The Intergalactic and Circumgalactic Medium surrounding Star-Forming Galaxies at Redshifts $2 < z < 3$

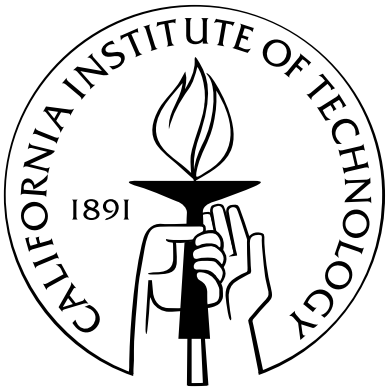
Thesis by

Gwen C. Rudie

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2013

(Defended May 17, 2013)

To those who are curious —
most importantly my parents,
who taught me to always ask why

Acknowledgments

I owe a tremendous debt of gratitude to so many people who have shaped my time at Caltech and my education before I ever reached the hallowed halls of Robinson. Perhaps my greatest debt, and the one I feel least equipped to repay, is to my advisor, Chuck Steidel. Chuck, I can never adequately thank you for taking me on as a first-year student, for trusting me with a project you care so deeply about, and for giving me the space and time to understand, appreciate, and eventually internalize the breadth of the field and the details of my work. It has been my great privilege and honor to be a part of your "group," to learn the craft you so artfully practice. Thank you for your six years of support, patience, and mentoring. I know no person with as great scientific creativity and intuition as you, and I hope that this work is something of which you can be proud.

Along the way, I was fortunate to collect two other scientific mentors, Alice Shapley and Max Pettini, who have steadied me through the rocky moments during graduate school and have taught me a great deal about science and also about friendship. Max and Alice, it has been my great pleasure to have had you both as scientific mentors and collaborators. Max, I'm not quite sure how you did it, but somehow you always seemed to know when I was feeling the most hopeless or frustrated. Your well timed emails — usually just checking in — are responsible in no small way for the completion of this work. And your reminders that the work I was doing was important was sometimes the only thing that kept me "vpfitting" away. I will always be grateful for your friendship, support, patience, and teaching. My visits to Cambridge are among my most warm memories from my Ph.D. years. My thanks also to Pauline for always being so welcoming, opening your home in Cambridge to your stressed out American colleagues!

Alice, I also have you to thank for keeping me sane. I don't fully understand why you agreed to meet me for our monthly dinners during my second year, even offering to schlep over to Pasadena for half of them! I don't know what I could have possibly offered you such that you were willing to commit so much of your time. But I am forever thankful to you for playing the role of my academic big sister, for teaching me the ropes, and for answering my thousands of inane questions about galaxies, graduate school, and the "group." Having someone so accomplished as one of Chuck's former students should have been intimidating but ended up being one of the greatest benefits of signing on with this group. I am forever thankful for your commitment, patience and friendship during these years.

I would be remiss not to thank some of my earlier scientific mentors. I owe more than can possibly be expressed to my first scientific advisor, Dartmouth professor Rob Fesen. Rob fostered in me his great love

of astronomy and of observing. Rob, you showed me what it meant to be an astronomer and instilled in me the confidence to go out and become one myself. Your humor, lightheartedness, and unwavering positivity made my early years in astronomy some of my most enjoyable, and I am immeasurably grateful for your mentorship and support, which have thankfully continued long after my graduation from Dartmouth. I am very proud to have been your student and even more thankful for your friendship.

I am grateful also to Elizabeth Wehner, who gave me my first introduction to IRAF and to data analysis during a summer job at UW Madison. Thank you for investing so much time in a bumbling first year student! Thanks also to Stella Kafka, my REU director from CTIO, who admitted me to the program and supported me scientifically throughout it. Stella also proved to be an excellent nurse and stand-in parent when I was injured in Chile. I am grateful for your mentoring and support and also for your friendship, all of which have extended long past the few months of my internship. I'd also like to thank my first observing partner, Dartmouth graduate student Danny Milisavljevic, for contributing humor, lightheartedness, and friendship to my early astronomical endeavors.

I am similarly grateful to my Dartmouth professor from first year physics, Barrett Rogers. I'm thankful not only for the excellent education you gave me in physics, but also for your expression of confidence in my abilities to complete a major in physics. You loaned me the confidence needed for me to pursue this degree, and for that I will always be grateful. I'd also like to express my gratitude to my most dedicated high school teachers, Greg Fleenor, Molly Spier, and Joan Dutter, who were responsible for my introduction to written argument, calculus, and physics. You all taught me as much about commitment, character, and creativity as the subjects of your lectures, and hopefully I am a better scientist and human being for your efforts.

In my 6 years at Caltech, I have benefited greatly from many of the staff at both Caltech and at Keck Observatories. Many thanks to Gita Patel, Gina Armas, Althea Keith, Judith Mack, and Sandy Lester for their cheerful overseeing of so many of the details of my time at Caltech. You all have not only made my life easier, but also happier, and I thank you for that. Thanks especially to Judy McClain, who has done so, so many things for me during these years and has always been supportive, kind, and concerned for my well-being. I'll miss seeing you often and miss our chats. My sincere thanks also to the unequaled computer help group in the Astronomy department. Patrick Shopbell, Anu Mahabal, and Jose Calderon, thank you all for your patient assistance. You are in no small way responsible for any degree of productivity that occurs in our department, and I am very grateful for the help you have given me in these six years. I'll miss not only your extreme competence and abilities, but also your humor and kindness. This Ph.D. would not have been possible without the competence and dedication of the staff of the Keck Observatory. I am grateful especially to Greg Wirth and Marc Kassis for their persistent commitment and their unfailing affability. Returning to Caltech, special thanks go to Efrain Hernandez, whose frequent visits and chats always brightened my day. I don't know what I did to deserve your friendship, but I am better for it, and for that I am grateful. I was fortunate during my second year to have met by chance Michelle Vine, whose extreme kindness, loving personality, and disarming openness have greatly enriched my time at Caltech. You, Michelle, are a rare and

wonderful friend.

I was privileged to be part of the MOSFIRE instrument team during my time at Caltech. Thanks to the whole team, but especially John Cromer, Hector Rodriguez, Krisitin Kulas, Ryan Trainor, and Greg Mace who brought humor to some otherwise very serious meetings. Thanks also to Ian McLean for never failing to celebrate the accomplishments of the team. I'm grateful especially to Nick Konidaris, who taught me literally everything I know about optics, and who shared many memorable conversations with me and with Brian Siana during my early education as a graduate student. I'm thankful for his humor and for his friendship.

A thesis is not possible without a thesis committee, and I am very thankful for mine: Chuck, Chris Hirata, Lynne Hillenbrand, Chris Martin, Juna Kollmeier, and candidacy committee member Wal Sargent. Wal was taken from us too early to sit on my defense committee, but my time at Caltech was much more enjoyable for his dry humor and none-too-serious presence. It is a great honor to be his academic granddaughter, and I miss him very much. Chris Hirata, thanks for chairing both committees, for always having an open door — making yourself available for questions and advice. I, and I believe all the graduate students, very much appreciate your commitment to the students during your time as option rep. I will be very sad next year when I return to Caltech and don't see you in the halls. Lynne, thank you for all of your advice and assistance over the years. It was great fun being the grad rep with you as chair and I am thankful for all of the opportunities I have had to interact with you over my years at Caltech. Thanks to Juna, not only for stepping in as a committee member, but also for her advice and help during my Ph.D. I'm looking forward to closer collaboration (and more chats) during my upcoming time at Carnegie. Though not on my committee, thanks nonetheless go to John Johnson. Thanks always for having an open door and for taking the time to listen and to argue with me. I appreciate very much your advice and your friendship — and I will miss you next year when you move to Harvard.

I was fortunate to arrive at Caltech with a great group of classmates. Drew Newman, Vera Guscevic, Mike Anderson, Tucker Jones, Sam Lee, Tim Morton, Elisabeth Krause, and Krzysztof Findeisen: thanks for your humor, help, and support - especially during the trials and turbulence of first year. Thanks especially to Vera and Tim for fun times playing cards, chatting over dinner, and planning a wedding. It's been my great privilege over the years to share an office with Laura Perez, Varun Bhalerao, Kunal Mooley, Jackie Villadsen, and Sebastian Pineda. Thanks especially to Laura and Varun who took me in second year when I had nowhere to sit in Robinson, provided me with wise and timely advice more times than I can recall, and were generally supportive, loyal, and wonderful friends — more than I deserved! I'm grateful also to Sarah Miller, a wonderful friend and scientific co-conspirator. Our chats and brain-storming sessions were always great fun — and the dinners with you and Oliver King were even better! Thanks also to my other constant dinner companions, Sonja Graves and Michelle Selvans. I'm so grateful to have met you both at Caltech and am looking forward to many more years of friendship! I have been lucky in my final years at Caltech to gain new friends in Allison Strom, Antonija Oklopcic, Mislav Balokovic, and Ben Montet. You guys have made the last two years a lot of fun — and I'm glad to be staying in Pasadena for a few more years of galavanting!

Over the years, I was lucky to benefit from the friendship of postdocs Annika Peter, Samaya Nissanke, and Carrie Bridge whose persistence, positivity, and generosity have been an inspiration and a true example of the strength and depth of women in science. Many thanks to postdoc Brian Siana who patiently took me under his wing, answering many of the obvious and occasionally more complicated questions of an early graduate student. Brian, thank you also for confidence in me and for your friendship. On my visits to Cambridge, I was lucky to gain the collaboration and friendship of Ryan Cooke. He and wife Alis have been such a welcome and enriching addition to my academic life.

I'm grateful for all of Chuck's former graduate students who accepted me into their clan and brought me up to speed. Thanks especially to Alice, Naveen, and Milan who were particularly welcoming and helpful. "Group" meetings have been more interesting and enjoyable when Jean-Rene Gauthier, Brian Siana, Yuichi Matsuda, or Nick Konidaris were present. Special acknowledgement goes to my closest collaborator and very good friend Olivera Rakic. I'm inexpressibly thankful for your encouragement and unwavering support during these years. I'm so sad to no longer have you as a collaborator but very blessed to have you as a very dear friend. I've been very grateful to have Ryan Trainor and Allison Strom as collaborators and friends. You guys have both brightened my days (and observing runs!) Ryan, I am always amazed and inspired by your unending humor and positivity, something I aspire to emulate — if even in some very small way. Allison, it has been such fun getting to know you and teaching you the little that I can about this field. I look forward to many more scientific and real-world adventures with you.

Words cannot express the depth of my gratitude or my love for Drew Newman. I would have never made it through these six years without your love, support, and kindness — all of which are so much more than I deserve. I have you to thank for nearly all of the joy during these years. You are truly my better half.

And finally, I have the impossible task of thanking my parents Cindy and Alan, and my brother Ian. Without their support surely none of this would have been possible. Mom and Dad, I don't know how I have been so lucky to have such loving, loyal, brilliant, and unendingly kind parents. You all instilled in me very early the love and wonder I have for science. You always pushed me to be my best possible self, cheered me when I succeeded, and loved me seemingly more when I failed. My greatest hope is that this effort, and the life that I lead, make you proud. And that even with the miles of space between us, and the sparsity of the time when we can be together, I hope that you know the depth of meaning you have brought to my life — both as my parents and as my friends. I thank you most of all.

Abstract

We present measurements of the spatial distribution, kinematics, and physical properties of gas in the circumgalactic medium (CGM) of $2.0 < z < 2.8$ UV color-selected galaxies as well as within the $2 < z < 3$ intergalactic medium (IGM). These measurements are derived from Voigt profile decomposition of the full Lyα and Lyβ forest in 15 high-resolution, high signal-to-noise ratio QSO spectra resulting in a catalog of ~ 6000 H I absorbers.

Chapter 2 of this thesis focuses on H I surrounding high-z star-forming galaxies drawn from the Keck Baryonic Structure Survey (KBSS). The KBSS is a unique spectroscopic survey of the distant universe designed to explore the details of the connection between galaxies and intergalactic baryons within the same survey volumes. The KBSS combines high-quality background QSO spectroscopy with large densely-sampled galaxy redshift surveys to probe the CGM at scales of ~ 50 kpc to a few Mpc. Based on these data, Chapter 2 presents the first quantitative measurements of the distribution, column density, kinematics, and absorber line widths of neutral hydrogen surrounding high-z star-forming galaxies.

Chapter 3 focuses on the thermal properties of the diffuse IGM. This analysis relies on measurements of the ~ 6000 absorber line widths to constrain the thermal and turbulent velocities of absorbing "clouds." A positive correlation between the column density of H I and the minimum line width is recovered and implies a temperature-density relation within the low-density IGM for which higher-density regions are hotter, as is predicted by simple theoretical arguments.

Chapter 4 presents new measurements of the opacity of the IGM and CGM to hydrogen-ionizing photons. The chapter begins with a revised measurement of the H I column density distribution based on this new absorption line catalog that, due to the inclusion of high-order Lyman lines, provides the first statistically robust measurement of the frequency of absorbers with H I column densities $14 \lesssim \log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 17.2$. Also presented are the first measurements of the column density distribution of H I within the CGM ($50 < d < 300$ pkpc) of high-z galaxies. These distributions are used to calculate the total opacity of the IGM and IGM+CGM and to revise previous measurements of the mean free path of hydrogen-ionizing photons within the IGM. This chapter also considers the effect of the surrounding CGM on the transmission of ionizing photons out of the sites of active star-formation and into the IGM.

This thesis concludes with a brief discussion of work in progress focused on understanding the distribution of metals within the CGM of KBSS galaxies. Appendix B discusses my contributions to the MOSFIRE instrumentation project.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The study of the diffuse gas in between galaxies has a long history that is intimately connected with the history of astronomy at Caltech. Our understanding of the structures that imprint the dark bands against bright blue background sources has evolved significantly in the years since 1962 when Maarten Schmidt used the mighty 200-inch Hale Telescope on Palomar Mountain to record the first spectrum of the radio source 3C 273 and correctly identified the strange spectral emission features it presented to be the Balmer lines of the hydrogen atom, redshifted by 15.8% (Schmidt 1963). This discovery led to others such as the first robust determination in 1980 by Wal Sargent and collaborators that the dark absorption features seen against all high-z QSOs were due to absorption from intergalactic gas - gas at great distances from the QSO itself. Tireless efforts by these greats of the field, as well as many others, have paved the way for joining our understanding of the diffuse gas that lies between galaxies - and the formation and evolution of galaxies themselves. That is the topic of this thesis.

1.1 The Intergalactic Medium

The highly-ionized, diffuse gas that lies in the space between galaxies is called the intergalactic medium (IGM). Because the majority of the gas is diffuse, with densities of $n_{\rm H} \approx 10^{-5} - 10^{-3}$ cm^{-3} (see e.g. Schaye 2001), it is particularly challenging to study in emission. Therefore, most of our understanding of the IGM comes from the detection of gas through absorption line spectroscopy. In principle, any background source of light can be used to study the IGM, however, due to their extreme luminosities, flat spectral slopes, and relatively featureless intrinsic spectrum, QSOs have been used for most general IGM studies, although Gamma Ray Bursts (GRBs, see e.g. Jensen et al. 2001; Prochaska et al. 2007; Fynbo et al. 2009) and background galaxies (Adelberger et al. 2005a; Steidel et al. 2010; Rubin et al. 2010) have also been employed in various studies.

The most prominent IGM features in the spectrum of high-redshift sources are due to the small fraction of the gas which is neutral; H I in the IGM imprints upon the spectrum of background sources a series of resonant absorption features due to Lyman series transitions from gas at a large range of redshifts. Collectively, these

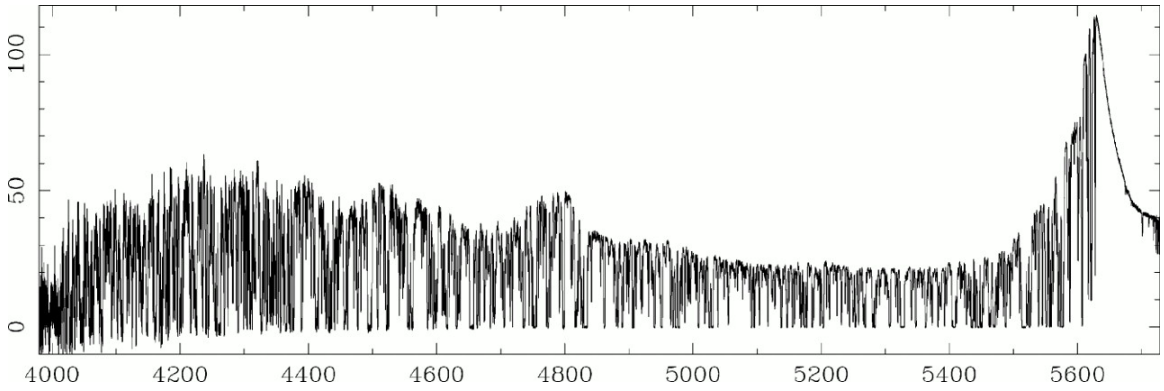


Figure 1.1 A high-resolution spectrum of the QSO Q1422+23 at z=3.62 taken with Keck HIRES. Figure from Rauch (1998). The broad emission feature at 5600Å is Lyα emission from the QSO. The feature at 4800Å is due to O VI emission. The absorption features at wavelengths shortward of Lyman α are predominantly due to H I absorption in the Lyman α forest.

features are commonly referred to as the Lyman α forest. Figure 1.1 shows a high-resolution spectrum of a high-z QSO and the foreground Lyα forest it traces.

Observed with high-resolution spectrographs, the absorption features in the Lyα forest separate into discrete lines (see e.g. Carswell et al. 1984) with shapes that are well characterized by Voigt profiles. These features can then be fit in order to obtain the parameters of the cloud of gas giving rise to the absorption: namely its redshift, column density, and velocity width. The velocity width, as discussed at length in Section 2.7 and Chapter 3, provide a measure of the kinetic energy in the gas through a combination of thermal and turbulent broadening. The Lyα forest imprinted upon the spectrum of each high-z QSO has hundreds of discrete absorption systems, and so by carefully studying a small number of high-quality QSO spectra, one can learn a great deal about the nature of the Lyman α forest (Sargent et al. 1989; Petitjean et al. 1993; Davé & Tripp 2001; Kim et al. 2002; Penton et al. 2004; Janknecht et al. 2006). This thesis implements the same line-fitting techniques used by previous authors on QSO spectra with improved wavelength coverage and signal-to-noise ratios, to revisit the statistics of the Lyman α forest in Chapters 3 and 4. In Chapter 2, we apply our understanding of the IGM to the space around galaxies - the circumgalactic medium (CGM).

The current theoretical understanding of the IGM suggests that the gaseous density fluctuations present within the IGM simply mirror the large scale density fluctuations present in the underlying dark matter (Tytler 1987; Cen et al. 1994; Hernquist et al. 1996; Schaye 2001). This "cosmic web" of gas sheets and filaments (see Figure 1.2) is therefore expected to be a high-fidelity tracer of the density distribution. For this reason, it is straightforward to see that some parts of the Lyman α forest, those with higher density, would coincide with the overdensities in which galaxies form and evolve. Other evidence pre-dating the modern theory of the IGM has also suggested a link to galaxies - the presence of metals.

In addition to the Lyman α forest, there are many other transitions that appear in QSO spectra which arise from a variety of metallic transitions of C, O, Si, N, Fe, Mg, etc. These transitions encompass a wide range of ionization states, from neutral to highly ionized species. Typically the neutral and low-ions (e.g.

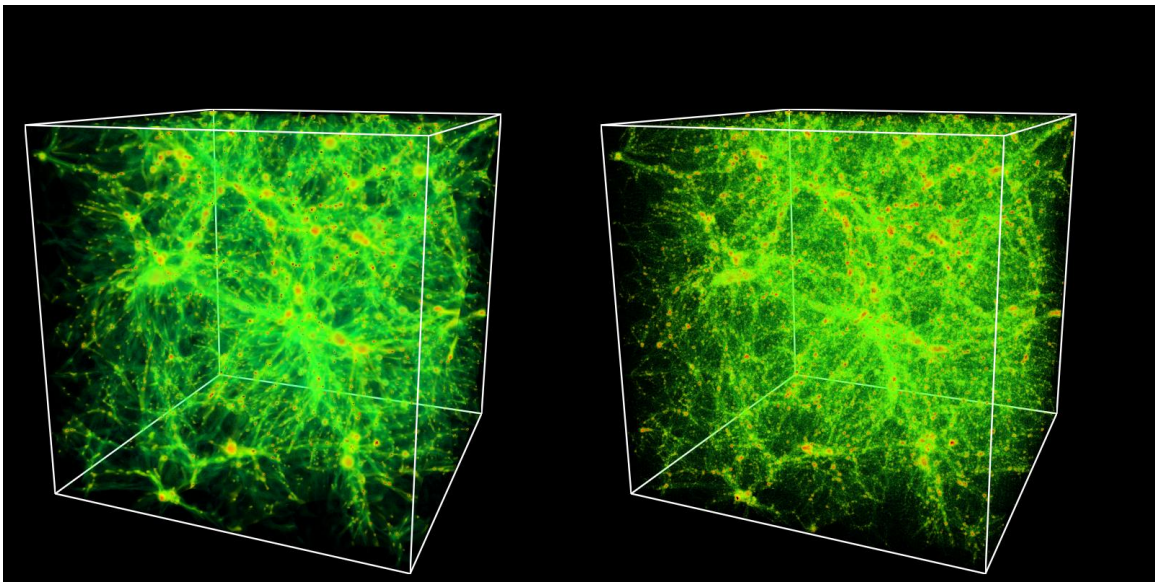


Figure 1.2 Simulations of the "cosmic web." The cubes are 100 h^{-1} Mpc on a side seen at $z = 0$. The left cube shows the gas density while the right cube shows the dark matter density. Figures from Cen & Ostriker (2006).

O I, C II, Si II) appear at the same redshifts as the systems with very high H I column density ($N_{\rm HI}$) inside which they are shielded from the UV background. Higher ionization species (e.g. C IV, O VI) are commonly found in both high-$N_{\rm HI}$ and intermediate-$N_{\rm HI}$ gas as they do not require that their surrounding H I cloud be self-shielding.

The existence of metal absorption in intergalactic space did not go un-noticed, and early investigators of the IGM often commented on the necessary connection between IGM metal absorbers and galaxies (Young et al. 1979; Sargent et al. 1980). In 1978, Boksenberg & Sargent found the first direct evidence for their association. In the spectrum of 3C 232, they discovered calcium absorption lines at the same redshift as the $z \simeq 0.005$ spiral galaxy NGC 3067 separated by 2 arcminutes from the QSO. The galaxy had previously been observed in 21 cm, and the Ca II H and K lines showed the same velocity offset as the neutral H I. Since that time, there have been many additional studies of the presence of H I and various metallic species in the space surrounding galaxies and this thesis presents the first truly statistical study of circumgalactic gas at high redshift using line fitting to derive the physical properties of the CGM.

1.2 The Formation and Evolution of Galaxies

While significant progress has been made in reproducing the observed large-scale structure of the universe with our adopted concordance cosmology, there remain many unsolved problems related to processes acting on smaller scales. One of the most fundamental puzzles is how star formation is regulated. A clear prediction of ΛCDM is the rapid growth of mass within halos, requiring the accretion of both dark matter and gas.

However, it appears that on all scales, from molecular clouds to full galaxies, the conversion of gas into stars is inefficient (Schmidt 1959; Kennicutt 1998). There is growing evidence that star formation may be most efficient in dark matter halos of characteristic mass $M_{halo} \approx 10^{12}$ $M_{\odot}$ (Moster et al. 2010) which seems to be relatively constant with time (Behroozi et al. 2012). But even in such halos, less than 15% of baryons have been converted into stars by the present day.

Because it far outweighs the baryonic mass, dark matter governs the initial gravitational collapse of structure, but dark matter alone cannot explain the observed properties of galaxies. Notably, ΛCDM N-body simulations produce a steep power-law mass function of dark matter halos in contrast to the observed Schechter (1976) function characteristic of the stellar mass distribution of galaxies which has a much shallower slope at low mass and an exponential decline at high mass (e.g., Moster et al. 2010, see Figure 1.3). There must be some baryonic physics that suppresses star formation at all scales, and these processes must be more effective at low and high mass.

The often-invoked solution to this problem is feedback: the injection of energy or momentum into the ISM of a galaxy from the combined effects of star formation, supernovae, and AGN. Clearly, the large discrepancy between the halo and stellar mass function requires that this feedback ejects or continuously heats a large fraction of the ISM to prevent the overproduction of stars. Order-of-magnitude calculations suggest that sufficient energy is released in the combined explosions of massive stars or the build-up of a central black hole to unbind a large fraction of the ISM. Unfortunately, understanding the detailed physical processes by which these localized events couple to the cold ISM to produce galaxy-scale outflows remains elusive despite substantial theoretical efforts (e.g., Springel & Hernquist 2003; Dalla Vecchia & Schaye 2008; Murray et al. 2010; Davé et al. 2011; Faucher-Giguère & Quataert 2012; Hopkins et al. 2012). Fortunately, different feedback mechanisms are expected to produce unique signatures in the gas surrounding galaxies such as specific chemical abundance patterns. And because these mechanisms depend on different physical drivers (radiation pressure, thermal or kinetic SNe feedback, AGN feedback), they are also expected to have specific dependancies on galaxy properties (such as metallicity, star-formation rate, and mass).

