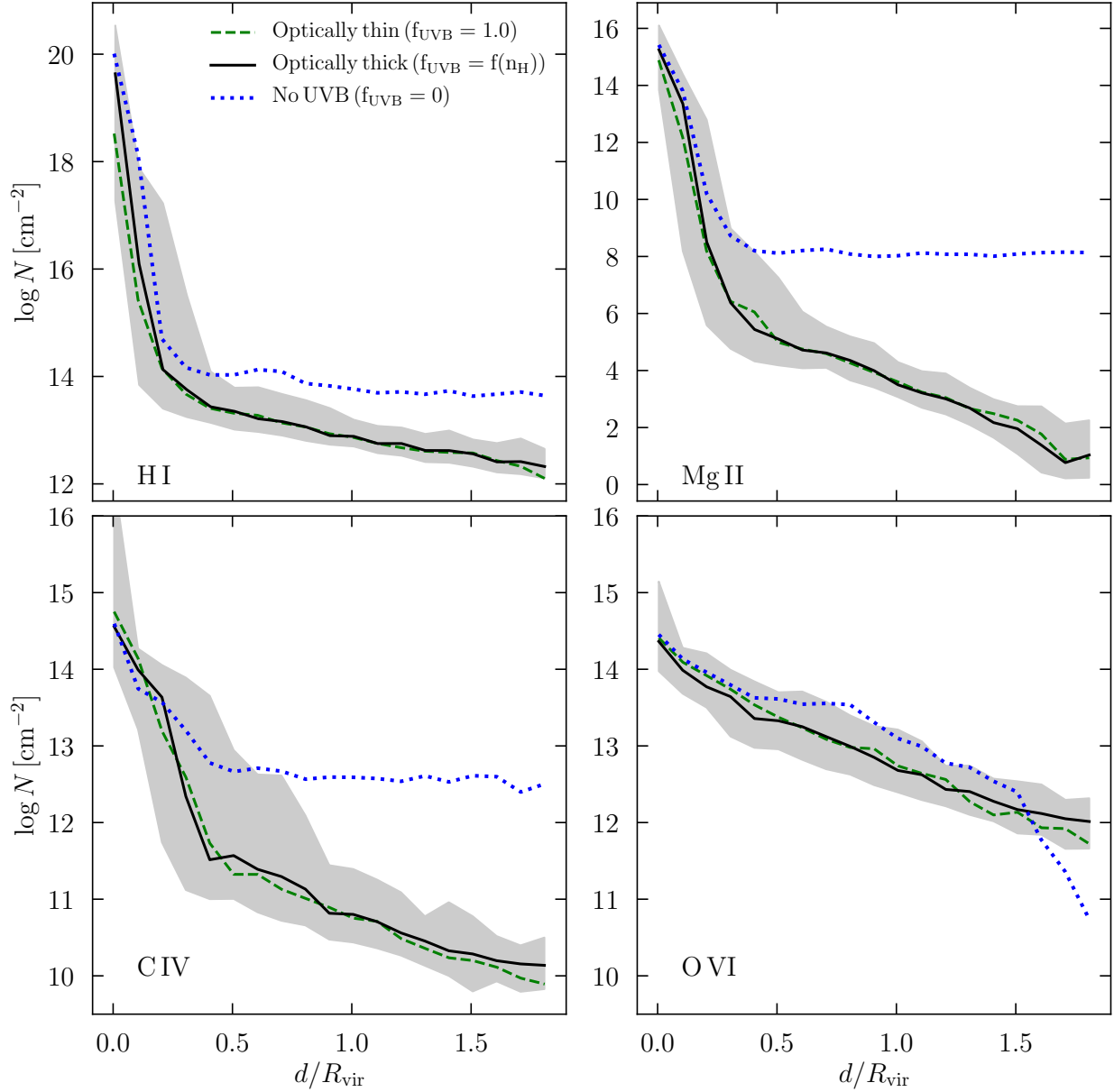


Figure 4.5: Column density profiles of HI, Mg II, CIV and OVI computed for the different treatment of ionization radiation (see legend and caption of Figure 4.4). The comparison of different ionization models show that photo-ionization from the UVB extinguishes an abundant of low and intermediate ions at distances beyond $\approx 0.2R_{\text{vir}} \approx 65$ kpc. On the other hand, photo-ionization does not have a strong effect on the column density of OVI, showing that it is primarily collisionally ionized. The model that takes into account CGM self-absorption generally increases the abundance of low ions.

where $n_{\text{H,ssh}}$ is a free parameter calibrated for effects of self-shielding, which depends on the UVB. For our choice of baseline UVB model (Haardt & Madau, 2012, HM12) and our analysis of the simulation snapshot at $z = 0$, we have adopted $n_{\text{H,ssh}} = 5.1 \times 10^{-4} \text{cm}^{-3}$ (table 2 in Rahmati et al., 2013). We then run a grid of CLOUDY (Ferland et al., 2013) models with varying ratio (f_{UVB}), density, and temperature in order to compute the modified ion fraction (e.g., $f_{\text{HI}} = f_{\text{HI}}(n_{\text{H}}, T, f_{\text{UVB}})$). For a given simulation cell with n_{H} and T , the ionization fraction is uniquely determined since f_{UVB} is also a function of n_{H} .

Figure 4.4 and 4.5 show the resulting spectra and the overall abundance of ions from three ionization models with varying degree of contribution from photoionization: optically thin ($f_{\text{UVB}} = 1$), taking into account CGM self-absorption ($f_{\text{UVB}} = f_{\text{UVB}}(n_{\text{H}})$), and no photoionization ($f_{\text{UVB}} = 0$). Figure 4.5 shows that taking into account attenuation of the background by self-absorption of the CGM does not affect the median column density profiles of ions substantially, as the profiles for $f_{\text{UVB}} = 1$ and $f_{\text{UVB}} = f_{\text{UVB}}(n_{\text{H}})$ at larger radii are nearly identical. This is because the covering fraction of dense and cold gas is small. At small radii ($r/R_{\text{vir}} \lesssim 0.2R_{\text{vir}}$) and high densities, accounting for self-absorption boosts the column densities of low ions by a factor of a few to ten (see the O I, Fe II, Mg II and C II panels in Fig 4.4). This difference is smaller than the scatter of the profiles for different LOS for this object.

The OVI profile is not sensitive to the CGM self-absorption corrections because in our simulations OVI absorbing gas *mostly* arises in hotter and less dense collisionally ionized gas. This is reflected in the relative insensitivity of the OVI column density profile to the presence of UVB radiation. HI, CIV, and Mg II ion column densities, however, are very different for the $f_{\text{UVB}} = 0$ case at $r/R_{\text{vir}} \gtrsim 0.2R_{\text{vir}}$ and thus show significant contribution from photoionization at these radii. Photoionization is significant for low ions for these large radii due to their low ionization potential and the decrease in optically thick gas with increasing radius. This also illustrates that photoionization shapes the exponential form of the column density profiles in Figure 4.5, as discussed in Liang et al. (2016).

4.3 A Bayesian Approach for Extracting Physical Properties of the CGM

Physical properties of the CGM in galaxy formation simulations are sensitive to models of star formation and feedback (e.g., Hummels et al., 2013; Suresh et al., 2015; Faucher-Giguère et al., 2015; Liang et al., 2016). For example, Liang et al. (2016) shows that metallicity of halo gas can change by a factor of 30 with a variation of star formation and feedback parameters. An advantage of combining synthetic spectrum generation tools with insights from the simulations is to make mock observations and directly compare quantities that are derived from them with the actual properties of the gas in the simulations. This can test the techniques used in observations, as well as make comparisons with simulations more meaningful.

Without proper treatment of upper limits, observational analyses are often plagued by a limited number of detections of ions to constrain physical properties of the absorbing gas. Lower limits similarly prohibit precise constraints of metallicity, which usually occur in saturated Ly α line in the absence of Lyman series due to limited wavelength coverage. In this section, we present a Bayesian framework for deriving physical properties of the absorbing gas, specifically focusing on systems with poor constraints of column densities. Both saturated and undetected absorption lines make modelling the underlying physical properties very uncertain since they require some modelling of the probability distribution of the lower/upper limits themselves.

Recently, a number of studies explored ways to maximize the utility of non-detections in observational analyses (e.g., Crighton et al., 2015; Fumagalli et al., 2016) by modelling the corresponding posterior probability distribution function (PDF). In particular, the PDF is assumed to be flat for column density below the upper limit and is joined by a one-sided Gaussian with constant standard deviation above the limit. However, the actual shape of the posterior distribution is determined by the nonlinear function through which parameters

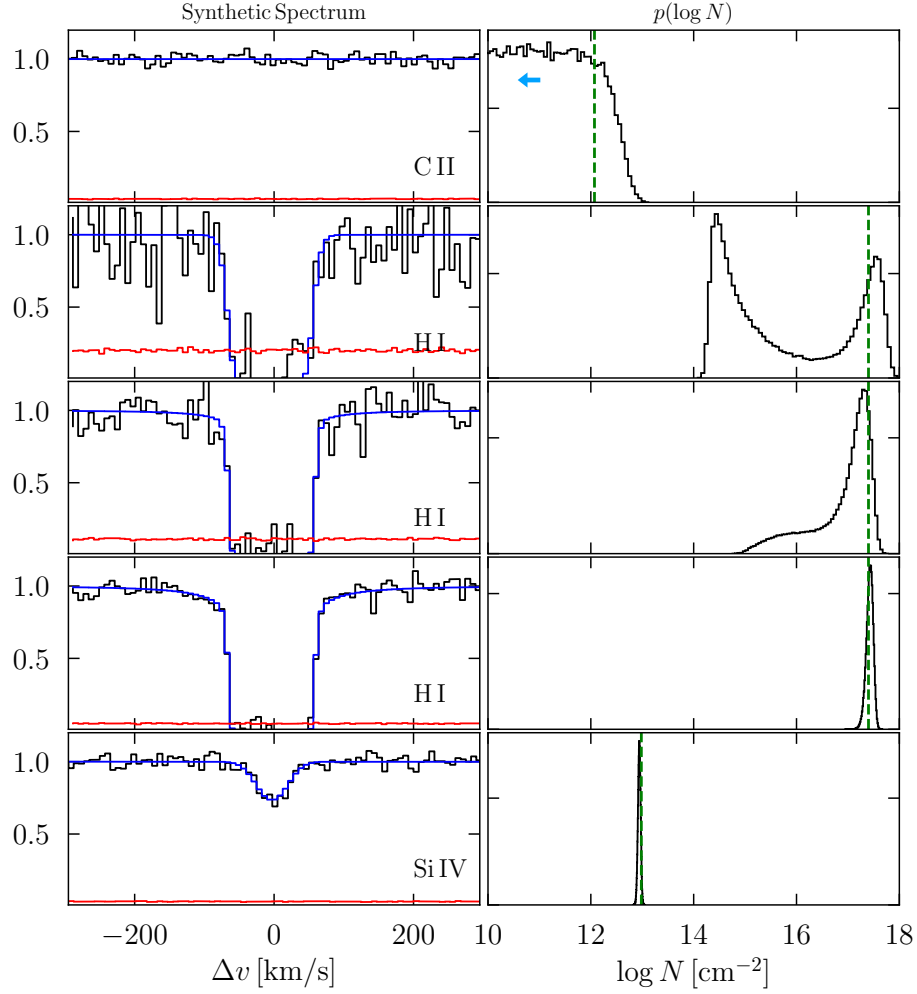


Figure 4.6: Each row shows synthetic spectrum (with continuum $C(\Delta v) = 1$) around a particular absorption line derived from an LOS through simulation along with the best fit model Voigt profile resulting from our Bayesian fitting (left panel) and the derived posterior probability distribution of the column density of the corresponding. The true column densities are indicated in the right panel by the green vertical dashed lines. We have selected five representative examples of absorption lines that show different degrees to which the lines were detected: from a non-detection (top row) to high signal-to-noise detection (two bottom rows). The same saturated H I line is repeated three times in the middle panels with varying S/N ($= 5, 10, 20$) showing that the resulting fit can be non-Gaussian in low S/N spectra due to the degeneracy between N and b -parameter. For the non-detection/upper limit constraints in C II $\lambda 1334$, the lack of decrement in the flux implies that column density of C II greater than $\sim 10^{13} \text{ cm}^{-2}$ is disfavored by the data. In contrast, the Si IV is well measured with a narrow Gaussian column density PDF. These probability distributions can then be used for ionization modelling even though C II line is not formally “detected” or that the H I line is saturated.

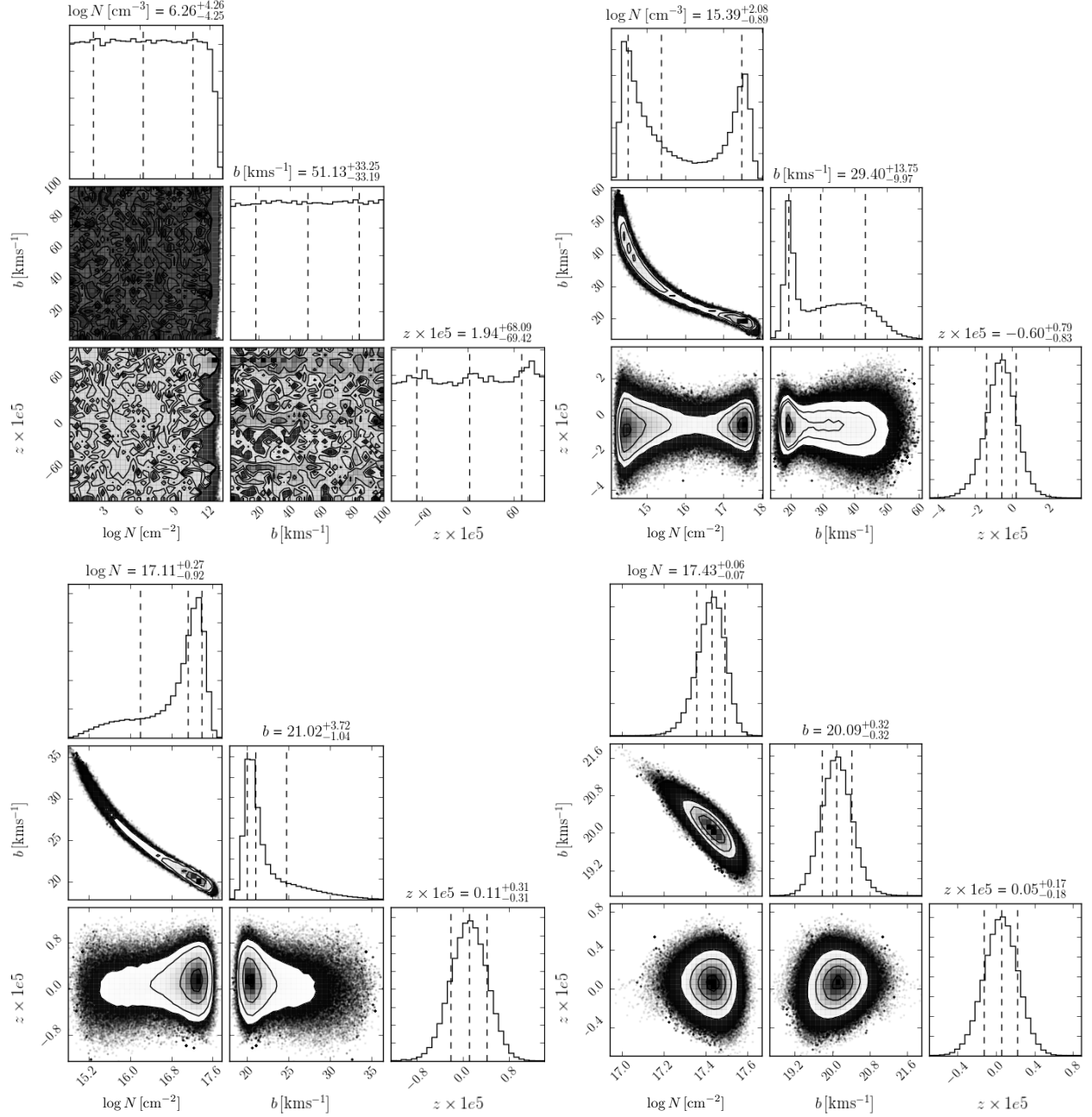


Figure 4.7: These four corner plots show the joint posterior distributions corresponding to absorption lines in Figure 4.6. The top left panel shows the posteriors of the non-detected C II line. The joint probability distributions show noisy fluctuation without a preferred location in the parameter space. The marginalized probability of column density drops precipitously for $\log N/\text{cm}^{-2} > 12$, where high column density is strongly disfavored by the data. The other three panels show the posteriors of the saturated H I line with corresponding S/N of 5, 10 and 20. The elongated shape in the joint probability shows the degeneracy between b and N in a saturated regime. As expected, the constraints improve as S/N of the spectra increases. The corner plots are made using publicly available code `corner.py` (Foreman-Mackey, 2016).

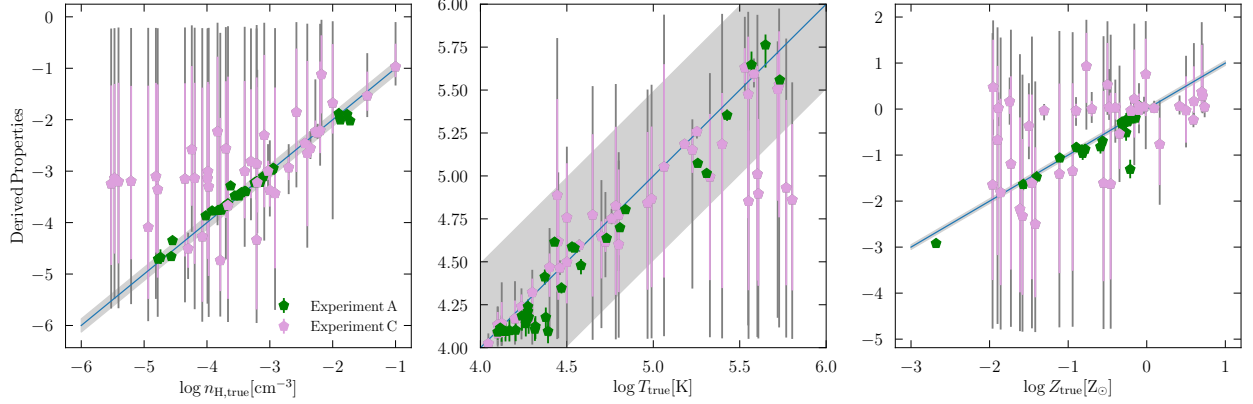


Figure 4.8: Comparison between derived properties and the “true” HI mass-weighted averages of each quantity along corresponding LOS in simulations. Experiment A (green points) represent values derived using the integrated column density along the entire LOS, which may include absorption from multiple absorbers. The column densities in experiment A have uncertainties of 0.1 dex. Generally, the derived values are close to the HI-weighted properties along the LOS. Experiment B (not shown) shows similar results as A. The gray bands show the typical intrinsic scatter along these LOS. The pink points in experiment C show how well one can retrieve the physical properties of absorbers when column densities constraints are poor (see the upper limit constraint of C II in Fig 4.6 for an example). The pink error bars show the ranges that enclose 68% and the gray error bars enclose 99% of probability of the posterior distributions. Naturally, the constraints of the derived parameters are poorer due to the lack of constraining power from upper limits PDF. Nevertheless, this demonstrates a bias-free, formally correct way to retrieve physical properties of the CGM when there’s limited information available.

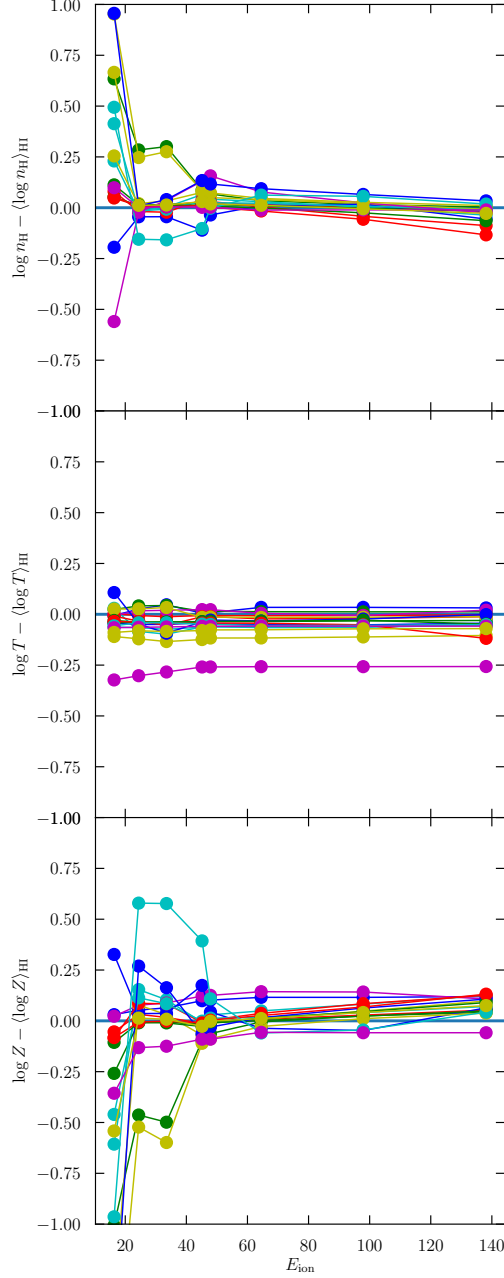


Figure 4.9: Experiment B: comparison of how well ionization modelling can retrieve the HI density weighted physical properties by including a set of ions with increasing ionization energy. Each line connected by color points represent a single LOS. The color points at a given E_{ion} means that the ionization modelling includes ions *up to* the specified E_{ion} . For example, points at $E_{\text{ion}} = 16.345$ eV means HI ($E_{\text{ion}} = 13.598$ eV), O I (13.618), Mg II (15.035) and Si II (16.345) are included in the ionization modelling. The typical error bars are larger with a low number of ions (not shown), and all derived properties are roughly consistent with the input values. Surprisingly, the derived physical properties approach the HI mass-weighted values by including higher ions.

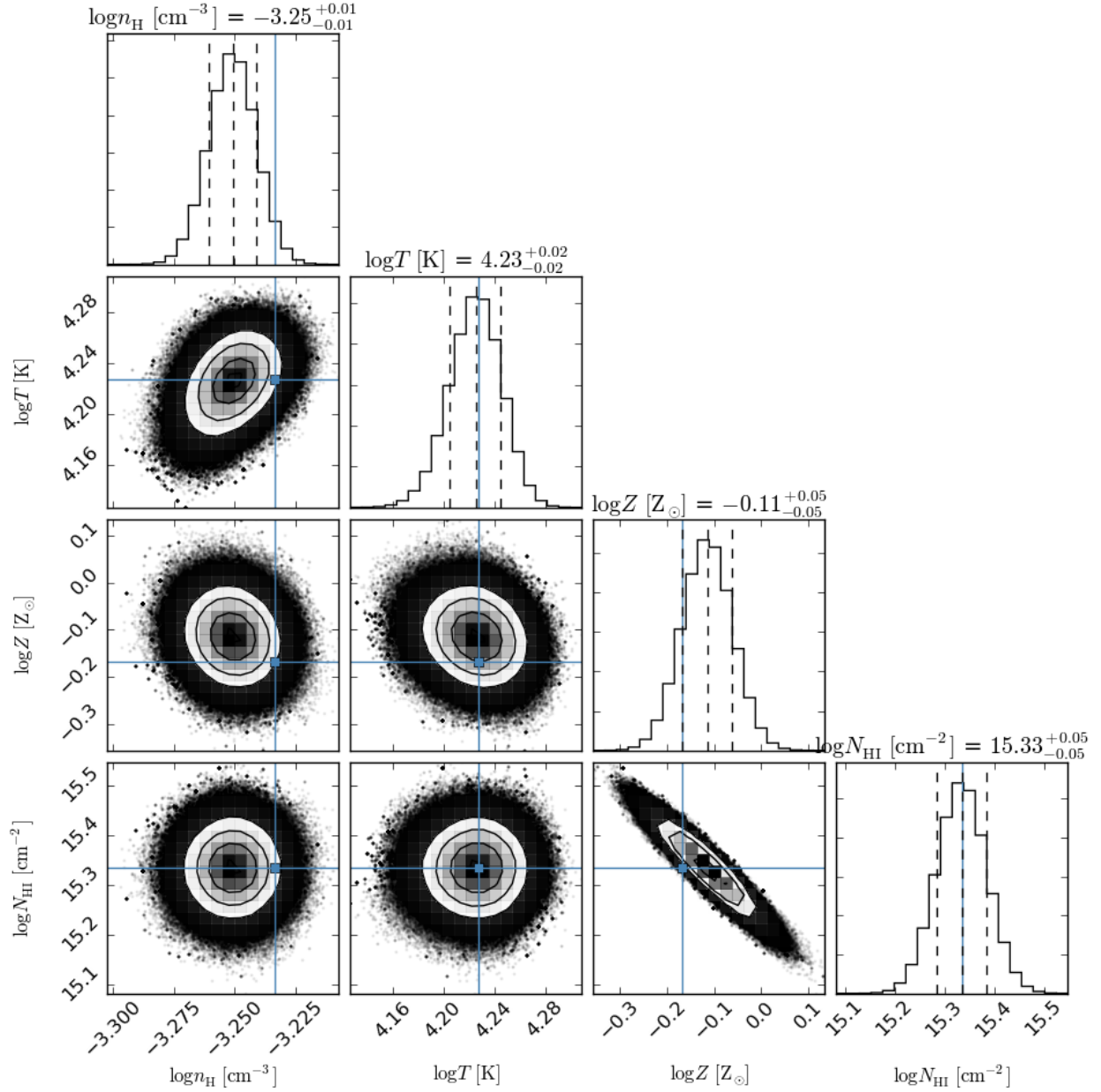


Figure 4.10: An example of posterior PDFs of the parameters in experiment B, in which all low to high ions with $E_{\text{ion}} \leq 140$ eV (up to O VI, with $\log N_{\text{OVI}}/\text{cm}^{-2} = 14.06$) are included. As shown, the derived properties are fairly consistent with the underlying H I-weighted values.