1.2.1 Outflows

There now exists substantial evidence that galactic winds are exceedingly common in star-forming galaxies across all spectroscopically observable redshifts (see Veilleux et al. 2005, for a review). Heckman (2002) observed that such winds are present in any galaxies with a star-formation surface density that exceeds ~ 0.1 $M_{\odot}$ yr^{-1} kpc^{-2}. In the most local galaxies, the star-burst generated winds from galaxies such as M82 (Bland & Tully 1988; Heckman et al. 1990; Devine & Bally 1999; Lehnert et al. 1999) have been observed and studied in great detail from radio (Kronberg et al. 1985) to TeV (Lacki et al. 2011) wavelengths. At higher redshifts, the most common evidence for outflows seen in the spectra of nearly all star-forming galaxies are blue-shifted absorption features. Such features are observed in transitions due to gas at a variety of ionization states. Because these absorption features are seen against the continuum light of the galaxy (which is due to

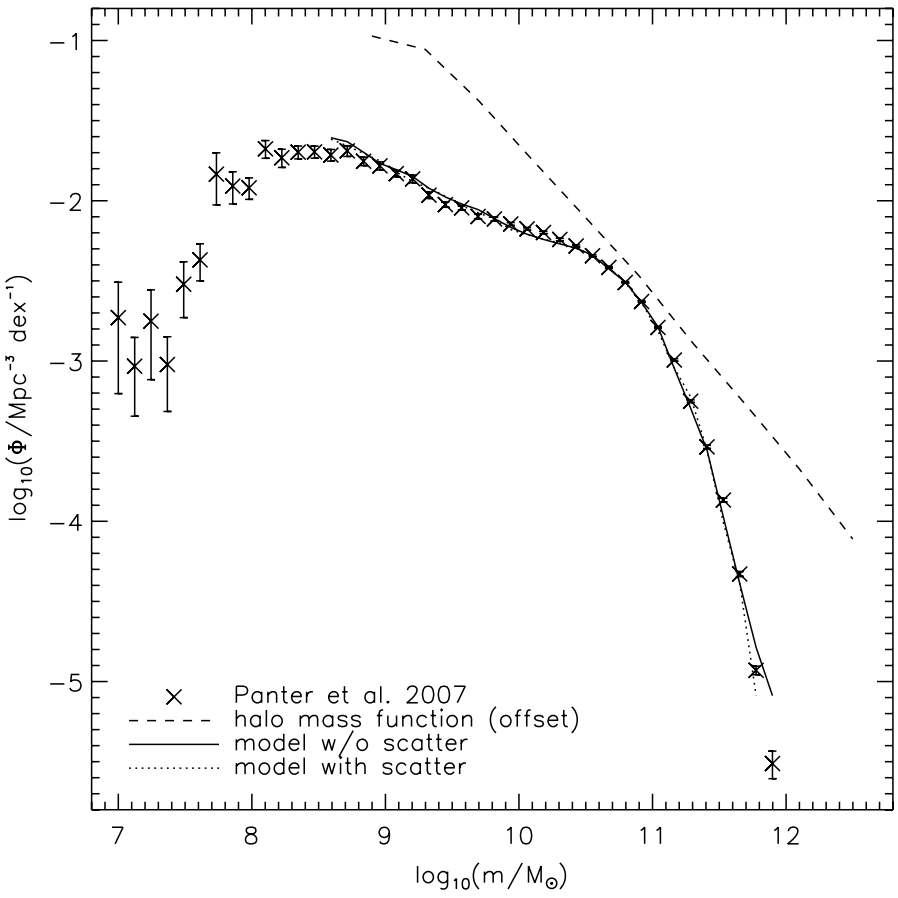


Figure 1.3 The mass function of galaxies as measured from the Sloan Digital Sky Survey (SDSS) by Panter et al. (2007) is shown with symbols. The straight dashed line is the mass function of dark matter halos produced by ΛCDM N-body dark matter simulations, offset by a factor of 0.05. The extreme discrepancy in both the shape and the normalization of these two distributions leads one to infer that some baryonic process impedes the formation of stars within all halos, and that these processes must be more efficient within the lowest and highest mass halos. Figure from Moster et al. (2010).

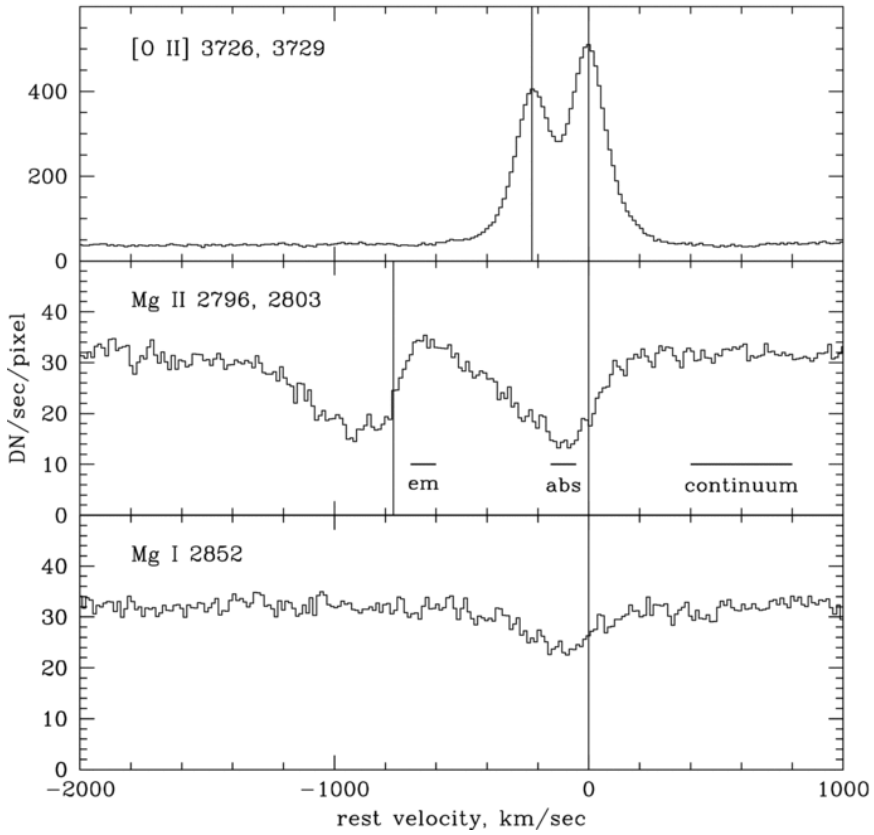


Figure 1.4 Portions of a stacked spectrum of 1406 $z \sim 1.4$ galaxies surrounding the [O II] $\lambda\lambda 3726.0, 3728.8$, Mg II $\lambda\lambda 2795.5, 2802.7$, and Mg I $\lambda 2852.1$ spectral features. Here, the redshift of the [O II] doublet, which arrises from emission within star-forming regions, is taken as the systemic redshift of the galaxy. The Mg II and Mg I absorption are seen systematically blue shifted, suggesting outflows are common in such systems. Figure from Weiner et al. (2009).

star light), the gas that is causing the absorption features must lie between the observer and the stars. As the features are seen blue shifted, the gas is coming toward the observer and away from the stars, and hence the gas is outflowing from the system. A study of $0.042 < z < 0.16$ ultra-luminous infrared galaxies (ULIRGs) by Martin (2005) found 330 km s^{-1} blue-shifted Na I absorption in 15 out of 18 galaxies studied. A study of an average spectrum of over 1400 $z \sim 1.4$ galaxies drawn from the DEEP2 survey found Mg II absorption blue shifted by $300-500$ km s^{-1} (Weiner et al. 2009, see Fig 1.4). Studies of UV color selected galaxies (such as those studied in this thesis) at $2 < z < 3$ (Steidel et al. 2010) and $z \gtrsim 3$ (Pettini et al. 2001; Shapley et al. 2003) rely on far-UV interstellar absorption features such as low ions Si IIλ1260, O I+ Si IIλ1303, C IIλ1334, Si IIλ1526, Fe IIλ1608 and Al IIλ1670 and high-ions O VI$\lambda\lambda$10321038, N V $\lambda\lambda 1238, 1242$, Si IV $\lambda\lambda 1393, 1402$, C IV $\lambda\lambda 1548, 1550$. All of these features are seen blue-shifted compared to tracers of stellar emission, such as C IIIλ1176, or of H II regions, such as [O II]$\lambda\lambda 3726, 3728$, [O III]$\lambda\lambda 4958, 5006$, and H$\alpha\lambda$6563. Jones et al. (2012) recently extended similar analyses to galaxies at $z \sim 4$, showing that winds were common during this early epoch as well.

Our understanding the physical drivers of such winds is comparatively poor. At intermediate redshift $z \sim 1$, galaxies drawn from the DEEP2 survey that were viewed face-on were found to have stronger absorption features present in their spectra than those galaxies viewed at higher inclination angles (Kornei et al. 2012), suggesting that galaxies as early as $z \sim 1$ may exhibit biconical outflows similar to those observed in the local universe. They also found a trend between the star-formation rate surface density and the speed of the outflowing winds. Wind speeds have also been shown to increase with increasing stellar mass (Erb et al. 2012), as well as the star-formation rate (Martin 2005) although few of these trends are seen at high significance in large data sets.

1.2.2 Accretion

Another important component to understanding the discrepancies between the halo and stellar mass functions is the process of mass accretion. This subject has gained considerable attention in the last decade, especially from theorists, as the theoretical understanding of this problem has significantly changed. The previous paradigm, largely the result of simplified calculations that relied on assumptions of spherical symmetry, supposed that all gas as it fell into the potential well of a galaxy would be shocked to the viral temperature of the halo. In this scenario, the efficiency with which star formation would proceed in such a halo would depend significantly on the cooling time for gas heated to temperatures well above 10^5 K (Rees & Ostriker 1977). More recent work by Birnboim & Dekel (2003) considered the stability of such shocks and found that when gas cooling within the halo is efficient compared to the rate of gas infall, the shocks' pressure support fails, resulting in less heating of infalling gas within the halo.

Other recent efforts that have relied on the interpretation of large hydrodynamic simulations have shown that the morphology of accreting gas is far from spherical (Kereš et al. 2005, see Figure 1.5). Instead, simulations of cosmological volumes seem to universally show that galaxies are fed by the filaments and sheets that make up the classic "cosmic web" - the same structures thought to give rise to the IGM absorbers. For low-mass and low-z galaxies which may be imbedded in a filament, this picture is not appreciably different from spherically symmetric inflow; however, massive high-z galaxies in these simulations are generally found at the intersections of such filaments. Further, these studies have shown that the gas infalling along such filaments may be highly collimated and therefore may escape many of the heating processes (such as shocks) which may occur in the halos of such galaxies and may affect more diffuse gas (Kereš et al. 2005; Ocvirk et al. 2008). The properties and presence or absence of such structures, referred to as "cold flows," have been invoked to solve such problems as the source of angular momentum in disk galaxies (Brooks et al. 2009; Agertz et al. 2009; Stewart et al. 2013), the evolution of the cosmic star formation with redshift (van de Voort et al. 2011a), the quenching of star formation within massive galaxies (Dekel & Birnboim 2006, 2008), and the origin of the Milky Way's high velocity clouds (HVCs, Kereš & Hernquist 2009).

While this theory is indeed attractive for solving a diverse array of outstanding problems in astrophysics, it suffers from two principal weaknesses: (1) most simulations which show cold flows have rudimentary (or

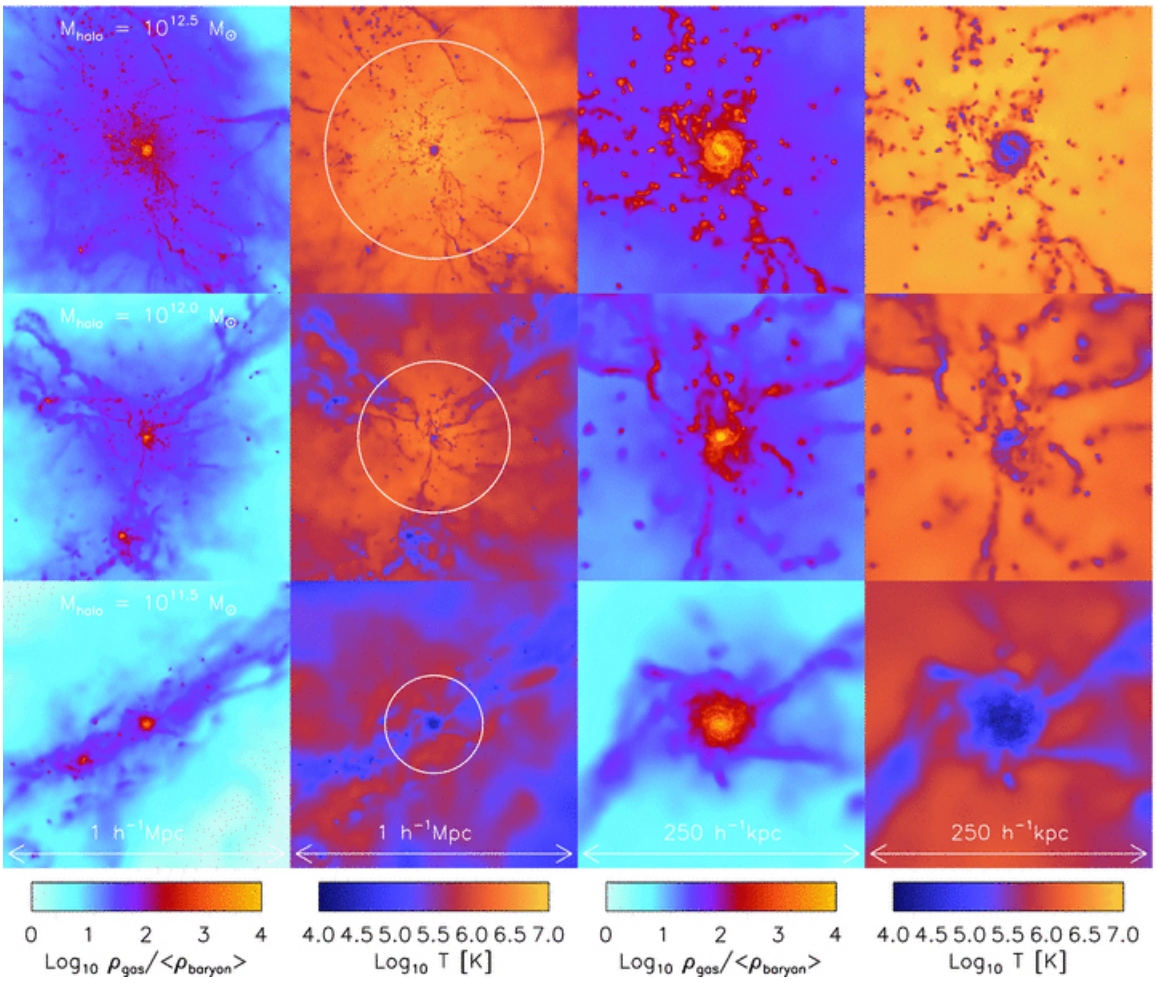


Figure 1.5 The gas overdensity (first and third columns) and temperature (second and fourth columns) in $z = 2$ halos of three different masses drawn from cosmological simulations. The white circles give the locations of the virial radius of the halo. Note the presence of cool, dense, filamentary gas, especially in halos of $M < 10^{12}$ $M_{\odot}$. Figure from van de Voort et al. (2011b).

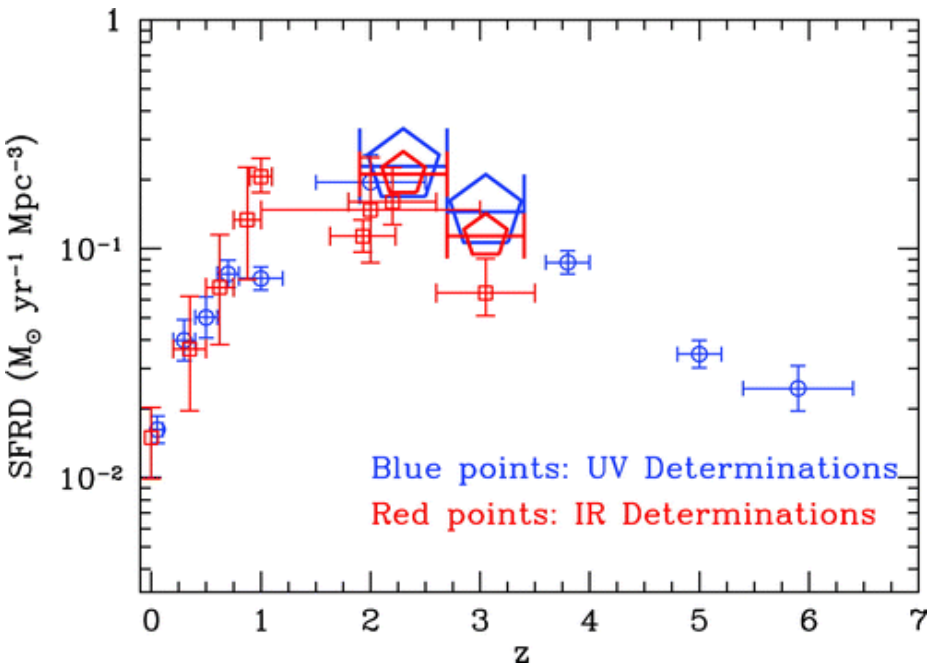


Figure 1.6 The star formation rate density of the universe as a function of redshift. Note that there is a broad peak at redshifts $2 < z < 3$. Figure from Reddy et al. (2008).

no) prescriptions for feedback and outflows which are likely to significantly affect the physical properties of infalling gas (Faucher-Giguère et al. 2011; Stewart et al. 2011a; Nelson et al. 2013) and (2) there currently exists relatively little observation evidence of cold flows and therefore essentially no observational constraints on their nature.

Arguably, the best epoch in which to study both baryonic outflows and accretion is $2 < z < 3$ during the peak of cosmic star formation (Madau et al. 1996; Hopkins & Beacom 2006; Reddy et al. 2008, see Figure 1.6) and super-massive black hole growth (Richards et al. 2006). At these redshifts, spectroscopic observations of star-forming galaxies universally exhibit signatures of strong outflowing winds (Pettini et al. 2001; Shapley et al. 2003). At the same time, the baryonic accretion rate onto galaxies is predicted to be near its peak (e.g., Faucher-Giguère et al. 2011; van de Voort et al. 2011a). As such, we expect the signatures of baryonic flows to be most readily observable at $2 < z < 3$.

1.3 The Circumgalactic Medium

Absorption line spectroscopy of bright background sources provides a uniquely sensitive probe of the physical properties and chemical composition of gas in close proximity to galaxies - gas that traces processes such as galactic winds and baryonic accretion.

1.3.1 The Low-Redshift CGM

There have been a plethora of searches for galaxies associated with low-redshift ($z \lesssim 1$) strong MgII, CIV, and Lyα systems. In Lyα, it seems generally that strong HI absorption is common near gas-rich (Ryan-Weber 2006; Chen & Mulchaey 2009) or bright galaxies of various types (Chen et al. 1998), and that the strength (Chen et al. 1998; Bowen et al. 2002; Wilman et al. 2007) or prevalence (Lanzetta et al. 1995; Morris & Jannuzi 2006) correlates with distance from the galaxy albeit with non-unity covering fraction (Chen et al. 1998). However, there is little consensus as to how directly related the absorption is to the galaxies - some favor the majority of the HI resulting directly form large scale structure (Penton et al. 2002), while others suggest a more direct connection with galaxies (Chen et al. 2001b). Recent results from the *Cosmic Origins Spectrograph* (COS) on the *Hubble Space Telescope* (HST) by Thom et al. (2012), suggest high column-density H I is commonly found within 150 kpc of both star-forming and passive systems where they define passive galaxies as those with specific star-formation rates (sSFR, SFR/M_*) <10^{-11} $M_\odot$ yr^{-1}. The majority of the gas is observed at velocities consistent with it being bound to the system, while the measured line widths suggest the gas is cool with $T_{\rm gas} < 2 \times 10^5$ K (the virial temperature of such halos is $T_{\rm vir} > 10^6$ K).

From the low-z metal line studies there is some agreement that strong Mg II absorbers[1] as well as C IV systems are commonly found near (and may or may not be physically associated with) bright ($L \gtrsim 0.1L^\star$) galaxies (Bergeron & Boissé 1991; Steidel et al. 1994, 1997) with distances up to 200 kpc (Zibetti et al. 2007) and non-unity covering fractions (Churchill et al. 2005; Gauthier et al. 2010). The properties of the galaxies and the physical origin of the gas seen is still debated, with some evidence that the strength of absorption may correlate with recent star formation (Zibetti et al. 2007; Bouché et al. 2007; Ménard et al. 2009; Nestor et al. 2010; Chelouche & Bowen 2010), previous generations of star formation (Steidel et al. 2002), stellar mass (Steidel et al. 1994; Chen et al. 2010), and/or distance from the galaxy in question (Lanzetta & Bowen 1990; Chen et al. 2001a; Kacprzak et al. 2011; Nielsen et al. 2012; Churchill et al. 2013). In all likelihood, all of these elements are at work. A more recent study by Werk et al. (2012) considers the low and intermediate ions within the COS-Halos program, finding significant evidence of bound, cool, metal-enriched gas within the CGM of both actively star-forming and passive galaxies.

Bordoloi et al. (2011) studied the distribution of Mg II absorbers surrounding ~ 4000 galaxies at $0.5 < z < 0.9$ using background galaxies as probes. They found that bluer (presumably star-forming) galaxies had a much higher equivalent width of Mg II at small distances compared to redder galaxies. They also found that within 50 kpc, systems probed along the disk axis of the system showed weaker absorption than those probed at small azimuthal angles, see Figure 1.7. They interpret this as evidence that strong Mg II systems more commonly occur in biconical outflows. However, Kacprzak et al. (2011) considered a much smaller sample using background QSOs, reaching the opposite conclusion - that Mg II is more commonly found in coplanar geometries. Stewart et al. (2011b) suggested that orbiting coplanar CGM gas could be a detectable signature of cold streams, even suggesting in Stewart et al. (2013) that such streams could be the principal

[1] $W_0^{MgII} \gtrsim 0.3$ has been shown to trace optically thick H I gas with $N_{\rm HI} \gtrsim 10^{17}$cm^{-2} (Steidel 1992).

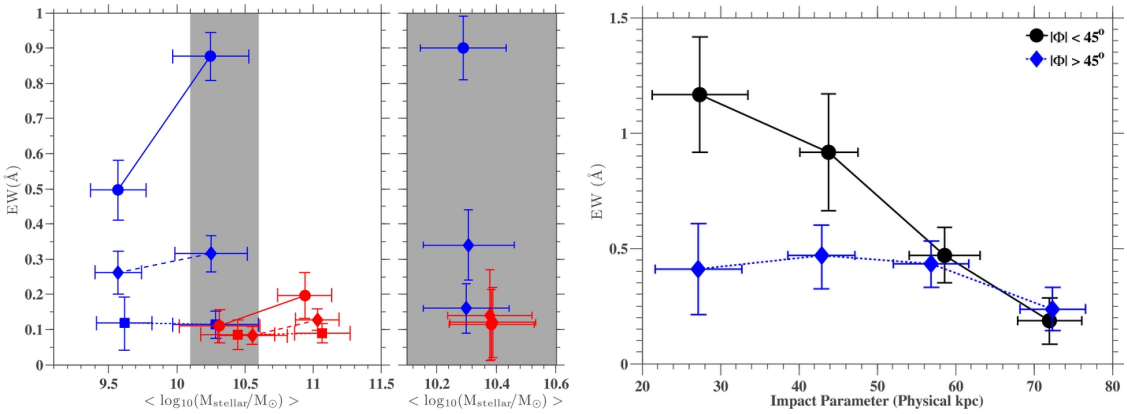


Figure 1.7 The dependence of the equivalent width (W_0) of Mg II absorption in the CGM of $0.5 < z < 0.9$ galaxies in the zCOSMOS survey. *On the Left:* The dependance of the strength of absorption with stellar mass for blue and red galaxies in different bins of galactocentric distance. Notably bluer galaxies showed larger W_0 than redder galaxies at fixed impact parameter. At the smallest distances, the strength of absorption correlates with stellar mass, especially in blue galaxies. *On the Right:* The azimuthal dependence of the fall of of equivalent with of Mg II absorption as a function of galactocentric distance from disk galaxies. The stronger W_0 seen at small azimuthal angles suggest that strong Mg II absorption could be associated with wind material in a biconical outflow. Figures from Bordoloi et al. (2011).

source of angular momentum within disk galaxies.