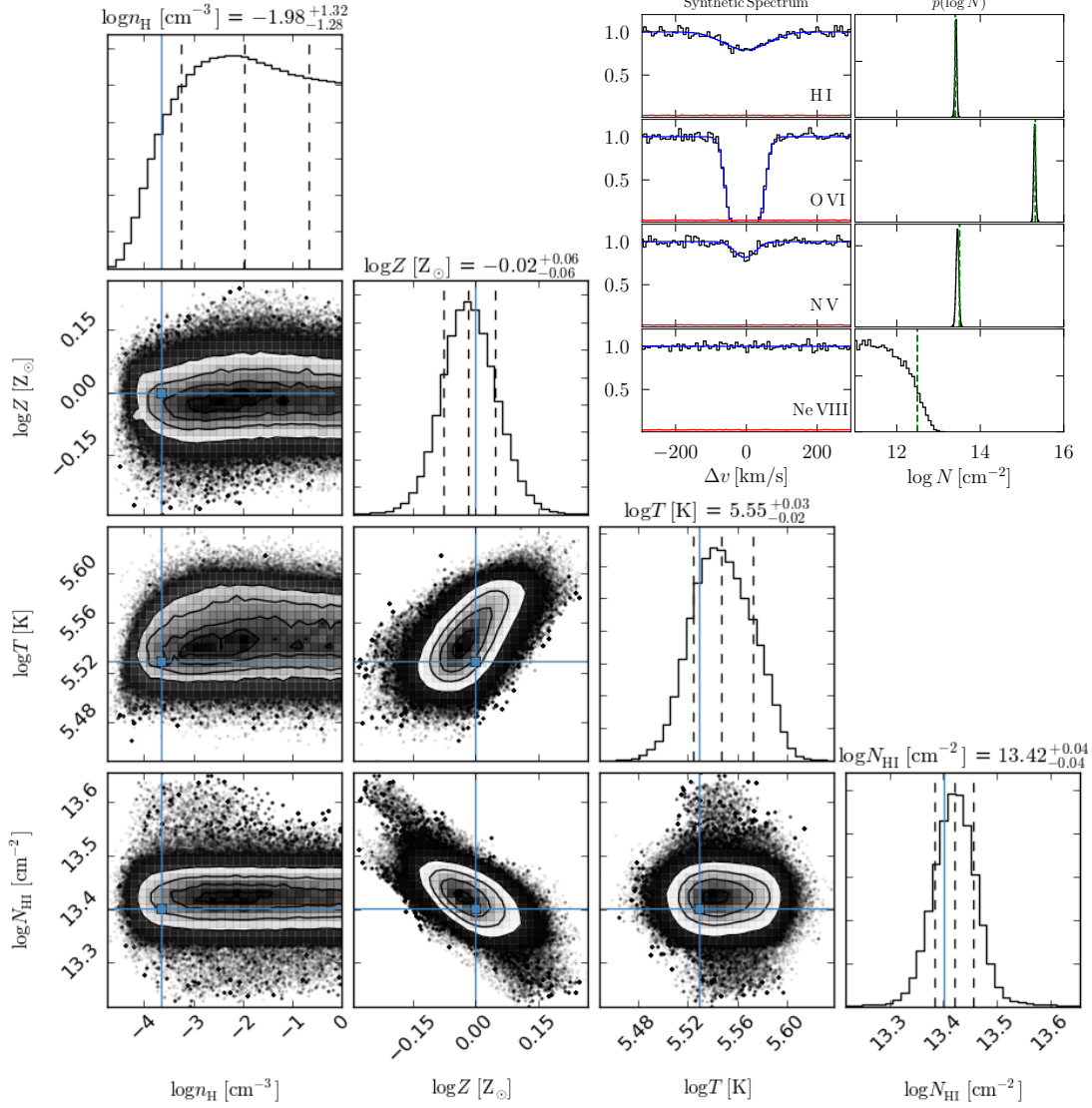


Figure 4.11: MCMC sampled posterior distributions of best-fit parameters of an ionization model (including photo- and collisional ionization) derived from constraints of column density in the form of non-Gaussian probability distributions including “non-detected” metal lines. In this example, column density constraints come from H I, O VI, N V and Ne VIII, where Ne VIII is a “non-detection” similar to C II constraints in Figure 4.6. The corresponding input mock spectrum and derived column density PDF is shown top right. Although with the poor column density constraints with Ne VIII, this figure shows one can formally recover the true parameters (shown by blue crosses) in this mock observation to the degree that the data allow. If N V becomes an upper limit when S/N of the spectrum is poorer, the corresponding constraints of temperature and metallicity will be poorer. Explicit priors can also be set to limit the wide range of allowed density based on our physical understanding of O VI-producing gas. Note that the probability distribution of H I column density from Voigt profile fit acts as a prior for the ionization modelling. In other words, the posterior of H I shown here is a re-sampled version of the posterior sampled from Voigt profile fitting.

enter into a model of observables and, partly, through assumption about prior distributions. The latter may have a significant effect on the posterior distribution when the constraining power of the data is weak, which is exactly the regime of “non-detections.” As is well known, and as our results below illustrate, the posterior distribution in this regime can have a wide range of shapes, that are generally not consistent with the standard assumptions.

In our approach, we take advantage of knowing the underlying properties of the gas in the simulation and ability to generate synthetic spectra to illustrate how a Bayesian approach can be used to avoid making assumptions about the posterior distributions by directly constructing them from the data or making explicit and more justifiable assumptions through the priors based on plausible physical conditions at hand. In addition, a Bayesian approach allows a uniform treatment of both undetected and detected (including lower limits) absorption lines. In particular, the PDFs of detections need not be Gaussian since the degree of saturation will simply be reflected in the shape of the posterior PDF.

Our pipeline starts with extraction of column densities of various ions using absorption lines in the spectrum and then uses the derived column density PDFs to constrain the density, temperature, and metallicity of the absorbing gas. In both steps, we use the Bayesian approach, combined with Monte Carlo Marko Chain (MCMC) sampling as we outline below.

4.3.1 Bayesian Voigt Profile Fitting

We introduce a Bayesian approach to Voigt profile fitting in this section. We refer readers to our companion paper (Liang & Kravtsov, 2017) where we describe in details the implementation and other useful utilities (e.g., continuum model) in a PYTHON package, `bayesvp`. The package is publicly available online¹.

Briefly, the posterior distribution for fitting a Voigt profile ($\vec{\theta} = \{N, b, z\}$) to the spectrum

1. <https://github.com/cameronliang/bayesvp>

(flux and its uncertainties; $\vec{F}$) can be written as:

$$\pi_{\text{VP}}(\vec{\theta}|\vec{F}) = \mathcal{L}_{\text{VP}}(\vec{F}|\vec{\theta})p_{\text{VP}}(\vec{\theta}), \quad (4.7)$$

where $\pi_{\text{VP}}(\vec{\theta}|\vec{F})$ is the product of the likelihood $\mathcal{L}_{\text{VP}}(\vec{F}|\vec{\theta})$ with the prior of the parameters $p_{\text{VP}}(\vec{\theta})$. The total likelihood is the product of all likelihoods given the individual (uncorrelated) velocity intervals of the spectrum:

$$\mathcal{L}_{\text{VP}}(\vec{F}|\vec{\theta}) = \prod_i \ell_i(F_i|\vec{\theta}) \quad (4.8)$$

The likelihood for flux F_i at a spectral pixel i , $\ell(F_i|\vec{\theta})$, represents the distribution of the fluctuations about the continuum, which can be measured in a spectral region free of absorption lines. For simplicity, it is usually assumed to be Gaussian (or Poisson distributed in photon limited cases). As an example, if ℓ is a Gaussian, the form is:

$$\ell_i(F_i|\vec{\theta}) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left[-\frac{(F_i(\vec{\theta}) - F_i)^2}{2\sigma_i^2} \right] \quad (4.9)$$

Here, we emphasize that the posterior $\pi_{\text{VP}}(\vec{\theta})$ does not make any distinctions among detections, lower limits, and upper limits. By definition and design, the probability π_{VP} for a combination of $\{N, b, z\}$ is simply constrained by the spectral data. The posterior form is determined on the fly and is not assumed. As mentioned above, we derive model parameters from MCMC chains using publicly available MCMC sampler **emcee** (Goodman & Weare, 2010; ForemanMackey et al., 2013) and **kombine** (Farr & Farr, 2015). Note that the initializations of walkers do not affect the PDF as long as the chains have converged and the priors encompass the true values of model parameters. For convergence criteria, we employ the Gelman-Rubin (GR) indicator (Gelman & Rubin, 1992) with $\text{GR} \leq 1.005$ (see the implementation details in Liang & Kravtsov, 2017). In addition, we select the known number of Voigt profile components in our tests for simplicity. Nevertheless, the

number of components (i.e the best model) can be identified using Bayesian evidence criteria implemented in our package (Liang & Kravtsov, 2017).

We now turn our attention to example fits that bracket cases of non-detections and detections, including saturated lines. The top panel of Figure 4.6 shows a very weak C II line buried inside a relatively noisy spectrum and the corresponding marginalized column density probability distribution given the data. Clearly, column density greater than 10^{13} cm^{-2} would result in absorption larger than the noise level allows. The posterior sampled by the MCMC chain disfavors larger values of column density. Therefore, the PDF for non-detection is flat for values $< 10^{13} \text{ cm}^{-2}$ and drops precipitously beyond that.

The second and third rows of panels of Figure 4.6 show when lines are detected but the derived column densities are uncertain. Most notably, the column density PDF of saturated H I line is bimodal due to the correlation between the Doppler parameter and the low S/N of the spectra. These two modes often are separated by two dex in column densities ($\sim 10^{15} \text{ cm}^{-2}$ and $\sim 10^{17} \text{ cm}^{-2}$). Note that it is difficult for a traditional χ^2 minimization grid search to capture the full degeneracy and locate both modes since it can be trapped in a χ^2 local minimum. This could result in a dramatic under- or overestimate of gas metallicity by orders of magnitudes if only one of these modes is adopted. Finally, this bimodality reduces to a single mode with if S/N of the spectrum is high enough. The bottom two panels show examples of well-detected absorption lines with Gaussian posteriors of the column density.

4.3.2 *Bayesian Ionization Modelling*

From the posteriors of column densities to posteriors of physical properties

Given the posterior distributions of column density of low and intermediate ions derived for each ion from the spectra, we can use the Bayesian approach to constrain parameters of a single-phase ionization model. As an illustration, we apply such analysis for an ionization model that uses two basic assumptions: (1) the absorbing gas is ionized by a combination of

photo and collisional ionization and is in ionization equilibrium and (2) absorbing gas can be modelled as a plane parallel slab described by a single density, temperature, and metallicity. These assumptions can, of course, be relaxed in more complicated models, which would, however, require more observational constraints in the form of more ion column density measurements.

For simplicity, we generate a grid of single-phase ionization models with parameters $\vec{\alpha} = \{n_H, T, Z/Z_\odot, N_{\text{HI}}\}$ using **CLOUDY** (Ferland et al., 2013). We assume the gas follows solar pattern. Note that we include N_{HI} as a parameter in the model (see also Crighton et al., 2015; Fumagalli et al., 2016), such that $N_k = N_k(\vec{\alpha})$ is the column density of the k th metal ion. Note also that the thickness of the plane-parallel slab in the model is not a free parameter, as it is determined by a combination of model parameters, $\Delta S = \frac{N_{\text{HI}}}{f_{\text{HI}}(n_H, T)n_H}$, where $f_{\text{HI}}(n_H, T)$ is the ionization fraction of H I.

The data constraining the model are a set of derived posterior distributions of column densities, $\vec{D} = \{\pi_{\text{FeII}}(N), \pi_{\text{SiII}}(N), \pi_{\text{SiIII}}(N), \dots\}$. Specifically, $\pi_k(N)$ is the probability distribution as a function of column density N for metal ion k marginalized over the b parameter and redshift z . Recall that these posterior distributions are extracted using MCMC sampling, as described in section 3.1, and can be of any form depending on how well absorption lines of a given ion are detected. We note that *all* co-spatial ionic species can be included in the ionization modelling, as long as there exists a transition within the wavelength coverage (and without contamination from systems of other redshifts). Within the framework described below the combination can also include detection of varying quality including upper and lower limits. We assume that the column density constraints for different ions are independent of each other.

The posterior of the ionization model parameters $\vec{\alpha}$, given the column density probability distributions, can be written as:

$$\pi_{\text{ion}}(\vec{\alpha}|\vec{D}) = \mathcal{L}_{\text{ion}}(\vec{D}|\vec{\alpha})p(\vec{\alpha}), \quad (4.10)$$

where the prior is simply:

$$p(\vec{\alpha}) = p(n_{\text{H}})p(T)p(Z/Z_{\odot})\pi_{\text{HI}}(N_{\text{HI}}). \quad (4.11)$$

Note that N_{HI} is constrained using the measured probability distribution π_{HI} in the Voigt profile fitting process. We adopt uninformative priors for $\log n_{\text{H}}$, $\log T$, and $\log Z/Z_{\odot}$ such that all values are equally probable within a given range.

The total likelihood $\mathcal{L}_{\text{ion}}$ is the product of individual likelihood π_k at N_k for k metal ions of a fixed set of parameters $\vec{\alpha}$:

$$\mathcal{L}(\vec{D}|\vec{\alpha}) = \prod_k \pi_k(N_k(\vec{\alpha})). \quad (4.12)$$

In other words, if one measures the column density probability distributions of Si II, Si IV and Fe II, $\mathcal{L}_{\text{ion}} = \pi_{\text{FeII}}(N(\vec{\alpha})) \times \pi_{\text{SiII}}(N(\vec{\alpha})) \times \pi_{\text{SiIII}}(N(\vec{\alpha}))$. In practice, we evaluate the probability for π_k at a given N_k through interpolated version of the MCMC sampled $\pi_k(N(\vec{\alpha}))$.

Mock Observations and Experiments

In the following, we conduct three sets of mock observations and experiments (A, B and C) to explore the relationship between the derived physical properties of absorbers and the underlying properties in the simulations through ionization modelling. Briefly, Experiment A tests if we can recover the true properties over an entire LOS by including the total integrated column density despite the number of velocity absorption components. Experiment B tests if one can include higher ionization species with the low ions. Experiment C tests the recoverability of physical properties in poor column density constraints (i.e., with an abundance of lower limits).

In experiment A, we broadly divide LOS into two types. The first set of LOS contains low and intermediate ions (i.e., HI, CII, CIII, CIV, SiII, SiIII, SiIV, OI, and FeII). The second

type of LOS probes the coronal hot gas with a negligible abundance of low metal ions. We, therefore, include HI and high ions only (e.g., NV, OVI, and NeVIII). Since it is often difficult to distinguish the spatial information along the LOS based on velocity information in real spectra, we are interested in testing whether the LOS HI - weighted properties can be recovered despite the number of velocity/absorption components. Therefore, we simply test the recoverability of the physical properties during the ionization modelling process by assuming the column densities are well measured with uncertainties of $\sigma_{\log N} = 0.1$ dex for all ions. We do so by directly integrating the number density in our simulations in this exercise. This implies that we include all absorption/velocity components in the spectra. Nevertheless, in Figure 4.8, we find that the derived values generally follow the HI mass-weighted properties along the LOS. Occasionally, catastrophic failures occur when there is more than one density peak are present in the LOS (shown by data points with small error bar but large deviation from the truth values). This is because the ratio of the sum of metals and HI is not equivalent to the ratio of metal and HI in each density peak. In other words, if these structures overlap in velocity space where column density contribution cannot be separated, single phase ionization model would fail to capture properties of both structures. It is also worth pointing out the deviation at low-temperature regime (near 10^4 K) in Figure 4.8 is due to the prior that prohibits MCMC walkers to explore temperature lower than 10^4 K. The posterior distribution is therefore asymmetric; the median is artificially offset despite having the most probable temperature at 10^4 K. This can systematic shift can be fixed by including CLOUDY models below 10^4 K and relaxing the priors.

It is common to only include ions that are co-spatial to extract properties using a single-phase model. This is because a single-phase model predicts sensible properties only if ions arise from the same density and temperature. For experiment B, in contrast, we test explicitly what properties one would derive if we include column densities of ions which are not co-spatial. Note that we only consider isolated peaks in contrast to the discussion of multiple peaks above. We do so by extracting small spatial segments (typically few kpc across) of

LOS containing a single structure with varying density and temperature (such as the density peak in the left panel of Figure 4.2). For each density structure, we perform a series of ionization fit by cumulatively including column densities of ions with increasing ionization energy, therefore tracing more and more diffuse gas at the interface of the dense absorber and the halo. For example, we first include 4 low ions with $E_{\text{ion}} \leq 16.345$ eV (i.e H I, O I, Mg II and Si II). Next, we add the ion with higher ionization potential (C II with $E_{\text{ion}} = 24.383$ eV) to test how the derived properties change, and so on up to O VI with $E_{\text{ion}} = 138.11$ eV. Somewhat surprisingly, Figure 4.9 shows that by including ions with increasing E_{ion} , the derived properties all converge to the H I mass-weighted properties over the density structure similar to experiment A. An example of a corner plot that includes both low and high ions are shown in Figure 4.10, where the recovered parameters are consistent with the underlying LOS values. Given that the derived properties are HI-mass weighted along a peak, the derived values slowly change depending on the ions included in the analysis. In other words, including low ions only in the modeling will give density, temperature, and metallicity closest to the highest density region. Including additional high ions will drive the derived density lower since high ions occur in the outskirts of the structure where density is lower.

Note that our finding here is in contrast to our usual expectation where O VI should not be included in the ionization modelling (e.g., Churchill et al., 2015). It is conceivable that what we show here can break down in higher resolution simulations where the morphology can be much more complex. It is also important to note that we use the same underlying ionization model that produce the ions abundances in the inverting process. In reality, we might not have full knowledge of the ionization process in observed ionic abundances which can jeopardize the retrieval of absorbers physical properties. Other sources in addition to the UVB may provide additional ionizing photons that change the relative abundance of ions (e.g. the highly ionized gas O VI discussed in Werk et al., 2016).

Finally, we design experiment C to illustrate the uniform treatment of detections and non-detections. To avoid possible confusion from experiment A and B, we generate an

idealized single phase model with a random combination of density, temperature, metallicity and velocity dispersion. A noise-free synthetic spectrum can then be created to represent such isothermal clouds. We then add noise to the synthetic spectra to the level ($S/N = 20$) that weak lines are not detectable in the traditional sense (i.e., by eye). The primary motivation for doing so is to test the utilities of upper limits. This is because detections in observed spectra can be scarce. It is especially useful to include proper treatment of upper limits when the number of detections is less than the ionization model parameters. Figure 4.8 shows ionization modelling results for 30 LOS with varying quality of column density constraints. In some cases where there is a mix of detections and non-detections, the posteriors are similar to that shown in Figure 4.11 when constraints are poor and are similar to Figure 4.10 when constraints are good.

The corner plot of Figure 4.11 shows the posterior distributions of density, temperature, and metallicity of the hot halo gas. Here, we use “detected” ions (H I, O VI, and N V) and one upper limits constraint (Ne VIII) as input for a four-parameter model. Despite the poor constraint in Ne VIII column density, the temperature and metallicity of the gas are recovered. Although constraint on the density is poor, this test shows that one can formally recover the true parameters to the degree that the data allow. The large error bars of the derived properties from these coronal LOS illustrate that the absorption spectrum, in this case, has little constraining power. Explicit priors can also be set to limit the any unphysical values of model parameters based on our understanding of the absorbing gas physical condition (e.g., unphysically large density for O VI-producing gas).

4.4 Conclusions

We explored the multiphase structure of the CGM probed by synthetic spectra through a cosmological zoom-in galaxy formation simulation. We use simulations to produce synthetic absorption spectra containing lines of different ionic species. In addition, we present a Bayesian method for modelling a combination of absorption lines to derive physical properties

of absorbers regardless of the signal-to-noise or whether a particular line is actually detected. In other words, the method treats detections (including lower limits) and non-detections (upper limits) in a uniform manner. Our main findings can be summarized as follows.

1. We show that in our simulations low ions (e.g., H I, Mg II), arise from the densest and coldest regions of absorbers, while intermediate and high ions arise in the interface regions between cold gas and hot halo, as shown in Figure 4.2 (see also Churchill et al., 2015).
2. In the lines of sight passing through the cold gas, absorption lines of low, intermediate and high ions are present in the spectrum and overlap in velocity space. This is because the velocity across a typical absorber does not vary significantly over sub-kpc scale in our simulations. In lines of sight that passing the hot halo, only H I and high ions are present with significant column densities. In particular, we find detectable O VI and O VII at $\log N \sim 13 - 14, \text{cm}^{-2}$ in our simulations. In such lines of sight, the absorption lines are typically broad due to the complex velocity fields across the entire halo (Figure 4.2 and Figure 4.3; see also Savage et al., 2011; Werk et al., 2016).
3. We find that attenuation of the UVB flux by neutral hydrogen in the CGM increases the abundance of low ions by a factor of a few to an order of magnitude. This difference, however, is smaller than LOS to LOS scatters. A comparison between photo- and collisional ionization models show that low and intermediate ions in our simulations are mostly photoionized, while high ions (e.g., OVI) is primarily collisionally ionized. In addition, photoionization from the UVB becomes more important at the outskirts of the halo ($> 0.25R_{\text{vir}}$) for low ions due to the decreasing covering fraction of dense gas. The decreasing density and constant UVB together are responsible for shaping the exponential form of the column density profiles of low ions (Figure 4.4 and Figure 4.5).
4. We present a Bayesian method for Voigt profile fitting that treats detections and non-

detections (upper and lower limits) in a uniform manner and produces posterior of column densities, b-parameters, and redshifts, reflecting the quality of the absorption lines. We also use Bayesian approach to derive density, temperature, and metallicity of absorbers from using these posteriors.

5. We show that the derived properties match closely the corresponding H I mass-weighted averages along the LOS. While this works well for an isolated structure with internal varying density, it also implies that single-phase model would fail catastrophically in multiple density peaks, especially if they overlap in velocity space such that their column density contribution are not well separated. Finally, we also show that constraints on H I and O VI together with upper limits on Ne VIII or N V, can in principle provide constraints on the properties of the hot halo (Figure 4.6 - 4.11).

The methods developed in this study can be used to make comparisons between simulations and observations more realistic using synthetic spectra. The results of our analysis indicate density, temperature, and metallicity of the CGM absorbing gas can be reliably recovered in a Bayesian ionization modelling presented here.

CHAPTER 5

ON THE MODEL OF CIRCUMGALACTIC MIST: THE IMPLICATIONS OF CLOUD SIZES IN GALACTIC WINDS AND HALOS

5.1 Introduction

The conceptualization of gaseous halos around galaxies arose from a hypothesis by Spitzer (1956): the observed cold gas at large distances away from the Milky-Way galactic plane (Münch & Zirin, 1961) must be pressure supported by a hot corona. This idea was later incorporated into the Lambda Cold Dark Matter (Λ CDM) paradigm where intergalactic gas accretes onto dark matter potential wells and is shocked heated to the virial temperature of the halo (e.g., White & Rees, 1978). Radiative cooling is thought to play an important role subsequently (Binney, 1977; Silk, 1977; Rees & Ostriker, 1977) and leads to multiphase and dynamic gaseous halos (Mo & Miralda-Escude, 1996; Maller & Bullock, 2004).

Observational evidence of multiphase galactic winds and halo gas has accumulated over the years in support of this picture. In the early days, the Mg II doublet $\lambda\lambda 2976, 2803$ has primarily been used for probing the connection of heavy elements with host galaxies at moderate redshifts (e.g., Bergeron & Boissé, 1991; Churchill, 1996; Steidel et al., 2002; Chen et al., 2010a). This is owing to the strong oscillator strength and relatively long wavelength of the Mg II doublet where it can be observed in the optical regime at intermediate redshifts. At the same time, galaxies with similar redshifts can be observed relatively easily using ground-based telescopes, such as the Sloan Digital Sky Survey (SDSS; York et al., 2000).

The launch of the *Hubble Space Telescope* (HST) and the Cosmic Origin Spectrograph (COS; Green et al., 2012) brought a renaissance to the galaxy-gaseous halo connection and characterizations. The ultraviolet (UV) wavelength coverage and the high sensitivity of COS allow one to measure absorption lines of a broad range of ions, probing a wide range

of physical conditions of the multiphase CGM. These include early studies probing neutral hydrogen using $\text{Ly}\alpha$ $\lambda 1215$ (Lanzetta et al., 1995; Chen et al., 1998, 2001b), heavy elements with C IV $\lambda\lambda$ 1548,1550 (Chen et al., 2001a), and highly ionized O VI (Chen & Mulchaey, 2009; Prochaska et al., 2011).

In recent years, there has been an accelerated progress thanks to a combination of large HST and ground-based observing programs that probe a broad range of ions at once (e.g., H I, Si II, Si III, Si IV, C II, C IV, O VI). For example, the COS-Halo program (Tumlinson et al., 2011; Werk et al., 2013) probes a set of star-forming and quiescent galaxies within 160 kpc. The authors show the CGM harbors a significant fraction of the baryons (Werk et al., 2014). The COS-GASS program explores the connection between the interstellar medium (ISM) and the CGM by combining with HI 21 cm line observations (Borthakur et al., 2015). While these studies focus on the inner regime of the CGM of some specific types of galaxies, Liang & Chen (2014), in contrast, used 195 galaxy-QSO pairs to probe CGM with a broad range of stellar mass of the galaxies ($10^7 - 10^{12} M_{\odot}$) and out to large distances (500 kpc). With a stringent limit, they show that detection rates of low ions (e.g., Si II) drop quickly beyond a small fraction of the virial radius ($0.4 - 0.5 R_{\text{vir}}$), while intermediate ions (e.g., Si IV, C IV) are detected up to $\sim 0.7 R_{\text{vir}}$. Johnson et al. (2015b) extended this study to highly ionized gas where they found continuous detection of O VI up to $\sim R_{\text{vir}}$.

One of the natural explanations for this enrichment of the CGM is via galactic winds (e.g., Muratov et al., 2015). Galactic outflows are commonly observed in both the local and high-redshift Universe (e.g., Shapley et al., 2003; Veilleux et al., 2005). In addition to the hot phase observed through X-ray emitting gas (e.g., Strickland & Heckman, 2009), the outflows can be observed via self-absorption of Na I λ 5890,5896 and Mg II doublets (e.g., Heckman et al., 2000; Weiner et al., 2009) and warm ionized gas (e.g., Lehnert et al., 1999; Heckman et al., 2017). In some cases, even molecular gas can also be detected (e.g., Veilleux et al., 2009; Beirão et al., 2015; Leroy et al., 2015).

These analyses have provided a comprehensive characterization of multiphase gas in a

broad range of galaxies. The task from the theoretical side has been to understand the implications of these observations and incorporate the governing physical processes into our modeling of galaxy formation. The current generation of galaxy formation models appears to capture many of the relevant physical processes. State-of-the-art simulations can now produce galaxies with realistic stellar properties (Feldmann et al., 2011; Sales et al., 2012; Aumer et al., 2013; Stinson et al., 2013; Hopkins et al., 2014; Agertz & Kravtsov, 2015; Schaye et al., 2015). This success in the theoretical community can be mainly attributed to our improved understanding of star formation and feedback processes, and the corresponding improved subgrid recipes in our simulation codes. A general conclusion is that galactic winds are believed to play crucial roles in shaping the galaxy as a whole. It has been shown that star formation feedback can drive large-scale galactic winds to regulate star formation. In doing so, the hot winds can provide the necessary heating to balance the global cooling of gaseous halos.

Despite the success of recent simulations in producing realistic stellar properties of galaxies, matching the CGM properties proves to be difficult. Simulations can explain some CGM properties, such as the bimodality of O VI column densities and specific star formation rate (sSFR) of the host galaxies (Tumlinson et al., 2011; Oppenheimer et al., 2016). However, most simulations have been unable to produce enough cold/warm gas in the CGM, despite a diversity of feedback processes implemented in them (e.g., Stinson et al., 2012; Ford et al., 2013; Hummels et al., 2013; Shen et al., 2013; Suresh et al., 2015). On the other hand, a variation of large feedback energy in Liang et al. (2016) leads to a more realistic CGM by launching cold gas into the halo; however it does so by destroying their central galaxy. It is also thought that cosmic rays feedback may help launch the cold gas into large distances (e.g., Socrates et al., 2008; Booth et al., 2013; Salem et al., 2014), although it has been shown that these initial attempts were not yet enough (e.g., Liang et al., 2016; Salem et al., 2016; Butsky & Quinn, 2018).

We are motivated by many factors to consider new solutions to the problem of under-

producing the cold/warm phase of the CGM in simulations. These factors include (1) the ubiquitous detections of cold gas in halos of nearly all types of galaxies, (2) small sizes of cold gas, which are both observationally and theoretically motivated, and (3) the common underproduction despite a broad range of simulation codes, star formation and feedback processes and implementations. In fact, it is becoming clear that galaxy simulations with cosmological volume lack the appropriate spatial resolution to resolve the sizes of cold gas. This is because these simulations traditionally focus their computational power on refining the densest region of galaxies. As a result, the galactic halo of low-density gas is resolved only down to the order of $\sim$ kpc. By contrast, it has been shown observationally that the cold gas is in the order of $\sim$ pc (e.g., Rauch et al., 1999; Prochaska & Hennawi, 2009; Crighton et al., 2015; Lan & Fukugita, 2017). Similarly, theoretical models and small-scale simulations focused on the galactic winds and thermal instabilities have long predicted pc-size cold cloudlets under various conditions (e.g., Burkert & Lin, 2000; Sánchez-Salcedo et al., 2002; Audit & Hennebelle, 2005; Cooper et al., 2009; McCourt et al., 2018). Models that consider condensation of cold gas from hot halo have also been quite successful in explaining observations of intracluster medium (e.g., McCourt et al., 2012; Gaspari et al., 2012; Li et al., 2015; Voit et al., 2015b; Gaspari et al., 2017; Fogarty et al., 2017)

In summary, the challenges and tantalizing hints from this line of observational and theoretical studies point towards the need to consider small-scale structure of cold gas in the circumgalactic context. Thus, we explore a new multiphase model, the circumgalactic mist or CGm, and its implications for the multiphase nature of galactic winds and gaseous halos. In a series of papers, we aim to investigate the CGm model on explaining a multitude of observations, including the cloud sizes, the gas content at various phases, their kinematics, and more.

In this paper, we focus on the production of cold gas due to thermal instability in a set of small-scale simulations with an array of physical processes to study the resulting cloudlet sizes. We are interested in the spatial distribution the CGM in various phases, in particular,

the presence of observed small cold clouds in the inner CGM, and the lack thereof at larger radii. We leave the survival of cold cloudlets and the kinematics of the circumgalactic mist in a future study.

The structure of the paper is organized as follows. In section 5.2, we discuss the process of thermal instability due to runaway radiative cooling with a focus on the resulting sizes of cold clumps. We provide a simple semi-analytic model to describe the evolution of isobaric cooling and cloudlet sizes. In section 5.3, we use a set of MHD simulations to study an episode of the birth and death of cold gas and their sizes in the presence of a hot wind. We systematically introduce physical processes, such as radiative cooling/heating, thermal conduction and magnetic fields to study their effects on the cloudlet sizes or the lack thereof. In section 5.4, we discuss the implication of the characteristic cloudlet sizes predicted by the model, showing that the inner CGM may be composed of an upward of millions of cold clumps with sizes in the order of parsecs. Consequently, the model also predicts the spatial extent of ions probing the cold, warm and hot phase of the CGM. Throughout this paper, we have used the term cloudlet as a shorthand for characteristic parsec-sized regions of cold gas with increased density.

5.2 Semi-Analytic Consideration

Much of the previous analyses of thermal instability has focused on the process of runaway cooling that leads to a multiphase medium at a final equilibrium state (e.g., Parker, 1953; Field, 1965; Hennebelle & Péroult, 1999; Burkert & Lin, 2000). Famous two-phase and three-phase models of the interstellar medium (ISM) were born out of the studies of thermal instabilities (Field et al., 1969; McKee & Ostriker, 1977). Cloud formation and compressible turbulence due to isochoric and isobaric cooling have been an area of active research for many years as well (e.g., Kritsuk & Norman, 2002; Sánchez-Salcedo et al., 2002; Audit & Hennebelle, 2005).

We begin our study with theoretical expectations based on semi-analytic calculations. In

this section, we consider the nonlinear evolution of thermal instability via fragmentation and monolithic condensation. However, we are most interested in the final characteristic size of cloudlets from which thermal instabilities bring.

5.2.1 *Fragmentation in thermal instability*

It is well known that a region of perturbed density of an initial scale $r_{\text{cl},i}$ will undergo condensation if the cooling time is large compared to the corresponding sound crossing time, $t_{\text{cool}}/t_{\text{cross}} > 1$ (e.g., Field, 1965; Burkert & Lin, 2000). In contrast, regions of density perturbations with $t_{\text{cool}}/t_{\text{cross}} < 1$ cool more quickly than thermal motion can equilibrate pressure; large pressure gradients can develop. In turn, these pressure forces lead to a more dynamic evolution of the condensation (e.g., Field, 1965; Meerson, 1996; Sánchez-Salcedo et al., 2002).

Most recently, McCourt et al. (2018) argued that a faster way to reach the equilibrium state is via fragmentation, similar to the process of the Jeans' instability in gravitational collapse. This thermal fragmentation process ensues when the communication across a cloudlet by thermal motion is slow compared to the cooling time t_{cool} . In other words, this shattering process stops when thermal pressure is able to smooth out perturbations faster than their growths:

$$\frac{r_{\text{cl}}}{c_s} \equiv t_{\text{cross}} \leq t_{\text{cool}} \quad (5.1)$$

where r_{cl} is the size of the perturbed density region (a cloudlet) and c_s is the sound speed.

Therefore, one can write down the expected cloudlet size at the end of the fragmentation process based on this ansatz. Given a metallicity Z_{cl} , sizes of cloudlets live on a density-temperature phase space (Figure 5.2) that satisfy the following condition:

$$r_{\text{cl}} \leq c_s(T_{\text{cl}})t_{\text{cool}}(n_{\text{cl}}, T_{\text{cl}}, Z_a) \quad (5.2)$$

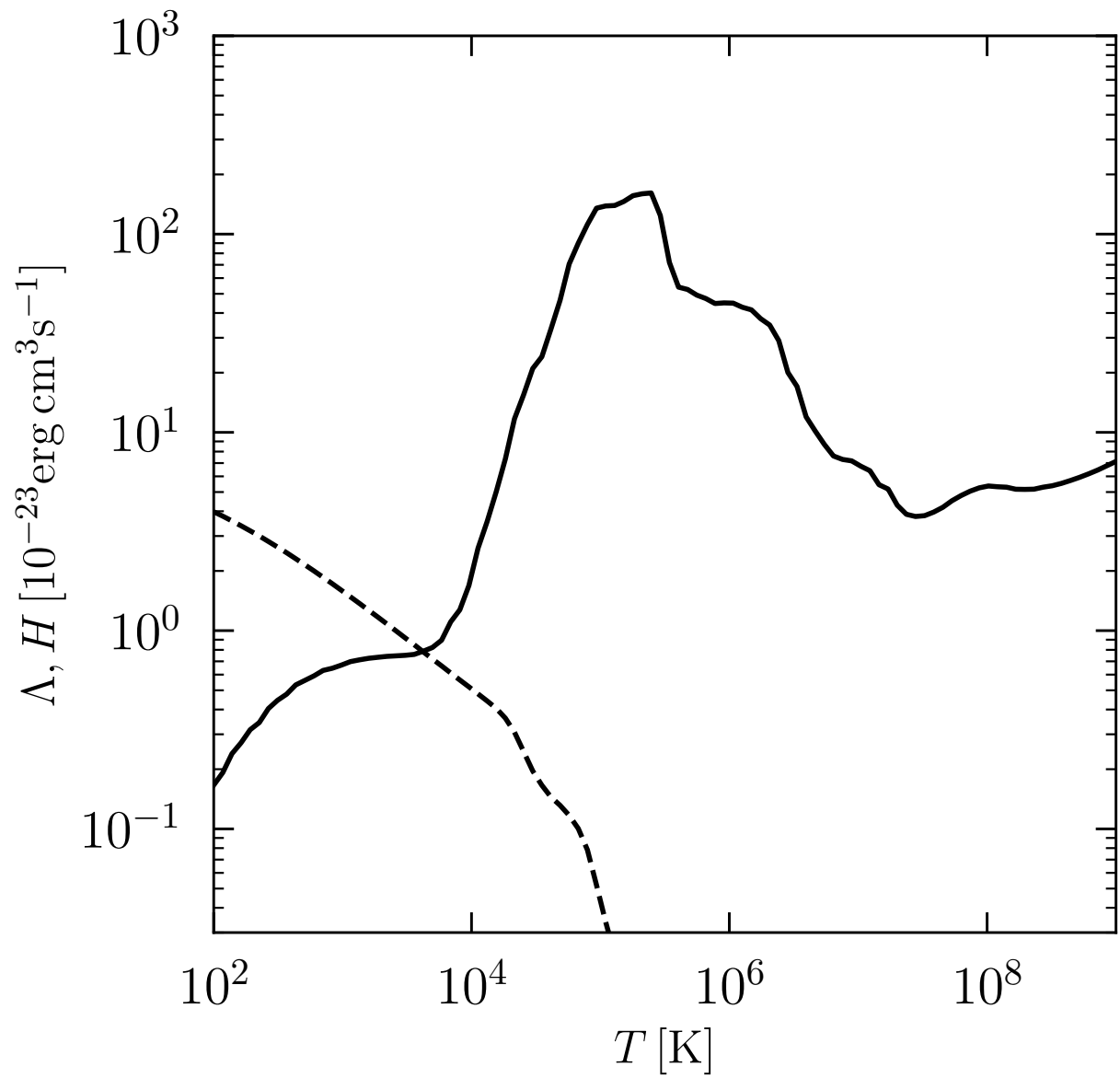


Figure 5.1: Cooling (solid) and heating (dashed) curves used throughout this work at $Z/Z_{\odot} = 0.3$ (Gnedin & Hollon, 2012).

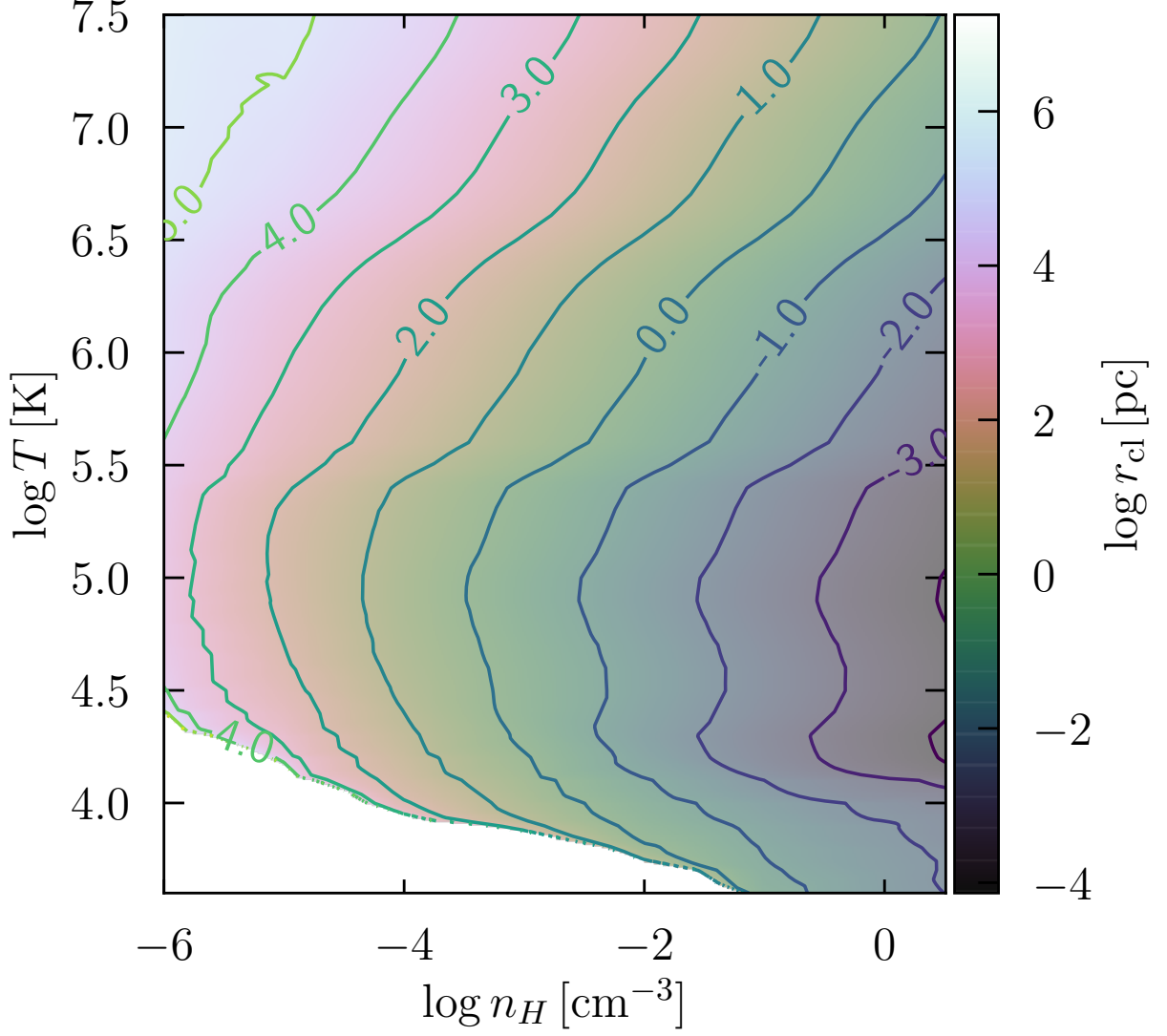


Figure 5.2: The size of cloudlets under the ansatz, $r_{\text{cl}} = c_s(T)t_{\text{cool}}(n, T, Z)$, as a function of density and temperature. The contours are constant cloudlet sizes in parsecs at $Z = 0.3Z_{\odot}$. The contours of different sizes converge to the lower boundary, which is set by the balance of heating and cooling. In other words, the constraint of the cloudlet size disappears as $t_{\text{cool}} \rightarrow \infty$. At this point, the problem becomes scale-free. Note that at the typical condition of observed warm gas ($T < 10^{4.5}\text{K}$ and $n_{\text{H}} \sim 10^{-3}\text{cm}^{-3}$), the expected cloudlet size is of the order of 1 pc or less.

The equality predicts the allowed cloudlet sizes at an evolutionary boundary between fragmentation and monolithic collapse as a function of the temperature, density, and metallicity. We can see in Figure 5.2 that the cloudlet sizes can be very small. At a typical physical condition of the observed cold/warm gas with $T < 10^{4.5}\text{K}$ and $n_{\text{H}} \sim 10^{-3}\text{cm}^{-3}$, the expected size is in the order of 1 pc or less. However, it is worth pointing out that the growth rate is a function of spatial scale. In fact, the perturbation can only grow as fast as pressure gradient pushes them around on a time scale of t_{cross} .

The contours of constant cloudlet sizes in Figure 5.2 converge to a lower boundary set by the balance of heating and cooling. At this locus, the cooling time is so long that it no longer limits the size of the cloudlet (i.e., $t_{\text{cool}} \rightarrow \infty$).

Models of cloud formation that describe the size of cold gas have mainly focused on the competition of cooling time and sound crossing time at the hot phase (e.g., Hennebelle & Pérault, 1999; Burkert & Lin, 2000). As noted by McCourt et al. (2018), the difference here is that the final cloudlet size depends on the cooling time and the sound speed evaluated near the temperature at which $c_s t_{\text{cool}}$ is minimized:

$$r_{\text{cl}} \approx \min(c_s t_{\text{cool}}) \quad (5.3)$$

Figure 5.3 shows the length scale of $c_s t_{\text{cool}}$ as a function of temperature. The corresponding temperature at which Eq. 5.3 is minimized happens at $T_{\text{min}} \approx 10^{4.3}\text{K}$. It is worth pointing out that the parcel of dense gas will continue to cool as long as its cooling time is not formally infinite. However, the increase of cooling time is dramatic with $T < T_{\text{min}}$ in the regime we are considering. Therefore, we assume the cooling and density growth are negligible after this point; the final temperature of the cloudlet is close to this minimum temperature, $T_{\text{cl,f}} \approx T_{\text{min}}$, as we will discuss in what follows.

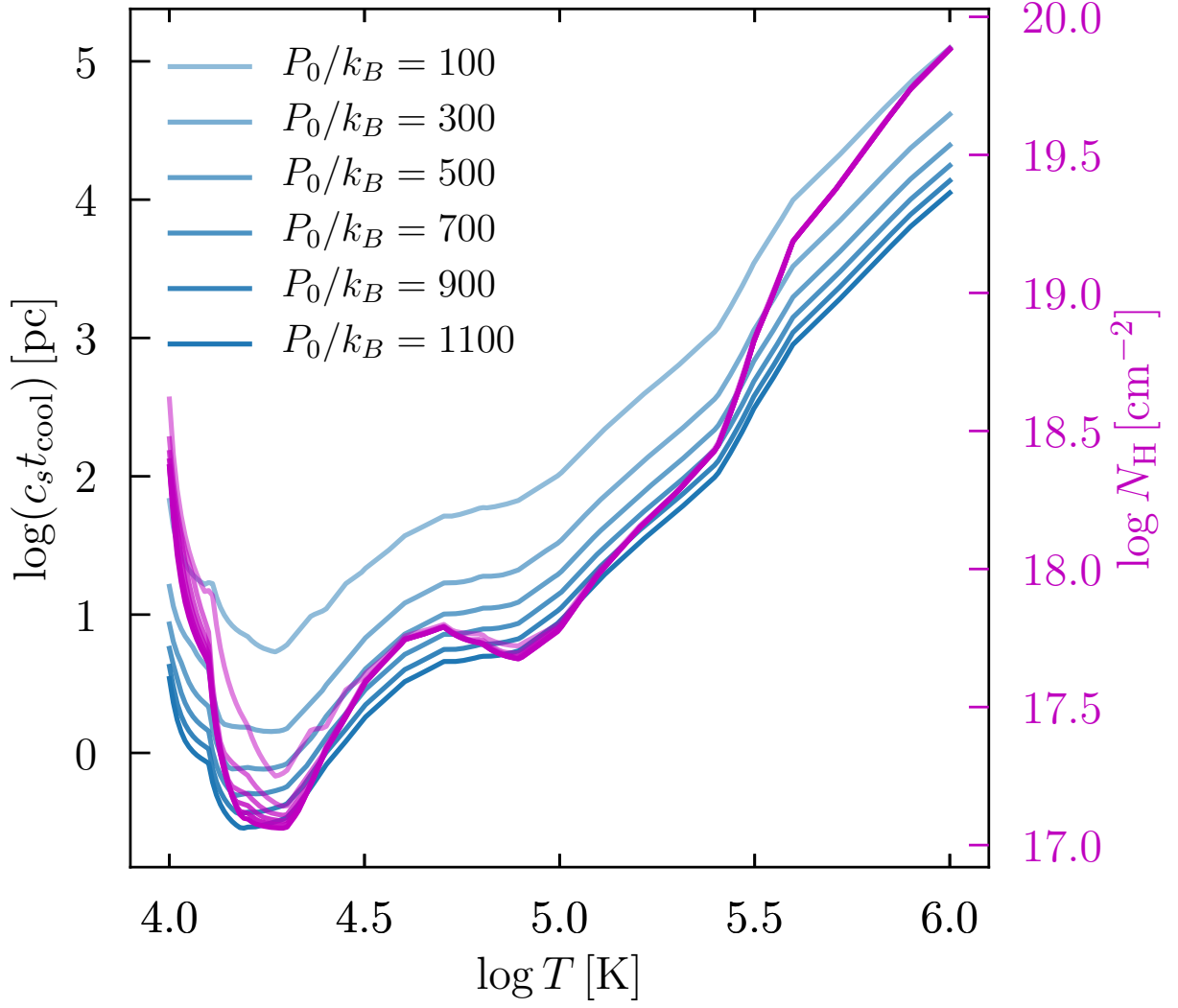


Figure 5.3: The cloudlet sizes as function of temperature for different ambient pressures. The gradient of blue lines shows higher background pressure is able to compress cloudlets more, leading to smaller sizes. The corresponding column densities of cloudlets are represented by the series of magenta lines of different opacities. As shown, the column densities are insensitive to the change of background pressure.

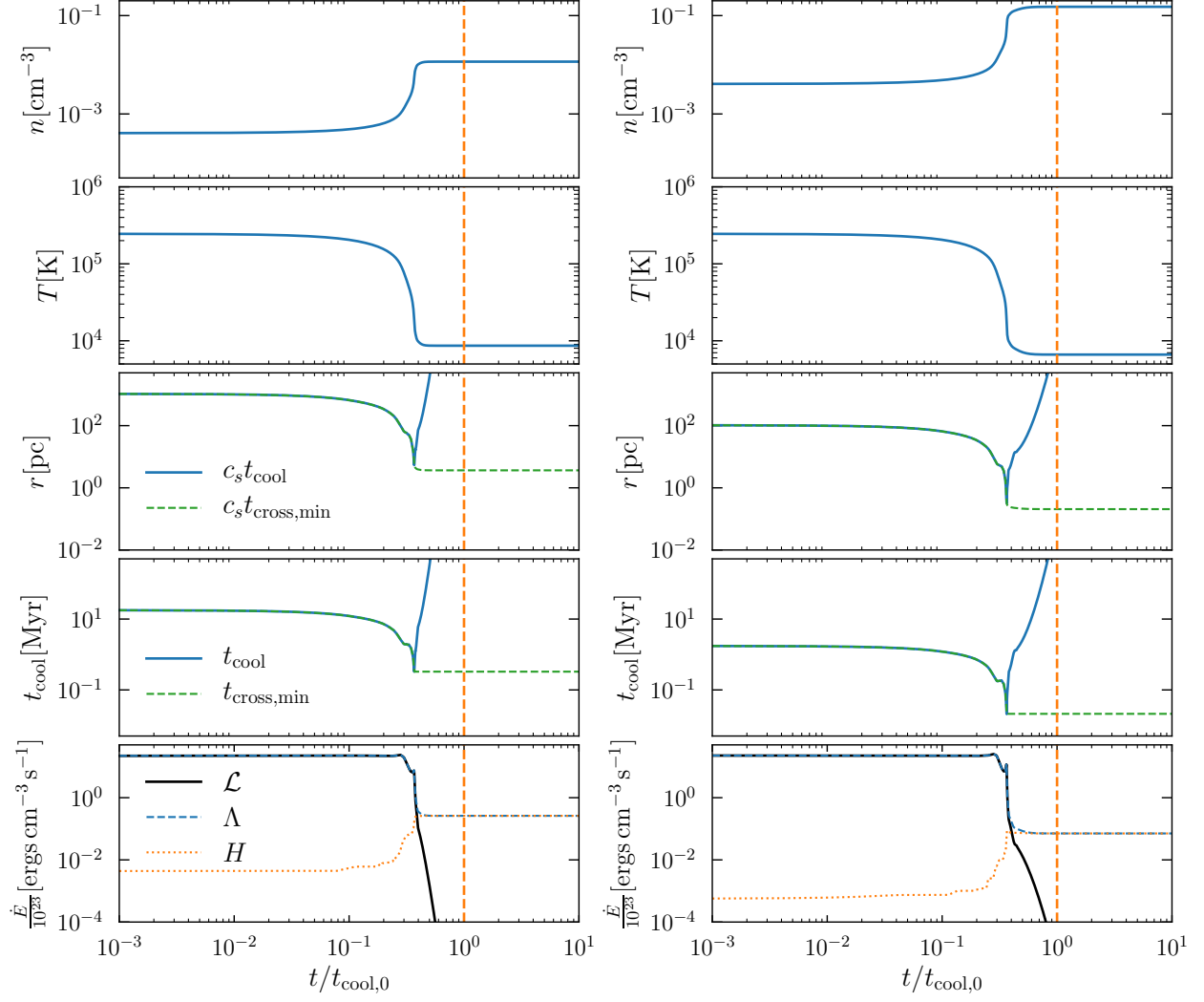


Figure 5.4: The evolution of the cloudlet properties given the background pressure with $P/k_B = 100$ (left panel) and $P/k_B = 1000$ (right panel). As shown, cloudlet quickly cools and condenses within a cooling time of the initial state $t_{\text{cool},0}$. Upon reaching the minimum of the cooling time, the radius of the cloudlet follows $r_{\text{cl}} = c_s t_{\text{cross,min}}$ as t_{cool} departs and grows as the net cooling rate approaches zero. The final size of the cloudlet depends on the external pressure as shown, where $r_{\text{cl}} \approx 0.2$ pc for $P/k_B = 1000$, and $r_{\text{cl}} \approx 3$ pc for $P/k_B = 100$.

5.2.2 *Isobaric condensation mode in thermal instability*

In this subsection, we study the evolution of thermal instability by considering an initial uniform perturbed density region with a size $r_{\text{cl},i}$, such that with $t_{\text{cool}} = t_{\text{cross}}$. In this regime, the perturbed density will undergo isobaric condensation.