One resolution to this discrepancy posits that absorbers with $W_0 > 1$ Å may result from outflows, while weaker absorbers may be due to accretion (Kacprzak & Churchill 2011). Considering blue star-forming galaxies with associated Mg II absorption with $W_0 > 0.1$ Å, Kacprzak et al. (2012) found that the distribution of azimuthal angles is in fact bimodal, with a tendency for gas to lie projected along the plane of the disk or along the opening angle of the disk, with relatively little detected gas at intermediate angles, again supporting the claim that Mg II absorption in low-z systems likely traces both outflowing and accreting/orbiting material. A recent study of $z < 1$ Lyman Limit systems (LLS[2]) from the general IGM found a bimodality in the metallicity distribution of high-$N_{\rm HI}$ absorbers, again suggesting that absorbers in the low-z universe may trace multiple phenomena (Lehner et al. 2013).

Another ion of particular interest is O VI. The presence of O VI is thought to indicate current activity (such as accretion shocks and outflows) in the CGM because gas at temperatures that maximize O VI/O ($T \approx 10^{5.5}$ K) has a very short cooling time. O VI may also be a unique tracer of the boundary between the hot wind fluid observed locally in X-rays and the cool 10^4 K gas observed in Hα emission and lower-ionization absorbers believed to be entrained in the fluid. In another recent study using COS, Tumlinson et al. (2011) measured the distribution of O VI absorbers within 150 kpc in the CGM surrounding 42 galaxies with $0.10 < z < 0.32$ as part of the COS-Halos program. They found a remarkable bimodality in the observed column density of O VI depending on the level of current star formation in the galaxy (see Figure 1.8). With simple geometric and ionization constraints, they argue that the mass in the CGM within 150 kpc of star-forming low-z galaxies is

[2]LLSs are H I absorber with $N_{\rm HI} > 17.2$.

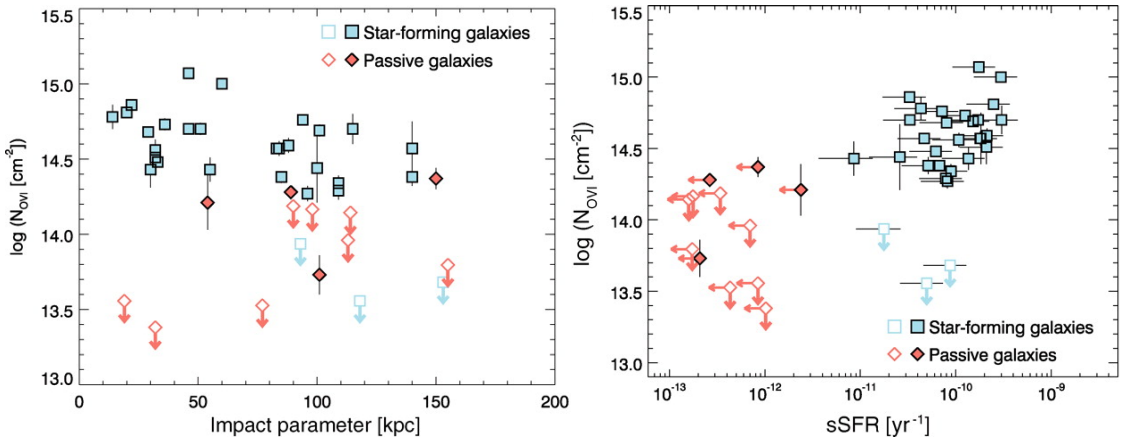


Figure 1.8 The column density of O VI in the CGM of 42 $0.10 < z < 0.32$ galaxies from the COS-Halos program. *On the Left:* N_{OVI} plotted as a function of the impact parameter, the distance between the galaxy and the QSO line of sight at the redshift of the galaxy. There is no strong trend detected within this sample, in contrast to the behavior of other lower-ionization species which typically fall off more rapidly with distance. *On the Right:* N_{OVI} plotted as a function of the specific star formation of the galaxy. Notably galaxies which are less actively forming stars (red diamonds) show significantly lower N_{OVI} than those with ongoing star formation (blue squares). Figures from Tumlinson et al. (2011).

comparable to and may exceed the mass within the ISM of the galaxy.

1.3.2 The High-Redshift CGM

While these low-redshift studies can give us a suggestion of the relationship between galaxies and their surrounding gas, high-z studies are perhaps more likely to provide robust evidence of their association as the peak of gas accretion (Faucher-Giguère et al. 2011) and of star formation (thought to trigger galactic outflows) both occurred above $z \sim 1$ (see Madau et al. 1996; Hopkins & Beacom 2006; Reddy et al. 2008).

At higher redshifts, the data have historically been more limited due to the difficulty of high-redshift galaxy spectroscopy. The most systematic searches were performed by Adelberger et al. (2003, 2005a). This work, a precursor to the project presented here, was a systematic look at absorption mainly from H I and C IV in 13 fields along 23 lines of sight to background QSOs. These sightline surveys were paired with $2 < z < 4$ Lyman Break Galaxy (LBG) surveys, allowing for the first glimpse at the distribution of gas surrounding star-forming galaxies at high redshift, and motivated by a hope of detecting unequivocal evidence of galactic winds.

Briefly, Adelberger et al. (2003) analyzed the flux decrement due to H I in redshift ranges near the locations of galaxies. In 8 lines of sight from $2.5 \lesssim z \lesssim 3.5$ they found excesses of H I were common out to $\sim$7 comoving Mpc (2 physical Mpc) and a lack of H I very close to these galaxies (< 0.7 comoving Mpc, 200 physical kpc). C IV absorption was observed out to 3.5 comoving Mpc (900 physical kpc). The cross-correlation of C IV systems with LBGs and the auto-correlation of LBGs were found to be similar, suggesting

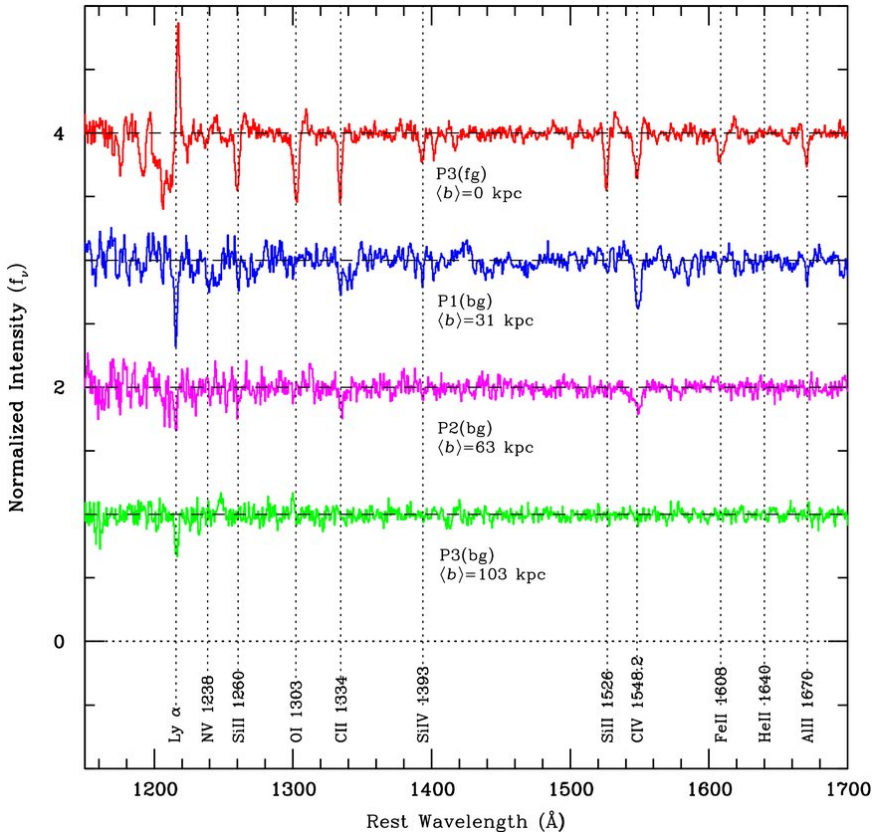


Figure 1.9 Stacked galaxy spectra probing gas surrounding $2 < z < 3$ galaxies. The red curve shows the average spectrum of the foreground galaxies. In this spectrum, the location of the absorbing gas with respect to the stars within that galaxy is known, and so the absorption signatures are unambiguous evidence of outflowing kinematics. The blue, magenta, and green spectra are coadditions of the background galaxy spectra within the specified radial distances, stacked in the rest frame of the foreground galaxy. Absorption signatures in this data reflect that gas content of the CGM at various galactocentric distances. The ratios of the Si II doublets in the top two spectra suggest that those transitions are saturated; however, they do not hit zero flux. This indicates that the gas has non-unity covering of the background light. Thus, the fraction of the absorbed flux at the center of the line is a measure of the covering fraction of the gas. Figure from Steidel et al. (2010).

LBGs and C IV absorbers may be causally related. The strongest C IV absorbers were so highly correlated with LBGs, the authors concluded that they must be the same object.

Adelberger et al. (2005a) focused on $2 \lesssim z \lesssim 3.5$ galaxies and absorption systems. They found that strong C IV systems ($N_{C\,IV} >> 10^{14}$ cm^{-2}) with velocity extents in excess of 200 km s^{-1} were commonly found within 40 proper kpc of a star-forming galaxy, and that absorption with $N_{C\,IV} \simeq 10^{14}$ cm^{-2} was found out to ~ 80 proper kpc. Additionally they found that the correlation length of $N_{C\,IV} \gtrsim 10^{12.5}$ cm^{-2} was similar to that of the autocorrelation length of the galaxies themselves. Strong Lyα absorption was found surrounding most LBGs, but a lack of strong absorption was found in $1/3$ of the galaxies in the survey. The authors attempted to correlate the observed IGM properties with galaxy properties although no meaningful conclusions could be reached due to the small size of their catalogue of galaxies with known properties near sight lines.

A recent exhaustive study of the demographics of gas absorption within 125 kpc surrounding $2 < z < 3$ star-forming galaxies and its link to galactic winds was presented by Steidel et al. (2010) using 512 galaxy pairs with small angular separations. They studied the distribution of absorbing gas associated with the CGM of the foreground galaxy seen in the spectra of the foreground galaxies themselves along with that seen in the spectra of background galaxies. This method is advantageous because foreground-background galaxy pairs can be found in much greater number and at much smaller angular separation that galaxy-QSO pairs. Also, using the spectrum of the galaxy itself as a probe of the absorption provides information about the kinematics of the gas with respect to the stars (in this case the blue shift shows the gas is outflowing). However, one does not know at what distance the gas lies from the position of the stars. The background galaxy spectra, on the other hand, provide a known projected distance from the galaxy, however, the relative velocity of the gas no longer provides a clear constraint on the relative kinematics of the gas. But by combining the two methods, one can attempt to trace the motions and distribution of gas within the CGM.

One further complication in interpreting the background galaxy data arrises from the fact that galaxies (unlike background QSOs) have a non-negligible angular size. Because of this, the absorption signatures in the spectra of background galaxies trace the gas distribution on $\sim$kpc scales in the CGM of the foreground galaxies. As such, the equivalent width of an absorber (W_0) traces a combination of the column density, velocity spread, and covering fraction of the foreground gas.

Steidel et al. (2010) found a decline in the equivalent width of H I and metal line absorption as a function of distance from the foreground galaxies. Because the absorbers are found to be saturated (see Figures 1.9 and 1.10), this decline is interpreted as evidence for the decreasing covering fraction of the gas as a function of distance as well as a decline in the line of sight velocity spread of the gas with distance. They present a simplified model that suggests one can explain both the absorption features in the galaxy spectra as well as those at larger distance within the CGM (as traced by the background galaxy spectra) with a simple outflow model in which the highest velocity gas is seen at larger radii, and thus at lower covering fractions.

Unfortunately, the utility of faint background galaxies in probing foreground gas falls off as a function of distance due to a lack of sensitivity. Chapter 2 of this thesis extends the study of the CGM of these galaxies to much larger distances using background QSOs whose spectra can probe much weaker absorption systems with high sensitivity.

1.4 This Thesis

This thesis presents the results of a full Voigt profile decomposition of the $2.0 < z < 2.8$ Lyα and Lyβ forest in the spectra of 15 high-z QSOs. This effort has resulted in the largest high-z catalog of H I absorbers to date which increases by an order of magnitude the number of forest absorbers fit with the added constraint of higher-order Lyman series lines. This constraint allows for much more precise measurements of the $N_{\rm HI}$ and widths of absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2}, whose Lyα transitions are saturated. We have used this

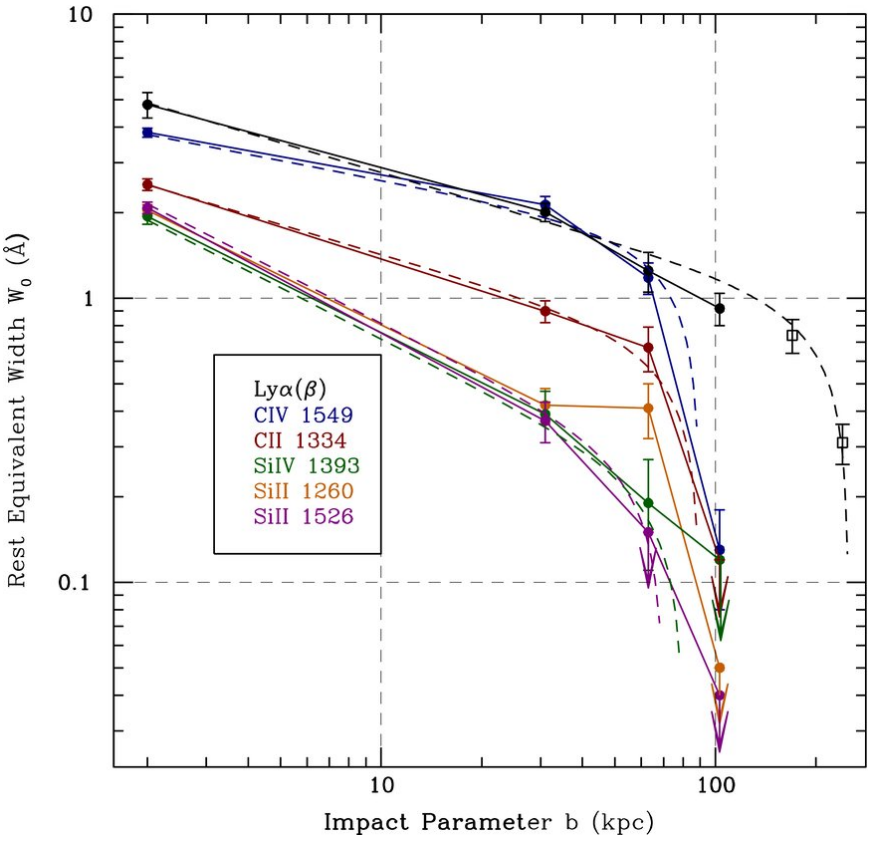


Figure 1.10 The equivalent width (W_0) of neutral hydrogen and a variety of metal absorption lines within the CGM of $2 < z < 3$ star-forming galaxies as a function of impact parameter. The first point at small impact parameters is measured from the red "down the barrel" galaxy spectrum shown in Figure 1.9. The points at larger distances are measured from the spectra of background galaxies. Note that the two Si II transitions at 0 and 31 kpc have the same W_0. As these two transitions have different oscillator strengths, their matched strengths mean that the transition is saturated. The fall off of W_0 with impact parameter is thought to be due to the decline in the covering fraction of the gas and in the range of line of sight velocities present in the the absorbing gas. Figure from Steidel et al. (2010).

new catalog to complete the first study of the distribution and physical properties of H I within the CGM surrounding a statistically representative sample of $2 < z < 3$ galaxies. This work is presented in Chapter 2. Chapter 3 considers the line widths of absorbers drawn from this catalog to measure the temperature-density relation within the $2 < z < 3$ IGM. Chapter 4 presents the revised statistics of the Lyα forest that can be derived from this new analysis, and for the first time considers several classic IGM statistics as applied to the CGM of galaxies. It concludes with a new measurement of the mean free path of hydrogen-ionizing photons through the IGM and the IGM+CGM. In an epilogue (Chapter 5), I present the current status of my next project. This early analysis concerns the distribution of metals in the CGM of these same galaxies. Appendix B discusses my contributions to the MOSFIRE instrumentation project.

A note on the contents of this thesis: The contents of Chapter 2 were published in Rudie et al. 2012, ApJ, 750, 67. Chapter 3 appeared in Rudie et al. 2012, ApJL, 757, L30. Chapter 4 is published in Rudie et al. 2013, ApJ, 769, 146.

Chapter 2

The Gaseous Environment of High-z Galaxies: Precision Measurements of Neutral Hydrogen in the Circumgalactic Medium of $z \sim 2-3$ Galaxies in the Keck Baryonic Structure Survey

This chapter was previously published as Rudie et al. 2012, ApJ, 750, 67 with coauthors Charles C. Steidel, Ryan F. Trainor, Olivera Rakic, Milan Bogosavljević, Max Pettini, Naveen Reddy, Alice E. Shapley, Dawn K. Erb, and David R. Law.

Abstract

We present results from the Keck Baryonic Structure Survey (KBSS), a unique spectroscopic survey of the distant universe designed to explore the details of the connection between galaxies and intergalactic baryons within the same survey volumes, focusing particularly on scales from ~ 50 kpc to a few Mpc. The KBSS is optimized for the redshift range $z \sim 2-3$, combining S/N$\sim$ 100 Keck/HIRES spectra of 15 of the brightest QSOs in the sky at $z \simeq 2.5-2.9$ with very densely sampled galaxy redshift surveys within a few arcmin of each QSO sightline. In this paper, we present quantitative results on the distribution, column density, kinematics, and absorber line widths of neutral hydrogen (H I) surrounding a subset of 886 KBSS star-forming galaxies with $2.0 \lesssim z \lesssim 2.8$ and with projected distances ≤ 3 physical Mpc from a QSO sightline. Using Voigt profile decompositions of the full Lyα forest region of all 15 QSO spectra, we compiled a catalog of ~ 6000 individual absorbers in the redshift range of interest, with $12 \leq \log(N_{\rm HI}) \leq 21$. These are used to measure H I absorption statistics near the redshifts of foreground galaxies as a function of projected galactocentric distance from the QSO sightline and for randomly chosen locations in the intergalactic medium (IGM) within the survey volume. We find that $N_{\rm HI}$ and the multiplicity of velocity-associated H I components increase rapidly with decreasing galactocentric impact parameter and as the systemic redshift of the galaxy is approached. The strongest H I absorbers within $\simeq 100$ physical kpc of galaxies have $N_{\rm HI} \sim 3$ orders of magnitude higher than those near random locations in the IGM. The circumgalactic zone of most significantly enhanced H I absorption is found within transverse distances of $\lesssim 300$ kpc and within ± 300 km s^{-1} of galaxy systemic redshifts. Taking this region as the defining bounds of the circumgalactic medium (CGM), nearly half of absorbers with $\log(N_{\rm HI}) > 15.5$ are found within the CGM of galaxies meeting our photometric selection criteria, while their CGM occupy only 1.5% of the cosmic volume. The spatial covering fraction, multiplicity of absorption components, and characteristic $N_{\rm HI}$ remain significantly elevated to transverse distances of $\sim$2 physical Mpc from galaxies in our sample. Absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2} are tightly correlated with the positions of galaxies, while absorbers with lower $N_{\rm HI}$ are correlated with galaxy positions only on $\gtrsim$ Mpc scales. Redshift anisotropies on these larger scales indicate coherent infall toward galaxy locations, while on scales of ~ 100 physical kpc peculiar velocities of $\Delta v \simeq \pm 260$ km s^{-1} with respect to the galaxies are indicated. The median Doppler widths of individual absorbers within 1-3 $r_{\rm vir}$ of galaxies are larger by $\simeq 50\%$ than randomly chosen absorbers of the same $N_{\rm HI}$, suggesting higher gas temperatures and/or increased turbulence likely caused by some combination of accretion shocks and galactic winds around galaxies with $M_{\rm halo} \simeq 10^{12}$ M$_\odot$ at $z \sim 2-3$.

2.1 Introduction

Hydrogen, comprising three quarters of the baryonic mass of the universe, is the principal component of all luminous objects in the universe. It is the fuel source for stars and therefore for star formation. Thus, in order to understand the formation and evolution of galaxies, one must understand and be able to trace the inflow and outflow of this fuel.

There exist very poor observational constraints on the movement of baryons in and out of galaxies. At high redshift in star-forming systems, it has been argued that the outflow rate must be similar to the star-formation rate (Pettini et al. 2000) and that the inflow rate must be similar to the combined star-formation rate and outflow rate (Erb 2008; Finlator & Davé 2008).

Recently there has been a flurry in the theoretical literature predicting the prevalence of accretion of cold gas ($\log(T) \lesssim 4.5-5.5$ K) onto high-z galaxies via filamentary "cold flows" (Birnboim & Dekel 2003; Kereš et al. 2005; Ocvirk et al. 2008; Brooks et al. 2009; Faucher-Giguère et al. 2011; van de Voort et al. 2011a,b). In this model, the baryons stream into galaxies along the filamentary structure of the cosmic web, accreting onto galaxies without experiencing virial shocks. A wide range of predictions has been made concerning the efficiency of the transport of this material into galaxy halos, as well as its role in fueling ongoing star formation (van de Voort et al. 2011b; Faucher-Giguère et al. 2011). Further, there may be substantial suppression of the cold accretion rate caused by galaxy-scale mass outflows, evidence for which is commonly observed in the spectra of high-z star-forming systems (Pettini et al. 2001; Shapley et al. 2003; Adelberger et al. 2005a; Steidel et al. 2010). Mapping the gas distribution surrounding galaxies is critical to constraining these models (Faucher-Giguère & Kereš 2011; Fumagalli et al. 2011), and would be a significant step toward understanding and quantifying the exchange of baryons between the sites of galaxy formation and the nearby intergalactic medium (IGM).

There has been a large amount of recent theoretical examination of the nature of IGM absorbers and their relation to galaxies using simulations. Lyα is believed to broadly trace the filamentary large-scale structure (Cen et al. 1994; Zhang et al. 1995; Miralda-Escudé et al. 1996; Hernquist et al. 1996; Rauch et al. 1997; Theuns et al. 1998; Davé et al. 1999; Schaye 2001) although there are indications (Barnes et al. 2011) that winds could blow spatially extended halos of gas which may have recently been observed both in absorption (Steidel et al. 2010) and in Lyα emission (Steidel et al. 2011). There seems to be general agreement that galactic winds are responsible for metal absorbers in the IGM. Booth et al. (2010) suggest that mostly low mass ($M_{DM} \lesssim 10^{10}$ $M_\odot$) galaxies must be responsible for the pollution, while Wiersma et al. (2010) suggest only half of the metals would originate from galaxies with $M_{DM} \lesssim 10^{11}$ $M_\odot$. Wiersma et al. (2010) also studied the history of the ejection of these metals and found that half of the metals observed at redshift 2 were ejected during the time between $2 < z < 3$. Using cosmological "zoom-in" simulations Shen et al. (2011) found a "Lyman Break"-type galaxy could distribute metals to 3 virial radii by $z = 3$. Simcoe (2011) recently considered this problem observationally, finding that indeed 50% of the metals observed in the IGM

at $z \sim 2.4$ were placed there since $z \sim 4.3$, i.e. in 1.3 Gyr. Oppenheimer & Davé (2008) and Oppenheimer et al. (2010) studied the fate of winds using cosmological simulations and found that while galactic winds are likely responsible for the metallic species seen in the IGM, much of the outflowing gas may be bound to galaxies and may fall back in. In their simulations the recycling timescale scaled inversely with mass because winds emanating from more massive galaxies experienced greater hydrodynamic drag due to the increased abundance of dense IGM surrounding them. Further, in these simulations the largest source of gaseous fuel for star formation after $z \sim 1$ was recycled wind material.

To date, systematic attempts to jointly study high-z galaxies and their intergalactic environs have been made by Adelberger et al. (2003, 2005a) (see also Crighton et al. 2011). These studies focused primarily on H I and C IV absorption in the spectra of background QSOs whose sightlines probed regions covered by $2 < z < 4$ Lyman break galaxy (LBG) surveys. This work allowed for a first glimpse of the distribution of diffuse gas surrounding forming galaxies at high redshift, and, perhaps more tantalizing, evidence for interactions between the IGM and galaxies during the epoch when galaxies are expected to be most active. Generally, Adelberger et al. (2003) and Adelberger et al. (2005a) found excess H I absorption out to $\approx 5-6h^{-1}$comoving Mpc (cMpc) of galaxies [$\sim$2 physical Mpc (pMpc) at $\langle z \rangle = 3.3$ using the cosmology adopted in this paper]. C IV systems were found to correlate strongly with the positions of galaxies suggesting a causal connection. Unfortunately, these papers could not consider physical properties of the gas such as its column density or temperature because the data were not of sufficiently high quality to perform Voigt profile analysis. As such these papers focused on the transmitted flux which could be applied to a wider range of data qualities.

In this work, we provide high-accuracy analysis of the spectral regions surrounding 886 high-z star-forming galaxies as seen in absorption against the spectra of background hyper-luminous QSOs using data drawn from the Keck Baryonic Structure Survey (KBSS; Steidel et al, in prep). The KBSS was specifically designed to allow for the observation of gas absorption features surrounding high-redshift star-forming galaxies, providing unique insight into the IGM/galaxy interface at high redshift. The size and quality of the KBSS sample allow us to map the distribution and properties of gas near to individual star-forming galaxies with direct physical parameters such as the column density as opposed to proxies such as the equivalent width. This paper is the first in a series designed to study the physical properties of star-forming galaxies at high redshift using Voigt profile analysis of this data sample. The complementary analysis presented by Rakic et al. (2011b) describes the pixel statistics of the QSO spectra from the KBSS and the correlation of H I optical depth with the positions and redshifts of galaxies.

In §2.2 we discuss the galaxy and QSO data and present the Voigt profile analysis of the QSO spectra in §2.3. The distributions of H I absorbers as a function of velocity, impact parameter, and 3D distance are presented in §2.4. In §2.5 we consider the geometric distribution of the gas using the covering fraction and incidence of absorbers. §2.6 focuses on 2D "maps" of the absorber distribution around galaxies. In §2.7 we analyze the velocity widths of individual absorbers and their correlation with the proximity of galaxies. We discuss the results and their possible interpretation in §2.8 with a brief summary of the paper and our

conclusions in §2.9.

Throughout this paper we assume a Λ-CDM cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. All distances are expressed in physical (proper) units unless stated otherwise. We use the abbreviation pkpc and pMpc to indicate physical units, and ckpc or cMpc for co-moving units. At the mean redshift of the galaxy sample ($\langle z \rangle = 2.3$), 300 pkpc is 210h^{-1}pkpc (physical) or $\simeq$ 700h^{-1}ckpc; the age of the universe is 2.9 Gyr, the look-back time is 10.9 Gyr, and 8.2 pkpc subtends one arcsecond on the sky.

2.2 Observations

The data presented in this paper are drawn from the Keck Baryonic Structure Survey (KBSS) and include a large sample of rest-UV (2188) and rest-optical (112) spectra of UV-color selected star-forming galaxies at $\langle z \rangle \sim 2.3$. These galaxies were photometrically selected to lie in the foreground of one of 15 hyper-luminous QSOs in the redshift range $2.5 \leq z_{QSO} \leq 2.85$ for which we have obtained high-resolution, high signal-to-noise ratio (S/N) echelle spectra.

The redshift range of this survey has important significance in the history of the universe – it coincides with the peak of both universal star formation (see Reddy et al. 2008) and super-massive black hole growth (see Richards et al. 2006). Spectroscopic observations of star-forming galaxies during this epoch commonly exhibit signatures of strong outflowing winds (Pettini et al. 2001; Shapley et al. 2003; Adelberger et al. 2005a; Steidel et al. 2010). At the same time, the baryonic accretion rate onto galaxies is predicted to be near its peak at $z \sim 2.5$ (e.g., van de Voort et al. 2011b; Faucher-Giguère et al. 2011; van de Voort et al. 2011a.) Thus, the signatures of galaxy formation within the IGM should be at their peak as well.

The redshift range $2.0 \lesssim z \lesssim 2.8$ offers a number of practical advantages as well: first, the rapidly-evolving Lyα forest has thinned enough to allow measurements of individual H I systems and to enable the detection of important metallic transitions falling in the same range (notably, O VI $\lambda\lambda$1031, 1036); second, the astrophysics-rich rest-frame far-UV becomes accessible to large ground-based telescopes equipped with state-of-the-art optical spectrographs; third, the rest-frame optical spectrum is shifted into the atmospheric transmission windows in the near infrared, allowing for observations of a suite of diagnostic emission lines (Hα, Hβ, [O II] $\lambda\lambda$3726,3729, [O III]$\lambda\lambda\lambda$4363,4959,5007, and [S II] $\lambda\lambda$6717,6732) neatly packed into the J, H, and K bands.

2.2.1 The Galaxy Sample

The KBSS galaxy sample, a subset of which is used in this paper (Figure 2.1), is described in detail by Steidel et al (2012); here we present a brief summary. Galaxies were selected for spectroscopy using their rest-frame UV colors (i.e., LBGs) according to the criteria outlined by Steidel et al. (2004) and Adelberger et al. (2004) for $z \sim 2$ and by Steidel et al. (2003) for $z \sim 3$. In combination, these criteria have been devised to efficiently select star-forming galaxies over the redshift range $1.5 \lesssim z \lesssim 3.5$. In total, the KBSS galaxy sample includes

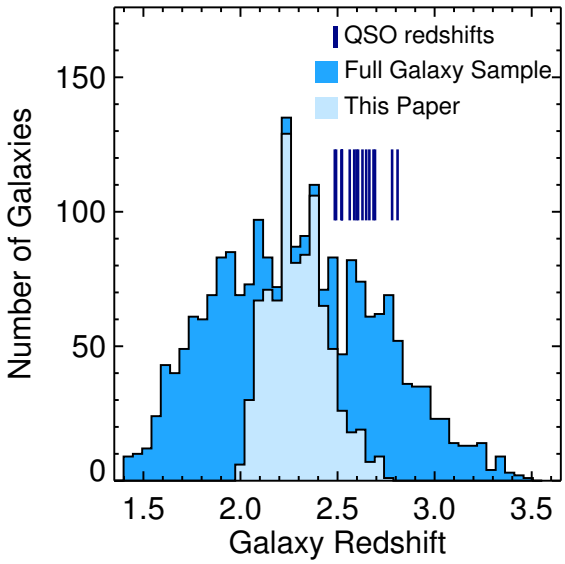


Figure 2.1 The redshift distribution of the 2188 galaxies in the full KBSS sample (dark shaded histogram) and of the subset of 886 galaxies used for the analysis in this paper (light shaded histogram), which are those with redshifts high enough that the corresponding wavelength of the Lyβ transition is observed in the relevant QSO spectrum, and with $z < z_{\rm QSO} - 3000$ km s^{-1}. The dark vertical line segments mark the redshifts of the 15 KBSS QSOs.

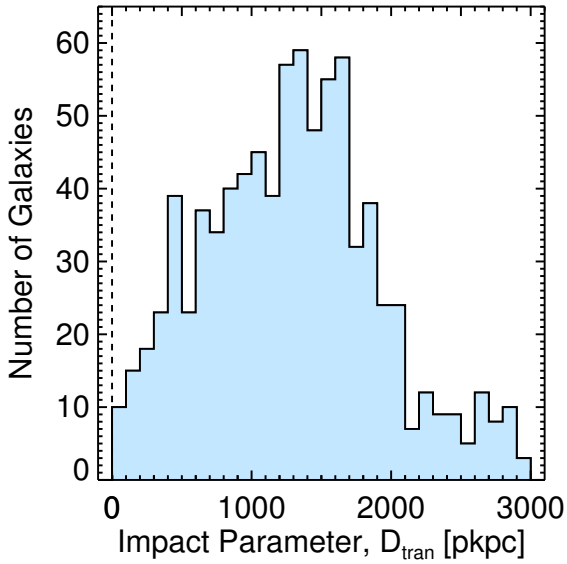


Figure 2.2 The number of galaxies as a function of physical impact parameter $D_{\rm tran}$ for the sample which has appropriate redshifts to be used in this work (light shaded histogram in Figure 2.1; see Table 2.1). The decline in the number of galaxies at $D_{\rm tran} \approx 2$ pMpc is due to the typical survey geometry of the KBSS fields as described in §2.4.2.1.

$R \simeq 1000$ ($\sim$5-10Å) rest-UV spectra of 2188 star-forming galaxies obtained at the W.M. Keck Observatory using the Keck 1 10m telescope and the blue arm of the Low Resolution Imaging Spectrometer (LRIS-B; Oke et al. 1995; Steidel et al. 2004.)

Most of the observations were conducted using a 400 lines mm^{-1} grism (blazed at 3400 Å in first order) in combination with a dichroic beamsplitter sending all light shortward of $\sim$ 6800 Å into the blue channel, where the wavelength coverage for a typical slit location was $\simeq 3100-6000$ Å with a resolving power $R \simeq 800$ using $1''.2$ wide slits (Steidel et al. 2010). Some of the spectra were obtained using the d560 beamsplitter (beam divided near 5600 Å) together with a 600 lines mm^{-1} grism (4000 Å blaze), typically covering $\simeq 3400-5600$ Å with a resolving power $R \simeq 1300$. Observations obtained after July 2007 made use of the Cassegrain Atmospheric Dispersion Corrector, thus minimizing the effects of differential atmospheric refraction over the spectral range of interest. Wavelength calibration was accomplished using Hg, Cd, Zn, and Ne lamps, with zero point corrections based on night sky emission lines on each individual exposure.[1]

The exposure times allocated to individual galaxies ranged from 1.5-7.5 hours depending on the number of separate masks containing the same target; typically galaxies were observed on either 1 or 2 masks, each mask receiving a total integration of 1.5 hours. Further details on the selection, observing strategy, and data reduction for KBSS galaxies are presented elsewhere (Steidel et al. 2010, Steidel et al. 2012).

The rest-UV spectra of LBGs are dominated by the continuum emission of O and B stars, over which are superposed numerous resonance absorption lines of metallic ions and H I. The H I Lyα line at 1215.67 Å may be seen in emission or absorption (and often in both). The absorption features arise in cool interstellar gas in the foreground of the OB stars; they are most commonly observed to be blue-shifted by 100 – 800 km s^{-1} with respect to the systemic velocity of the stars, as measured from either rest-frame optical nebular emission lines or stellar photospheric lines in stacked spectra (see Pettini et al. 2001; Shapley et al. 2003; Adelberger et al. 2005a; Steidel et al. 2010) or from the redshift-space symmetry of Lyα absorption in the nearby IGM (Rakic et al. 2011a). Common lines observed include: O VI $\lambda\lambda$1031,1036, Si II λ1260, λ1526, Si II+O I λ1303 (blend),Si III λ1206, Si IV $\lambda\lambda$1393,1402, N V $\lambda\lambda$1238,1242, C II λ1334, C III λ977, and C IV $\lambda\lambda$1548,1550. The profile of the Lyα emission or absorption line is modulated by the optical depth of the material closest to the systemic velocity of the stars, which has been shown to correlate most significantly with the baryonic mass (Steidel et al. 2010) and the physical size (Law et al., in prep) of the galaxy.

The sample used in this study includes 886 galaxies within the redshift range where at minimum Lyα and Lyβ are observed in the HIRES spectrum and with redshifts placing them at least 3000 km s^{-1} blue-ward of the redshift of the QSO. The latter criterion was selected to avoid proximate systems that originate within material ejected from the QSO itself and/or the region affected by its ionizing radiation field.

A typical galaxy in the spectroscopic survey has a bolometric luminosity of $\sim 2.5 \times 10^{11}$ L$_\odot$ (Reddy et al. 2008, 2011), a star-formation rate (SFR) of ~ 30 M$_\odot$ yr^{-1} (Erb et al. 2006b), a stellar age of ~ 0.7 Gyr (Erb

[1] An extensive discussion of the LRIS/KBSS wavelength calibration errors (including slit illumination and atmospheric dispersion) is given in Steidel et al. (2010).

et al. 2006c), and a gas-phase metallicity of $\simeq 0.5\ Z_{\odot}$ (Erb et al. 2006a). The galaxies inhabit dark matter halos of average mass $\sim 10^{12}\ M_{\odot}$ (Adelberger et al. 2005b; Conroy et al. 2008; Trainor & Steidel 2012; Rakic 2012) and average dynamical masses of $\sim 7 \times 10^{10}\ M_{\odot}$ (Erb et al. 2006c) and generally exhibit dispersion-dominated kinematics (Law et al. 2009). The luminous parts of the galaxies are dominated by baryons, typically half stars and half cold gas (Shapley et al. 2005; Erb et al. 2006c), with half-light radii of ~ 2 pkpc (Law et al. 2011). The spectroscopic sample includes objects with apparent magnitudes $\mathcal{R} \leq 25.5$, where $\mathcal{R}$ is equivalent to m_{AB}(6830 Å). At the mean redshift of the sample ($\langle z \rangle = 2.30$) the faint limit corresponds to a galaxy of $0.25L^{*}_{UV}$ (Reddy & Steidel 2009).

The redshift distribution of the galaxy and QSO sample is presented in Figure 2.1, and the distribution of physical impact parameters, $D_{\rm tran}$, between the galaxies and the QSO lines of sight is shown in Figure 2.2.

At present, 112 galaxies in the full KBSS sample have been observed spectroscopically in the near-IR using using NIRSPEC (McLean et al. 1998) on the Keck II telescope. 87 of these galaxies lie in our chosen redshift interval. The NIRSPEC target selection, data, and reductions are discussed in Erb et al. (2006b,c). The NIR redshifts, generally based on the Hα emission line, are estimated to be accurate to ~ 60 km s^{-1} or $\sigma_z \simeq 0.0007$ at $z \sim 2.3$.

2.2.2 Measured and Calibrated Redshifts

Because the most prominent features in the UV spectra of star forming galaxies are not at rest with respect to a galaxy's systemic redshift, $z_{\rm gal}$, corrections must be applied to avoid substantial systematic redshift errors. The velocity peak and centroid of the Lyα emission line, when present, tend to be redshifted with respect to $z_{\rm gal}$ by several hundred km s^{-1}, while the strong UV absorption features ($z_{\rm IS}$) tend to be similarly blue-shifted with respect to $z_{\rm gal}$ (Shapley et al. 2003; Adelberger et al. 2003; Steidel et al. 2010). These observations are generally interpreted as strong evidence for the presence of galaxy-scale outflows.

Here we adopt estimates of galaxy systemic redshifts ($z_{\rm gal}$) computed in the manner proposed by Adelberger et al. (2005a) and later updated by Steidel et al. (2010) and Rakic et al. (2011a). Adelberger et al. (2005a) and Steidel et al. (2010) analyzed the subset of the UV sample for which both rest-UV and rest-optical spectra had been obtained. They measured the average offset between redshifts defined by Hα emission versus $z_{\rm Ly\alpha}$ and $z_{\rm IS}$ to estimate average corrections. The Hα line traces the ionized gas in star-forming regions and is therefore a reasonable proxy for the systemic velocity of the stars, which are more difficult to measure due to the weakness of the UV photospheric absorption lines. Rakic et al. (2011a) used the QSO and galaxy data set presented here and calibrated velocity offsets appropriate for various classes of LBGs by insisting that the average IGM Lyα absorption profiles should be symmetric with respect to galaxy redshifts. In both cases, the offsets represent those of the ensemble while in reality there is some scatter between individual objects even if their spectral morphology is similar. However, as we will demonstrate, the adopted $z_{\rm gal}$ must be generally quite accurate in order to produce the trends described below.

The formulae used for estimating $z_{\rm gal}$ from $z_{\rm Ly\alpha}$ and $z_{\rm IS}$ measurements are reproduced below. For galaxies

with Hα-based redshifts (87/886), we set $z_{\rm gal} = z_{\rm H\alpha}$. For galaxies which have measured $z_{\rm IS}$ with or without the presence of Lyα emission (691/886),

$$z_{\rm gal,IS} \equiv z_{\rm IS} + \frac{\Delta v_{\rm IS}}{c}(1+z_{\rm IS}) \, , \tag{2.1}$$

where

$$\Delta v_{\rm IS} = 160 \text{ km s}^{-1} \tag{2.2}$$

is the velocity shift needed to transform the observed redshift into its systemic value, $z_{\rm IS}$ is the measured redshift from the centroids of interstellar absorption lines, and $z_{\rm gal,IS}$ corresponds to the estimated systemic redshift of the galaxy.

For galaxies which have redshifts measured *only* from Lyα in emission (90/886), we compute the redshift as

$$z_{\rm gal,Ly\alpha} \equiv z_{\rm Ly\alpha} + \frac{\Delta v_{\rm Ly\alpha}}{c}(1+z_{\rm Ly\alpha}), \tag{2.3}$$

where

$$\Delta v_{\rm Ly\alpha} = -300 \text{ km s}^{-1} \tag{2.4}$$

is the velocity shift needed to transform the observed redshift into its systemic value, $z_{\rm Ly\alpha}$ is the measured redshift from Lyα, and $z_{\rm gal,Ly\alpha}$ is adopted systemic redshift.

For galaxies with measurements of both $z_{\rm IS}$ and $z_{\rm Ly\alpha}$, we verify that $z_{\rm IS} < z_{\rm gal,IS} < z_{\rm Ly\alpha}$. If the corrected absorption redshift is not bracketed by the two measured redshifts (18/886 galaxies)[2], then we use the average of $z_{\rm IS}$ and $z_{\rm Ly\alpha}$:

$$z_{\rm gal} \equiv \frac{z_{\rm Ly\alpha} + z_{\rm IS}}{2}. \tag{2.5}$$

The residual redshift errors have a significant impact on our ability to interpret the kinematic information in the data; thus, their amplitude will be important to consider in the examination of the line-of-sight distribution of H I. Steidel et al. (2010) found this method generally corrects the redshifts to within ~ 125 km s^{-1} of the systemic velocity.

2.2.3 QSO Observations

The 15 hyper-luminous ($m_V \simeq 15.5-17$) QSOs in the center of the KBSS fields (Table 2.1) were observed with the High Resolution Echelle Spectrometer (HIRES; Vogt et al. 1994) on the Keck I telescope. All available archival data for these 15 QSOs have been incorporated, including data taken with UVES (Dekker et al. 2000) on the VLT for Q2206-199 and Q2343+125. We obtained additional HIRES observations in

[2]In these cases, $z_{\rm IS}$ and $z_{\rm Ly\alpha}$ have very similar values and therefore taking their average generally provides a systemic redshift with less potential error than in those galaxies for which an average shift (as in equations 2.1 and 2.3) is required.

Table 2.1. KBSS Central QSOs and Foreground Galaxy Samples

Name	RA	Dec	z_{QSO}[a]	λ_{min}[b]	z_{gal} range	N_{gal}[c]	S/N Lyα[d]	S/N Lyβ [d]
Q0100+130 (PHL957)	01:03:11.27	+13:16:18.2	2.721	3133	2.0617– 2.6838	47	77	50
HS0105+1619	01:08:06.4	+16:35:50.0	2.652	3230	2.1561– 2.6153	53	127	89
Q0142−09 (UM673a)	01:45:16.6	−09:45:17.0	2.743	3097	2.0260– 2.7060	65	71	45
Q0207−003 (UM402)	02:09:50.71	−00:05:06.5	2.872	3227	2.1532– 2.8339	46	82	55
Q0449−1645	04:52:14.3	−16:40:16.2	2.684	3151	2.0792– 2.6470	50	73	41
Q0821+3107	08:21:07.62	+31:07:51.17	2.616	3239	2.1650– 2.5794	37	50	33
Q1009+29 (CSO 38)	10:11:55.60	+29:41:41.7	2.652[e]	3186	2.1132– 2.6031	36	99	58
SBS1217+499	12:19:30.85	+49:40:51.2	2.704	3098	2.0273– 2.6669	43	68	38
HS1442+2931	14:44:53.67	+29:19:05.6	2.660	3152	2.0798– 2.6237	46	99	47
HS1549+1919	15:51:52.5	+19:11:04.3	2.843	3165	2.0926– 2.8048	54	173	74
HS1603+3820	16:04:55.38	+38:12:01.8	2.551[f]	3181	2.1087– 2.5066	37	108	58
Q1623+268 (KP77)	16:25:48.83	+26:46:58.8	2.5353	3126	2.0544– 2.4999	133[g]	48	28
HS1700+64	17:01:00.6	+64:12:09.4	2.751	3138	2.0668– 2.7138	110[g]	98	42
Q2206−199	22:08:52.1	−19:43:59.7	2.573	3084	2.0133– 2.5373	45	88	46
Q2343+125	23:46:28.30	+12:48:57.8	2.5730	3160	2.0884– 2.5373	84[g]	71	45

[a]The redshift of the QSO

[b]The minimum wavelength covered in the HIRES QSO spectrum

[c]The number of galaxies in our LRIS survey with $D_{tran} < 3$ pMpc and spectroscopic redshifts in the correct range given in the previous column.

[d]The average signal to noise ratio per pixel of the QSO spectrum in the wavelength range pertaining to CGM Lyα and Lyβ absorption.

[e]The redshift of this QSO was revised after the fitting of the HIRES spectrum was completed. The redshift assumed for the $\Delta v = -3000$ km s^{-1} is 2.6395

[f]The redshift of this QSO was revised after the fitting of the HIRES spectrum was completed. The redshift assumed for the $\Delta v = -3000$ km s^{-1} is 2.5420

[g]The photometry and spectroscopy in fields Q1623, HS1700, and Q2343 cover larger areas than a single LRIS footprint. As such, they represent the fields in which we sample the galaxy-IGM connection at $D_{tran} > 2$ pMpc.

order to reach a uniformly high S/N ratio over the spectral range of primary interest, 3100 – 6000 Å. The final HIRES spectra have $R \simeq 45,000$ (FWHM$\simeq$ 7 km s^{-1}), S/N $\sim 50-200$ per pixel, covering at least the wavelength range 3100 – 6000Å with no spectral gaps. The significant improvement in the UV/blue sensitivity of HIRES resulting from a detector upgrade in 2004 enabled us to observe Lyβ λ1025.7 down to at least $z = 2.2$ in all 15 KBSS sightlines, and to significantly lower redshifts in many (see Table 1.) The additional constraints provided by Lyβ (and in many cases, additional Lyman series transitions) allow for much more accurate measurements of H I for $N_{\rm HI}= 10^{14}-10^{17}$ cm^{-2} (see §2.3); this is particularly important since these column densities are typical of H I gas in the CGM at these redshifts (see §2.5).

The QSO spectra were reduced using T. Barlow's MAKEE package which is specifically tailored to the reduction of HIRES data. The output from MAKEE is a wavelength-calibrated[3] extracted spectrum of each echelle order, corrected for the echelle blaze function and transformed to vacuum, heliocentric wavelengths. The spectra were continuum-normalized in each spectral order using low-order spline interpolation, after which the normalized 2-D spectrograms were optimally combined into a single one-dimensional, continuum-normalized spectrum resampled at 2.8 km s^{-1} per pixel.

Three of the KBSS QSO spectra contain a damped Lyman α (DLA) system ($N_{\rm HI}> 10^{20.3}$ cm^{-2}), and three contain a sub-DLA ($10^{19} \lesssim N_{\rm HI}< 10^{20.3}$ cm^{-2}). Special care must be taken in the continuum fit to the regions surrounding these systems, as the damping wings of the absorption lines extend for thousands of km s^{-1} from line center. In these regions, we carefully fit a Voigt profile to the core of the absorber and adjusted the original continuum fit so that the Voigt profile produced a good fit to the damping wings, as shown in the top panel of Figure 2.3. The final Voigt profile fit is divided into the true spectrum, resulting in a new spectrum where the wings of the DLA have been removed, as illustrated in the bottom panel of Figure 2.3. The re-normalized spectrum can then be used to fit additional absorption systems superposed on the damping wings.

The redshifts of the QSOs are measured from rest-frame optical emission lines using lower-dispersion NIR spectra. The details of this procedure and the expected errors in the QSO redshifts are reported upon in Trainor & Steidel (2012). The precise QSO redshifts do not affect our analysis.

2.3 Analysis of QSO Absorption Spectra

The process of accurately measuring H I in the Lyα forest of QSO spectra is complicated by the saturation of moderately strong absorbers and the blending of H I features with other H I or with lines of metallic species that happen to fall in the forest region.

Our analysis includes a full Voigt-profile decomposition of the Lyα forest from the lowest redshift for which Lyβ is available in each spectrum up to 3000 km s^{-1} blue-ward of the QSO redshift; Table 1 shows the relevant redshift range used for each QSO in the sample. The cut-off at the high redshift end is to avoid

[3]The wavelength calibration of the HIRES spectra introduce negligible error into our analysis. The HIRES spectra have calibration errors less that 0.5 km s^{-1}.