While the Jeans' instability is driven by the local gravity of the collapsing objects, the size of gaseous cloudlets due to runaway cooling depends on the competition between its internal pressure and the ambient pressure. As a region of overdensity cools, internal pressure support is lost. The ambient pressure (if higher) squeezes the overdensity further to maintain pressure equilibrium. This isobaric compression continues until heating and cooling are balanced, a pressure equilibrium state is reached at the end (e.g., see Figure 5.9). From this, we can simply rewrite the final density of cloudlet as $n_{\text{cl},f} = P_h / (k_B T_{\text{cl},f})$, where P_h is the pressure of the ambient hot gas.

The final temperature $T_{\text{cl},f}$ and the minimum of the cooling time depend on the cooling function, via the metallicity of the gas in which the cloudlet condenses out of (i.e., $Z_{\text{cl}} = Z_h$). Therefore, it is useful to express the cloudlet properties (density, temperature, and size) entirely as functions of their environment, P_h and Z_h .

As discussed above, one can consider a cloudlet that satisfies the condition given by Eq. (5.2) with uniform perturbed density of size $r_{\text{cl},i}$ in an idealized setting. This idealized parcel of gas would not continue to shatter into smaller fragments in the absence of perturbations of smaller scales. Nevertheless, the cloudlet will still undergo runaway condensation due to radiative cooling ($\dot{e} = n^2 \mathcal{L}$) as long as there exists net cooling $\mathcal{L} = \Lambda - H > 0$. Using the fact that the final size of a cloudlet is tied to its thermal properties, we can solve for it by considering its energy evolution. In this idealized setting, we assume thermal conduction is suppressed owing to the ubiquitous detections of cold gas in hot halos. Thus, the accounting of the thermal energy of the cloudlet is:

$$\frac{de_{\text{cl}}}{dt} = -n_{\text{cl}}^2(\Lambda - H) + \frac{P_{\text{cl}}}{n_{\text{cl}}} \frac{dn_{\text{cl}}}{dt} \quad (5.4)$$

The first term on the right-hand side is the loss of internal energy by radiative cooling; the second term is the work done by the background through pressure forces. If we consider isobaric compression over a time interval dt , the internal energy of the cloudlet does not change due to the supply from the environment ($\frac{de_{\text{cl}}}{dt} = 0$). This condition would be sustained if there is an adequate amount of heating (e.g., from galactic feedback processes) to balance the energy loss of the hot bath.

The evolution of the cloudlet density thus follows this simple equation:

$$\frac{dn_{\text{cl}}}{dt} = \frac{n_{\text{cl}}^3}{P_{\text{cl}}}(\Lambda - H) \quad (5.5)$$

Formally, the cloudlet reaches its final state when it can no longer grow in density ($dn/dt = 0$). Equivalently, the equilibrium state is when the radiative loss approaches the photo-heating (i.e., $\mathcal{L} = \Lambda - H \rightarrow 0$). As mentioned above, the subsequent growth in density is negligible after the cloudlet reaches a critical temperature since the cooling time grows steeply for $T < T_{\text{cl,f}}$. Therefore, we simply can approximate the final size using Eq. (5.2) at $T_{\text{cl,f}}$. In Figure 5.3, we show $c_s t_{\text{cool}}$ as a function of temperature and ambient pressure. For the chosen metallicity at $Z = 0.3Z_{\odot}$, the minimum size near the end of this evolutionary stage is at $T_{\text{cl,f}} \approx 10^{4.3}$ K.

Note the uptick of the size, in Figure 5.3, using $c_s t_{\text{cool}}$ for $T < T_{\text{cl,f}}$ no longer applies as t_{cool} departs from t_{cross} . Obviously, there are no reasons for a cloudlet to expand so that its crossing time can continue matching its cooling time. In reality, the cloudlet has, by definition, a size of $r_{\text{cl}} = c_s(T)t_{\text{cross}}(T)$ as it cools. In a sense, we are exploiting the fact that the background hot gas attempts to reach pressure equilibrium in the shortest cooling time of a cloudlet by using Eq. (5.3), which allows one to independently approximate $t_{\text{cross,min}}$ with $t_{\text{cool,min}}$.

In Figure 5.4, we show the evolution of thermodynamic properties of cloudlets by directly solving Eq. (5.5) given a background pressure and a fixed metallicity at $Z = 0.3Z_{\odot}$. As shown, the cooling time (and thus $c_s t_{\text{cool}}$) hovers at nearly the same value up to tens of

percent of the initial cooling time $t_{\text{cool},0}$ when the density of the gas is small. As soon as the density grows large enough, the bulk of cooling happens very quickly as cooling time drops precipitously. However, this stage of cooling does not last very long because the heating rate, which depends only on temperature, quickly approaches cooling rate at low temperature.

As shown in Figure 5.4, it is also clear that higher background gas pressure will lead to denser, colder and smaller cloudlets. The final size is independent of the initial temperature and density of the cloudlet. It is also worth noting that the cloudlet properties reach the same equilibrium very quickly at about a cooling time $t_{\text{cool},0}$ evaluated at the initial state, even if the initial conditions are changed.

This model due to thermal fragmentation and/or isobaric cooling provides an interesting prediction: the column density of the cloudlet has a characteristic value at $N_{\text{H}} \approx 10^{17} \text{cm}^{-2}$ (see also Voit & Donahue, 1990; McCourt et al., 2018). The minimum cloudlet size is $r_{\text{cl}} = c_s t_{\text{cool}} \propto f(T)/n_{\text{cl}}$, where $f(T)$ is the combination of the temperature dependence from the sound speed and cooling function. The column density is $N_{\text{H}} \approx n_{\text{cl}} r_{\text{cl}} \propto f(T)$. Therefore, as long as the temperature of the final state does not change very much with metallicity, the resulting column density will not differ dramatically. Using a realistic cooling curve at $Z = 0.3Z_{\odot}$, we show that lines correspond to different pressures collapse onto nearly the same region, with minima between $10^{17} - 10^{17.5} \text{cm}^{-2}$ (Figure 5.3). Note that the column density in Figure 5.3 is calculated with the cloudlet diameter instead, $N_{\text{H}} = 2n_{\text{cl}}r_{\text{cl}}$, to compare cross section of cold gas in simulations in the section 3.

5.3 Simulations

In this section, we use a suite 10 two-dimensional MHD simulations to study an episode of birth and death of cold gas in the presence of hot winds in the circumgalactic context. We focus on the spatial scale of cold gas from a combination of thermal instabilities and the shock-cloud interactions. We also vary input physical processes systematically, such as cooling, (an)isotropic conduction, and magnetic fields, to study how they affect the resulting

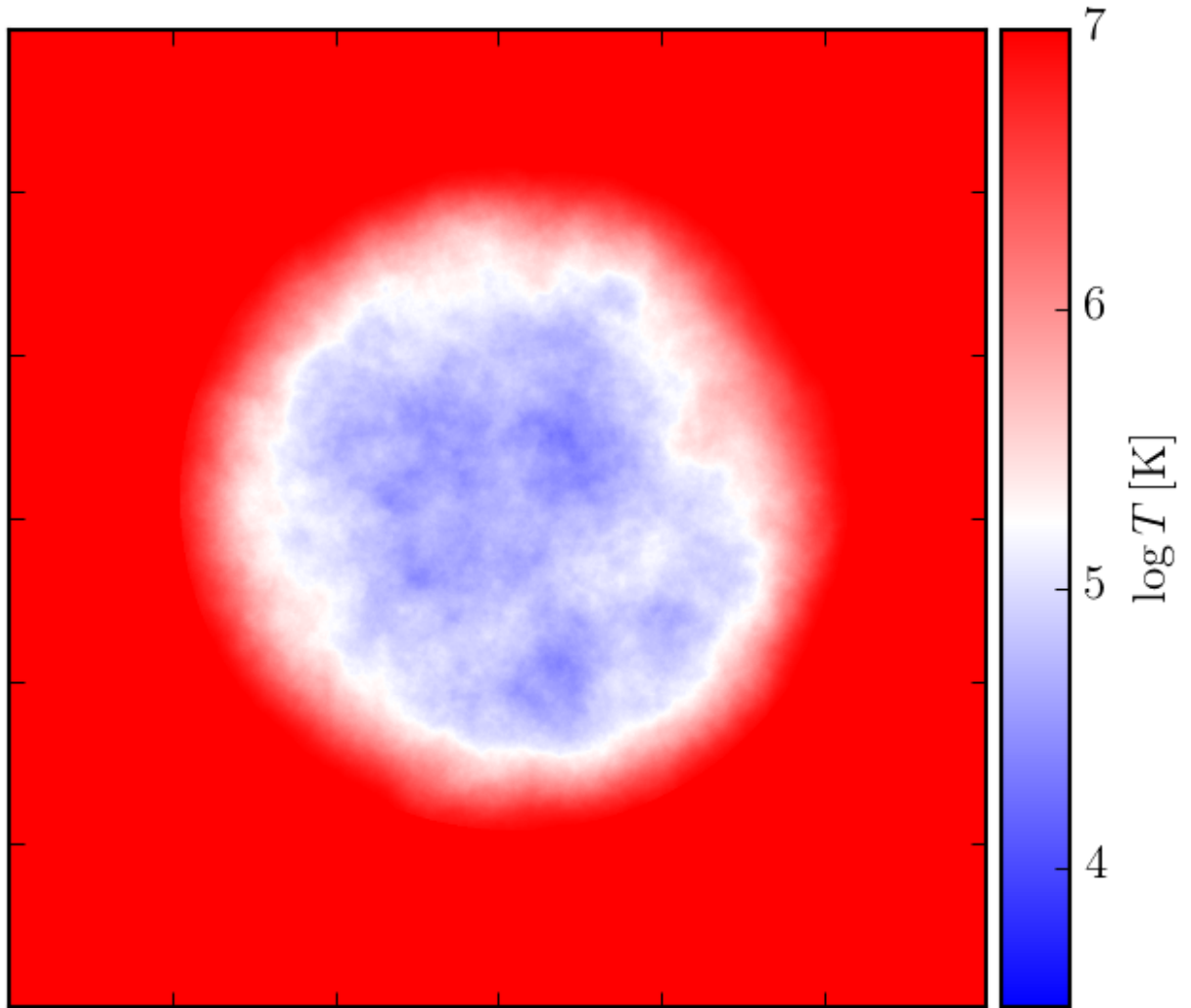


Figure 5.5: Initial condition of the density region with isobaric density perturbation with radius $\sim 150\text{pc}$. The region has a median temperature of $\sim 10^5\text{K}$ and is tapered by a gaussian on the edges that smoothly transitions into the background temperature at $T_h = 10^7\text{K}$ and density $n_h = 10^{-4}\text{cm}^{-3}$.

fragments of cold gas. We check our theoretical expectation of cloudlets by a comparison with their column densities across simulations.

5.3.1 *Physics and initial conditions in simulations*

We use the High-Speed Combustion and Detonation (HSCD) code (Khokhlov, 1998) for our simulations. The implementations of radiative cooling and adaptive mesh refinement are identical to those in the Adaptive Refinement Tree code (ART; Kravtsov et al., 1997), which HSCD shares a common ancestor. We solve a set of compressible ideal MHD equations using a third order finite difference, weighted essentially non-oscillatory (WENO) scheme with magnetic field divergence cleaning (Dedner et al., 2002; Mignone et al., 2010). We also implement isotropic and anisotropic thermal conduction following the first-order super-time-stepping scheme as described in Meyer et al. (2014).

The effects of thermal conduction, radiative cooling and heating are described in the energy equation:

$$\frac{\partial e}{\partial t} + \nabla \cdot \left(e + P + \frac{|\mathbf{B}|^2}{8\pi} \right) \mathbf{v} - (\mathbf{v} \cdot \mathbf{B}) \frac{\mathbf{B}}{4\pi} = n^2(\Gamma - \Lambda) + \nabla \cdot \mathbf{F}_c \quad (5.6)$$

Γ and Λ are the heating and cooling functions. We adopt a nominal value of $Z = 0.3Z_\odot$ for the cooling function for all simulations. The heating and cooling functions are shown in Figure 5.1 (Gnedin & Hollon, 2012). $\nabla \cdot \mathbf{F}_c$ is the net heat flux contribution from thermal conduction that smoothly transitions from the classical heat flux to the saturated regime:

$$\mathbf{F}_c = \frac{F_{\text{sat}}}{F_{\text{sat}} + |\mathbf{F}_{\text{class}}|} \mathbf{F}_{\text{class}} \quad (5.7)$$

where F_{sat} is the saturated heat flux when the mean free path of electrons is large compare to the local temperature scale length such that the transfer of heat is less efficient. Following Cowie & McKee (1977), we compute F_{sat} as

$$F_{\text{sat}} = 5\phi\rho c_s^3 \quad (5.8)$$

where we have chosen $\phi = 0.3$ as recommended by Balbus & McKee (1982). The second term $\mathbf{F}_{\text{class}}$ is the classical heat flux that depends on the temperature gradients and magnetic fields. We compute the heat flux in its anisotropic form using the Spitzer formula (Spitzer, 1962; Braginskii, 1965):

$$\mathbf{F}_{\text{class}} = \kappa_{\parallel} \mathbf{b}(\mathbf{b} \cdot \nabla T) + \kappa_{\perp} [\nabla T - \mathbf{b}(\mathbf{b} \cdot \nabla T)] \quad (5.9)$$

where $\mathbf{b} = \mathbf{B}/|\mathbf{B}|$ is the direction of the magnetic fields. The coefficients $\kappa_{\parallel}$ and $\kappa_{\perp}$ are the thermal conductivities parallel and perpendicular to the fields (Spitzer, 1962). For completeness, the parallel conductivity is

$$\kappa_{\parallel} = \varepsilon \times 20 \left(\frac{2}{\pi} \right)^{3/2} \frac{(k_B T)^{5/2} k_B}{m_e^{1/2} e^4 \ln \Psi} \quad (5.10)$$

where ε is a coefficient of order unity (≈ 0.4) that reduces the Spitzer conductivity, $\kappa_{\parallel} = \varepsilon \kappa_{\text{SP}}$, to an effective conductivity due to thermoelectric effects. In addition, $\ln \Psi$ is the coulomb logarithm:

$$\ln \Psi = 29.7 + \ln \left(\frac{T/10^6 \text{K}}{\sqrt{n/\text{cm}^{-3}}} \right) \quad (5.11)$$

and the perpendicular conductivity has a smaller magnitude, and is defined as

$$\kappa_{\perp} = \frac{8(\pi m_H k_B)^{1/2} n_H^2 e^2 c^2 \ln \Psi}{3|\mathbf{B}|^2 T^{1/2}} \quad (5.12)$$

To study the resulting sizes of cold clumps in a more realistic setting, we are interested in a region with varying density structure instead of a single spherical cold cloud. To do so, we first initialize the entire simulation domain in pressure equilibrium at the hot gas density and temperature. We then add density perturbations (or equivalently entropy perturbations) that follow a random Gaussian field modulated by a power spectrum $P(k) \propto k^{-3}$. We then

Table 5.1: Simulation suite

Simulations	cooling	conduction	magnetic field $\beta = P_{\text{th}}/P_B$	Δx_{res} (pc)
$\beta_{\text{ad},\infty}$			∞	0.37
$\beta_{\text{fid},\infty}$	✓		∞	0.37
β_{∞}	✓		∞	0.73
β_{1000}	✓		1000	0.73
β_{100}	✓		100	0.73
β_{10}	✓		10	0.73
$\beta_{c,\infty}$	✓	✓	∞	0.37
$\beta_{c,100}$	✓	✓	100	1.46
$\beta_{c,10}$	✓	✓	10	1.46
$\beta_{c,1}$	✓	✓	1	1.46

All simulations have the same initial condition with hot bath temperature at $T_h = 10^7$ K. The initial cloud is set up with $T \approx 10^{4.3} - 10^{6.5}$ K. Magnetic fields are initialized parallel to the hot winds with velocity $v_h \approx 1.4c_s$.

truncate the highest density region following McCourt et al. (2018) to avoid seeding of small-scale structures due to rapid cooling via its sensitive dependence with density $\dot{e} \propto n^2$. With small cold clumps emerging from thermal instabilities in many similar studies (e.g., Audit & Hennebelle, 2005; Cooper et al., 2009; McCourt et al., 2018), we do not expect our results to be strongly dependent on the details of the initial conditions.

To embed a fractal cloud in the hot bath, we select a spherical region of the perturbed density field with a radius of 150 pc tapered by a Gaussian (Figure 5.5). This fractal cloud is initially warm with a range of temperature from $\sim 10^{4.3} - 10^{6.5}$ K with a median temperature at $\sim 10^5$ K. We have also adopted a metallicity of the entire simulation domain at $Z = 0.3Z_{\odot}$, which provides a maximum cooling rate near the median temperature. The hot wind is initialized with velocity $v_h \approx 1.4c_s$, where c_s is the sound speed of the ambient gas at 10^7 K. The high-temperature hot bath at 10^7 K cools inefficiently at its density $n_{\text{H}} = 10^{-4}\text{cm}^{-3}$, which allows us to separate the cooling timescale of the targeted warm gas that will undergo thermal instability. A variation of the hot bath temperature at 10^6 K does not qualitatively change our results.

Table 5.1 lists the suite of simulations in this study. All of the simulations have the same initial condition as described above besides the variations of the input physical processes,

such as cooling, thermal conduction, and magnetic fields. All simulations are labeled with $\beta = P_{\text{th}}/P_{\text{B}}$ as a parameterization for the initial strength of the magnetic field, ranging from $\beta = \infty$ (no magnetic fields) to $\beta = 1$. For a hot bath of 10^7 K, $\beta = 1$ corresponds to a magnetic field strength of $\sim 2\mu\text{G}$, which is typical of observations in star formation galaxies (see Beck, 2015). The subscript “ad” indicates an adiabatic simulation without cooling ($\beta_{\text{ad},\infty}$); and the subscript “c” indicates thermal conduction is included in the simulations. The initial strength of the magnetic field is also subscripted for convenience. The initial direction of the magnetic fields is parallel with the hot winds for all of the simulations.

5.3.2 *The evolution of cold gas in simulations*

We begin this section with a qualitative description of the evolution of cold gas in our simulations. In particular, we discuss and compare with the relevant literature and focus on the destruction and fragmentation process and the resulting cloudlet sizes.

Adiabat: scale free evolution

We begin with the adiabatic simulation ($\beta_{\text{ad},\infty}$). Early numerical studies of the shock-cloud interactions have mainly focused on the competition of acceleration and destruction of spherical clouds. It is well known that the evolution is self-similar in purely hydrodynamic systems where the problem depends only on the density contrast after scaling out the Mach number (e.g., Klein et al., 1994; Nakamura et al., 2006). It is also well known that the Kelvin-Helmholtz and Rayleigh-Taylor instabilities disrupt the cloud in just a few cloud crushing times, where it is defined as:

$$t_{\text{cc}} = \chi^{1/2} \frac{r_{\text{cl}}}{v_h} = \left(\frac{n_{\text{cl}}}{n_h} \right)^{1/2} \frac{r_{\text{cl}}}{v_h} \quad (5.13)$$

where n_h and v_h is the density and velocity of the hot winds. The adiabatic simulation can be thought as a baseline comparison for our study of the mist-like gas. From Figure 5.7, it

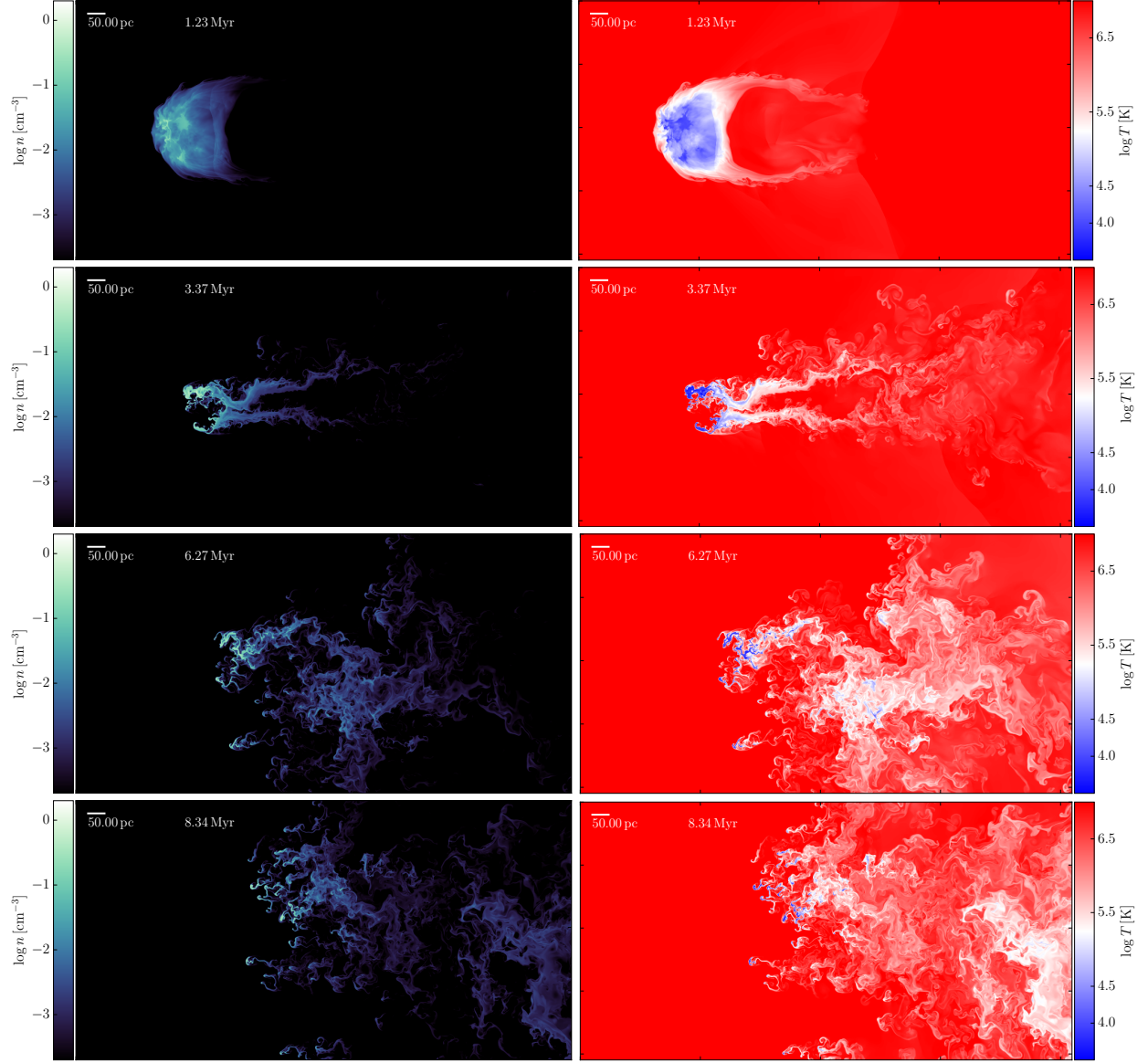


Figure 5.6: The evolution of the density (left column) and the temperature (right column) in the fiducial simulation with cooling only ($\beta_{\text{fid}} = \infty$). As shown, the initial warm gas cools quickly to $T \sim 10^4 \text{K}$. The cold gas then fragments into pc-size cloudlets through internal dynamics due to cooling and the interaction with the hot winds.

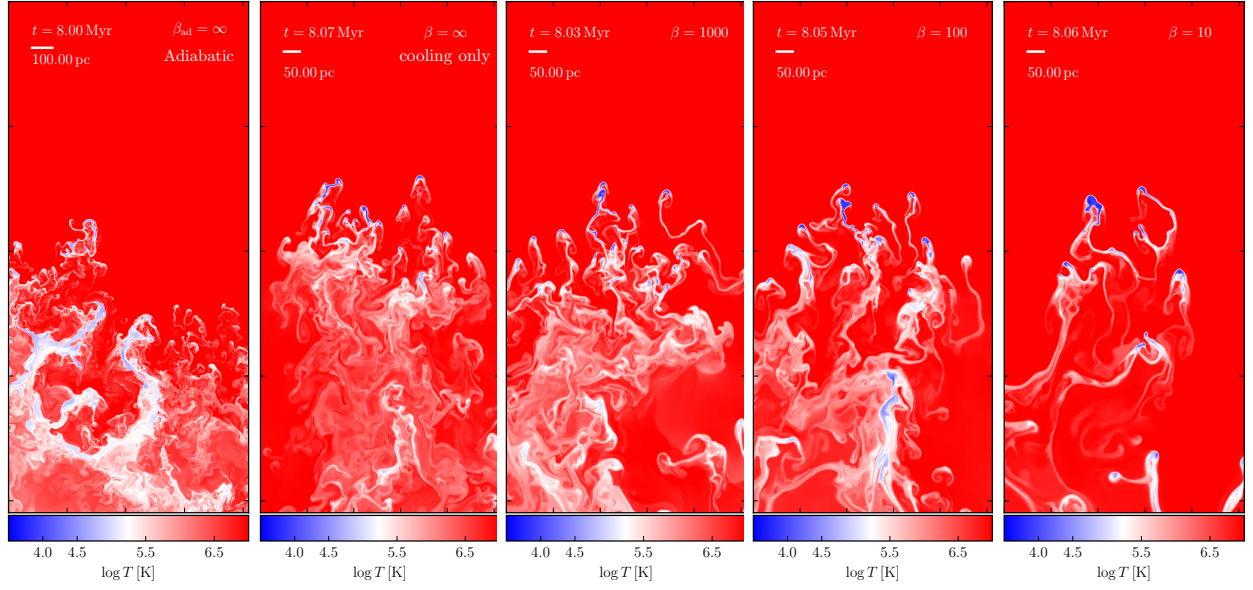


Figure 5.7: Temperature maps of simulations with the same initial conditions except the strength of the initial magnetic fields ranging from $\beta = 10 - \infty$. These simulations do not include thermal conduction. The left panel shows the adiabatic run where no cold clumps are present. This is in stark contrast to those with radiative cooling, where they produce similar cold clumps with $N_{\text{H}} \sim 10^{17.5} \text{cm}^{-2}$. Larger magnetic field strengths produce more filamentary structure (e.g., $\beta = 10$ in the right panel).

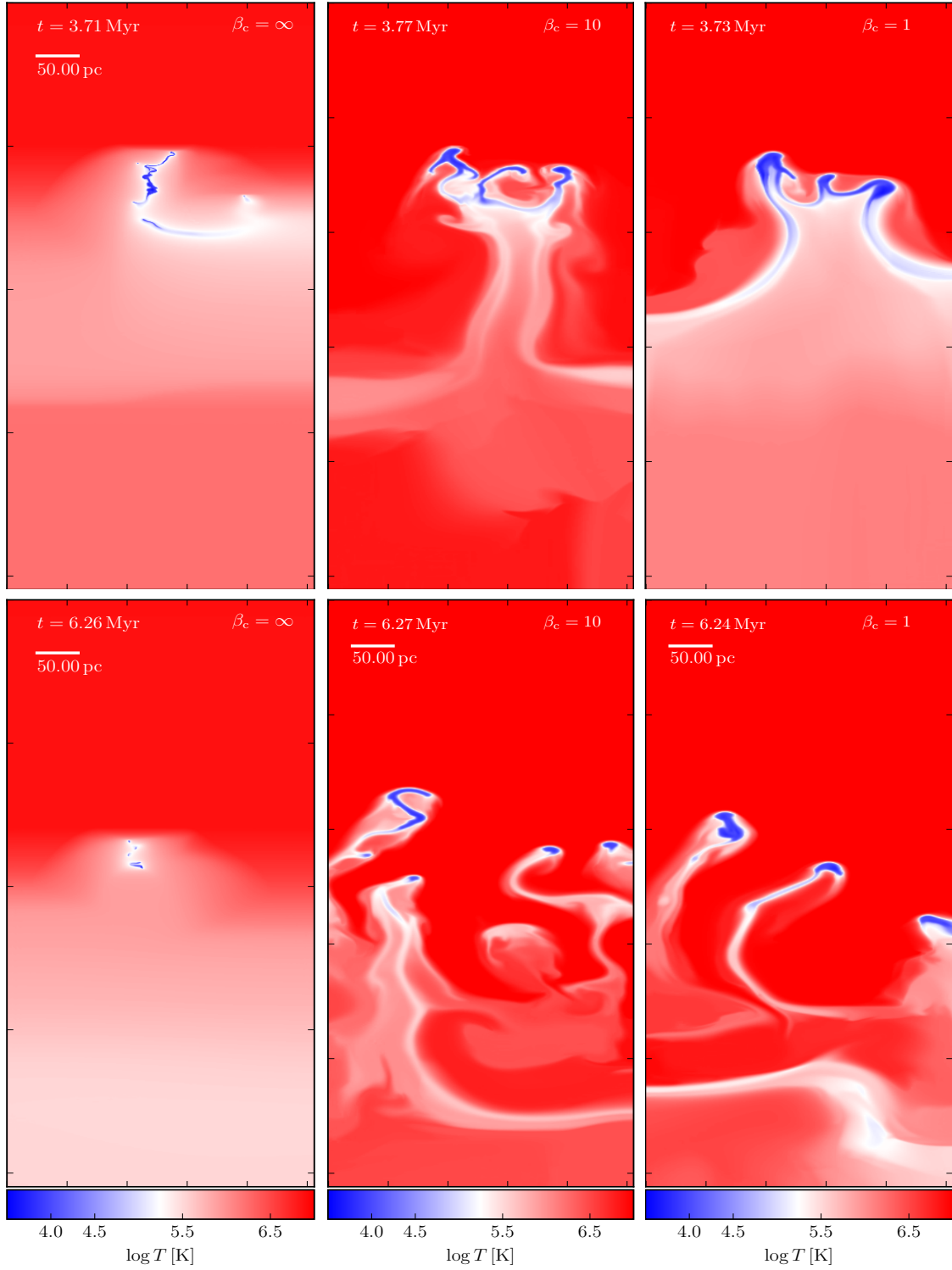


Figure 5.8: Simulations with thermal conduction at two representative time steps. The left panel ($\beta_c = \infty$) shows a run with cooling and isotropic thermal conduction. The middle and right columns show the evolution with an increased magnetic field strengths, which prolong the survival of cold gas.

is clear that the purely hydrodynamical case does not exhibit any characteristic cold and dense cloudlets compared to all the other simulations that include radiative cooling.

Radiative cooling: cold clumps from thermal instability

The effects of radiative cooling has also been analyzed in many papers in the context of entrainment of cold gas by hot winds (e.g., Mellema et al., 2002; Melioli et al., 2005; Cooper et al., 2008; Scannapieco & Brüggen, 2015; Schneider et al., 2018) and formation of multiphase ISM via thermal instabilities (e.g., Kritsuk & Norman, 2002; Audit & Hennebelle, 2005).

As radiative cooling is introduced to the simulations, it sets a qualitative different path for the evolution of the fractal cloud in our simulations. Most notably, a large fraction of cold gas is condensed out of the warm medium very quickly before the shock has a chance to break apart and penetrate through the more diffuse warm/hot gas (top left panel in Figure 5.6). After some time, roughly two large clumps of cold gas formed from the initial density perturbations. Meanwhile, a long tail of warm gas follows as the gas is stripped by the hot winds. Eventually, a combination of internal dynamics due to the lost of thermal pressure and the hot winds are able to break apart the cold clumps through the path of least resistance, which is regions of slightly lower density (see also Cooper et al., 2009; Schneider & Robertson, 2017). Similar to results in our paper, fragmentation via thermal instabilities allow small clumps of cold gas to survive longer (Mellema et al., 2002; Fragile et al., 2004; McCourt et al., 2018), though they eventually mix with the ambient hot gas. We find similar conclusion as Cooper et al. (2009), where the number of cold clumps increases with increasing resolution of their simulations.

Cooling and thermal conduction: elongation and evaporation

The evolution changes very drastically as isotropic thermal conduction is added to the simulation ($\beta_{c,\infty}$; see left column in Figure 5.8). Initially, the warm gas cools as in the simulations

with cooling only. At the same time, a conduction front propagates from outside in due to temperature gradients with the hot bath. At this point, most of the density and temperature fluctuations are erased except the densest cold gas. Although our initial condition of the cloud morphology is drastically different (fractal cloud), the morphology of cold gas at late times is similar to the findings in Brüggen & Scannapieco (2016). That is, the cold gas is shaped into an elongated filament by a combination of thermal conduction and the hot winds (see Figure 5.8). This is because any cold gas structures perpendicular to winds are most susceptible to complete destruction. Thus, the cold gas that is shielded from behind survives for the longest. Naturally, this leads to a minimization of the cross-section area of cold gas with the hot winds, leaving an elongated column of materials at late times (see also Figure 5 in Brüggen & Scannapieco, 2016). This minimization of cross-sectional area and the high column of mass makes it more difficult to accelerate by ram pressure of the hot winds since the acceleration $a_{\text{ram}} \propto \frac{A}{m_{\text{cl}}}$. Furthermore, this column of cold gas evaporates in just 7 Myr, much more quickly than its counterpart without conduction ($\gtrsim 10$ Myr with cooling only). Ubiquitous detection of cold gas in galactic winds and halos suggests that thermal conduction may be suppressed.

Cooling, magnetic fields, and conduction: filamentary cold gas

It is natural to include magnetic field in this problem since it not only prolongs the lifetimes of cold gas to help solve the entrainment problem (e.g., Mac Low et al., 1994; McCourt et al., 2015; Zhang et al., 2017), it also shapes the morphology and size of the resulting cold gas from thermal instabilities (e.g., Shin et al., 2008). In this study, we run a set of simulations with a range of initial magnetic field strength. The weak magnetic field simulation ($\beta = 1000$) essentially show the same number of fragments and very similar evolution as the simulation with cooling only. Very quickly, the initial fields evolve and are draped around individual cloudlets as they form. The hot winds and draping of magnetic fields lead to a substantial increase in magnetic energy density around the cold gas. At

late times, the interface of cold and hot gas is stabilized by strong magnetic fields, which suppresses fragmentation and mixing (e.g., Shin et al., 2008). This also provides some protection for the cold gas against thermal conduction and preserves large temperature gradients. The cold gas generally survives for a longer period of time compared to simulations without magnetic fields. Investigating the competition between the destruction by hot winds and conduction, and the protection by magnetic fields is out of the scope of this paper and deserves its adequate attention; we will leave this to a future study.

As we increase the initial strength of the magnetic fields, its energy density around the cold gas increases large enough that they start to play a dynamical role (e.g., Dursi & Pfrommer, 2008). As shown in Figure 5.7 and Figure 5.8, the number of cloudlets decreases with increasing magnetic field strength since it becomes more difficult to break apart the cold gas with the protection of the fields. The morphology becomes more and more filamentary. The tail of warm gas also becomes less turbulent and more coherent.

5.3.3 *Column density of cloudlet comparisons in simulations*

To directly verify the model ansatz that the size of cloudlet follows $r_{\text{cl}} = \min(c_s t_{\text{cool}})$, it requires high resolution simulations to resolve the gas clumps at a scale of $\sim 0.1 - 1$ pc. We expect decreasing clump sizes and increasing density with increasing spatial resolution if they are not converged in simulations. In an ideal situation, one should run simulations with varying resolution and prove convergence of cloudlet sizes. Given the constraints of computational expense, it is difficult to do so directly. Instead, we consider the column density per cloudlet, $N_{\text{H}} \sim 2r_{\text{cl}}n_{\text{cl}}$, since the size and the density goes in the opposite direction with a change of the resolution. This is essentially equivalent to a statement of mass conservation. That is, the column density should also be insensitive to the resolutions in our simulations. In addition, the model predicts a characteristic value of column density at $N_{\text{H}} \sim 10^{17-17.5} \text{ cm}^{-2}$, which is insensitive to the background pressure. Therefore, we will measure column density per cloudlet in all of our simulations. Nevertheless, we run a convergence

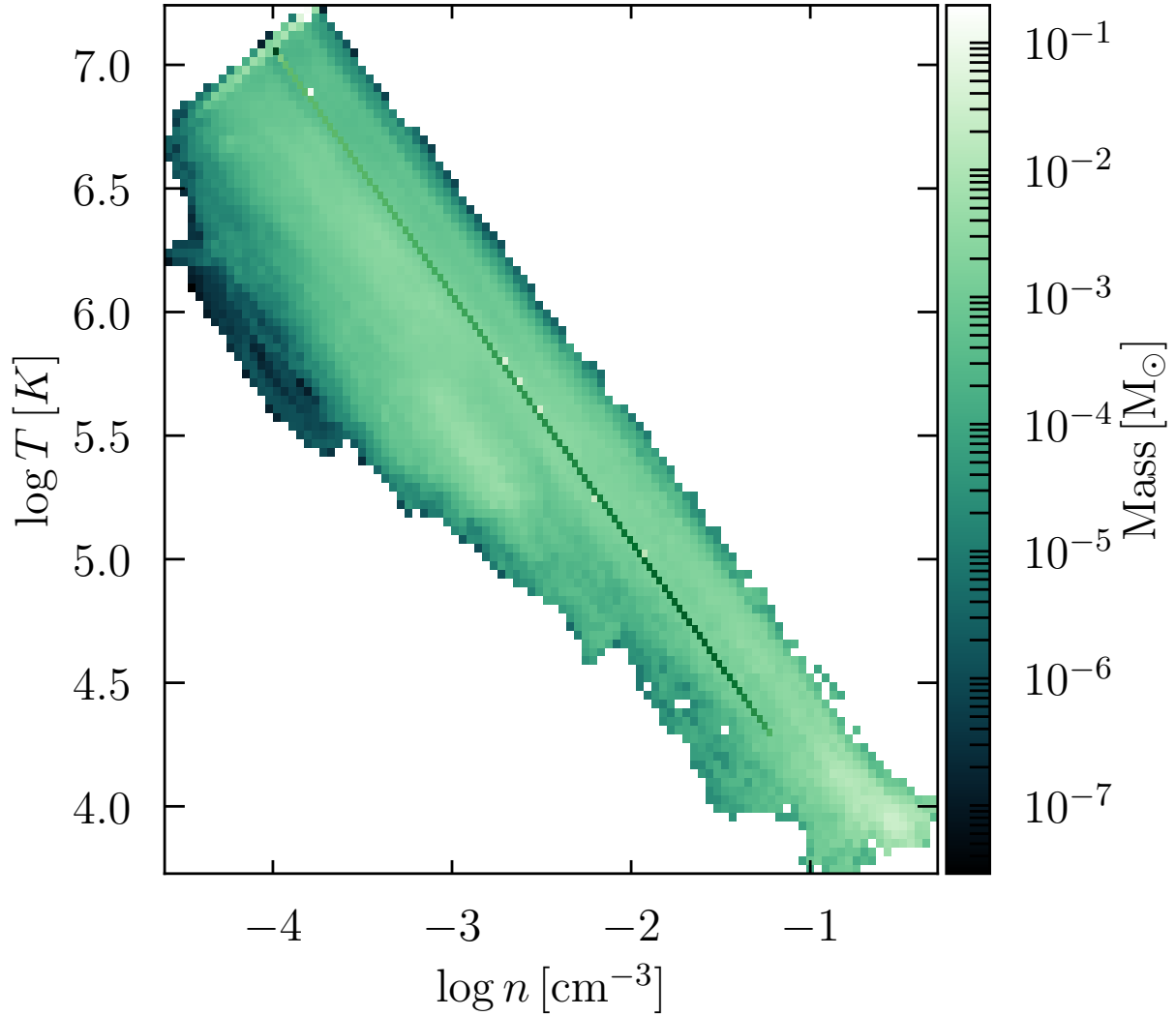


Figure 5.9: Temperature and density distribution of all the computational cells of the cooling only simulation. The dark green line shows the initial distribution where the simulation is initialized in pressure equilibrium. The broader distribution shows the gas maintains approximate pressure equilibrium with the ambient hot gas.

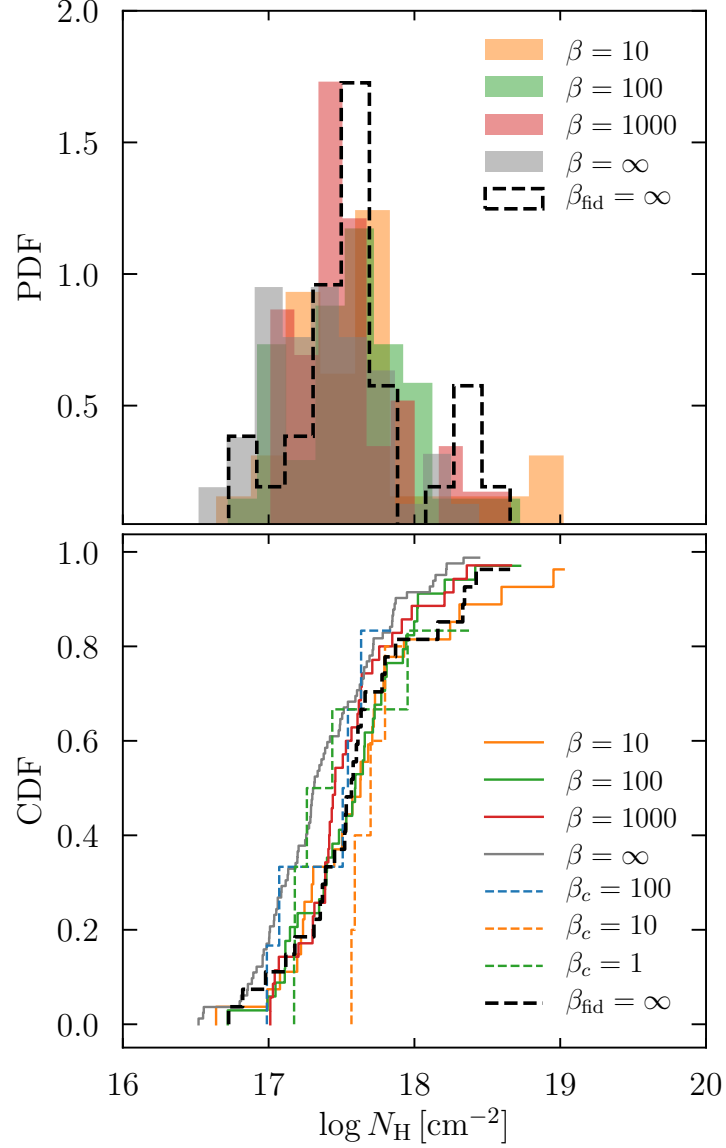


Figure 5.10: The top panel shows column densities distribution for simulations without thermal conduction. Simulations with conduction are better shown without binning in the cumulative distribution in the bottom panel, since there is only a small number of cloudlets. As shown, the medians of these distributions are all located near $N_{\text{H}} \approx 10^{17.5} \text{cm}^{-2}$, as predicted by the model. Simulations with large initial magnetic field strengths (e.g., $\beta_c = 1 - 10$) show the cloudlet column densities are slightly higher.

study of column density of cloudlets using cooling only simulations with two resolutions ($\Delta x = 0.37, 0.74$ pc) given the computational power constraints we have. We find that the column density does not change dramatically (see comparisons in Figure 5.10). While one can argue that convergence has not been satisfactorily demonstrated, for the purposes of our discussion of pc-scale cold gas in many kpc-scale gaseous halo, these differences are unimportant.

To measure the column density of a cloudlet, we first extract regions of cold gas using a density threshold of $n \geq 0.01 \text{ cm}^{-3}$. We find that the exact value of this density threshold does not change the result very much by varying its value from $10^{-3} - 10^{-2} \text{ cm}^{-3}$. After locating the cloudlets, it is easy to measure the column density by summing density over a line of sight. However, the morphology of cold gas is not always close to spherical. In fact, the cold gas is more filamentary in the presence of magnetic fields. The column density thus depends on the orientation of the lines of sight. Therefore, we consider the median of the column densities distribution by sampling all orientation as a representative value for a cloudlet.

Figure 5.10 shows the distribution of column densities of cloudlets from all of the simulations listed in Table 5.1. The top panel shows the probability density function (PDF) of simulations with cooling and magnetic fields. Simulations with conduction are not included in the PDF since there are only a small number of cloudlets ($\sim 3 - 5$). Instead, it is better shown in the cumulative distribution function (CDF) without the need of binning. The CDF also includes the other simulations without conduction for comparison. As shown, all column densities are distributed around $N_{\text{H}} \sim 10^{17.5} \text{ cm}^{-2}$. As mentioned above, strong magnetic fields alter the dynamics of cold gas. We see hints in our simulations that stronger fields lead to bigger cloudlets due to magnetic tension. This is reflected in the small shifts towards larger cloudlet column densities in simulations with large initial magnetic fields (i.e $\beta_c = 1$, see Figure 5.10). However, as Dursi & Pfrommer (2008) pointed out, the dynamical process results from magnetic draping is an inherently three-dimensional problem. A detailed ex-

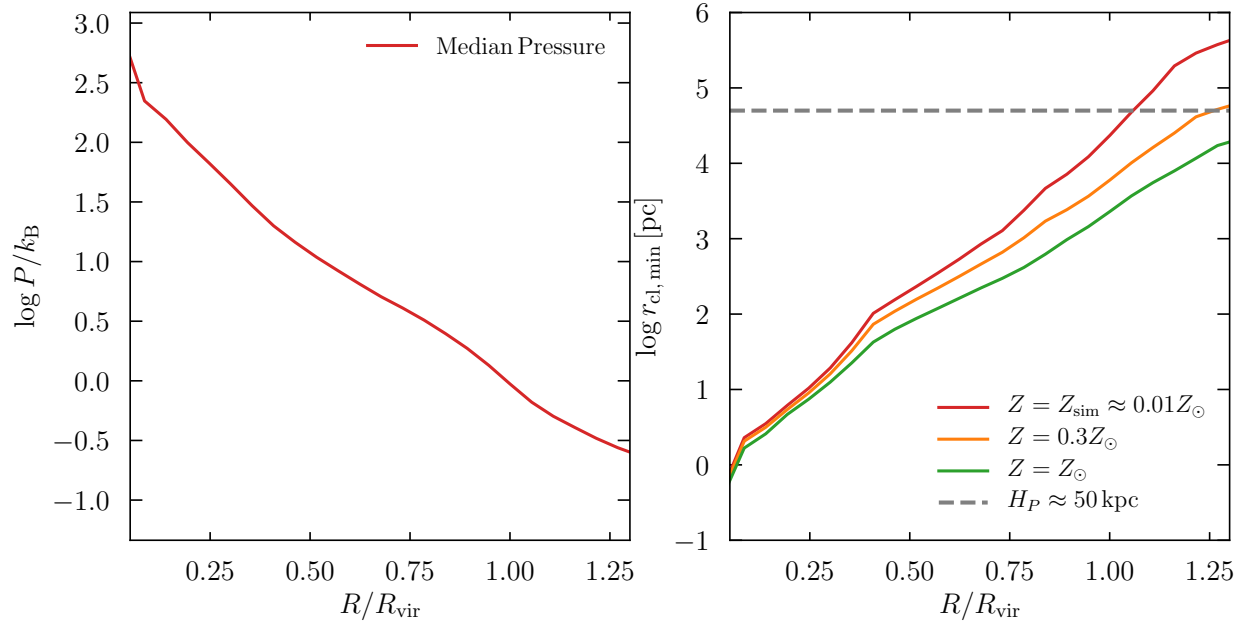


Figure 5.11: Left panel: simulation median pressure profile of L_* like galaxy at redshift $z = 0$ with $M_{\text{vir}} = 10^{12} M_{\odot}$ (Agertz & Kravtsov, 2015; Liang et al., 2016). Right panel: corresponding allowed cloudlet sizes, $r_{\text{cl,min}} = \min(c_{\text{st}} t_{\text{cool}})$, as a function of radius based on the pressure profile. Different color curves show the dependence of metallicity. The higher the metallicity, the more efficient the gas can cool, which leads to smaller cloudlet sizes. Note that the sharp boundary of detections of low and intermediate ions at $0.5 - 0.7 R_{\text{vir}}$ (Liang & Chen, 2014) corresponds to similar radii where small cloud sizes are small, in the order of 100 pc or less. In addition, the pressure profile on the left can be approximated by an exponential profile, $P/k_B \propto \exp[-(r/H_P)]$, where $r = R/R_{\text{vir}}$ and $H_P \approx 50 \text{ kpc}$ is pressure scaleheight. We expect the assumption of pressure equilibrium to derive the cloudlet size breaks down roughly at the size of the pressure scaleheight.

ploration of the dynamical importance of magnetic fields in this scale requires a parameter study in three-dimensional simulations in the future.

5.4 Implications of cloud sizes in the Circumgalactic context

We now turn our attention to discuss the implications of cloudlet sizes on the larger context, namely the global picture of the gaseous halos around galaxies.

5.4.1 *The circumgalactic mist in galactic halos*

To paint a physical picture of the circumgalactic mist, let us consider a parcel of hot halo gas with size L , density n_h , and temperature T_h . We are particularly interested in the multiphase nature of the gas as a function of these background properties. Let us assume that this parcel of hot gas is located at some galactocentric radius R , where its cooling time is short compared to the global free-fall time:

$$t_{\text{cool}}/t_{\text{ff}} < \xi \quad (5.14)$$

where ξ is a parameter that controls whether the parcel of gas can cool and condense. It has been shown halo gas can undergo thermal instability and precipitate at the mass scale of galaxy clusters if $\xi \lesssim 10$. (e.g., McCourt et al., 2012; Voit et al., 2015b). The cycles of condensation have been shown to be promoted by active galactic nuclei (AGN) feedback (e.g., Gaspari et al., 2013; Li et al., 2015; Meece et al., 2015). While there exists indirect evidence for this same precipitation limit for lower mass halos (Voit et al., 2015a), we nevertheless will treat this as a free parameter. Given a fixed ξ , we can find that a new hot gas density $n_{h,s}$, such that it is thermally stable. This can be achieved with the reverse condition, $t_{\text{cool}}/t_{\text{ff}} \geq \xi$, which implies¹,

$$n_{h,s} \leq \frac{3k_{\text{B}}T_h}{\xi\Lambda t_{\text{ff}}} \quad (5.15)$$

Thus, the gas left over from condensation has a long cooling time due to its relatively low density. Again, the global balance of energy accounting can be maintained and supported by feedback processes and photoionization. While we are interested in the cold gas out of the condensation, this hot gas density profile has its own implication. In fact, Voit et al. (2018) derived a hot density profile using $\xi \geq 10$ and showed X-ray luminosity L_X of halos obeys this precipitation limit across a large range of mass scales, from clusters, galaxies groups to

1. We have assumed equilibrium between electrons and the ions, $n_e = n_i = n_{h,s}$.

massive ellipticals (Cavagnolo et al., 2009; Sun et al., 2009; Werner et al., 2012, 2014).

Let us now consider the cooling phase of this parcel of gas. First, we assume some density perturbations $\delta n/n_h$ with scale less than L , such that a fraction of this parcel of gas can undergo thermal instability with fragmentation and contraction. At this stage, we are not concerned with how this portion of the mass distributes itself over a range of temperature over time. The mass of this cooling phase simply is the difference between the original total mass and the leftover stable hot gas:

$$M_{\text{cooling}} = L^3 \mu m_H (n_h - n_{h,s}) \quad (5.16)$$

where μ and m_H is the mean molecular weight and mass of the hydrogen atom, respectively. Note that $M_{\text{cooling}} = 0$ if the gas is already stable with $n_h - n_{h,s} \leq 0$, which is the case at large radii when $t_{\text{cool}} > 10t_{\text{ff}}$ (e.g., Voit et al., 2015a). The available cooling materials are limited in the inner regime, and provide a natural explanation for decreasing covering fraction of cold gas with increasing radius (Liang & Chen, 2014). However, this does not mean cloudlets can exist as long as $t_{\text{cool}}/t_{\text{ff}} < 10$ because the background pressure is important as well.