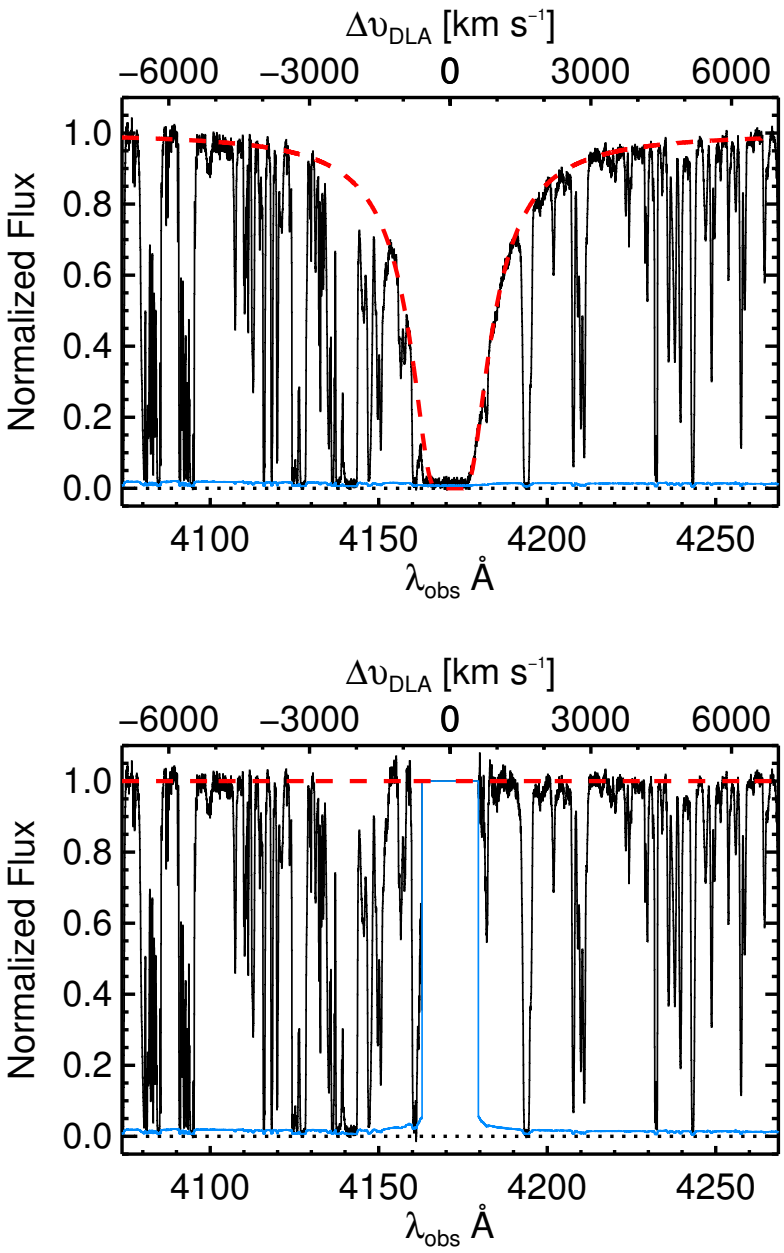


Figure 2.3 A demonstration of our treatment of the continuum surrounding damped-Lyα systems. *Top*: In black, the continuum-normalized HIRES spectrum of Q2343+125 showing $\pm$7000 km s^{-1} surrounding the DLA. The (red) dashed line corresponds to the Voigt profile of the DLA centered at $z = 2.4312$ with $\log(N_{\rm HI})$ = 20.4. Shown in the light (blue) curve is the error spectrum. *Bottom*: The HIRES spectrum of the same QSO with the DLA profile divided out. The new error spectrum accounting for the DLA profile division is shown by the light (blue) curve. The new continuum (with the DLA divided out) is shown in the dashed (red) curve.

H I systems which could be ejected from and/or ionized by the QSO itself; the low-redshift cut is necessary due to the high frequency of systems with $N_{\rm HI} \gtrsim 10^{14.5}$ cm^{-2} that will be saturated in Lyα[4]. For saturated systems, the fit to Lyα is degenerate between an increase in column density or an increase in the width of the line, $b_{\rm d}$. The best way to resolve this degeneracy is to measure higher-order Lyman lines where decreasing oscillator strengths allow accurate $N_{\rm HI}$ determination.

Simultaneous fits were made to as many Lyman lines as were both (1) available in the observed spectral range and (2) needed to measure an unsaturated and uncontaminated profile in the highest-order line. The exact number of higher-order Lyman lines used therefore depended upon both the redshift of the absorber and the degree of contamination in the spectral region containing the higher-order Lyman series absorption features. Higher-order lines, whether saturated or unsaturated, were used whenever doing so provided additional constraints on the overall fit.

Example Voigt profile fits to the H I absorption in regions surrounding the systemic redshift of five galaxies from our sample are shown in Figure 2.4. Note that for galaxies with redshifts significantly larger than $z \sim 2.2$, many higher-order Lyman transitions can be measured.

To facilitate the fitting of the Lyman α forest, we developed a semi-automatic line-fitting code. Briefly described, the code works with ~ 1500 km s^{-1} sections of spectrum at a time[5], fitting to Lyα and as many higher-order lines as are accessible within the HIRES spectrum at the redshift of the H I systems being fit. The algorithm first searches for systems by cross correlating a template hydrogen absorption spectrum (i.e., a single non-saturated H I absorption component) with the HIRES spectrum. Peaks in the cross correlation are taken as initial estimates of the centroids of absorption lines. We fit Gaussians to these lines to estimate column densities and Doppler parameters. Residual absorption features (i.e., those inconsistent with Lyα) in the Lyα portion of the spectrum are assumed to be metal lines. Residual absorption in the higher-order Lyman series sections of the spectra are assumed to be lower-redshift H I systems. The fits begin with Lyα at the high-redshift end of the range so that their higher-order Lyman absorption can be flagged as a known contaminant for fitting lower-redshift H I absorbers.

Once estimates of the locations, column densities, and Doppler parameters of all the absorption lines are complete, they are input into the χ^2 minimization code VPFIT[6] written by R.F. Carswell and J.K. Webb. VPFIT simultaneously fits all transitions of H I as well as the specified contaminating lines. The results are checked by eye, alterations made where the fit is inappropriate, and the process is repeated iteratively until a good fit (reduced $\chi^2 \approx 1$) is achieved. At this point, the multiple sections of spectrum are spliced together until a full fit to the forest is achieved.

It should be noted that the Voigt profile fit to a spectrum does not represent a unique solution. In this work, we fit each set of absorbers with the minimum number of components, adding additional components

[4]Note that the exact saturation point for the Lyα line and for any other line depends on the line width (i.e. the Doppler parameter) and thus the internal properties of the absorbing cloud.

[5]Note, the choice of 1500 km s^{-1} was made largely out of convenience. This velocity window is the largest window which can be easily displayed to check the goodness of fit and for which the number of components allows for a reasonable VPFIT run time.

[6]http://www.ast.cam.ac.uk/ rfc/vpfit.html; © 2007 R.F. Carswell, J.K. Webb

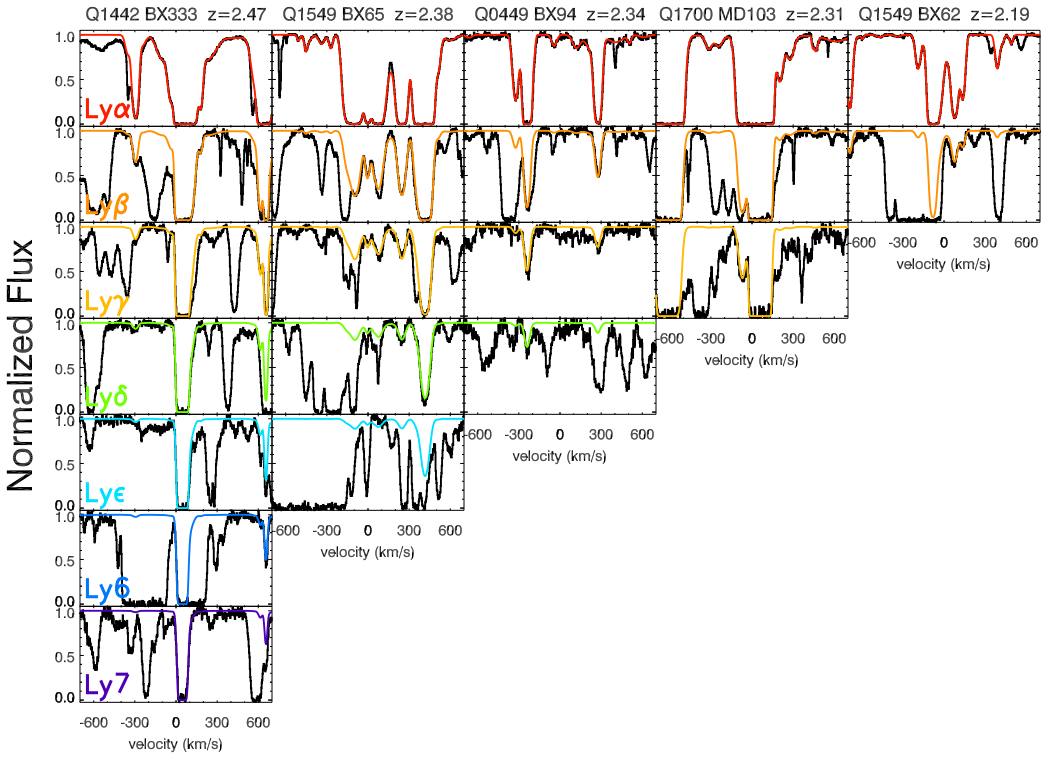


Figure 2.4 Example fits to the QSO data surrounding the redshifts of galaxies in our sample. Displayed in black are the continuum-normalized HIRES spectra showing the Lyman series transitions within $\pm$700 km s^{-1} of the systemic redshift of 5 galaxies with redshifts as indicated. Over-plotted in color are Voigt profile decompositions for H I absorption systems within $\pm$700 km s^{-1} of the galaxy redshift. Successive rows illustrate the fit to Lyα, Lyβ, etc. Absorption in the HIRES spectra that does not appear in the colored fit corresponds to absorption from metallic species or from H I at a redshift far from that of the galaxy. Note that for all galaxies in our sample, the QSO spectra cover the Lyα and Lyβ transitions near the galaxy redshifts, and for the higher-redshift galaxies, many more transitions in the Lyman series can be measured.

only when they significantly improve the χ^2. Median errors in $N_{\rm HI}$ reported by VPFIT are 0.07 dex for absorbers with $13 < \log(N_{\rm HI}) < 14$ and 0.03 dex for absorbers with $14 < \log(N_{\rm HI}) < 16$; however in many cases the systematic errors will exceed these values. The largest source of error in our fits to low column-density systems is uncertainty in the continuum level; for high column-density absorbers, it is the possibility of unrecognized sub-component structure.

The complete decomposition of the Lyα forest in these 15 lines of sight includes Voigt profile fits to > 5900 distinct H I absorption systems with $N_{\rm HI} > 10^{12.0}$ cm^{-2}, making it the largest absorber catalog ever compiled at these redshifts. It increases by an order of magnitude the number of intermediate $N_{\rm HI}$ absorbers measured with the additional constraint of higher-order Lyman lines.

2.4 Circumgalactic H I

In the following sections, we discuss the statistics of individual H I absorption systems with respect to the redshifts and transverse positions of galaxies. We do not uniquely "assign" each absorber to a specific galaxy or vice versa. Instead we rely on comparisons between the absorption measured close to galaxies with that typical of "random" locations[7] in the IGM. This allows us to quantify the significance of any apparent correlation with galaxies.

In principle, it would be preferable to compare absorbers found close to galaxies with those found in IGM locations known to be far from galaxies; however, because the galaxy sample is spectroscopically incomplete compared to our photometrically selected targets,[8] we do not measure the redshifts of all galaxies in our survey volume. As a consequence, we do not know which locations in 3D space *do not* have a nearby galaxy. Thus, we can only compare locations near to galaxies with random locations in the IGM (irrespective of the positions of galaxies).

In order to reproduce the absorber distributions for a typical place in the IGM, we compiled a catalog of 15,000 random locations (in both redshift space and on the plane of the sky). The redshifts are drawn from the actual galaxy redshift list and therefore reproduce the typical IGM absorption associated with the redshift ranges covered by our galaxy sample. With each redshift, we also associate a randomly drawn QSO and an impact parameter to the QSO sightline from the real galaxy impact parameter list. We can then study the distribution of absorption systems around these 15,000 random locations and thereby understand how the presence of a galaxy alters the distributions.

Below we consider the distribution of H I surrounding galaxies: first along the line of sight, then on the plane of the sky, and finally as a function of 3D distance. These measurements are used to determine the relevant velocity and transverse scales of circumgalactic H I.

[7] A "random" location, as described later in this section, is in effect a random redshift from our galaxy redshift catalog and an independent random position on the sky from our galaxy position catalog.

[8] Our photometric sample is also incomplete at the faintest apparent magnitudes (see Reddy et al. 2008), and no attempt is made to include galaxies with $\mathcal{R} > 25.5$.

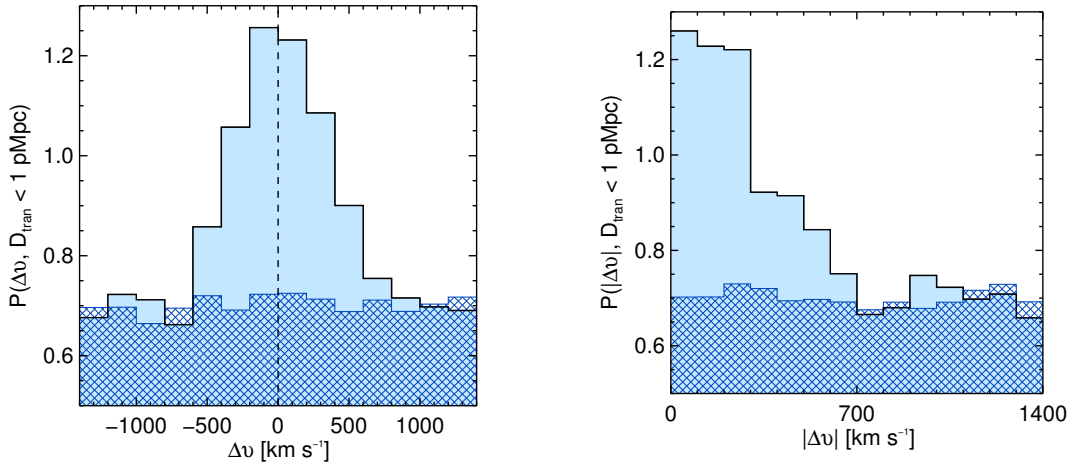


Figure 2.5 The velocity-space distribution of H I absorption systems with respect to the systemic redshift of galaxies, normalized by the number of galaxies in the sample. Absorbers with $\log(N_{\rm HI}) > 13$ and within 1 pMpc of the sightline to a QSO are included. The solid histogram represents the distribution of H I around galaxies, whereas the hatched histogram represents the average absorber density near randomly-chosen redshifts drawn from our galaxy redshift distribution.

2.4.1 Velocity-Space Distribution of H I Near Galaxies

The properties of H I gas near galaxies in our sample can be quantified in several ways. First, we consider the line of sight velocity distribution of absorbers relative to the redshifts of galaxies. Shown in Figure 2.5 is the velocity distribution of all absorbers with $N_{\rm HI} > 10^{13}$ cm^{-2} within 1400 km s^{-1} in redshift and 1 pMpc in projected distance from a galaxy. We define the velocity offset, Δv, of an absorber,

$$\Delta v \equiv \frac{\left(z_{\rm abs} - z_{\rm gal}\right) c}{1 + z_{\rm gal}} \tag{2.6}$$

where $z_{\rm abs}$ is the absorption system redshift and $z_{\rm gal}$ is the adopted systemic redshift of the galaxy from §2.2.2. With this definition, absorbers blue-shifted with respect to galaxies have negative Δv.

We define $P(\Delta v, D_{\rm tran})$ as the number of absorbers per galaxy at a given Δv and within the specified range of impact parameters $D_{\rm tran}$. The solid histogram in Figure 2.5 represents the distribution of H I around galaxies, whereas the hatched histogram shows the average number of absorbers expected relative to randomly chosen redshifts, as described above. The number of absorbers is clearly higher near galaxies, with an excess peaking near $\Delta v = 0$ (the galaxy systemic redshift) and confined to $\Delta v \simeq \pm 700$ km s^{-1}.

Figure 2.6 shows a similar pattern; in this case, each absorption system is weighted in proportion to $\log(N_{\rm HI})$ such that high-column density systems contribute more significantly to the histogram. Each absorber with $\log(N_{\rm HI}) > 13$ contributes $[\log(N_{\rm HI}) - 13]$ to the histogram, which is normalized by the number of galaxies considered. Taken together, Figures 2.5 and 2.6 indicate that there is an increase near galaxies of both the number and the column density of H I absorbers. The narrow peak of the Δv distribution has an apparent half-

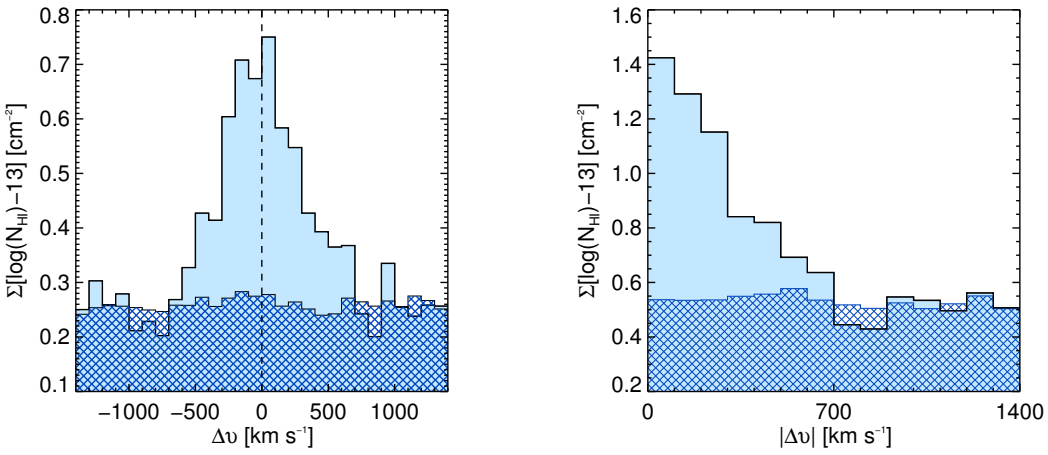


Figure 2.6 As in Figure 2.5, where the histogram is $N_{\rm HI}$-weighted as described in the text.

width of $\sim$300 km s^{-1}, while the full excess extends to $\simeq \pm$700 km s^{-1} (most clearly shown in the right-hand panel of Figure 2.6).

As has been argued by Shapley et al. (2003); Adelberger et al. (2003); Steidel et al. (2010); Rakic et al. (2011a), these velocity distributions are also useful for checking our redshift calibration; their symmetry about Δv= 0 km s^{-1} is a sensitive probe of systematic errors in our galaxy systemic redshift calibration, while the width of the distribution provides an upper limit on the random errors.[9] A Gaussian fit to Figure 2.5 yields a mean $\Delta v = 18 \pm 26$ with a standard deviation of 308 ± 30 km s^{-1}.

An alternative method of quantifying the column density dependence in Figure 2.6 is to examine the distribution of $N_{\rm HI}$ as a function of Δv. Figure 2.7 shows the median $N_{\rm HI}$ as a function of Δv for all absorbers within $D_{\rm tran} \leq$ 300 pkpc of a galaxy[10]. The asterisks indicate the median value of $N_{\rm HI}$ for each bin in velocity space. The dark vertical lines represent the dispersion in the median computed through the bootstrap method[11]. The value of the median column density across all velocity bins for the random distribution is shown as the dark horizontal bar; the light shaded contours are the 1-σ bootstrapped dispersions in the median of the random sample, where we consider samples of the same size as those from the real distribution. Note that there is an enhancement by a factor of $\simeq 3-10$ in the median $N_{\rm HI}$ out to $|\Delta v| \simeq 300$ km s^{-1} relative to galaxy redshifts. However, for $|\Delta v| > 300$ km s^{-1}, the measurements are consistent with random places in the IGM.

Shown in Figure 2.8 are the individual measurement of $N_{\rm HI}$ as a function of Δv for all absorbers within 1000 km s^{-1} of a galaxy within $D_{\rm tran} \leq$ 100 pkpc of a QSO sightline. Notably, the higher-$N_{\rm HI}$ absorbers cluster

[9] The width is an upper limit because it includes both redshift errors *and* any peculiar velocity of absorbers relative to galaxies.

[10] We choose 300 pkpc as our distance cut because the majority of the excess absorption is found within that zone, as described in §2.4.2.

[11] To bootstrap the dispersion in the median, we draw (with replacement) sets of data with the same size as the real distribution. For each set the median is computed; 100 such data sets are evaluated, and then we take the 1-σ symmetric bounds on the distribution of the medians, i.e. the 16th and 84th percentiles.

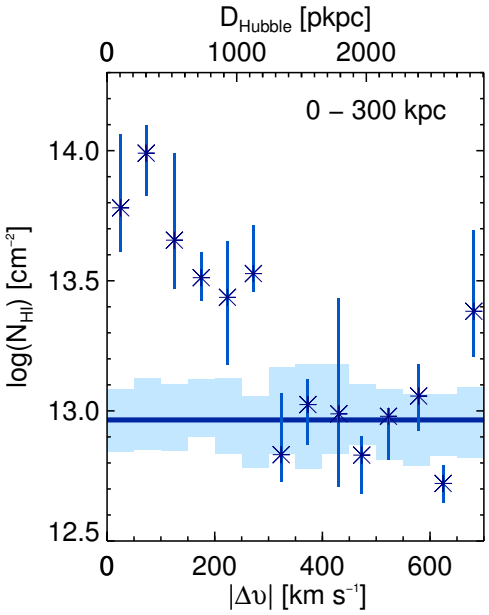


Figure 2.7 The median value of $N_{\rm HI}$ as a function of the velocity offset of absorbers with respect to those galaxies with impact parameters $D_{\rm tran} <$ 300 pkpc. Asterisks represent the median value of $N_{\rm HI}$, dark vertical bars are the 1-σ dispersion in the median determined via the bootstrap method. The dark horizontal line is the median value of $N_{\rm HI}$ in the random distribution. The light shaded boxes are the bootstrapped symmetric 1-σ dispersion in the median values of the samples drawn from the random distribution. The top x-axis shows the conversion between velocity offset and Hubble distance. See Equation 2.7 and §2.4.3 for further discussion.

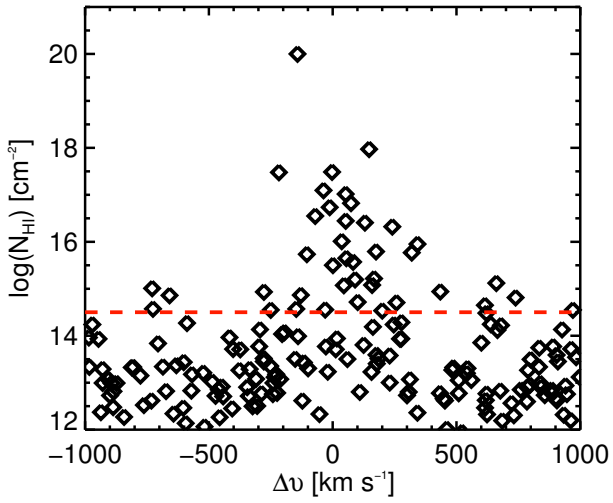


Figure 2.8 $N_{\rm HI}$ as a function of Δv for absorbers with $D_{\rm tran}<$ 100 pkpc. The dashed (red) line marks the position of $\log(N_{\rm HI}) = 14.5$. The significance of this column density threshold is discussed in §2.5.

strongly near the galaxy systemic redshift with a full width of $\sim \pm 300-400$ km s^{-1}.

In summary, Figures 2.5 – 2.8 clearly illustrate an enhancement in both the column density and number of absorption systems near the systemic velocities of galaxies. The most significant enhancement of column density is seen within $\pm$ 300 km s^{-1} (Figure 2.7), but there is a higher number of absorbers out to at least $\pm$ 700 km s^{-1} (Figures 2.5 and 2.6).

2.4.2 Transverse Distribution of Absorbers

In addition to the strong velocity alignment of absorption systems with the systemic redshift of nearby galaxies, there is also a significant increase in the column densities, $N_{\rm HI}$, of individual absorbers with decreasing projected (or transverse) distance between the galaxy and the line of sight, $D_{\rm tran}$. Here, we suppose that, as an ensemble, these galaxies show similar circumgalactic absorption signatures. Therefore, because galaxies fall at various discrete impact parameters from the QSO line of sight, we can combine the information from each galaxy to make a sparsely sampled map of the absorption as a function of $D_{\rm tran}$ relative to the ensemble galaxy.