As in Spitzer (1956), we assume the cold cloudlets are pressure confined by the hot gas. One can then calculate the mass of an individual cloudlet, since $n_{\text{cl,f}}$ and r_{cl} are both functions of the ambient pressure (see section 5.2):

$$m_{\text{cl}} = \frac{4\pi}{3} r_{\text{cl,f}}^3 \mu m_H n_{\text{cl,f}} \quad (5.17)$$

In the inner region of the CGM, a large fraction of the cooling mass M_{cooling} would have reached $T_{\text{cl,f}} \sim 10^4\text{K}$ very quickly in the form of a cloudlet (section 5.2). In this simplified picture of spherical cloudlets, we can also estimate the total number, $\mathcal{N}_{\text{cl}}$, within this region of size L :

$$\mathcal{N}_{\text{cl}} = \frac{M_{\text{cooling}}}{m_{\text{cl}}} = \frac{3}{4\pi} \left(\frac{n_h - n_{h,s}}{n_{\text{cl,f}}} \right) \left(\frac{L}{r_{\text{cl}}} \right)^3 \quad (5.18)$$

It also useful to consider the number of cloudlets per volume, which is independent of size of the domain L ,

$$\eta_{\text{cl}} = \frac{\mathcal{N}_{\text{cl}}}{L^3} = \frac{3}{4\pi} \left(\frac{n_h - n_{h,s}}{n_{\text{cl},f}} \right) \left(\frac{1}{r_{\text{cl}}} \right)^3 \quad (5.19)$$

From this, we can also calculate the average spacing between cloudlets:

$$\langle d_{\text{cl}} \rangle = \eta_{\text{cl}}^{-1/3} \approx \left(\frac{n_{\text{cl},f}}{n_h - n_{h,s}} \right)^{1/3} r_{\text{cl}} \quad (5.20)$$

The number of cloudlets depends strongly on the size of the individual cloudlet. For the inner region of the CGM, as shown in Figure 5.11, the size of the cloudlets is in the order of $r_{\text{cl}} \sim 0.1$ to a few parsec. Given a region with $L = 1$ kpc within $\sim 0.25R_{\text{vir}}$, the ratio $\left(\frac{L}{r_{\text{cl}}} \right)^3 = 10^9 - 10^{12}$. If we take a typical value of $n_h \sim 10^{-4} \text{cm}^{-3}$ and $n_{\text{cl},f} \sim 10^{-1} \text{cm}^{-3}$, the ratio of the densities is roughly $(n_h - n_{h,s})/n_{\text{cl},f} \sim n_h/n_{\text{cl},f} \sim 10^{-3}$. With the assumption of spherical cloudlets, there exists at least in the order of $\sim 10^6 - 10^9$ clumps within a 1 kpc region! The average inter-cloudlet spacing is in the order of 30 cloudlet radii. In this picture, the cloudlets forms a kind of mist, resembling terrestrial clouds and mist.

McCourt et al. (2018) calculated the expected voluming filling and covering fractions of cloudlets (see their Figure 1). For completeness, we can compute these quantities now with an expression for the number density of cloudlets (Eq. 5.19). The volume filling factor is simply the ratio of the total volume occupied by the cloudlets over the volume of the region,

$$f_V = \mathcal{N}_{\text{cl}} \left(\frac{r_{\text{cl}}}{L} \right)^3 = \eta_{\text{cl}} r_{\text{cl}}^3 \approx \frac{n_h - n_{h,s}}{n_{\text{cl},f}} \quad (5.21)$$

Therefore, the volume filling factor of the mist is actually quite small at $\sim 0.1\%$ of the total volume L^3 in the inner CGM. However, this does not imply the covering fraction is small. To get an estimate, we can start by assuming cloudlets are spherical with a cross sectional area, $\sigma_{\text{cl}} = \pi r_{\text{cl}}^2$. If they do not shadow each other, the covering fraction integrated over some length scale, say L , is equivalent to the formula for optical depth:

$$f_A = \frac{\sigma_{\text{tot}}}{A} = \frac{\mathcal{N}_{\text{cl}}\sigma_{\text{cl}}}{A} = \frac{\eta_{\text{cl}}LA\sigma_{\text{cl}}}{A} = \eta_{\text{cl}}\sigma_{\text{cl}}L \quad (5.22)$$

where A is some area of interest for calculating the covering fraction. With the expression for cloudlet number density, η_{cl} , the covering fraction is:

$$f_A \approx \frac{n_h - n_{h,s}}{n_{\text{cl},f}} \frac{L}{r_{\text{cl}}} \quad (5.23)$$

Therefore, the covering fraction can quickly reach unity for $r_{\text{cl}} \sim 1$ pc over a length scale of 1 kpc. This is in concordant with observations of Mg II absorbers close to the galaxies where the covering fraction is very high (Nielsen et al., 2013; Chen et al., 2010a; Huang et al., 2016), even if the volume filling factor can be small (e.g., Hennawi et al., 2015). In fact, the large number of cloudlets may also affect the degree of resonant scattering with Ly α photons and have observable signatures in their emission spectra (e.g., Gronke et al., 2017). We also note the morphology of cold gas needs not be spherical. The assumption and simplification to sphericity provide a simple way to calculate a physical picture of the mist. In reality, the morphology can be more complicated. For example, the cold gas can be filamentary in the presence of magnetic fields as shown in this paper. This is especially true in the picture that magnetic field lines are draped around cold gas to prolong lifetimes of cold gas within galactic outflows and the inner halo.

5.4.2 Spatial extents of low, intermediate and high ions from the galactic profile of cloud sizes

In Spitzer (1956), it was proposed that the observed cold gas (Na I and Ca II; Münch & Zirin, 1961) must be pressured confined to explain their existence at large ($\sim$ few kpc) distances away from the galactic plane. Otherwise, the number density of such ions will decrease as they expand; the atoms will be ionized into a higher state. The absorption by these low ions will, therefore, be very slight. It is natural to consider what happens over

larger distances where the pressure profile does indeed drop exponentially. The CGm model naturally predicts the change in density and sizes of resulting cloudlets from the pressure profile.

The abundance of low, intermediate and high ions peak at a range of preferred density, or ionization parameter (U) given a fixed photoionization background. In other words, there will be more ionizing photons per atom (higher U) at larger distances. Therefore, if the gas is primarily photo-ionized, it will be elevated to higher and higher ionization states as the cloudlet size increases and its corresponding density drops (see similar discussion in Liang et al., 2016, 2017). This in turns provides predictions for the spatial extent of commonly observed low, intermediate and high ions.

The left panel of Figure 5.11 shows the pressure profile of hot gas from a cosmological “zoom-in” simulation of a Milky-Way size galaxy at $z = 0$ (Agertz & Kravtsov, 2015; Liang et al., 2016). The halo mass at this time is $M_h \approx 10^{12} M_\odot$ with a virial radius of $R_{\text{vir}} \approx 260 \text{ kpc}$. Given the description in section 5.2, we can calculate the corresponding *allowed* cloudlet sizes as a function of distance given the ambient pressure, as seen in the right panel of Figure 5.11. Since the final cloudlet size depends also on how effectively the gas can cool through the cooling function Λ , we show different color lines bracketing a range of sizes from metallicity effects, with $Z/Z_\odot = 0.01, 0.3$ and 1 . While in reality there exists pressure variation within the halo, the lines in the right panel of Figure 5.11 represents the cloud sizes from the median pressure at a given radius. In addition, the pressure profile can be approximated by an exponential profile, with $P/k_B \propto \exp[-(r/H_P)]$, where $r = R/R_{\text{vir}}$ and $H_P \approx 50 \text{ kpc}$ as the pressure scaleheight. This implies that the pressure equilibrium assumption, at the very least, breaks down above this scale, which puts a limit on the applicability of the model. This occurs near the virial radius of the halo for this particular L_* galaxy.

The cloudlet sizes r_{cl} and the densities (n_{cl} , not shown) span many orders of magnitude within the halo. It is interesting to point out the characteristic sizes at a few different radii.

First, the cloudlets are tiny with $\lesssim 1$ pc at small distances where the background pressure is the highest (see also fragment sizes in Audit & Hennebelle, 2005). Second, observed cloud sizes derived from cold/warm absorbing gas generally fall between tens to hundreds of parsecs (e.g., Rauch et al., 1999; Prochaska & Hennawi, 2009; Crighton et al., 2015; Stern et al., 2016; Lan & Fukugita, 2017). The ~ 100 pc scale of cloudlets corresponds to distances at $0.5R_{\text{vir}}$ or ~ 130 kpc. This is in concordant with the observed spatial extent of Mg II, Si II and C II (Chen et al., 2010a; Werk et al., 2013; Nielsen et al., 2013; Liang & Chen, 2014; Huang et al., 2016). In other words, not everywhere within the galactic halo can the mist exist (i.e., small, cold, and dense cloudlets).

Similarly, intermediate ions that trace warm gas preferred a range of slightly lower density and larger cloud sizes. This allows ions like C IV, and Si IV to be present at larger distances from the host galaxy. Indeed, these have been observed up to $\sim 0.7R_{\text{vir}}$ (Liang & Chen, 2014). These authors also showed differential covering fractions from low to high ionization state for a fixed specie (e.g., Si II, Si III and Si IV).

Beyond $\sim 0.7R_{\text{vir}}$, the cloudlet sizes increase to many kiloparsecs, approaching the size of the hot halo. In this limit, only high ions that trace the hot gas can survive. Photoionization may not play a big role for high ions since the ionization energy is high (e.g., $E_{\text{ion}} = 138$ eV for O VI). Instead, O VI tracing this component of the warm/hot gas is most likely collisionally ionized. However, the ionization mechanism of O VI is being debated from both observational and theoretical perspectives (e.g., Thom & Chen, 2008; Tripp et al., 2008; Werk et al., 2016; Faerman et al., 2017; Mathews & Prochaska, 2017; Liang et al., 2017; Stern et al., 2018; Nelson et al., 2018). Nevertheless, QSOs absorption-line observations show that O VI is detected up to R_{vir} (e.g., Johnson et al., 2015b). The abundance of O VI observed in the outskirt implies that the gas somehow sustains hot despite its high cooling rate at $10^{5.5}$ K. The question of how the hot gas remains in this state has to do with the global thermal balance from various heating mechanisms. As mentioned in the introduction, the subject of feedback processes and its regulation of star formation and the hot halo are

being investigated (e.g., Oppenheimer et al., 2016; Fielding et al., 2017).

5.5 Discussions and Conclusions

Throughout this paper, we have loosely used the terms, circumgalactic mist and cloudlet, to paint a general picture of cold gas in galactic outflows and the inner CGM owing to their similarities with condensation in thermal instabilities. This analogy with terrestrial clouds goes back many decades in the literature. In fact, condensation of cold gas has been explicitly compared with the van der Waals gas (Zel'Dovich & Pikel'Ner, 1969). However, it is worth pointing out that there are no natural boundaries in which one can easily define the discreteness of a cloudlet. While the underlying language that describes the physics of the CGM is mathematical (e.g., hydrodynamics and turbulence), the choice of words is only used to paraphrase the general properties of the cold gas. In reality, the model describes a region of the CGM with local density variations at which thermal instability can be excited within this compressible turbulence. Afterward, the competing timescales imprint a spatial scale ($c_s t_{\text{cool}}$) for cold gas. We show in this paper thermal instability can be triggered by a combination of supersonic galactic winds and initial density perturbations in a hot bath. Just as terrestrial clouds and mist can form when a region of hot and humid air is hit with a cold front, an episode of cold gas can form within a compressible turbulence induced by supersonic hot winds.

The abundant observations of cold gas in the inner region of the CGM suggest that it is a dominant evolutionary phase of the galaxies. In fact, the cold phase of the CGM is not only ubiquitous across cosmic times (e.g., Steidel et al., 2010; Chen, 2012; Werk et al., 2013; Liang & Chen, 2014), but also across galaxy types (e.g., Prochaska et al., 2014; Bordoloi et al., 2014; Huang et al., 2016; Heckman et al., 2017). The origin of this dynamic yet steady multiphase CGM is not yet fully understood. There are competing scenarios for the origins and implications of cold gas. For example, can cold gas be directly launched from the ISM into the halo by entrainment with the help of magnetic fields (McCourt et al., 2015;

Zhang et al., 2017)? Fragmentation may additionally prolong the lifetime of cold gas due to inefficient mixing and allow cloudlets to be more easily entrained. At the same time, cold gas can be formed directly within the galactic winds and halos via thermal instability as presented in this paper. A third option for the formation of cold gas can be mass-loaded winds through adiabatic cooling (see also Thompson et al., 2016; Schneider et al., 2018). Thus, it is worth exploring any observables, such as kinematics signatures, that could distinguish the different scenarios.

While the simulations in this paper represent an episode of the birth and death of cold gas, the analytic CGM model describes, in a sense, a steady state of the multiphase CGM. The spatial extent depends mainly on the ambient pressure profile. The abundance of cold halo gas can fluctuate over many episodes of this process; it can also depend on other processes such as AGN feedback in more massive systems, which fluctuates in the order of many millions of years (e.g., Li et al., 2015). However, comparison between observations of low ($z \approx 0$) and high redshifts ($z \approx 2$) show remarkable similarities over a much longer timescale (e.g., Steidel et al., 2010; Chen, 2012; Prochaska et al., 2013; Liang & Chen, 2014). Galaxy simulations that reproduced the observed low- z CGM also suggest that the CGM evolves very little over cosmic time once it is properly normalized in units of the dark matter halo scale radius r_s (Liang et al., 2016). The unchanging nature of the CGM over cosmic time is somewhat remarkable and puzzling because it implies that the spatial extent of cold gas depends little on the star formation rates of galaxies where it is an order of magnitude higher in the cosmic high noon at $z \approx 2$ compared to today. To understand this, we must model self-consistently in our galaxy simulations the processes of star formation, feedback and thermal instabilities that lead to a multiphase CGM over cosmic time.

It is clear that the current generation of galaxy simulations do not resolve the million of gas clumps within $\lesssim 0.5R_{\text{vir}}$ with a typical resolution of $\Delta x_{\text{res}} \sim 1$ kpc in the halo. This is because current state-of-the-art simulations focus on their adaptive refinement criteria at the highest density regime. The common underproduction of cold halo gas by cosmological

galaxy simulations may be due to the poor resolution in the halo. The profiles of cloud sizes in Figure 5.11 and Eq. 5.19 suggest that future galaxy simulation should address the mismatch of cloud sizes and their spatial resolutions, in order to correctly predict the clumping factor and consequently cooling of the halo gas. It implies that the resolution in the halo within $\sim 0.5R_{\text{vir}}$ must be in the order of 100 pc or less to match the observed abundance of cold gas (e.g., Chen et al., 2010a; Nielsen et al., 2013; Werk et al., 2013; Liang & Chen, 2014; Huang et al., 2016). An alternative is to employ some types of subgrid models that incorporate the process of fragmentation and thermal instabilities, especially for galaxy simulations with large cosmological volume (e.g., Vogelsberger et al., 2014; Schaye et al., 2015). Resolving this process may also introduce stronger coupling between the multiphase CGM and the central star-forming sites and lead to a better understanding of their symbiotic relationship.

In summary, we have presented a new multiphase model of the CGM, the circumgalactic mist. Our conclusions can be summarized as follows:

1. We considered the process of fragmentation and isobaric condensation modes in thermal instability and showed that the resulting cloudlet sizes are very small ($\sim \text{pc}$) in the galactic hot halo environment. We provided a semi-analytical prescription to derive the cloudlet sizes entirely based on background hot halo properties (section 5.2).
2. Using a set of MHD simulations, we discuss the process of fragmentation and the birth of cold gas in general. We show that the resulting cold gas conforms to the characteristic column density of $N_{\text{H}} \approx 10^{17} \text{cm}^{-2}$ as predicted by the $c_s t_{\text{cool}}$ ansatz (section 5.3).
3. Despite the initial setup of uniform and parallel magnetic fields, the variation of density structures within the initial fractal cloud lead to field lines tangled and draped around the later-formed cloudlets. Variation of magnetic field strength in the simulations shows magnetic fields alter the morphology of the cold gas. It can suppress thermal conduction and mixing with the hot winds, preserving strong temperature gradients.

4. We present a new multiphase model of the CGM based on the implications of small cloudlet sizes. We estimate the number of cloudlets in the inner region of the CGM is in the order of $\sim 10^{6-9}$ with an average inter-cloudlet spacing of ~ 30 cloudlet radii. Although the volume filling factor is small (10^{-3}), the covering fraction can reach unity if integrated over only a short $\sim \text{kpc}$ length scale. This provides a natural explanation for the ubiquitous detection of cold gas in the inner galactic halo.
5. We also discuss the spatial extent of cold, warm and hot gas as implied by the model. Regardless of the onset of thermal instability, the model naturally predicts that cold cloudlets can only exist up to ~ 100 kpc as the condensates are compressed to a density regime where it is favored by low ions (e.g., Mg II, Si II, C II). For larger distances, the pressure of the hot gas is relatively lower. Intermediate (e.g., C IV) and high ions (O VI) will, therefore, still thrive in these conditions (i.e., lower density and higher temperature).

The model as described in this paper represents a multiphase state of the CGM. The inner region is filled with cold, dense, and small gas clouds, resembling terrestrial mist or clouds. This stretches the relevant scales from 0.1 pc to hundreds of kpc in the CGM. The challenge of galaxy formation and evolution remains on the self-regulation of heating and cooling. This regulation and interplay between feedback and cooling of the hot halo gas can determine whether a galaxy will be star-forming or quenched and remains quiescent over a long timescale. The state of the galaxy should also be understood via the steady state of the CGM. Self-consistent galaxy simulations must incorporate the process of cold gas fragmentation with the interplay of star formation and feedback processes.

CHAPTER 6

SUMMARY AND CONCLUDING REMARKS

6.1 Summary of results

The central theme of this dissertation focuses on the spatial distribution of multiphase gas around galaxies. Chapter 2 presents the empirical mapping of the multiphase gas using absorption lines in spectra of background QSOs. I detected neutral hydrogen up to 500 kpc, well beyond the virial radius of the host dark matter halos. In contrast, heavy elements, signatures of enrichment, were detected only up a fraction of the virial radius. In addition, the differential spatial extent of cold and warm gas is observed. In Chapter 3, I use a set of cosmological zoom-in simulations of galaxies to compare with empirical constraints. The simulation analyses along with the comparison study show that the CGM profiles of ions are self-similar in mass for a given cosmic epoch when scaled with the virial radius. Furthermore, CGM profiles are similar across cosmic time when scaled in units of r_s . The simulations predict column density profiles are exponential in form, $N_{\text{ion}} \propto \exp(-d/h_s)$. The scale height h_s characterizes the spatial extent of a given ion. The differential spatial extent of photoionized gas can be summarized in a power law relation: $h_s \propto E^\alpha$. Chapter 4 describes a synthetic absorption spectrum pipeline, as a virtual telescope, to “observe” simulated galaxies and CGM. Finally, in Chapter 5, I ran a set of MHD simulations on the formation of cold cloudlets in presence of hot winds with varying physical processes, such as radiative cooling, thermal conduction, and magnetic fields. I show that thermal instability can lead to fragmented cold clumps, forming a multiphase medium. Magnetic field draping around these cloudlets help prolong their lifetimes and suppress mixing with the ambient hot gas. These cold clumps are expected in the order of $\sim \text{pc}$ in the inner region of the halo ($\lesssim 0.5R_{\text{vir}}$). I provide a new model of the circumgalactic medium based on this idea that cold gas seen in QSOs absorption lines are made of small cloudlets. It correctly predicts the spatial distribution of cold, warm and hot gas around galaxies based on the allowed cloud

sizes and gas density at a given radius.

6.2 Future outlook

The future of gaseous halos and its connection with galaxies remain an exciting field. New and exciting data will be coming in the next few years as new telescopes are coming online. These include space-based missions, such as the *James Webb Space Telescope*, and ground-based 30m class telescopes, like the Giant Magellan Telescope. The combination of these telescopes will yield many-folds of the current sample of galaxy-QSOs pairs. Observational constraints will no longer be plagued by small number statistics. Integral field spectrograph on these large telescopes will allow detailed studies. This will allow one to find the additional dependence of the CGM besides the host halo mass. For example, does it matter if the galaxy is or has just undergone a starburst episode? How does the star-forming rate affect the CGM? How would that change the state of the CGM? We have come a long way since 1956 when Lyman Spitzer first postulated the existence of a hot halo. Additional empirical data will be needed in the future to further polish our physical model of the gaseous halos and its connection with the host galaxies across cosmic time.

The natural extension of the work presented here is to explore dynamically the origin and fate of the circumgalactic mist. Circumstantial evidence presented in Chapter 5 show hints that the mist model may be able to explain multiphase gas in halos of a broad range of mass, from galaxy clusters, groups, giant elliptical and L^* -like galaxies. It is natural to extend the cosmological zoom-in simulations with this model in mind. In particular, addressing the mismatch of spatial resolution and the characteristic size of cold gas will be crucial in moving forward in this model. Our understanding of the multiphase CGM with star-forming and quiescent galaxies will benefit from such zoom-in cosmological galaxy simulations. It will allow one to make predictions with related to feedback processes that regulate the thermal state of the halo. The connection of AGN feedback and the CGM is particularly unexplored in giant elliptical galaxies with cosmological zoom-in simulations.

APPENDIX A

APPENDIX

A.1 Introduction

Absorption line spectroscopy is undoubtedly one of the most important tools for probing the physical properties of absorbing gas in a variety of astrophysical environments. Spectral analyses are ubiquitous in astrophysics and have a long and rich history in studies of stars, galaxies, gas, cosmology and more.

One of the basic tasks in such analyses is identification and characterization of spectral absorption lines. The most common way to do the latter is to fit the Voigt profile (VP) to the absorption profile and then use its parameters to extract basic properties of the absorbing gas, such as column density, Doppler parameter, and central velocity/redshift.

A number of codes for modeling absorption lines with the Voigt profile is available publicly. For example, `VPFIT` is a popular χ^2 -based code commonly used in the community over many years (Carswell et al., 1991). `autoVP` is an another χ^2 grid-search based code, an automated Voigt profile fitting procedure that chooses the best number of absorption components, designed for Ly α forest statistics (Davé et al., 1997).

While these codes have provided immense values to the community, there are fundamental limitations in methods based on χ^2 grid-search in the Voigt profile fitting. For example, such methods cannot constrain model parameters of non-detections in a statistically rigorous manner. The methods make very specific assumptions about the likelihood and the nature of the flux errors. In addition, it is challenging to capture the degeneracy between column density and Doppler parameter when absorption lines are saturated. Therefore, uncertainties of the best-fit parameters can be under- or overestimate estimated in a coarse grid of the parameter space. Furthermore, the grid search becomes prohibitively computationally expensive or impossible when fitting multiple absorption components because the number of parameters and thus the number of dimensions of parameter space is large. Thus, one

cannot reliably and rigorously find the best fit model for physical properties of the gas using the optimal number of absorption components. This can affect our physical interpretation of the kinematics of the gas (i.e., number of velocity components and the total dispersion) and the column density distribution function (for example, see Appendix in Gurvich et al., 2017).

Motivated by these problems, we developed a Bayesian approach to the Voigt profile fitting combined with efficient algorithms for sampling posterior distribution of parameter values. We introduce an implementation of such method in the python package `bayesvp`, which simultaneously addresses these issues and provides additional benefits. These include explicit control of the priors of parameters and providing constraints in a form of posterior distributions, which allows uniform treatment of detections and non-detections. Using efficient parallel sampling methods, `bayesvp` can fully explore high-dimensional parameter space efficiently, providing more accurate parameters uncertainties and their degeneracies. This allows us to reliably find the best number of absorption components and to include an additional continuum model. One can also explicitly compare models using Bayesian model comparison criteria. In section 2, we describe the Bayesian approach to absorption line fitting and our specific implementation of this approach in the `bayesvp` package. We include a short tutorial on the basic usage of the package in the Appendix.

A.2 The Package: `bayesvp`

A.2.1 *Posterior distribution and sampling*

In the Bayesian approach we constrain the posterior distribution, $\pi(\vec{\theta})$, of model parameters vector of column density, Doppler broadening factor, and redshift, $\vec{\theta} = \{N, b, z\}$, given the vector of spectral fluxes and their uncertainties, $\vec{F}$. The joint posterior distribution is simply

proportional to the product of likelihood $\mathcal{L}(\vec{F}|\vec{\theta})$ and the prior pdf of the parameters:

$$p(\vec{\theta}|\vec{F}) \propto \mathcal{L}(\vec{\theta}|\vec{F})p(\vec{\theta}), \quad (\text{A.1})$$

where the total likelihood $\mathcal{L}$ is the product of all likelihoods of the individual pixels within a spectral segment of interests:

$$\mathcal{L}(\vec{\theta}|\vec{F}) = \prod_i \ell_i(\vec{\theta}|F_i). \quad (\text{A.2})$$

Using methods, such as Markov Chain Monte Carlo (MCMC), that allow sampling of unnormalized distributions, we can sample the unnormalized posterior:

$$\pi(\vec{\theta}|\vec{F}) = \mathcal{L}(\vec{\theta}|\vec{F})p(\vec{\theta}), \quad (\text{A.3})$$

The individual likelihood ℓ_i of a given pixel i at F_i depends on the noise model. The specific noise characteristic depends on properties of telescopes and their instruments. In the limit of a large number of photons, we usually assume that ℓ_i is a Gaussian:

$$\ell_i(\vec{\theta}|F_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left[-\frac{(\hat{F}_i(\vec{\theta}) - F_i)^2}{2\sigma_i^2} \right] \quad (\text{A.4})$$

where σ_i is the uncertainty of the flux F_i at pixel i . However, the noise model can be modified and a different likelihood can be used for low photon counts (e.g., Poisson distribution). Although currently **bayesvp** implements only Gaussian likelihood, this can be easily changed to a different likelihood by supplying a different input log-likelihood function to the **Posterior** class in **Likelihood.py**. In **bayesvp**, we also provide an easy way to specify explicit priors based on knowledge of the specific problem in mind.

We sample the posterior distribution defined by eq. A.3 using an efficient sampling method. This is particularly useful in the Voigt profile fitting because it is common to

fit multiple absorption components simultaneously. The number of parameters can thus grow quickly, since each absorption component adds three parameters (e.g., N , b , z). In addition, many sampling methods can be easily parallelized, which allows efficient sampling of the high-dimensional parameter space. Due to blending between components and the relationship between Doppler parameters and column densities in the saturated regime, Voigt profile parameters can be highly correlated. The simplest Metropolis-Hasting MCMC algorithm that samples the parameter space with isotropic steps is thus not optimal. We thus use two MCMC sampling algorithms designed to efficiently sample narrow distributions of highly correlated parameters and in high dimensional spaces. In particular, `bayesvp` is set up to use the affine-invariant ensemble sampling algorithm of (Goodman & Weare, 2010) implemented in the `emcee` python package (ForemanMackey et al., 2013), and an additional powerful kernel-density-estimate based MCMC sampler `kombine` (Farr & Farr, 2015).

A.2.2 Voigt profile model

The Voigt profile is parameterized by the absorber column density N , Doppler parameter b , and redshift z . The normalized flux at frequency ν (or wavelength λ) is simply,

$$F(\nu|N, b, z) = \exp[-\tau(\nu)] \quad (\text{A.5})$$

where the optical depth τ at frequency ν of a system with (N, b, z) is:

$$\tau_i(\lambda|N_i, b_i, z_i) = N_i \frac{\pi e^2}{m_e c} f_{\text{osc}} \Phi(\lambda|b_i, z_i). \quad (\text{A.6})$$

and $\Phi(\nu)$ is the convolution of a Gaussian profile (due to the Doppler broadening) and a Lorentzian profile (due to pressure broadening). For computation speed, we compute the Voigt profile function $\Phi(\nu)$ via the real part of the Faddeeva function $w(x + iy)$ routine in the `SciPy` package:

$$\Phi(\nu|b, z) = \frac{\Re[w(x + iy)]}{\sqrt{\pi}\nu_D} \quad (\text{A.7})$$

where $x = \Delta\nu/\Delta\nu_D$ and $y = \Gamma/(4\pi\Delta\nu_D)$. Also, $\Delta\nu = \nu - \nu_0(1+z)^{-1}$, $\Delta\nu_D = b\nu/c$. In addition, $\sigma_0 = \frac{\sqrt{\pi}e^2}{m_e c^2}$ is the cross section.