We introduce two related statistics designed to trace the change in column density as a function of $D_{\rm tran}$. Recalling that $\pm$700 km s^{-1} encompassed the bulk of the "excess" absorption (§2.4.1), for each galaxy we define max($N_{\rm HI}$,700km s^{-1}) to be the value of $\log(N_{\rm HI})$ for the strongest absorber with $|\Delta v|<$700 km s^{-1} of the galaxy systemic redshift. The left panel of Figure 2.9 shows the median value of max($N_{\rm HI}$,700km s^{-1}) as a function of impact parameter. A second statistic is the logarithm of the *total* $N_{\rm HI}$, sum($N_{\rm HI}$), of all absorbers with $|\Delta v|<$700 km s^{-1}. The statistics of the median value of sum($N_{\rm HI}$,700km s^{-1}) versus $D_{\rm tran}$ is shown in the right-hand panel of Figure 2.9. Generally, the values of these two statistics are quite similar because the $N_{\rm HI}$

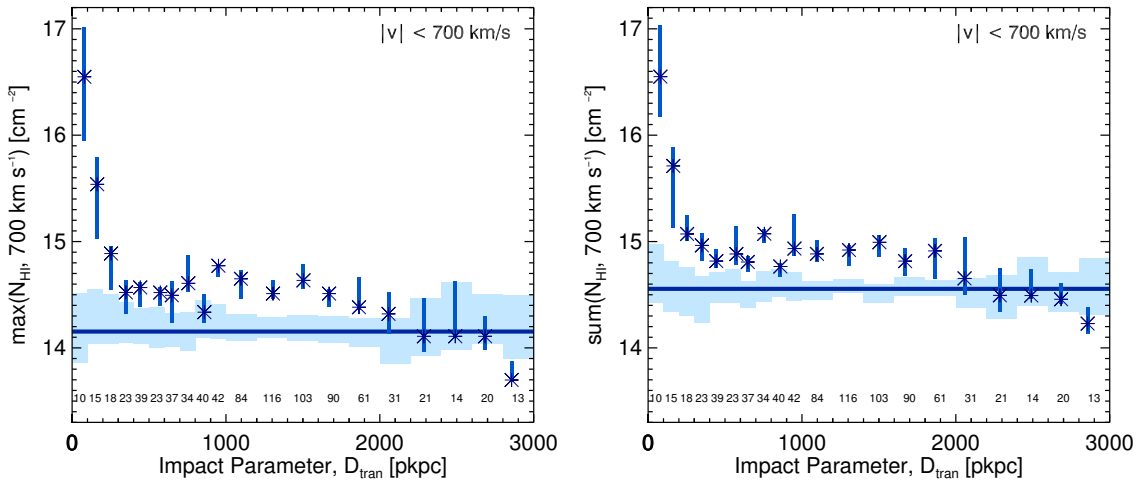


Figure 2.9 The log column densities of the strongest $N_{\rm HI}$ absorbers as a function of transverse distance. *On the left* we consider the max($N_{\rm HI}$) statistic, log($N_{\rm HI}$) of the single strongest absorber per galaxy with $|\Delta v|<$ 700 km s^{-1}. *On the right* is the sum($N_{\rm HI}$) statistic, the log of the sum of the $N_{\rm HI}$ of all the absorbers within $|\Delta v|<$ 700 km s^{-1}. Asterisks represent the median value of the considered statistic in a given bin of $D_{\rm tran}$. Dark vertical bars are their dispersions. The horizontal position of the asterisks represent the median $D_{\rm tran}$ of the galaxies in that bin. The number of galaxies in each bin is indicated at the bottom of the plot. The dark horizontal line is the median value drawn from the random distribution. The light shaded boxes are the bootstrapped symmetric 1-σ dispersion in the median values of the samples drawn from the random distribution. The bin size is 100 pkpc for absorbers with $D_{\rm tran}<$ 1 pMpc and 200 pkpc for those with $D_{\rm tran}>$ 1 pMpc. We increase in binning to reduce the shot noise in the bins at $D_{\rm tran} > 2$ pMpc which have fewer galaxies due to the limited field-of-view of LRIS.

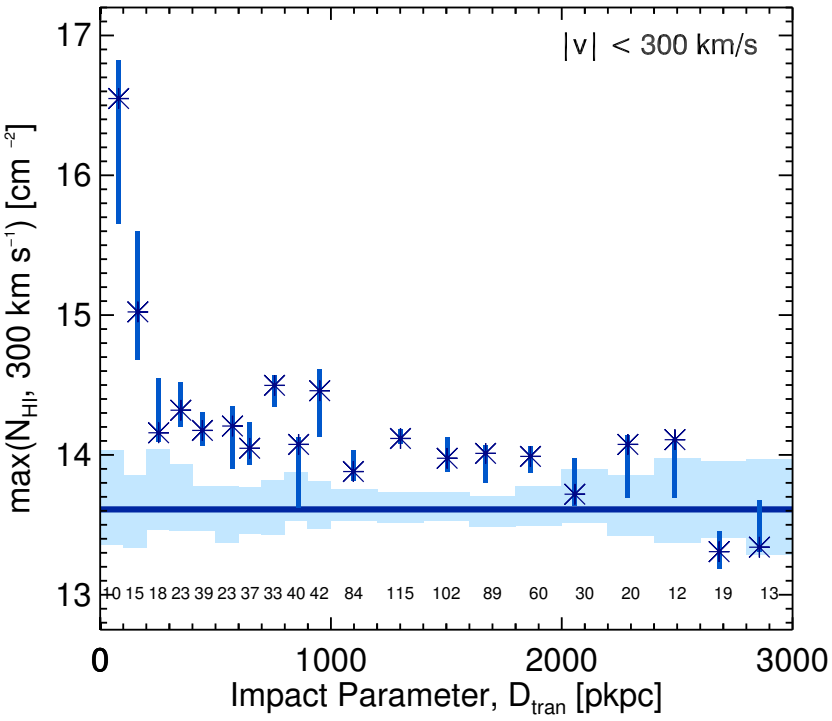


Figure 2.10 Same as Figure 2.9 but for the maximum column density absorber within $\pm$300 km s^{-1}. Changing the velocity interval considered with the max($N_{\rm HI}$) statistic has little effect on the observed trends.

in most velocity windows is dominated by the single highest-$N_{\rm HI}$ absorber. We consider the sum because it is most easily compared to results of numerical simulations, as it does not require the fitting of Voigt profiles to simulated data and can instead be compared to a simulation "collapsed" along the line of sight.

Figure 2.9 clearly demonstrates that both max($N_{\rm HI}$) and sum($N_{\rm HI}$) increase rapidly as one approaches a galaxy. In the bin corresponding to the smallest impact parameters, $D_{\rm tran}<$ 100 pkpc, the median value of max($N_{\rm HI}$) is more than two orders of magnitude higher than that of a random location. Moving outwards, the median value decreases with increasing $D_{\rm tran}$ to 300 pkpc, at which point the statistic "plateaus" and remains significantly higher than the random sample out to 2 pMpc. The plateau value in the galaxy-centric sample is max($N_{\rm HI}$)$\simeq 10^{14.5}$ cm^{-2}, while that of the random distribution is max($N_{\rm HI}$)$\simeq 10^{14.1}$ cm^{-2}. As we will discuss in §2.4.2.1, max($N_{\rm HI}$) begins to decline for $D_{\rm tran}>$ 2 pMpc.

A relevant question concerns the dependence of these statistics on the size of the velocity window considered. In Figure 2.9 we considered the maximum column density absorber within $\pm$700 km s^{-1} of the systemic velocity of each galaxy. This corresponds to the full width of the velocity distribution shown in Figure 2.6. However, it is clear from Figures 2.6 and 2.7 that the majority of the excess *strength* of absorption falls within $\pm$300 km s^{-1}, especially for those systems with small impact parameters. Figure 2.10 shows max($N_{\rm HI}$,300km s^{-1}); the trends are similar, though in the more restricted velocity window the peak on small

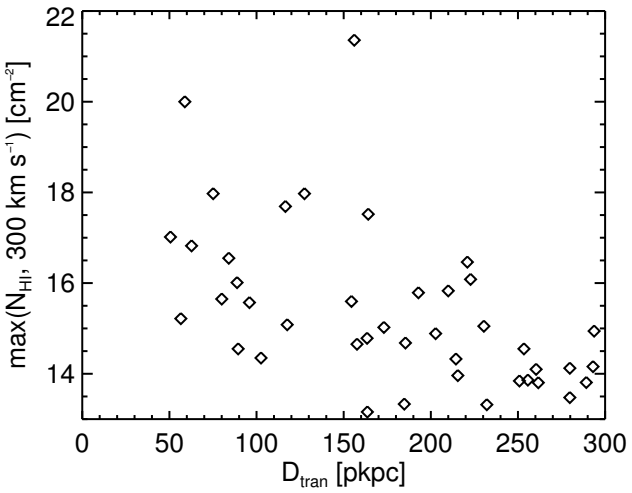


Figure 2.11 The individual measured values of max($N_{\rm HI}$,300km s^{-1}) for galaxies with $D_{\rm tran}<$ 300 pkpc.

scales is higher relative to random IGM locations–the median value in the first bin is 3 dex higher than the random redshift sample, and the extended floor of absorption is increased to at least .5 dex above the median of the random-redshift distribution. The more significant excess over random of the 300 km s^{-1} version is primarily due to the exclusion of unrelated absorbers at large velocity separation; however, we note that at $D_{\rm tran}>$ 1200 pkpc, the differential Hubble velocity associated with this distance along the line of sight is $>$ 300 km s^{-1}, meaning that for large $D_{\rm tran}$ it may be more appropriate to adopt max($N_{\rm HI}$,700km s^{-1}) as the relevant statistic. Regarding the sum statistic for the smaller velocity window (not shown), the value of the plateau and that of the random sample is $\sim$ 0.3 dex higher for sum($N_{\rm HI}$,300 km s^{-1}) than for max($N_{\rm HI}$,300 km s^{-1}), similar to the variation in the statistics shown in Figure 2.9.

As we will argue in §2.5 and §2.8.1, the velocity and spatial scales of 300 km s^{-1} and 300 kpc capture the most significant excess in both the column density and the number of absorbers near galaxies. In Figure 2.11 we provide the individual measurements max($N_{\rm HI}$, 300 km s^{-1}) for all galaxies in the sample with $D_{\rm tran}<$ 300 pkpc. Note the large intrinsic scatter in max($N_{\rm HI}$), even at fixed impact parameter. For the 10 galaxies with the smallest impact parameters ($D_{\rm tran}<$ 100 pkpc), the relevant portions of the QSO spectra within $\pm$1000 km s^{-1} of Lyα at $z_{\rm gal}$ are reproduced in Figure 2.12.

2.4.2.1 The Large-Scale Distribution of H I

We now consider the larger-scale distribution H I around galaxies. Unfortunately, the sampling of galaxies with $D_{\rm tran}>$ 2 pMpc is comparatively poor in our sample due to the survey geometry of most of the KBSS fields (typically 5.5 $\times$ 7.5 arcmin on the sky). This scale is imposed by the footprint of LRIS; with each field roughly centered on the bright QSO, the maximum observed impact parameter would be $\simeq$ 2.15 pMpc at $\langle z\rangle$ = 2.3. However, three out of 15 survey fields (see Table 2.1) were imaged with other instruments covering larger angular sizes and thus provide information on larger transverse scales. In Figures 2.9 and 2.10 we

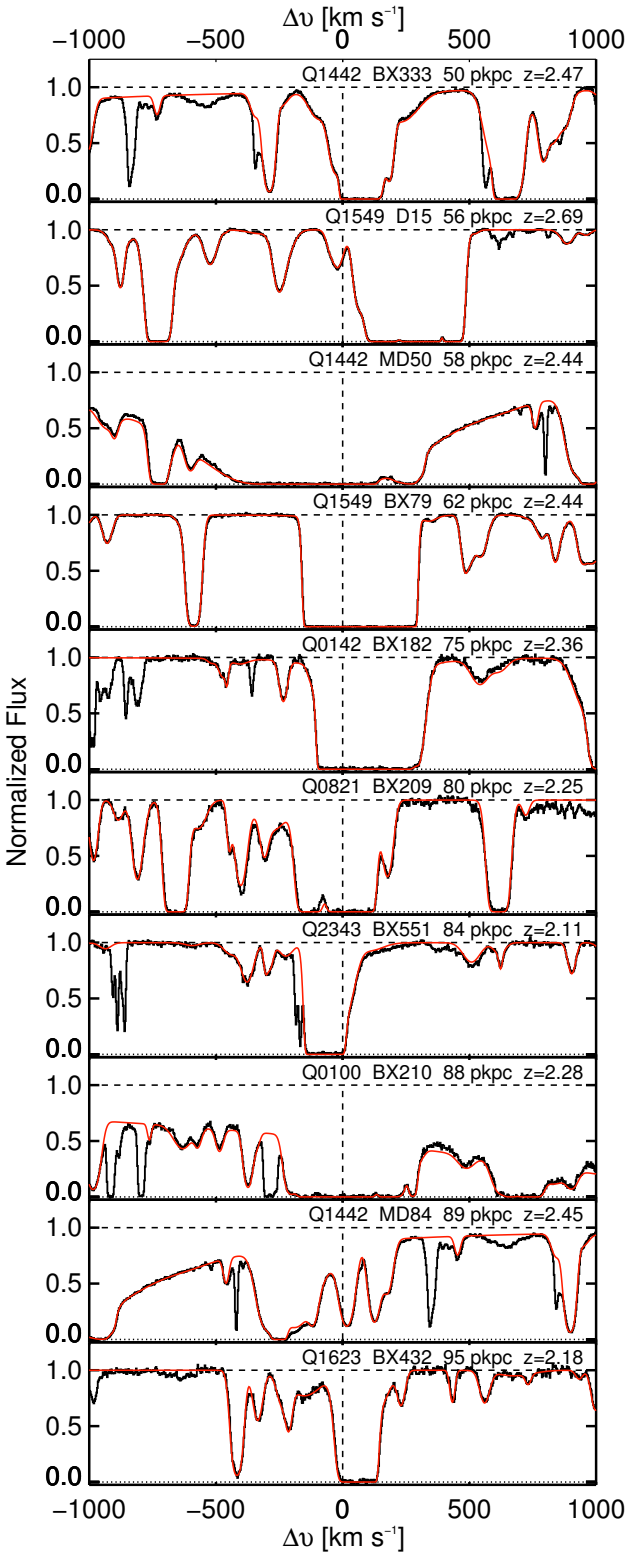


Figure 2.12 Lyα absorption within $\pm$1000 km s^{-1} of the systemic redshift of the 10 galaxies within $D_{\rm tran}<100$ pkpc of the line of sight to the QSO. The HIRES data are in black, while the red shows our Voigt profile decomposition of the H I absorption near the redshift of the galaxy. The continuum and zero level of the spectrum are shown in dashed and dotted lines respectively. The systemic redshift of each galaxy is marked by the vertical dashed line at 0 km s^{-1}. Note that the continuum is depressed in some of the spectral regions surrounding Q1442-BX333, Q1442-MD50, Q0100-BX210, and Q1442-MD84 by a DLA or sub-DLA near the galaxy redshift as described in §2.2.3.

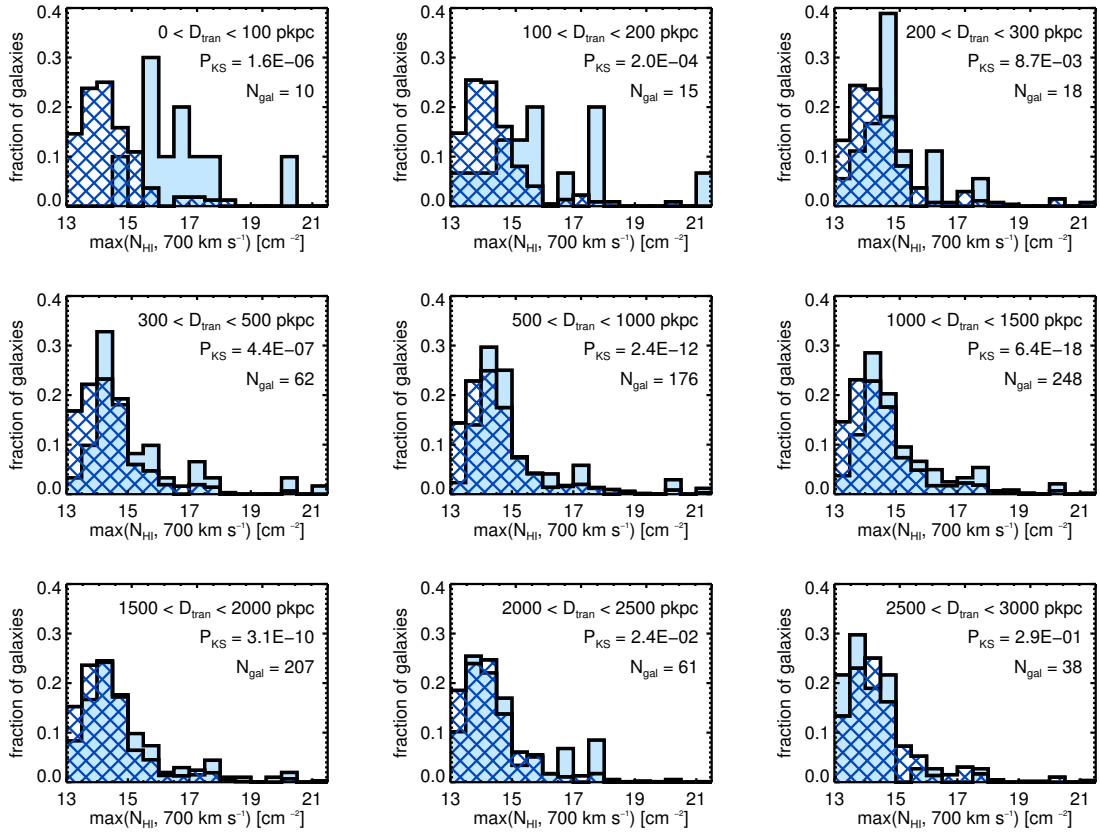


Figure 2.13 The max($N_{\rm HI}$,700km s^{-1}) statistic as a function of impact parameter (different panels). The hatched histograms are the values for the random sample, the solid histograms are the values for the real sample. These histograms quantify the variation from galaxy to galaxy of the max($N_{\rm HI}$) statistic at fixed impact parameter. $P_{\rm KS}$ is the probability that the two max($N_{\rm HI}$) sets were drawn from the *same* distribution. Notably, the last panel with the highest value of $D_{\rm tran}$ has the least significant departure from the random distribution.

use wider bin sizes for absorbers with $D_{\rm tran}>$ 1 pMpc in order to consider the large-scale distribution. This reduces the shot noise in the bins with $D_{\rm tran}>$ 2 pMpc. Figures 2.9 and 2.10 demonstrate that max($N_{\rm HI}$) remains higher than the global median value in the IGM (dark horizontal line) out to $\sim$2 pMpc, and then begins to decline. For larger $D_{\rm tran}$, the data suggest column densities at or below that of random places in the IGM.

Again, considering the degree of scatter in max($N_{\rm HI}$) at fixed impact parameter, Figure 2.13 shows max($N_{\rm HI}$,700km s^{-1}) for various bins in $D_{\rm tran}$, as indicated. The top row of panels correspond to $D_{\rm tran}<$ 300 pkpc, while the bottom two rows of panels consider larger impact parameters. Each panel shows the Kolmogorov-Smirnov probability that the two histograms are drawn from the same parent distribution. Notably, only the 2 $<D_{\rm tran}<$ 3 pMpc bins have a distribution of max($N_{\rm HI}$) consistent with the random sample.

Thus, we have shown that the column density of H I peaks sharply at the position of galaxies in the transverse direction, that the width of the peak is $\simeq$ 300 pkpc, and that there remains a significant excess of

H I gas to $D_{\rm tran} \simeq 2$ pMpc. In §2.8 we discuss the implications of these results.

2.4.3 3D Distribution of $N_{\rm HI}$

The 3D distance, $D_{\rm 3D}$, is computed using the quadrature sum of the physical impact parameter ($D_{\rm tran}$) and the line-of-sight distance calculated assuming the velocity differences Δv are due entirely to the Hubble flow,

$$D_{\rm Hubble}(\Delta v, z) \equiv \frac{\Delta v}{H(z)}. \tag{2.7}$$

The 3D distance is therefore

$$D_{\rm 3D}(\Delta v, z, D_{\rm tran}) \equiv \sqrt{(D_{\rm tran})^2 + \left(\frac{\Delta v}{H(z)}\right)^2}\ , \tag{2.8}$$

where $H(z)$ is given by

$$H(z) = H_0 \sqrt{\Omega_{\rm m}(1+z)^3 + \Omega_\Lambda} \tag{2.9}$$

such that $H(z = 2.3)$ in our cosmology is

$$H(z = 2.3) = 240 \text{ km s}^{-1} \text{ Mpc}^{-1}. \tag{2.10}$$

In this formalism, each absorber has a unique $D_{\rm 3D}$ with respect to a galaxy in the same field.

The 3D distance, due to its strong dependence on $|\Delta v|$, requires that absorbers have *both* small $D_{\rm tran}$ *and* very small values of $|\Delta v|$ in order to populate bins at small values of $D_{\rm 3D}$. As a result, $D_{\rm 3D}$ has the effect of isolating those absorbers *most likely* to be associated with the galaxy without imposing a velocity cut.

Figure 2.14 shows that the median $N_{\rm HI}$ stays above that of an average place in the universe out to $D_{\rm 3D} \simeq 3$ pMpc.[12] The decline of $N_{\rm HI}$ as a function of $D_{\rm 3D}$ is quite smooth, but again strongly peaked at the position of galaxies. The pixel analysis of the KBSS sample recently completed by Rakic et al. (2011b) studies in detail the 3D distribution of H I optical depths. There is excellent agreement between the optical depths measured in Rakic et al. (2011b) and the $N_{\rm HI}$ trends shown in Figure 2.14. The smoothness of the decline is caused by the shifting of absorbers with small $D_{\rm tran}$ and modest $|\Delta v|$ into bins at larger $D_{\rm 3D}$. Thus, the signal appearing in the inner 300 pkpc in Figure 2.9 is distributed across a larger number of bins in $D_{\rm 3D}$ as a result of the velocity distribution of the absorbers.

The nature of the velocity width of the excess absorption will be discussed at length in §2.6; however, it should be noted that all measurements of $|\Delta v|$ rely on the accuracy of the galaxy redshifts and in addition are affected by whatever peculiar velocities are present, whether due to random motion, inflows, or outflows.

[12] Note, this is consistent with the quadratic sum of the extent of the velocity and transverse distance excesses seen in previous sections.

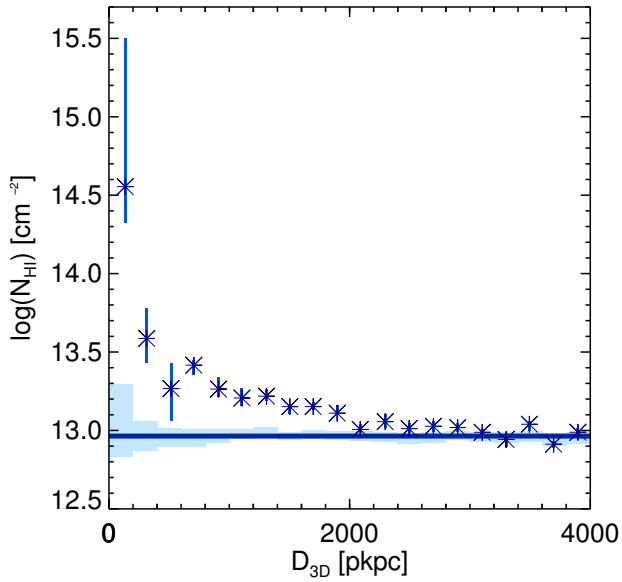


Figure 2.14 The median column density of all absorption systems within $\pm$1400 km s^{-1} of a galaxy as a function of the 3D distance between the absorber and the galaxy. The symbols have the same meaning as those in Figure 2.9. Note the steeply rising column densities at small D_{3D} and that the median value remains above that of a random location (horizontal bar) out to $\sim$3 pMpc.

2.4.4 Connection to Galaxy-Galaxy Pair Results

Steidel et al. (2010) presented a sample of 512 close angular pairs of galaxies with different redshifts (drawn primarily from the same KBSS catalogs used in the present paper), using the spectrum of the background galaxy to probe gas associated with that in the foreground. They were able to measure the strength of absorption from H I and several metallic species over a range in impact parameter $D_{\rm tran}$= 8 – 125 pkpc. The principal advantage of this method is that it allows for probes at very small angular separation ($\theta < 5''$)–obtaining statistical results for galaxies at such small separations from QSOs is difficult due both to the "glare" of the QSO and the relative rarity of QSO-galaxy pairs with very small separations (the smallest QSO sightline–galaxy separation is $D_{\rm tran}$= 50 pkpc or $\theta \approx 6''$) . As such, the galaxy pair results are highly complementary to the QSO sightline study described in this paper.

Using background *galaxies* instead of QSOs results in two important differences between these studies: first, background galaxies have a projected "beam size" of $\sim$ 1 kpc at the location of the foreground galaxy, whereas background QSOs have a projected beam of order $\sim$ 1 pc. As a result, the absorption seen against background galaxies measures a combination of the covering fraction of gas on kpc scales and the column density of absorbers. Further, background galaxies are faint and therefore only low-dispersion, low-S/N spectra can be obtained. Steidel et al. (2010) thus used stacks of background galaxy spectra shifted into the rest frame of the foreground galaxies to quantify the average absorption profile surrounding these galaxies. The lower-resolution spectra do not separate into individual components as would be found in individual

QSO spectra. As such, the galaxy pair method allows for the measurement of equivalent widths (W_0) only. As discussed by Steidel et al. (2010) interpretation of W_0 is complicated by its sensitivity to the covering fraction and velocity extent of the absorbing material, and its relative *insensitivity* to ionic column density.

Steidel et al. (2010) measured W_0 of H I, C IV, C II, Si IV, and Si II in bins of impact parameter ($D_{\rm tran}$). The large W_0, particularly at $D_{\rm tran} \lesssim 40$ pkpc, seemed to require high line-of-sight velocity spread in the gas– likely higher than could be easily explained by gravitationally-induced motions, but easily accounted for by the high velocities observed "down the barrel" in the spectra of LBGs, which typically reach $\Delta v \sim 800$ km s^{-1} for galaxies with properties similar to those in our sample.

Steidel et al. (2010) found that a simple geometric and kinematic model of outflows from LBGs could account simultaneously for the behavior of W_0 as a function of $D_{\rm tran}$ as well as the shape of the blue-shifted line profiles of strong interstellar absorption features observed along direct sightlines to the same galaxies. The presence of detected C II absorption to $D_{\rm tran} \sim 90$ pkpc was cited as evidence for a non-negligible covering fraction of H I gas having $N_{\rm HI} > 10^{17}$ cm^{-2} at such large galactocentric distances. Steidel et al. (2010) remarked that 90 kpc is very close to the expected virial radius $r_{\rm vir}$ for LBGs similar to those in the current sample. Our max($N_{\rm HI}$) statistic agrees well with this inference (Figure 2.9); in fact, we find a covering fraction $\simeq 30\%$ for absorbers with $N_{\rm HI} > 10^{17.2}$ cm^{-2} for $r \lesssim r_{\rm vir}$– see § 2.5.2.

In addition, Steidel et al. (2010) used the same HIRES spectra as this work to show that the W_0(Lyα) measured from the relevant portions of the high resolution spectra, averaged in the same way as the background galaxy spectra, are consistent with an extrapolation to $D_{\rm tran} \simeq 250$ pkpc of the trend seen in the galaxy-galaxy pairs results; a similar conclusion was reached by Rakic et al. (2011b), also using equivalent width analysis of the QSO spectra, where a smooth trend was noted out to $D_{\rm tran} > 1$ pMpc.

2.4.5 Comparison to Previous Studies at $z > 2$

Adelberger et al. (2003, 2005a) conducted the first systematic studies of high-z galaxies and their surrounding IGM using sightline surveys of the $2 < z < 4$ IGM paired with large LBG surveys designed to probe galaxies in the same volume. Adelberger et al. (2003) analyzed the transmitted flux in the Lyα forest of background QSOs, evaluated near the redshifts of survey galaxies. At the time, this was most easily accomplished using $z \sim 3$ LBGs and $z \simeq 3.5$ background QSOs. These authors did not attempt Voigt profile decompositions (as in the present work) but focused on transmitted flux because it made for easier comparisons to theory and because the spectra covered the Lyα transition only, making measurements of column density or optical depth difficult due to limited dynamic range for any $N_{\rm HI} \gtrsim 10^{14.5}$cm^{-2}. Based on 8 QSO sightlines covering $3 \lesssim z \lesssim 3.6$, Adelberger et al. (2003) found that excess H I absorption (i.e., lower transmitted flux than the average IGM at the same redshift) was present within $\lesssim 5$h^{-1}comoving Mpc (cMpc) of galaxies [1.7 physical Mpc (pMpc) at $\langle z \rangle = 3.3$ using the cosmology adopted in the present work], and an intriguing but not highly significant lack of H I very close to these galaxies (< 0.5h^{-1}cMpc or < 170 pkpc). C IV absorption was observed to be correlated with galaxy positions to out to 2.4h^{-1}cMpc (or 800 physical kpc). The cross-

correlation of C IV systems with LBGs was found to be similar to the LBG auto-correlation, suggesting that metal enriched IGM and galaxies shared the same volumes of space. The strongest C IV absorbers were so strongly correlated with LBG positions that the authors concluded that they must be causally connected to one another.

Adelberger et al. (2005a) extended similar studies to $1.8 \leq z \leq 3.3$, using a larger number of QSO sightlines and a wider range of QSO spectra for the analysis. Based once again on the transmitted flux statistics, the large-scale excess $N_{\rm HI}$ near galaxies was consistent with the results of Adelberger et al. (2003), but the "turnover" of H I absorption on the smallest scales was not confirmed using the larger sample at somewhat lower redshift. Generally, the transmitted flux was decreasing with galactocentric distance, though it was still the case that $\sim 30\%$ of the galaxies with the smallest impact parameters showed a lack of strong absorption, interpreted as evidence that the gas is clumpy. The results for C IV were extended, and it was shown that the correlation with galaxies grows increasingly strong as $N_{\rm CIV}$ increases; the correlation length of $N_{\rm CIV} \gtrsim 10^{12.5}$ cm^{-2} absorption systems was similar to that of the autocorrelation length of the galaxies. The authors attempted to correlate the observed IGM properties with galaxy properties, but no significant correlations were found given the relatively small sample of galaxies at small impact parameters.

There is no overlap in the data sample used in this paper with that of Adelberger et al. (2003) which focused on higher-redshift galaxies and QSOs. Three of the fields (Q1623, HS1700, and Q2343) included in our analysis were also included in Adelberger et al. (2005a) which considered similar redshifts to this work; however, we have increased the S/N of the QSO spectra and also added many galaxy redshifts to our catalogs for these fields since the earlier analysis. The most surprising result of the Adelberger et al. (2003) sample, especially in light of the work presented here, was the reported *deficit* of absorption found within 0.5 h^{-1} cMpc or 170 pkpc of galaxies. Adelberger et al. (2005a), with a larger data set, found that 1/3 of galaxies had relatively-little H I absorption (consistent with the pixel statistics of our sample presented in Rakic et al. 2011b), but that the majority of galaxies were associated with significant absorption. In our sample, only 1/21 galaxies with $D_{\rm tran} <$ 170 pkpc has a sum($N_{\rm HI}$, 300 km s^{-1}) value less than the median of the random sample. As such we concur with the argument of Adelberger et al. (2005a) suggesting that the reported deficit was due to the small sample size presented. Further, we note that a measure of $N_{\rm HI}$ shows that the majority of galaxies do have excess H I absorption in their surroundings.

On larger physical scales, Adelberger et al. (2003) and Crighton et al. (2011) considered the large-scale distribution of H I absorbers at $z \approx 3$ while Adelberger et al. (2005a) studied that at $z \approx 2.5$. All found evidence for increased absorption to $D_{\rm 3D} \approx$ 5-6 h^{-1} cMpc or ~ 2 pMpc. This scale is roughly consistent with our measurements of the 3D distribution of absorption presented in Figure 2.14, as well as with the trends in optical depth vs. $D_{\rm 3D}$ presented in Rakic et al. (2011b).

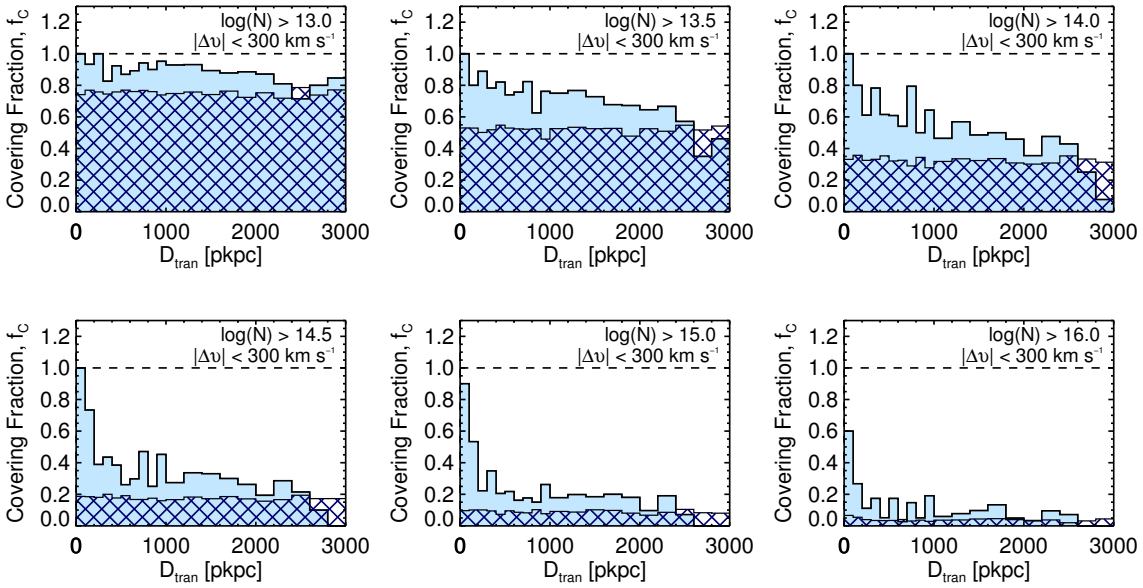


Figure 2.15 The covering fraction, f_c, of absorbers for various N_{HI} thresholds (different panels) as a function of D_{tran}. The solid histogram represents the fraction of galaxies with an absorber of a given N_{HI} or greater, within $\pm$300 km s^{-1}. The hatched histogram represents f_c at random locations in the IGM; the horizontal dashed line marks f_c = 1 (100% covering). The bin size is 100 pkpc for absorbers with $D_{tran} < 1$ pMpc and 200 pkpc for those with $D_{tran} > 1$ pMpc.

2.5 The Covering Fraction and Incidence of H I

As discussed in §2.4.1, the scale of the strongest correlation between N_{HI} and the positions of galaxies is found within 300 km s^{-1} of z_{gal}. Adopting this value as the characteristic velocity scale for circumgalactic gas, we can examine other useful measures of the gas distribution around galaxies. The covering fraction (f_c) and the incidence of absorbers (η_{abs}) are two ways of quantifying the geometry of the distribution as a function of both impact parameter and N_{HI}.

First, we define the covering fraction, $f_c(D_{tran}, N_0)$, as the fraction of galaxies in a bin of impact parameter, D_{tran}, that have an absorber within $|\Delta v| <$ 300 km s^{-1} with $N_{HI} > N_0$. This is equivalent to the geometric fraction of the area of an annulus centered on the galaxy that is covered by gas with $N_{HI} > N_0$ and $|\Delta v| <$ 300 km s^{-1}. This quantity measures the *variation* within the sample of the decline of N_{HI} as a function of D_{tran}.

A related quantity is the incidence of absorbers, η_{abs}, defined to be the number of absorbers *per galaxy* within a given range of N_{HI} and D_{tran}, and with $|\Delta v| <$ 300 km s^{-1}. Because this quantity can be greater than unity, it has larger dynamic range and so allows for a more-complete picture of the *average multiplicity* of absorbers at locations close to galaxies. Also, because we consider η_{abs} in differential bins of N_{HI}, it can be used to measure the degree to which absorbers of a given N_{HI} associate with galaxies.[13]

Figure 2.15 illustrates the dependence of f_c on D_{tran} for various thresholds of N_{HI} (different panels).

[13]To illustrate the difference between f_c and η_{abs}, suppose that all absorbers have the same N_{HI}. Then for a sample of two galaxies, one with 3 absorbers and the other with none, f_c=0.5 and η_{abs}= 1.5.

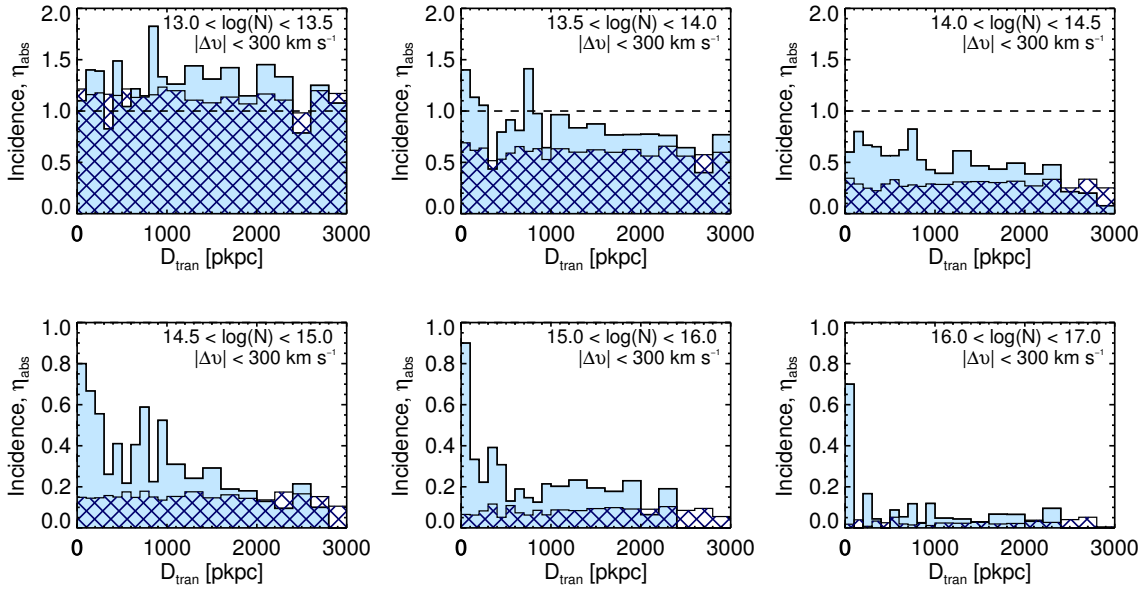


Figure 2.16 The incidence of H I absorbers, $\eta_{\rm abs}$, as a function of impact parameter, $D_{\rm tran}$. The solid histogram represents the mean number of absorbers per galaxy within $\pm$300 km s^{-1} at the given distance, whereas the hatched histogram represents the average incidence of absorbers near randomly-chosen redshifts. The different panels show various ranges of $N_{\rm HI}$. The bin size is 100 pkpc for absorbers with $D_{\rm tran} < 1$ pMpc and 200 pkpc for those with $D_{\rm tran} > 1$ pMpc. The incidence of absorbers exceeds the random distribution for $N_{\rm HI}$ $> 10^{13.5}$ cm^{-2} to $D_{\rm tran} > 2$ pMpc. Note that only absorbers with $\log(N_{\rm HI}) > 14.5$ show strong association with the positions of galaxies. Table 2.2 gives the data values for $f_{\rm c}$ and $\rm P_E$ determined from these distributions.

Within 100 pkpc, $f_{\rm c}> 0.5$ even for $N_{\rm HI}> 10^{16}$ cm^{-2}. Also, *every galaxy* has an absorber with $\log(N_{\rm HI}) > 14.5$ within 100 pkpc and $|\Delta v|< 300$ km s^{-1}. The distribution within 2.5 pMpc of lower-$N_{\rm HI}$ absorbers is relatively uniform, with $f_{\rm c}> 0.5$. However, at larger threshold $N_{\rm HI}$ (bottom three panels) we see that $f_{\rm c}$ is high ($\gtrsim 0.5$) only within $\sim$200 pkpc. We will return to our measured values of $f_{\rm c}$ in §2.5.2 and §2.5.3 where we compare our measurements to those made at low-z and also to results from numerical simulations.

Figure 2.16 shows the average incidence of absorbers, $\eta_{\rm abs}$, as a function of distance, $D_{\rm tran}$, and ranges of $N_{\rm HI}$ (different panels). Notably, in all panels (i.e., at all columns densities) the average $\eta_{\rm abs}$ is higher than random for $D_{\rm tran}< 2$ pMpc. For absorbers with $N_{\rm HI}> 10^{14.5}$ cm^{-2} (three bottom panels), there is a clear peak at small values of $D_{\rm tran}$. No similar peak is present in the distributions of absorbers with $N_{\rm HI}< 10^{14.5}$ cm^{-2} (top panels). Clearly absorbers with $N_{\rm HI}> 10^{14.5}$ cm^{-2} are more tightly correlated with the positions of galaxies than absorbers of lower column densities.

These distributions can also be used to determine the excess probability (over that of a random place in the IGM) of intersecting an absorber of a given column density, within a range of $D_{\rm tran}$ and within some $|\Delta v|$. The excess probability, $\rm P_E$, is defined through comparison to the random distribution:

$$\rm P_E = \frac{\eta_{abs} - \eta_{abs,random}}{\eta_{abs,random}} \tag{2.11}$$

Table 2.2. Absorber Incidence ($\eta_{\rm abs}$) and Excess Probability (P_E) for $|\Delta v|< 300$ km s^{-1}[a]

log($N_{\rm HI}$) range[b] [cm^{-2}]	$\eta_{\rm abs}$ $D_{\rm tran}< 0.1$ pMpc	P_E	$\eta_{\rm abs}$ $D_{\rm tran}< 0.3$ pMpc	P_E	$\eta_{\rm abs}$ $D_{\rm tran}< 2$ pMpc	P_E
13.0 —13.5	1.1± 0.3	-0.1± 0.3	1.3± 0.2	0.1± 0.1	1.33± 0.04	0.17± 0.04
13.5 —14.0	1.4± 0.4	1.0± 0.5	1.2± 0.2	0.8± 0.3	0.87± 0.03	0.44± 0.06
14.0 —14.5	0.6± 0.2	0.7± 0.7	0.7± 0.1	1.4± 0.4	0.53± 0.03	0.80± 0.09
14.5 —15.0	0.8± 0.3	4.4± 1.9	0.7± 0.1	3.4± 0.8	0.32± 0.02	1.03± 0.13
15.0 —16.0	0.9± 0.3	12.7± 4.6	0.4± 0.1	4.9± 1.4	0.22± 0.02	1.59± 0.20
16.0 —17.0	0.7± 0.3	38.2±14.8	0.2± 0.1	9.7± 3.4	0.06± 0.01	1.78± 0.40

[a]The values in this table refer to Figure 2.16

[b]log($N_{\rm HI}$) range: The range of $N_{\rm HI}$ considered; And the incidence, $\eta_{\rm abs}$, and excess probability, P_E considered over three ranges of $D_{\rm tran}$. The quoted uncertainties in $\eta_{\rm abs}$ and P_E are calculated assuming Poisson statistics.

Table 2.2 summarizes the measured values of $\eta_{\rm abs}$ and P_E for absorbers binned in $N_{\rm HI}$ for the velocity window $|\Delta v|< 300$ km s^{-1} and three distance cuts: $D_{\rm tran}< 100$ pkpc, $D_{\rm tran}< 300$ pkpc, and $D_{\rm tran}< 2$ pMpc.

2.5.1 Absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2}

We have shown above that absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2} appear to trace the positions of galaxies in our sample with high fidelity. The incidence of absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2} and $|\Delta v| < 300$ km s^{-1}, as shown in Figure 2.17, nicely encapsulates the "shape" of the CGM. Similar to Figure 2.9 for $D_{\rm tran} < 300$ pkpc, one sees rising values of $\eta_{\rm abs}$ as the galactocentric distance is reduced. From 300 pkpc $< D_{\rm tran} < 2$ pMpc, $\eta_{\rm abs}$ reaches a plateau value with $\eta_{\rm abs} \gtrsim 0.5$ [2 pMpc = 1.4 h^{-1} pMpc $\approx$ 4.6 h^{-1} cMpc (at $z = 2.3$)]. For $D_{\rm tran} > 2$ pMpc, $\eta_{\rm abs}$ drops to values consistent with the average IGM.

Recalling the quantity P(Δv,$D_{\rm tran}$), defined in §2.4.1 as the probability, per galaxy, of intersecting an absorber at a given Δv and within the specified range of $D_{\rm tran}$, here we consider the velocity distribution of absorbers with $N_{\rm HI}>10^{14.5}$ cm^{-2} in bins of $D_{\rm tran}$ as shown in Figure 2.18.

Fitting a Gaussian to the excess absorbers near galaxies compared to random gives a deviation ($\sigma_{\langle\Delta v\rangle}$) of 187 km s^{-1} again suggesting that the velocity scale of the excess is $\Delta v \approx \pm 300$ km s^{-1}.

Integrating the distribution in the first panel of Figure 2.18 corresponding to $D_{\rm tran} < 300$ kpc results in an incidence averaged over ± 350 km s^{-1} ($\sim 2\sigma$) and 300 pkpc, $\eta_{\rm abs} = 1.6$. Similarly, comparing the value of $\eta_{\rm abs}$ measured in the real and random distribution one infers that strong absorbers ($N_{\rm HI}> 10^{14.5}$) are > 4 times ($P_E = 4.1$) more likely to be found within 300 kpc and ± 350 km s^{-1} of a galaxy in our sample, than at a random place. The parameters of the Gaussian fits in Figure 2.18, as well as the inferred $\eta_{\rm abs}$ and P_E, are listed in Table 2.3.

In this section, we have shown that the covering fraction, f_c, of absorbers with log($N_{\rm HI}$) < 14.5 is roughly

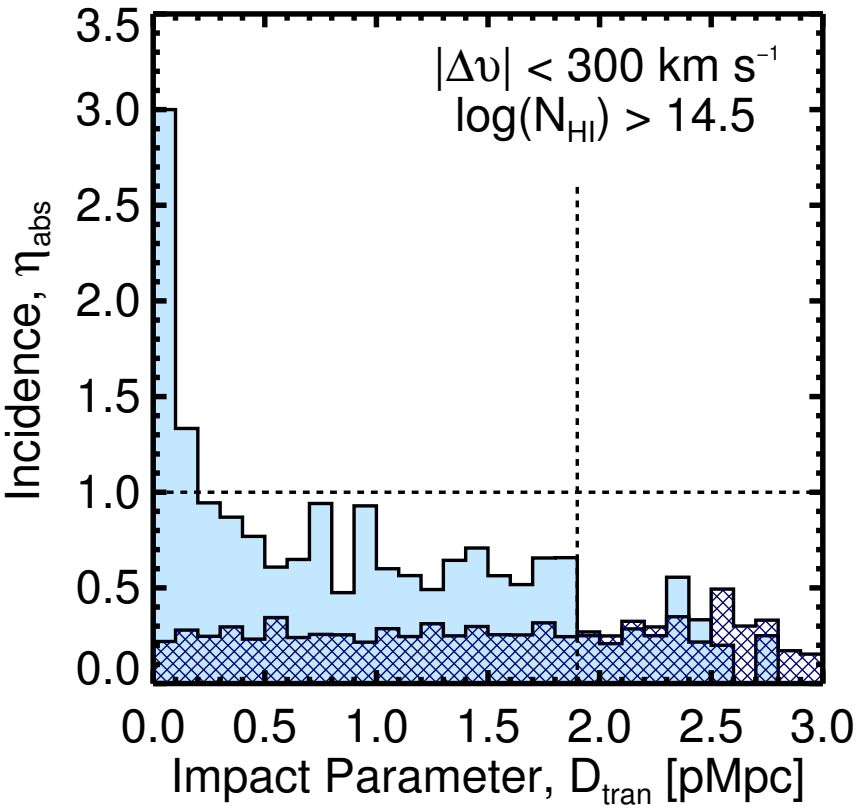


Figure 2.17 The incidence of absorbers, $\eta_{\rm abs}$, with $N_{\rm HI} > 10^{14.5}$ cm^{-2} as a function of impact parameter. The solid histogram represents the distribution of these high-$N_{\rm HI}$ absorbers for $|\Delta v| < 300$ km s^{-1}. The hatched histogram represents the average incidence of the same absorbers near randomly-chosen redshifts. The vertical dotted line indicates the distance at which the incidence of the real distribution becomes comparable with that of the random distribution, $D_{\rm tran} \approx 2$ pMpc. The dotted horizontal line marks $\eta_{\rm abs} = 1$.

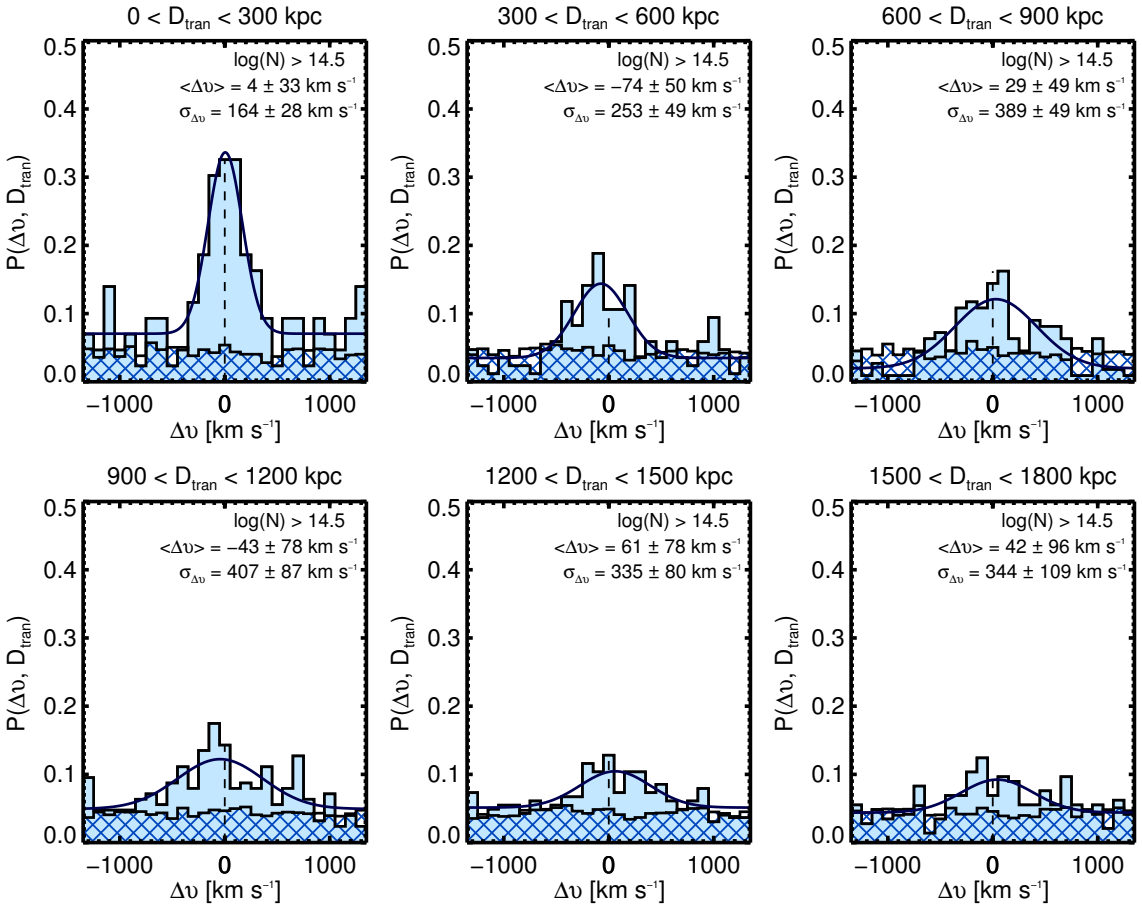


Figure 2.18 The velocity of absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2} in bins of $D_{\rm tran}$, as indicated. Note that $N_{\rm HI} > 10^{14.5}$ cm^{-2} absorbers are > 4 times more likely to be found within $\pm$300 km s^{-1} and 300 kpc of a galaxy than at a random place in the IGM. See Table 2.3 for $\eta_{\rm abs}$ and the excess probability for all of the panels.