Collectively, the set of atomic constants $\{e, m_e, f_{\text{osc}}, \nu_0, \Gamma\}$ are the charge and mass of the electron, oscillator strength, rest-frame frequency and damping coefficient of an atomic transition. We include atomic parameters from the UV to optical wavelength regimes compiled from Morton (2003). Any additional transitions can be easily added in the supplementary data file included in **bayesvp**.

A.2.3 Continuum model

We implement a polynomial continuum model in addition to the Voigt profile model fitting. The addition of the continuum model moves some of the error budget from systematic uncertainties to statistical uncertainties. The continuum is parametrized as a polynomial in **bayesvp**:

$$C(\lambda|a_i) = \sum_i a_i (\lambda - \bar{\lambda})^i + \bar{F} \quad (\text{A.8})$$

$$F_{\text{model}}(\lambda) = C(\lambda|a_i) \times F_{\text{Voigt}}(\lambda|N, b, z) \quad (\text{A.9})$$

where $\{a_i\}$ are the continuum parameters. Although **bayesvp** can fit a high order polynomial, in our experience for typical wavelength ranges used in absorption line fits, linear or quadratic polynomial are sufficient to model continuum accurately. Thus, we recommend fitting up to a quadratic polynomial (i.e., three continuum parameters). In any case, the degree of the polynomial should be considerably smaller than the number of spectrum pixels dominated by continuum.

Note that we have pivoted the polynomial fit from the observed wavelength λ to $(\lambda - \bar{\lambda})$, where $\bar{\lambda}$ and $\bar{F}$ are the median wavelength and flux of the input data. In Figure A.1, we show a test that includes a continuum model in a synthetic absorption line, where **bayesvp** successfully captures the correct input parameters.

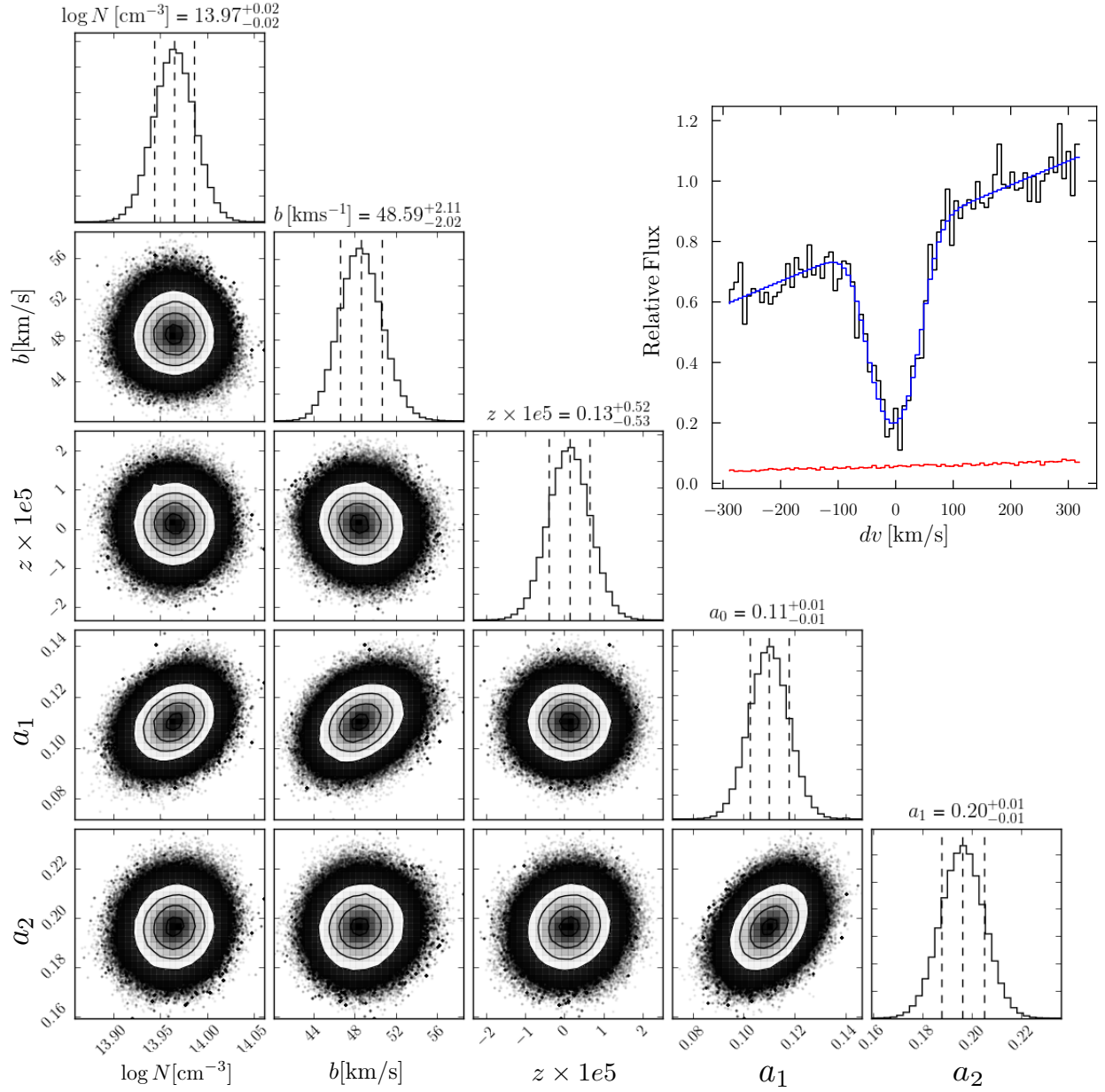


Figure A.1: Marginalized posterior distributions for parameters in a single component Voigt profile fit with a linear continuum model fit to a synthetic spectrum. The top right panel shows the best fit model relative to the input spectrum over the fitted wavelength interval.

A.2.4 Convergence criteria

When sampling a posterior, it is important to demonstrate the convergence of sample chains in order to trust the resulting posterior and statistics derived from it. `bayesvp` package provides functionality for convergence tests using the Gelman-Rubin indicator. The basic idea is that the chains should reach a state where the parameter distributions are stationary. To do so, the GR indicator compares the variances of parameters within and across independently sampled chains. We adopt the definition used in Gelman & Rubin (1992):

$$R = \frac{V}{W} = \frac{N_w + 1}{N_w} \frac{\sigma_+^2}{W} - \frac{N_s - 1}{N_w N_s} \quad (\text{A.10})$$

where W and V are the “within-chain” variance and “between-chain” variance. And σ_+^2 is the estimator for the true variance of a parameter. In addition, N_w and N_s are the number of independent chains (or “walkers”) and the number of steps (length of the individual chains), respectively. A perfect convergence would correspond to the “within-chain” and “between-chain” variances matching each other ($R = 1$). Figure A.2 shows an example the GR indicator as a function of the steps for the parameters in the continuum model fit shown in Figure A.1.

A.2.5 Model comparison

Observed absorption lines are frequently blended. The best number of absorption components for the observed profiles is not always clear. The column densities and Doppler parameters are strongly affected by the number of components that we choose to fit the data. Our physical interpretation of the absorbers may also change depending on the number of velocity components. One of the main advantages of a Bayesian approach is the ability to select the best fit model based on some objective criteria when it is not clear how many components one should choose. In `bayesvp`, we implement some measures for choosing the best fit models: the Bayesian information criterion (BIC) for a given model M (Ivezić et al.,

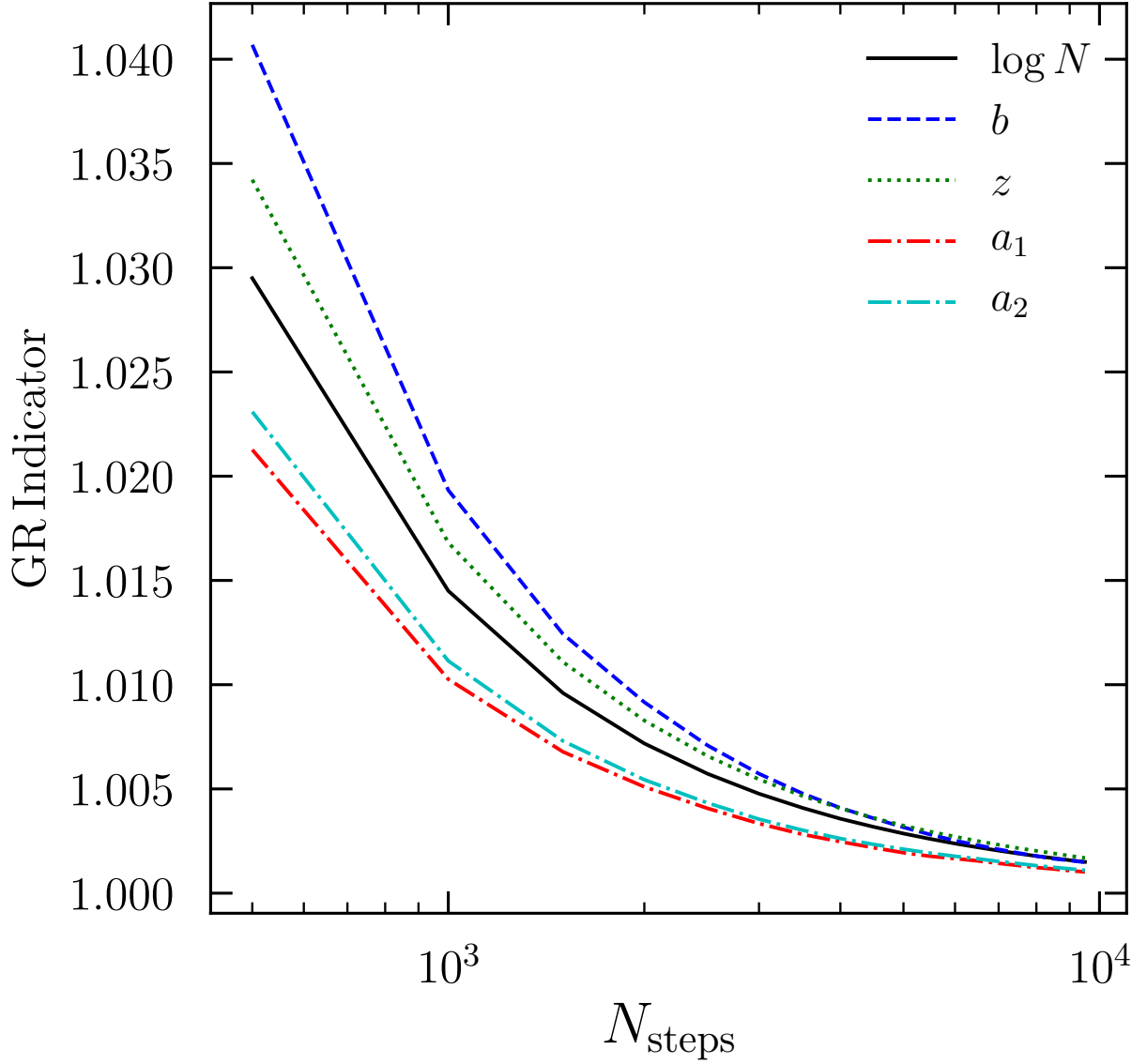


Figure A.2: Evolution of the Gelman-Rubin (GR) indicator for model parameters as a function of MCMC steps. A perfect convergence of a parameter corresponds a value of one. Users may choose a level of convergence based on a threshold of GR indicator to decide the number of MCMC steps necessary to run. Note also that parameters converge at different rates.

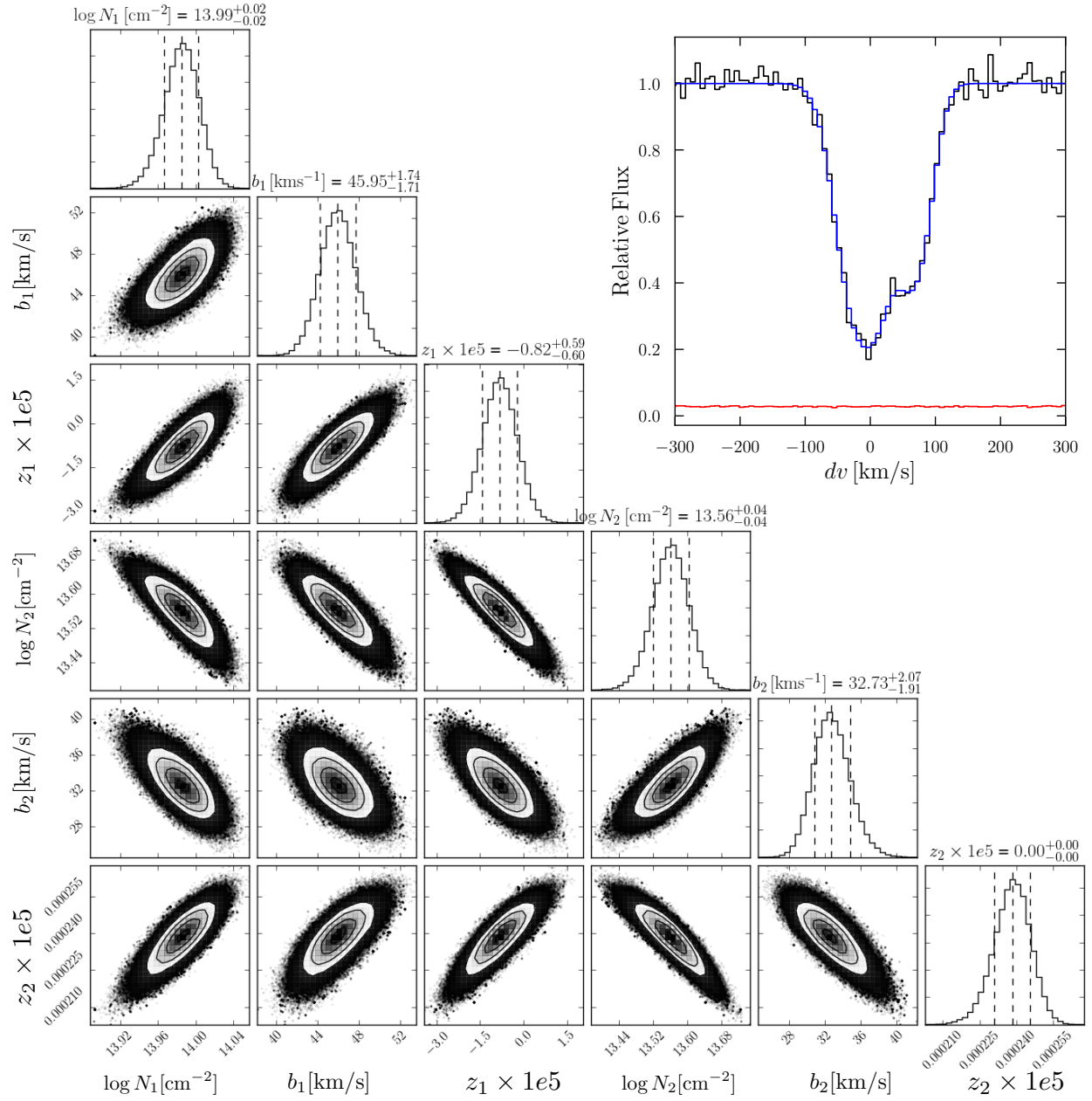


Figure A.3: This figure shows the posterior distributions Voigt profile fit of a synthetic spectrum. This shows that a 2-component model is a good fit to the data. Note also the parameters of the two Voigt profiles are correlated.

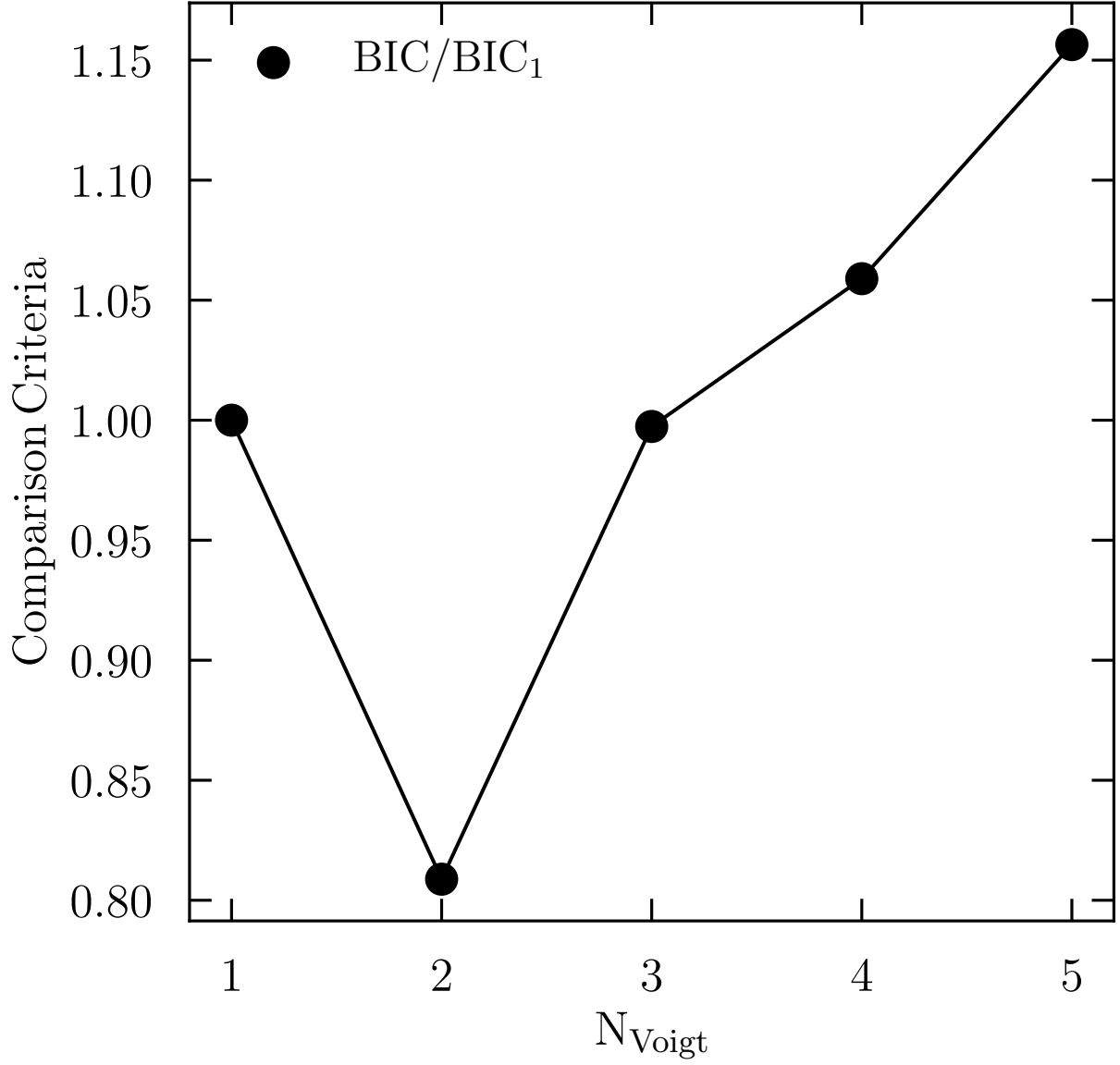


Figure A.4: Model comparison based on BIC as a function of the number of Voigt components (N_{Voigt}) for the spectrum shown in Figure A.3. Note that there are three free parameters in each component. The model of 2 components with the minimum value is preferred.

2014):

$$BIC = -2\ln(L^0(M)) + 2k + k \ln N \quad (\text{A.11})$$

where, $L^0(M)$ denotes the maximum value of the likelihood of model M , and k is the number of model parameters and N is the number of points in the MCMC chains. The objective for the best-fit model is to minimize BIC .

In general, odds ratio is also useful for comparison of two models M_1 and M_2 :

$$O_{21} = L(M_2)/L(M_1) \quad (\text{A.12})$$

For this measure, model 2 is preferred if $O_{21} \gg 1$. However, it is computationally expensive to compute $L(M)$ because it involves integrating the posterior distribution over all k model parameters, $L(M) = \int \pi(\vec{\theta}) d^k \theta$. Instead, we calculate $L(M)$ using the local density estimate of the points sampled by MCMC chains $\rho(\vec{\theta})$ and the posterior $\pi(\vec{\theta})$, i.e. $L(M) = N\pi(\vec{\theta})/\rho(\vec{\theta})$ (see section 5 in Ivezić et al., 2014, for more details.)

The basic idea for these criteria (e.g., BIC) is that the models with a larger number of parameters are penalized and have to compensate with the larger values of $L(M)$ or $L^0(M)$. As an example, in Figure 3 we show results of a two-component Voigt fit for a blended absorption profile (with two input absorption components) and the corresponding posterior distributions of the model parameters. Clearly, the two-component fit describes the observed flux quite well. The choice of a two-component fit is also supported by the comparison criteria shown in Figure A.4.

A.2.6 Uniform treatment for upper limits, lower limits and detections

Absorption line fits for weak lines in noisy spectra often result in unconstrained parameters. In the χ^2 based analyses, these are usually treated as upper or lower limits. For example, in observational analyses of the absorption profiles of the circumgalactic medium, lines of sights of large impact parameters usually result only in upper limits (e.g., Liang & Chen,

2014). The inclusion of such limits when fitting a model of the column density profile or comparing to simulation results (e.g., Liang et al., 2016) is challenging and requires strong assumptions about their posterior form.

Thus, there has been a substantial effort in trying to model the probability of the column density of upper-limit systems in order to maximize their utility, such as constraining the gas properties (e.g., density and metallicity; Crighton et al., 2015; Fumagalli et al., 2016; Stern et al., 2016). On the other hand, when the absorption lines are saturated, it is difficult to constrain the column density precisely due to its degeneracy with the Doppler parameter. In this case, lower limits are often quoted in the literature. An improvement over the lower limits is bracketed ranges using the absence of damping wings (e.g., Johnson et al., 2015b). In both of these cases, when the original data is available, this probability distribution can be directly measured as a marginalized posterior. In this case, no assumptions about the shape of the probability distribution are needed. We refer readers to Figure 6 in our companion paper (Liang et al., 2017) to see the diversity of the posterior distributions for the cases that would usually be identified as upper or lower limits and compared to the posterior distributions of “detections.”

A.3 Summary

We present a python package `bayesvp`, which implements a Bayesian approach for modeling absorption lines with the Voigt profile. The package can be used to simultaneously fit multiple absorption components to constrain column density, Doppler parameter, and redshifts of the absorbing gas. The parameter constraints are derived in the form of marginalized posterior distributions, which allows for uniform treatment of the non-detections (i.e., upper limits) and detections in subsequent analyses, model fits, and comparisons of observations with simulations. The `bayesvp` package is set up to use the affine-invariant MCMC method implemented in the MCMC `emcee` package and the kernel-density-estimate-based ensemble sampler `KOMBINE` to efficiently sample the posterior distributions of highly correlated param-

eters. It also provides other useful utilities, such as convolution with instrumental line spread function (LSF), explicit control of parameter priors, Bayesian model comparison criteria, and sampling convergence check. The `bayesvp` package is designed to have a simple user interface with a single configuration file that users can explicitly define the line spread function, the number of walkers, MCMC steps, and parallel threads. `bayesvp` is publicly available at <https://github.com/cameronliang/bayesvp>. In the Appendix below we present an example showing how to use the package to fit absorption lines.

REFERENCES

- Ackermann, M., Ajello, M., Allafort, A., & et al. 2013, *Science*, 339, 807
- Adelberger, K. L., Shapley, A. E., Steidel, C. C., et al. 2005, *ApJ*, 629, 636
- Agertz, O., & Kravtsov, A. V. 2015, *ApJ*, 804, 18
- . 2016, *ApJ*, 824, 79
- Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, *ApJ*, 770, 25
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARAA&A*, 47, 481
- Audit, E., & Hennebelle, P. 2005, *A&A*, 433, 1
- Aumer, M., White, S. D. M., Naab, T., & Scannapieco, C. 2013, *MNRAS*, 434, 3142
- Balbus, S. A., & McKee, C. F. 1982, *ApJ*, 252, 529
- Baldry, I. K., Driver, S. P., Loveday, J., et al. 2012, *MNRAS*, 421, 621
- Barai, P., Viel, M., Borgani, S., et al. 2013, *MNRAS*, 430, 3213
- Barnes, L. A., Haehnelt, M. G., Tescari, E., & Viel, M. 2011, *MNRAS*, 416, 1723
- Barton, E. J., & Cooke, J. 2009, *AJ*, 138, 1817
- Beck, R. 2015, *A&AR*, 24, 4
- Beck, R., & Krause, M. 2005, *Astronomische Nachrichten*, 326, 414
- Begelman, M. C., & Fabian, A. C. 1990, *MNRAS*, 244, 26P
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJ*, 770, 57
- Beirão, P., Armus, L., Lehnert, M. D., et al. 2015, *MNRAS*, 451, 2640
- Benson, A. J., Bower, R. G., Frenk, C. S., et al. 2003, *ApJ*, 599, 38

- Bergeron, J., & Boisse, P. 1986, *A&A*, 168, 6
- Bergeron, J., & Boissé, P. 1991, *A&A*, 243, 344
- Bernardi, M., Meert, A., Sheth, R. K., et al. 2013, *MNRAS*, 436, 697
- Binney, J. 1977, *ApJ*, 215, 483
- Bird, S., Haehnelt, M., Neeleman, M., et al. 2015, *MNRAS*, 447, 1834
- Birnboim, Y., & Dekel, A. 2003, *MNRAS*, 345, 349
- Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, *AJ*, 142, 31
- Blanton, M. R., & Roweis, S. 2007, *AJ*, 133, 734
- Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, *AJ*, 129, 2562
- Boksenberg, A., Danziger, I. J., Fosbury, R. A. E., & Goss, W. M. 1980, *ApJ*, 242, L145
- Boksenberg, A., & Sargent, W. L. W. 1978, *ApJ*, 220, 42
- Boksenberg, A., Sargent, W. L. W., & Rauch, M. 2003, *ArXiv Astrophysics e-prints*
- Booth, C. M., Agertz, O., Kravtsov, A. V., & Gnedin, N. Y. 2013, *ApJ*, 777, L16
- Bordoloi, R., Lilly, S. J., Knobel, C., & et al. 2011, *ApJ*, 743, 10
- Bordoloi, R., Tumlinson, J., Werk, J. K., et al. 2014, *ApJ*, 796, 136
- Borthakur, S., Heckman, T., Strickland, D., Wild, V., & Schiminovich, D. 2013, *ApJ*, 768, 18
- Borthakur, S., Heckman, T. M., Tumlinson, J., et al. 2015, *ApJ* submitted (arXiv:1504.01392)
- Borthakur, S., Heckman, T., Tumlinson, J., et al. 2016, *ApJ*, 833, 259

- Bowen, D. V., Blades, J. C., & Pettini, M. 1995, *ApJ*, 448, 634
- Braginskii, S. I. 1965, *Reviews of Plasma Physics*, 1, 205
- Brook, C. B., Stinson, G., Gibson, B. K., Wadsley, J., & Quinn, T. 2012, *MNRAS*, 424, 1275
- Brüggen, M., & Scannapieco, E. 2016, *ApJ*, 822, 31
- Bryan, G. L., & Norman, M. L. 1998, *ApJ*, 495, 80
- Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, *MNRAS*, 321, 559
- Burkert, A., & Lin, D. N. C. 2000, *ApJ*, 537, 270
- Butsky, I., & Quinn, T. R. 2018, *ArXiv e-prints*
- Cameron, E., & Driver, S. P. 2007, *MNRAS*, 377, 523
- Cantalupo, S., Arrigoni-Battaia, F., Prochaska, J. X., Hennawi, J. F., & Madau, P. 2014, *Nature*, 506, 63
- Caprioli, D., & Spitkovsky, A. 2014, *ApJ*, 783, 91
- Carswell, R. F., Lanzetta, K. M., Parnell, H. C., & Webb, J. K. 1991, *ApJ*, 371, 36
- Cavagnolo, K. W., Donahue, M., Voit, G. M., & Sun, M. 2009, *The Astrophysical Journal Supplement Series*, 182, 12
- Chen, H.-W. 2012, *MNRAS*, 427, 1238
- Chen, H.-W., Helsby, J. E., Gauthier, J.-R., et al. 2010a, *ApJ*, 714, 1521
- Chen, H.-W., Lanzetta, K. M., & Webb, J. K. 2001a, *ApJ*, 556, 158
- Chen, H.-W., Lanzetta, K. M., Webb, J. K., & Barcons, X. 1998, *ApJ*, 498, 77

—. 2001b, *ApJ*, 559, 654

Chen, H.-W., & Mulchaey, J. S. 2009, *ApJ*, 701, 1219

Chen, H.-W., & Tinker, J. L. 2008, *ApJ*, 687, 745

Chen, H.-W., Wild, V., Tinker, J. L., et al. 2010b, *ApJ*, 724, L176

Churchill, C. W. 1996, *ArXiv Astrophysics e-prints*

Churchill, C. W., Nielsen, N. M., Kacprzak, G. G., & Trujillo-Gomez, S. 2013, *ApJ*, 763, L42

Churchill, C. W., Vander Vliet, J. R., Trujillo-Gomez, S., Kacprzak, G. G., & Klypin, A. 2015, *ApJ*, 802, 10

Chynoweth, K. M., Holley-Bockelmann, K., Polisensky, E., & Langston, G. I. 2011, *AJ*, 142, 137

Colless, M., Dalton, G., Maddox, S., et al. 2001, *MNRAS*, 328, 1039

Cooksey, K. L., Kao, M. M., Simcoe, R. A., O’Meara, J. M., & Prochaska, J. X. 2013, *ApJ*, 763, 37

Cooper, J. L., Bicknell, G. V., Sutherland, R. S., & Bland-Hawthorn, J. 2008, *ApJ*, 674, 157

—. 2009, *ApJ*, 703, 330

Correa, C. A., Wyithe, J. S. B., Schaye, J., & Duffy, A. R. 2015, *arXiv:1502.00391*

Cowie, L. L., & McKee, C. F. 1977, *ApJ*, 211, 135

Crighton, N. H. M., Hennawi, J. F., Simcoe, R. A., et al. 2015, *MNRAS*, 446, 18

Cuesta, A. J., Prada, F., Klypin, A., & Moles, M. 2008, *MNRAS*, 389, 385

Davé, R., Hernquist, L., Weinberg, D. H., & Katz, N. 1997, *ApJ*, 477, 21

- Davis, J. D., Keeney, B. A., Danforth, C. W., & Stocke, J. T. 2015, ApJ submitted (arXiv/1506.04095)
- Dedner, A., Kemm, F., Kröner, D., et al. 2002, Journal of Computational Physics, 175, 645
- Dekel, A., & Silk, J. 1986, ApJ, 303, 39
- Dekel, A., & Woo, J. 2003, MNRAS, 344, 1131
- Diemand, J., Kuhlen, M., & Madau, P. 2007, ApJ, 667, 859
- Diemer, B., & Kravtsov, A. V. 2014, ApJ, 789, 1
- . 2015, ApJ, 799, 108
- Diemer, B., More, S., & Kravtsov, A. V. 2013, ApJ, 766, 25
- Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium
- Dubois, Y., & Teyssier, R. 2008, A&A, 477, 79
- Dursi, L. J., & Pfrommer, C. 2008, ApJ, 677, 993
- Efstathiou, G. 2000, MNRAS, 317, 697
- Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, ApJ, 136, 748
- Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, ApJ, 647, 128
- Faerman, Y., Sternberg, A., & McKee, C. F. 2017, ApJ, 835, 52
- Fall, S. M., & Rees, M. J. 1985, ApJ, 298, 18
- Farr, B., & Farr, W. M. 2015, in prep
- Faucher-Giguère, C.-A., Hopkins, P. F., Kereš, D., et al. 2015, MNRAS, 449, 987
- Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, 417, 2982

- Feldmann, R., Carollo, C. M., & Mayer, L. 2011, *ApJ*, 736, 88
- Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, *RMXAA*, 49, 137
- Field, G. B. 1965, *ApJ*, 142, 531
- Field, G. B., Goldsmith, D. W., & Habing, H. J. 1969, *ApJ*, 155, L149
- Fielding, D., Quataert, E., McCourt, M., & Thompson, T. A. 2017, *MNRAS*, 466, 3810
- Fogarty, K., Postman, M., Larson, R., Donahue, M., & Moustakas, J. 2017, ArXiv e-prints
- Ford, A. B., Davé, R., Oppenheimer, B. D., et al. 2014, *MNRAS*, 444, 1260
- Ford, A. B., Oppenheimer, B. D., Davé, R., et al. 2013, *MNRAS*, 432, 89
- Ford, A. B., Werk, J. K., Dave, R., et al. 2015, *MNRAS* submitted (arXiv:1503.02084)
- Ford, A. B., Werk, J. K., Davé, R., et al. 2016, *MNRAS*, 459, 1745
- Foreman-Mackey, D. 2016, *The Journal of Open Source Software*, 24
- ForemanMackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
- Fragile, P. C., Murray, S. D., Anninos, P., & van Breugel, W. 2004, *ApJ*, 604, 74
- Fukugita, M. 2004, in *IAU Symposium*, Vol. 220, *Dark Matter in Galaxies*, ed. S. Ryder, D. Pisano, M. Walker, & K. Freeman, 227
- Fumagalli, M., O’Meara, J. M., & Prochaska, J. X. 2016, *MNRAS*, 455, 4100
- Gaspari, M., Ruszkowski, M., & Oh, S. P. 2013, *MNRAS*, 432, 3401
- Gaspari, M., Ruszkowski, M., & Sharma, P. 2012, *ApJ*, 746, 94
- Gaspari, M., Temi, P., & Brighenti, F. 2017, *MNRAS*, 466, 677
- Gauthier, J.-R. 2013, *MNRAS*, 432, 1444

- Gauthier, J.-R., & Chen, H.-W. 2011, MNRAS, 418, 2730
- Gauthier, J.-R., Chen, H.-W., & Tinker, J. L. 2010, ApJ, 716, 1263
- Gelman, A., & Rubin, D. B. 1992, Statistical Science, 7, 457
- Gnedin, N. Y., & Hollon, N. 2012, ApJS, 202, 13
- Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, Vol. 5, No. 1, p. 65-80, 2010, 5, 65
- Gott, III, J. R., & Thuan, T. X. 1976, ApJ, 204, 649
- Green, J. C., Froning, C. S., Osterman, S., et al. 2012, ApJ, 744, 60
- Gronke, M., Dijkstra, M., McCourt, M., & Peng Oh, S. 2017, A&A, 607, A71
- Gurvich, A., Burkhart, B., & Bird, S. 2017, ApJ, 835, 175
- Haardt, F., & Madau, P. 2012, ApJ, 746, 125
- Hafen, Z., Faucher-Giguere, C.-A., Angles-Alcazar, D., et al. 2016, ArXiv e-prints
- Heckman, T., Borthakur, S., Wild, V., Schiminovich, D., & Bordoloi, R. 2017, ApJ, 846, 151
- Heckman, T. M., Lehnert, M. D., Strickland, D. K., & Armus, L. 2000, ApJS, 129, 493
- Hennawi, J. F., Prochaska, J. X., Cantalupo, S., & Arrigoni-Battaia, F. 2015, Science, 348, 779
- Hennawi, J. F., Prochaska, J. X., Burles, S., et al. 2006, ApJ, 651, 61
- Hennebelle, P., & Péroult, M. 1999, A&A, 351, 309
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
- Hopkins, P. F., Quataert, E., & Murray, N. 2011, MNRAS, 417, 950

- . 2012, MNRAS, 421, 3522
- Huang, Y.-H., Chen, H.-W., Johnson, S. D., & Weiner, B. J. 2016, MNRAS, 455, 1713
- Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, ApJS, 199, 26
- Hummels, C., Smith, B., & Silvia, D. 2016, ArXiv e-prints
- Hummels, C. B., Bryan, G. L., Smith, B. D., & Turk, M. J. 2013, MNRAS, 430, 1548
- Ivezić, Ž., Connolly, A., Vanderplas, J., & Gray, A. 2014, Statistics, Data Mining and Machine Learning in Astronomy (Princeton University Press)
- Johnson, S. D., Chen, H.-W., & Mulchaey, J. S. 2013, MNRAS, 434, 1765
- . 2015a, MNRAS, 452, 2553
- . 2015b, MNRAS, 449, 3263
- Joung, M. R., Bryan, G. L., & Putman, M. E. 2012, ApJ, 745, 148
- Kacprzak, G. G., Churchill, C. W., Steidel, C. C., & Murphy, M. T. 2008, AJ, 135, 922
- Kawata, D., & Mulchaey, J. S. 2008, ApJ, 672, L103
- Kennicutt, R. C., & Evans, N. J. 2012, ARAA&A, 50, 531
- Kereš, D., & Hernquist, L. 2009, ApJ, 700, L1
- Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160
- Khokhlov, A. 1998, Journal of Computational Physics, 143, 519
- Klein, R. I., McKee, C. F., & Colella, P. 1994, ApJ, 420, 213
- Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2009, ApJS, 180, 330
- Kravtsov, A., Vikhlinin, A., & Meshcheryakov, A. 2014, arXiv:1401.7329

- Kravtsov, A. V., Klypin, A. A., & Khokhlov, A. M. 1997, *ApJS*, 111, 73
- Kritsuk, A. G., & Norman, M. L. 2002, *ApJ*, 569, L127
- Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2009, *ApJ*, 699, 850
- Lan, T.-W., & Fukugita, M. 2017, *ApJ*, 850, 156
- Lanzetta, K. M., Bowen, D. V., Tytler, D., & Webb, J. K. 1995, *ApJ*, 442, 538
- Larson, R. B. 1974, *MNRAS*, 169, 229
- Lau, M. W., Prochaska, J. X., & Hennawi, J. F. 2016, *ApJS*, 226, 25
- Lehner, N., Howk, J. C., Tripp, T. M., et al. 2013, *ApJ*, 770, 138
- Lehnert, M. D., Heckman, T. M., & Weaver, K. A. 1999, *ApJ*, 523, 575
- Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, *ApJS*, 123, 3
- Leroy, A. K., Walter, F., Martini, P., et al. 2015, *ApJ*, 814, 83
- Li, Y., Bryan, G. L., Ruszkowski, M., et al. 2015, *ApJ*, 811, 73
- Liang, C., & Kravtsov, A. 2017, *ArXiv e-prints*
- Liang, C., Kravtsov, A., & Agertz, O. 2017, *MNRAS* submitted (arXiv/1710.00411)
- Liang, C. J., & Chen, H.-W. 2014, *MNRAS*, 445, 2061
- Liang, C. J., Kravtsov, A. V., & Agertz, O. 2016, *MNRAS*, 458, 1164
- Mac Low, M.-M., McKee, C. F., Klein, R. I., Stone, J. M., & Norman, M. L. 1994, *ApJ*, 433, 757
- Malkov, M., & Drury, L. O. 2001, *Reports on Progress in Physics*, 64, 429
- Maller, A. H., & Bullock, J. S. 2004, *MNRAS*, 355, 694

- Marasco, A., Debattista, V. P., Fraternali, F., et al. 2015, MNRAS, 451, 4223
- Martin, C. L. 1999, ApJ, 513, 156
- Martin, D. C., Chang, D., Matuszewski, M., et al. 2014, ApJ, 786, 106
- Mathews, W. G., & Prochaska, J. X. 2017, ApJ, 846, L24
- McCarthy, I. G., Frenk, C. S., Font, A. S., et al. 2008, MNRAS, 383, 593
- McCourt, M., Oh, S. P., O’Leary, R., & Madigan, A.-M. 2018, MNRAS, 473, 5407
- McCourt, M., Oh, S. P., O’Leary, R. M., & Madigan, A.-M. 2016, arXiv/1610.01164
- McCourt, M., O’Leary, R. M., Madigan, A.-M., & Quataert, E. 2015, MNRAS, 449, 2
- McCourt, M., Sharma, P., Quataert, E., & Parrish, I. J. 2012, MNRAS, 419, 3319
- McKee, C. F., & Ostriker, J. P. 1977, ApJ, 218, 148
- Meece, G. R., O’Shea, B. W., & Voit, G. M. 2015, ApJ, 808, 43
- Meerson, B. 1996, Reviews of Modern Physics, 68, 215
- Melioli, C., de Gouveia dal Pino, E. M., & Raga, A. 2005, A&A, 443, 495
- Mellema, G., Kurk, J. D., & Röttgering, H. J. A. 2002, A&A, 395, L13
- Meyer, C. D., Balsara, D. S., & Aslam, T. D. 2014, Journal of Computational Physics, 257, 594
- Mignone, A., Tzeferacos, P., & Bodo, G. 2010, Journal of Computational Physics, 229, 5896
- Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution
- Mo, H. J., & Miralda-Escude, J. 1996, ApJ, 469, 589
- More, S., Diemer, B., & Kravtsov, A. 2015, arXiv:1504.05591

- Morton, D. C. 2003, *ApJS*, 149, 205
- Moustakas, J., Coil, A. L., Aird, J., et al. 2013, *ApJ*, 767, 50
- Münch, G., & Zirin, H. 1961, *ApJ*, 133, 11
- Muratov, A. L., Keres, D., Faucher-Giguere, C.-A., et al. 2015, *MNRAS* submitted (arxiv:1501.03155)
- Murray, N., Ménard, B., & Thompson, T. A. 2011, *ApJ*, 735, 66
- Murray, N., Quataert, E., & Thompson, T. A. 2005, *ApJ*, 618, 569
- Muzahid, S. 2014, *ApJ*, 784, 5
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, 777, 18
- Nakamura, F., McKee, C. F., Klein, R. I., & Fisher, R. T. 2006, *ApJS*, 164, 477
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493
- Nelson, D., Kauffmann, G., Pillepich, A., et al. 2018, *MNRAS*, 477, 450
- Neto, A. F., Gao, L., Bett, P., et al. 2007, *MNRAS*, 381, 1450
- Nielsen, N. M., Churchill, C. W., & Kacprzak, G. G. 2013, *ApJ*, 776, 115
- Oppenheimer, B. D., & Davé, R. 2006, *MNRAS*, 373, 1265
- . 2008, *MNRAS*, 387, 577
- Oppenheimer, B. D., Davé, R., Katz, N., Kollmeier, J. A., & Weinberg, D. H. 2012, *MNRAS*, 420, 829
- Oppenheimer, B. D., & Schaye, J. 2013, *MNRAS*, 434, 1043
- Oppenheimer, B. D., Segers, M., Schaye, J., Richings, A. J., & Crain, R. A. 2017, *ArXiv e-prints*

- Oppenheimer, B. D., Crain, R. A., Schaye, J., et al. 2016, MNRAS, 460, 2157
- Pallottini, A., Gallerani, S., & Ferrara, A. 2014, MNRAS, 444, L105
- Parker, E. N. 1953, ApJ, 117, 431
- Pitts, E., & Tayler, R. J. 1997, MNRAS, 288, 457
- Prada, F., Klypin, A. A., Simonneau, E., et al. 2006, ApJ, 645, 1001
- Prochaska, J. X., & Hennawi, J. F. 2009, ApJ, 690, 1558
- Prochaska, J. X., Hennawi, J. F., & Simcoe, R. A. 2013, ApJ, 762, L19
- Prochaska, J. X., Lau, M. W., & Hennawi, J. F. 2014, ApJ, 796, 140
- Prochaska, J. X., Weiner, B., Chen, H.-W., Mulchaey, J., & Cooksey, K. 2011, ApJ, 740, 91
- Prochaska, J. X., Werk, J. K., Worseck, G., et al. 2017, ApJ, 837, 169
- Rahmati, A., Pawlik, A. H., Raicevic, M., & Schaye, J. 2013, MNRAS, 430, 2427
- Rasera, Y., & Teyssier, R. 2006, A&A, 445, 1
- Rauch, M., Haehnelt, M. G., & Steinmetz, M. 1997, ApJ, 481, 601
- Rauch, M., Sargent, W. L. W., & Barlow, T. A. 1999, ApJ, 515, 500
- Reddy, N. A., Pettini, M., Steidel, C. C., et al. 2012, ApJ, 754, 25
- Rees, M. J., & Ostriker, J. P. 1977, MNRAS, 179, 541
- Roškar, R., Teyssier, R., Agertz, O., Wetzstein, M., & Moore, B. 2014, MNRAS, 444, 2837
- Rudie, G. C., Steidel, C. C., Shapley, A. E., & Pettini, M. 2013, ApJ, 769, 146
- Rudie, G. C., Steidel, C. C., Trainor, R. F., et al. 2012, ApJ, 750, 67
- Rupke, D. S., Veilleux, S., & Sanders, D. B. 2005, ApJS, 160, 115

- Salem, M., & Bryan, G. L. 2014, MNRAS, 437, 3312
- Salem, M., Bryan, G. L., & Corlies, L. 2016, MNRAS, 456, 582
- Salem, M., Bryan, G. L., & Hummels, C. 2014, ApJ, 797, L18
- Sales, L. V., Navarro, J. F., Theuns, T., et al. 2012, MNRAS, 423, 1544
- Sánchez-Salcedo, F. J., Vázquez-Semadeni, E., & Gazol, A. 2002, ApJ, 577, 768
- Sargent, W. L. W., Boksenberg, A., & Steidel, C. C. 1988, ApJS, 68, 539
- Savage, B. D., Lehner, N., & Narayanan, A. 2011, ApJ, 743, 180
- Scannapieco, E., & Brüggen, M. 2015, ApJ, 805, 158
- Scannapieco, E., Thacker, R. J., & Davis, M. 2001, ApJ, 557, 605
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
- Schiminovich, D., Wyder, T. K., Martin, D. C., et al. 2007, ApJS, 173, 315
- Schiminovich, D., Catinella, B., Kauffmann, G., et al. 2010, MNRAS, 408, 919
- Schneider, E. E., & Robertson, B. E. 2017, ApJ, 834, 144
- Schneider, E. E., Robertson, B. E., & Thompson, T. A. 2018, ArXiv e-prints
- Schwartz, C. M., & Martin, C. L. 2004, ApJ, 610, 201
- Schwartz, C. M., Martin, C. L., Chandar, R., et al. 2006, ApJ, 646, 858
- Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
- Shen, S., Madau, P., Aguirre, A., et al. 2012, ApJ, 760, 50
- Shen, S., Madau, P., Guedes, J., et al. 2013, ApJ, 765, 89
- Shin, M.-S., Stone, J. M., & Snyder, G. F. 2008, ApJ, 680, 336

- Silk, J. 1977, *ApJ*, 211, 638
- Silk, J., & Rees, M. J. 1998, *A&A*, 331, L1
- Simcoe, R. A. 2011, *ApJ*, 738, 159
- Simcoe, R. A., Sargent, W. L. W., & Rauch, M. 2004, *ApJ*, 606, 92
- Sinnott, R. W. 1988, NGC 2000.0: The Complete New General Catalogue and Index Catalogues of Nebulae and Star Clusters by J. L. E. Dreyer
- Smith, B. D., Hallman, E. J., Shull, J. M., & O'Shea, B. W. 2011, *ApJ*, 731, 6
- Socrates, A., Davis, S. W., & Ramirez-Ruiz, E. 2008, *ApJ*, 687, 202
- Songaila, A. 2001, *ApJ*, 561, L153
- Spitzer, L. 1962, *Physics of Fully Ionized Gases*
- Spitzer, Jr., L. 1956, *ApJ*, 124, 20
- Steidel, C. C. 1990, *ApJS*, 72, 1
- Steidel, C. C. 1993, in *Astronomical Society of the Pacific Conference Series*, Vol. 49, *Galaxy Evolution. The Milky Way Perspective*, ed. S. R. Majewski, 227
- Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, *ApJ*, 717, 289
- Steidel, C. C., Kollmeier, J. A., Shapley, A. E., et al. 2002, *ApJ*, 570, 526
- Stern, J., Faucher-Giguère, C.-A., Hennawi, J. F., et al. 2018, *ArXiv e-prints*
- Stern, J., Hennawi, J. F., Prochaska, J. X., & Werk, J. K. 2016, *ApJ*, 830, 87
- Stinson, G. S., Brook, C., Macciò, A. V., et al. 2013, *MNRAS*, 428, 129
- Stinson, G. S., Brook, C., Prochaska, J. X., et al. 2012, *MNRAS*, 425, 1270

- Stone, J. M., & Norman, M. L. 1992, *ApJ*, 390, L17
- Strickland, D. K., & Heckman, T. M. 2009, *ApJ*, 697, 2030
- Strong, A. W., Porter, T. A., Digel, S. W., et al. 2010, *ApJ*, 722, L58
- Sun, M., Voit, G. M., Donahue, M., et al. 2009, *ApJ*, 693, 1142
- Suresh, J., Bird, S., Vogelsberger, M., et al. 2015, *MNRAS*, 448, 895
- Teyssier, R. 2002, *A&A*, 385, 337
- Thom, C., & Chen, H.-W. 2008, *ApJS*, 179, 37
- Thom, C., Tumlinson, J., Werk, J. K., et al. 2012, *ApJ*, 758, L41
- Thompson, T. A., Fabian, A. C., Quataert, E., & Murray, N. 2015, *MNRAS*, 449, 147
- Thompson, T. A., Quataert, E., Zhang, D., & Weinberg, D. H. 2016, *MNRAS*, 455, 1830
- Tilton, E. M., Danforth, C. W., Shull, J. M., & Ross, T. L. 2012, *ApJ*, 759, 112
- Tripp, T. M., Lu, L., & Savage, B. D. 1998, *ApJ*, 508, 200
- Tripp, T. M., Sembach, K. R., Bowen, D. V., et al. 2008, *ApJS*, 177, 39
- Trujillo-Gomez, S., Klypin, A., Colín, P., et al. 2015, *MNRAS*, 446, 1140
- Tumlinson, J., Peeples, M. S., & Werk, J. K. 2017, *ARAA&A*, 55, 389
- Tumlinson, J., Thom, C., Werk, J. K., et al. 2011, *Science*, 334, 948
- . 2013, *ApJ*, 777, 59
- Turner, M. L., Schaye, J., Steidel, C. C., Rudie, G. C., & Strom, A. L. 2014, *MNRAS*, 445, 794
- van de Voort, F., & Schaye, J. 2012, *MNRAS*, 423, 2991

- van de Voort, F., Schaye, J., Altay, G., & Theuns, T. 2012, MNRAS, 421, 2809
- Veilleux, S., Cecil, G., & Bland-Hawthorn, J. 2005, ARAA&A, 43, 769
- Veilleux, S., Rupke, D. S. N., & Swaters, R. 2009, ApJ, 700, L149
- Vogelsberger, M., Genel, S., Sijacki, D., et al. 2013, MNRAS, 436, 3031
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, Nature, 509, 177
- Voit, G. M., Bryan, G. L., O’Shea, B. W., & Donahue, M. 2015a, ApJ, 808, L30
- Voit, G. M., & Donahue, M. 1990, ApJ, 360, L15
- . 2015, ApJ, 799, L1
- Voit, G. M., Donahue, M., Bryan, G. L., & McDonald, M. 2015b, Nature, 519, 203
- Voit, G. M., Ma, C. P., Greene, J., et al. 2018, ApJ, 853, 78
- Wakker, B. P., & Savage, B. D. 2009, ApJS, 182, 378
- Wang, B. 1993, ApJ, 415, 174
- Weiner, B. J., Coil, A. L., Prochaska, J. X., et al. 2009, ApJ, 692, 187
- Werk, J. K., Prochaska, J. X., Thom, C., et al. 2013, ApJS, 204, 17
- Werk, J. K., Prochaska, J. X., Tumlinson, J., et al. 2014, ApJ, 792, 8
- Werk, J. K., Prochaska, J. X., Cantalupo, S., et al. 2016, ApJ, 833, 54
- Werner, N., Allen, S. W., & Simionescu, A. 2012, MNRAS, 425, 2731
- Werner, N., Oonk, J. B. R., Sun, M., et al. 2014, MNRAS, 439, 2291
- White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341
- Wild, V., Kauffmann, G., White, S., et al. 2008, MNRAS, 388, 227

- Wotta, C. B., Lehner, N., Howk, J. C., O'Meara, J. M., & Prochaska, J. X. 2016, *ApJ*, 831, 95
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579
- Zahid, H. J., Yates, R. M., Kewley, L. J., & Kudritzki, R. P. 2013, *ApJ*, 763, 92
- Zel'Dovich, Y. B., & Pikel'Ner, S. B. 1969, *Soviet Journal of Experimental and Theoretical Physics*, 29, 170
- Zemp, M. 2014, *ApJ*, 792, 124
- Zhang, D., Thompson, T. A., Quataert, E., & Murray, N. 2017, *MNRAS*, 468, 4801