Table 2.3. Incidence ($\eta_{\rm abs}$) and Excess Probability for $N_{\rm HI} > 10^{14.5}$ cm^{-2} [a]

$D_{\rm tran}$ [b] [pkpc]	$\langle\Delta v\rangle$ [km s^{-1}]	$\sigma_{\langle\Delta v\rangle}$ [km s^{-1}]	Δv window [km s^{-1}]	$\eta_{\rm abs}$	P_E
0–300	4±33	164± 28	-350 to 350	1.70±0.02	4.1
300–600	-74±50	253± 49	-550 to 450	1.07±0.01	1.19
600–900	29±49	389± 49	-750 to 850	1.31±0.01	0.80
900–1200	-43±78	407± 87	-850 to 750	1.60±0.01	1.14
1200–1500	61±78	335± 80	-650 to 750	1.20±0.01	0.76
1500–1800	42±96	344±109	-650 to 750	1.06±0.01	0.63

[a]The values in this table refer to Figure 2.18.

[b]$D_{\rm tran}$:The $D_{\rm tran}$ window considered; $\langle\Delta v\rangle$: the Gaussian velocity centroid and its associated error; $\sigma_{\langle\Delta v\rangle}$: the standard deviation of the Gaussian fit and its associated error; Δv window: the velocity window used to compute $\eta_{\rm abs}$ and P_E (comparable to $\pm 2\sigma$); $\eta_{\rm abs}$: the incidence, P_E: the excess probability. The quoted uncertainties in $\eta_{\rm abs}$ are calculated assuming Poisson statistics. The error in P_E is dominated by the variation in individual realizations of the random sample: typically the variation in P_E is ~ 0.05 except for the first bin ($0 < D_{\rm tran} < 300$ pkpc) where the variation is ~ 0.2.

uniform and > 0.5 out to $D_{\rm tran} \approx 2$ pMpc, but that those with higher $N_{\rm HI}$ only have $f_c > 0.5$ within $D_{\rm tran} < 200$ kpc (Figure 2.15). Consideration of the incidence, $\eta_{\rm abs}$, indicates that absorbers with $\log(N_{\rm HI}) > 14.5$ are more directly related to galaxies (Figure 2.16). Those absorbers nicely encapsulate the "shape" of the CGM (Figure 2.17) and are > 4 times more likely to be found within $|\Delta v| < 350$ km s^{-1} and $D_{\rm tran} < 300$ pkpc of a galaxy in our sample than at a random place in the IGM (Figure 2.18).

2.5.2 Covering Fractions: Comparison with Simulations

Motivated by the desire to predict the observational signatures of cold accretion streams, two recent theoretical papers have considered the covering fraction of absorbers of various $N_{\rm HI}$ surrounding galaxies. Faucher-Giguère & Kereš (2011) consider the covering fraction of Lyman Limit (LLS, $17.2 < \log(N_{\rm HI}) < 20.3$) and Damped Lyman Alpha (DLA, $\log(N_{\rm HI}) > 20.3$) absorbers originating within cold streams near two simulated star-forming galaxies at $z = 2$ using cosmological zoom-in simulations that do not include galactic winds. Since these authors explicitly *did not* consider galactic winds, their covering fractions are approximately[14] a lower limit on the expected values. They considered a galaxy slightly less massive than those in our sample ($M_{\rm DM} = 10^{11.5}$ $M_\odot$), as well as one comparable to those in this work ($M_{\rm DM} = 10^{12}$ $M_\odot$).

For comparison, we consider the fraction of sightlines at a given $D_{\rm tran}$ for which sum($N_{\rm HI}$,700 km s^{-1}) falls

[14]These values are only a lower-limit in the case that outflowing winds do not affect the gas distribution within the filamentary cold streams. van de Voort et al. (2011b) found that the wind prescription in simulations has little effect on the amount of gas that accretes onto halos but does affect the amount which is delivered into the ISM of the galaxy itself. Faucher-Giguère et al. (2011) argue that energetic winds can undergo hydrodynamical interactions with cold-streams removing some of the inflowing gas.

Table 2.4. F_c: Comparison with Faucher-Giguère & Kereš (2011)

Sample	$\log(N_{HI})$ [cm^{-2}]	$F_c(<1\ r_{vir})$	$F_c(<2\ r_{vir})$
$M_{DM} = 10^{12}\ M_\odot$	17.2−20.3	11%	···
	> 20.3	4%	···
$M_{DM} = 10^{11.5}\ M_\odot$	17.2−20.3	12%	4%
	> 20.3	3%	1%
This work[a]	17.2−20.3	30±14%	24±9%
	> 20.3	0^{+10}_{-0}%[b]	4±4%

[a]The F_c tabulated here from "this work" are the fraction of galaxies with a sum(N_{HI}) statistic in the column density range. This is subtly different from the f_c presented earlier. Note that while we use a velocity cut of $|\Delta v| < 700$ km s^{-1} here, the results would be the same if we used a ± 300 km s^{-1} window. Had we considered f_c as opposed to F_c, we would have calculated 20% for the LLS covering within both 1 and 2 r_{vir}. The uncertainties quoted for "this work" are 1-σ errors calculated assuming F_c follows a binomial distribution.

[b]We calculate the 1-σ upper limit on this non-detection as $1-(1-0.68)^{1/(1+n)}$ where $n = 10$ is the number of systems considered.

within the specified column density range. This is most akin to the results of Faucher-Giguère & Kereš (2011) as they measure the column density of absorbers after projecting their simulated cube onto a 2D plane. Note that here we use differential bins in $N_{HI}(17.2 < \log(N_{HI}) < 20.3)$, and since we are considering sum(N_{HI}) rather than max(N_{HI}), we use F_c to denote the associated covering fraction (to differentiate it from the f_c measured earlier in this section).

Our measurements are compared with the Faucher-Giguère & Kereš (2011) simulation results in Table 2.4. Particularly in the case of LLS gas, we find ~ 3 times higher F_c within r_{vir} and ~ 6 times higher within $2r_{vir}$ compared with the simulations. However, if the simulations results are treated as lower limits, then clearly they are consistent with the observations.

Fumagalli et al. (2011) also considered the covering fraction of H I surrounding 6 LBG-type galaxies at $z = 2-3$ using cosmological zoom-in simulations with galactic winds included (though they are relatively weak with their particular implementation). They consider absorbers within r_{vir} and $2r_{vir}$ for various thresholds in $\log(N_{HI}) : > 15.5, > 17.2, > 19.0$, and > 20.3 cm^{-2}. A comparison with the observations is given in Table 2.5. The models of Fumagalli et al. (2011) clearly under-predict the presence of $15.5 < \log(N_{HI}) < 17.2$ gas surrounding galaxies. Similar to Faucher-Giguère & Kereš (2011), they also seem to under predict the radial extent of gas with $N_{HI} \lesssim 10^{19}$ cm^{-2}.

Table 2.5. Covering Fraction: Comparison with Fumagalli et al. (2011)

Sample	$\log(N_{\rm HI})$ [cm^{-2}]	$F_{\rm c}$ ($<1\ r_{\rm vir}$)	$F_{\rm c}$ ($<2\ r_{\rm vir}$)
Fumagalli et al.	>15.5	38%	22%
	>17.2	16%	7%
	>19.0	6%	3%
	>20.3	3%	1%
This work[a]	>15.5	90±9%	68±9%
	>17.2	30±14%	28±9%
	>19.0	10±9%	8±5%
	>20.3	0^{+10}_{-0}%	4±4%

[a]Values in "this work" are computed as described in Table 2.4. Were we to consider $|\Delta v| < 300$ km s^{-1} rather than $|\Delta v| < 700$ km s^{-1}, only the $F_{\rm c}$ for $\log(N_{\rm HI}) > 15.5$ within 2 $r_{\rm vir}$ would change. The value for 300 km s^{-1} would be 60±10%.

2.5.3 Evolution of the CGM from $z \sim 2.3$ to $z \lesssim 1$

There have been numerous studies of the low-z galaxy-IGM connection. In the interest of brevity, below we compare our measured $f_{\rm c}$ with a small number of studies with comparable size and statistical power.

As discussed by Davé et al. (1999) and Schaye (2001), due to the expansion of the universe, the collapse of structure, and the evolution in the intensity of the metagalactic ionizing background, the $N_{\rm HI}$ associated with a fixed overdensity declines with redshift (e.g., an absorber at $z = 0$ would have $N_{\rm HI} \sim 20$ times lower than would be measured at the same overdensity at $z \approx 2.3$). As a consequence, absorbers with the same $N_{\rm HI}$ at different redshifts trace different structures, and have very different incidence rates. For example, there are no regions of the Lyα forest at $z \sim 2.3$ that have zero H I absorbers in a velocity window of $\pm$700 km s^{-1} – every galaxy in our sample can be associated with (generally) a large number of absorbers. This is not the case at low-z, where absorption-line spectra are sparsely populated and generally galaxies would be associated with very few (possibly no) H I absorbers. Similarly, the continued growth of structure and the rapid decline in the star-formation rate density toward $z \sim 0$ means that a typical low-z galaxy is expected to be in a very different place on its evolutionary sequence compared to a "typical" galaxy at $z > 2$. Large differences are also expected for gas fractions and baryonic accretion rates at low-z compared to high. Thus, one might reasonably expect circumgalactic gas to differ as well.

With these caveats in mind, we compare the covering fraction, $f_{\rm c}$, of the $\langle z \rangle = 2.3$ sample with results at low-z.

Table 2.6. H I Covering Fraction: Comparison with Low-z Studies[a]

Sample	$W_0(Ly\alpha)$ mÅ	$\log(N_{\rm HI})$ [cm^{-2}]	$D_{\rm tran}$ pkpc	$\lvert\Delta v\rvert$ km s^{-1}	$f_{\rm c}$
Chen et al. $0.1 < z < 0.9$	350	$\gtrsim 14$	<330	500	72%
This work $z \sim 2.3$	···	>14	<330	500	83±5%
Prochaska et al. $0.005 < z < 0.4$	50	$\gtrsim 13$	<300	400	96%
	300	$\gtrsim 14$	<300	400	70%
	50	$\gtrsim 13$	<1000	400	70%
	300	$\gtrsim 14$	<1000	400	38%
This work $z \sim 2.3$	···	>13	<300	400	100^{+0}_{-3}%[b]
	···	>14	<300	400	81±6%
	···	>13	<1000	400	95±1%
	···	>14	<1000	400	70±3%
Wakker & Savage $z < 0.017$	50	$\gtrsim 13$	<200	400	75%
	50	$\gtrsim 13$	< 400	400	81%
	50	$\gtrsim 13$	<1000	400	49%
	50	$\gtrsim 13$	<2000	400	48%
	50	$\gtrsim 13$	<3000	400	39%
	300	$\gtrsim 14$	<200	400	46%
	300	$\gtrsim 14$	<400	400	44%
	300	$\gtrsim 14$	<1000	400	24%
	300	$\gtrsim 14$	<2000	400	22%

Table 2.6—Continued

Sample	$W_0(Ly\alpha)$ mÅ	$\log(N_{\rm HI})$ [cm^{-2}]	$D_{\rm tran}$ pkpc	$\lvert\Delta v\rvert$ km s^{-1}	$f_{\rm c}$
	300	$\gtrsim 14$	<3000	400	20%
This work $z \sim 2.3$	···	> 13	< 200	400	100^{+0}_{-4}%[c]
	···	> 13	< 400	400	98±2%
	···	> 13	< 1000	400	95±1%
	···	> 13	< 2000	400	94±1%
	···	> 13	< 3000	400	94±1%
	···	> 14	< 200	400	92±5%
	···	> 14	< 400	400	82±5%
	···	> 14	< 1000	400	71±3%
	···	> 14	< 2000	400	64±2%
	···	> 14	< 3000	400	61±2%

[a]The distance binning is cumulative, so $f_{\rm c}$ is the covering fraction of gas with $D_{\rm tran} <$ the distance listed in the table and $|\Delta v| <$ the velocity listed in the table. The low-z samples generally do not include Voigt profile fits, but rather measure equivalent widths, W_0. The approximate value of $N_{\rm HI}$ associated with each W_0 limit is listed.The uncertainties quoted for "this work" are 1-σ errors calculated assuming $f_{\rm c}$ follows a binomial distribution.

[b]We calculate the 1-σ lower limit on this 100% detection as $(1-0.68)^{1/(1+n)}$ where $n = 43$ is the number of systems considered.

[c]We calculate the 1-σ lower limit on this 100% detection as $(1-0.68)^{1/(1+n)}$ where $n = 25$ is the number of systems considered.

2.5.3.1 The CGM at $0.1 < z < 0.9$

Chen et al. (2001c) considered 47 galaxies with $\langle z \rangle = 0.36$ and $D_{\rm tran} <$ 330 pkpc [15]. Their galaxy sample covered a wide range of properties, with 68% having $L_B > 0.25\ L_B^*$. The QSO spectra used were generally not of sufficient quality to allow for Voigt profile decompositions and thus the authors measured the equivalent width (W_0) of Lyα absorption. Chen et al. (2001c) considered absorbers to be associated with galaxies if $|\Delta v| \lesssim$ 500 km s^{-1} (Morris & Jannuzi 2006).

For $W_0 > 350$ mÅ ($N_{\rm HI} \gtrsim 10^{14}$ cm^{-2}) and $D_{\rm tran} <$ 330 pkpc, 34/47 galaxies in their sample had detectable Lyα absorption, corresponding to $f_{\rm c}$= 0.72. In the $\langle z \rangle = 2.3$ sample there are 48 galaxies with $D_{\rm tran} <$ 330 pkpc. Using the same velocity window $|\Delta v| <$ 500 km s^{-1}, we find 40/48 (83%) have Lyα absorption with $N_{\rm HI} \geq 10^{14}$ cm^{-2}, consistent with the measurements from Chen et al. (2001c). For clarity, the results from the

[15]The $D_{\rm tran}$ ranges quoted here have been adjusted from the cosmology considered in Chen et al. (2001c) to the cosmology used in this paper at the mean redshift of the sample, $\langle z \rangle = 0.36$.

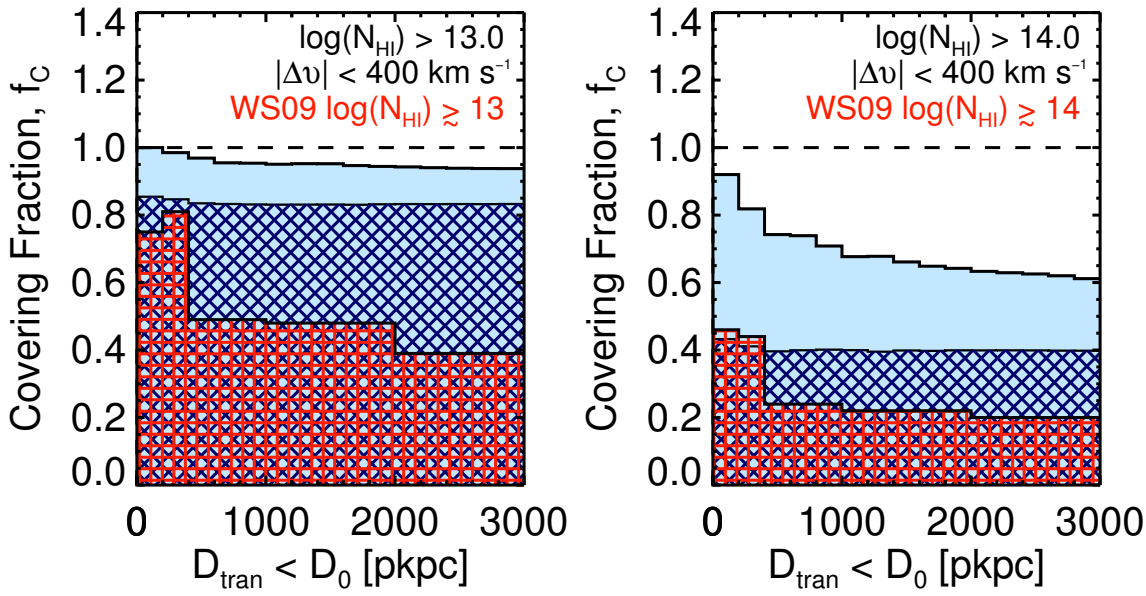


Figure 2.19 The covering fraction of high- and low-redshift absorbers in cumulative bins of $D_{\rm tran}$ (i.e., at each value of $D_{\rm tran}$, D_0, we consider all absorbers with smaller $D_{\rm tran}$, as opposed to those absorbers with $D_0 < D_{\rm tran} < D_1$.) In the (light-blue) solid histogram we plot our high-z data for f_c close to galaxies. The (dark-blue) cross-hatched histogram shows the distribution at random places in the high-z IGM. The (red) vertically-hatched histogram shows the distribution of f_c of low-z absorbers taken from Wakker & Savage (2009). Here we consider the same cut in $N_{\rm HI}$ at both epochs (W_0 = 50mÅ is equivalent to $N_{\rm HI} = 10^{13}$ cm^{-2} and similarly W_0 = 300mÅ is equivalent to $N_{\rm HI} = 10^{14}$ cm^{-2}).

low-z studies as well as our measurements are reproduced in Table 2.6.

2.5.3.2 The CGM at $0.005 < z < 0.4$

Prochaska et al. (2011) studied the association of $z < 0.2$ galaxies in 14 fields surrounding background QSOs. Similar to the Chen et al. (2001c) study, they generally relied on conversions of published W_0 measurements into $N_{\rm HI}$ using an assumed $b_{\rm d}$. In total, their galaxy survey included 37 galaxies with $D_{\rm tran}<$ 300 pkpc and 1200 galaxies in all with $\langle z \rangle = 0.18$. They use a velocity window of $|\Delta v|<$ 400 km s^{-1}. For the 26 galaxies with L > 0.1 L* within 300 pkpc, 25 had accompanying Lyα absorption with $N_{\rm HI}> 10^{13}$ cm^{-2}, or f_c=96%. In our sample, there are 43 galaxies with $D_{\rm tran}<$ 300 pkpc, all of which have a $N_{\rm HI}\geq 10^{13}$ cm^{-2} absorber within $|\Delta v|<$ 400 km s^{-1} (f_c=100%).

These values appear consistent; however, the covering fraction of $N_{\rm HI}\geq 10^{13}$ cm^{-2} gas is a very poor measure of the extent of the CGM at high redshift. For example, if we consider the same velocity window ($|\Delta v|<$ 400 km s^{-1}) and the same $N_{\rm HI}$ threshold ($N_{\rm HI}> 10^{13}$ cm^{-2}), we find f_c= 95% within 1 pMpc and f_c= 94% within 2 pMpc; i.e., there is no appreciable change in the f_c of gas of this $N_{\rm HI}$ with increasing distance. Prochaska et al. (2011) find f_c=70% for $D_{\rm tran} < 1$ pMpc in their sample.

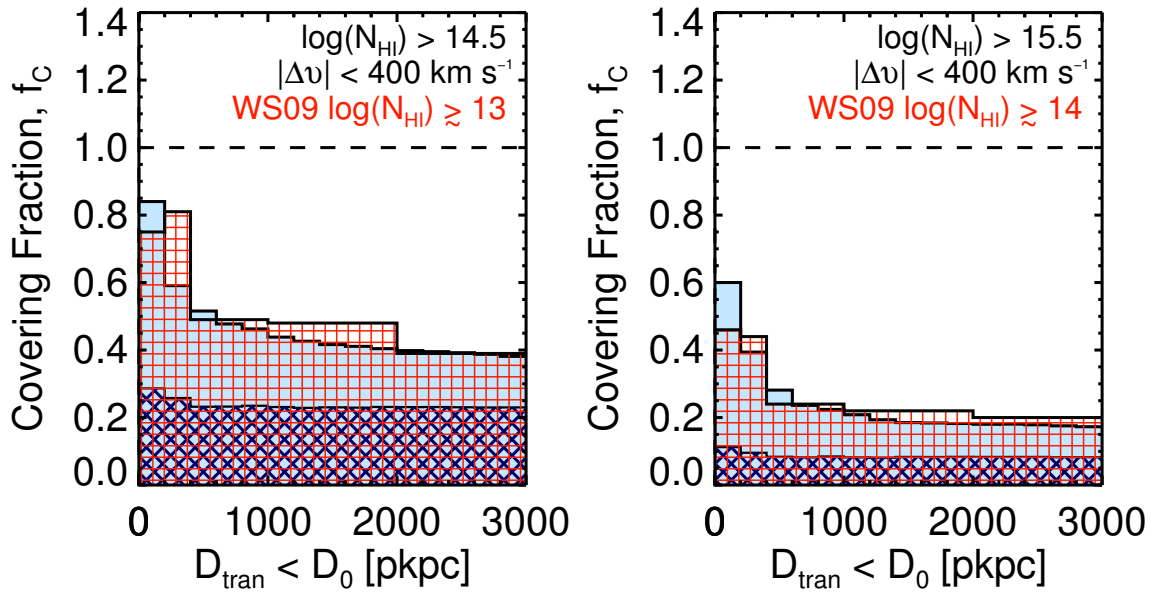


Figure 2.20 Same as Figure 2.19, but here we consider different $N_{\rm HI}$ cuts for the high-z sample. We have found the $N_{\rm HI}$ lower limit for the high-z sample which best reproduces the trend of the low-z $f_{\rm c}$ as a function of $D_{\rm tran}$. We can see that the trends of high-z absorbers of 1.5 dex higher $N_{\rm HI}$ are well matched to the distributions of low-z absorbers.

Measurements with a higher $N_{\rm HI}$ threshold have larger dynamic range; for $N_{\rm HI}> 10^{14}$ cm^{-2}, Prochaska et al. (2011) measure $f_{\rm c}$=70% for $D_{\rm tran}<$ 300 pkpc, and $f_{\rm c}$=38% for 1 pMpc. For our $\langle z\rangle \sim 2.3$ sample, we measure $f_{\rm c}$= 81% for $D_{\rm tran}<$300 pkpc and $f_{\rm c}$=70% for $D_{\rm tran}<$1 pMpc.

2.5.3.3 The CGM in the Local Universe: $cz < 6000$ km s^{-1}

Wakker & Savage (2009) (hereafter, WS09) studied the gaseous distribution around local galaxies ($cz < 6000$ km s^{-1}) along 76 lines of sight. The galaxy catalog considered included $\sim 20,000$ local galaxies. The absorber catalog included 115 intergalactic Lyα absorbers (i.e., not from the Galaxy) over the same redshift range. They calculate $f_{\rm c}$ in cumulative distance bins (e.g. $D_{\rm tran}<$ 200 pkpc).

For comparison to the WS09 sample, we use a velocity window $|\Delta v|<$ 400 km s^{-1} with respect to galaxy redshifts in our high redshift sample. From WS09 (Table 10) we use the values of $f_{\rm c}$ measured for the galaxy sample with L $>$ 0.25 L*.[16] They consider two classes of absorbers, those with $W_0 >$ 50mÅ ($N_{\rm HI}> 10^{13}$ cm^{-2}) and $W_0 >$ 300 mÅ ($N_{\rm HI} \gtrsim 10^{14}$ cm^{-2}). We make a direct comparison to our sample *with the same* $N_{\rm HI}$ *threshold* in Figure 2.19 and Table 2.6. Our measurements are shown in the solid histogram, the random high-z sample is plotted in the (dark-blue) cross-hatched histogram, and the WS09 sample is represented by the (red) vertically-hatched histogram. One can clearly see that at all $D_{\rm tran}$ and for both thresholds in $N_{\rm HI}$, the

[16]Our spectroscopic sample includes galaxies to $\mathscr{R}$ = 25.5. This is equivalent to L $>$ 0.25 L* at $z \approx 2.3$.

$z \sim 0$ sample shows significantly less covering than the high-z sample.

As mentioned above, absorbers of fixed $N_{\rm HI}$ are theoretically expected to trace higher-overdensity gas at low redshift than at high redshift. By considering various $N_{\rm HI}$ thresholds in our own data, we found that $f_{\rm c}$ of the low-z $W_0 > 50$mÅ($N_{\rm HI} > 10^{13}$ cm^{-2}) sample was most comparable to high-z absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2}; similarly, the WS09 $W_0 > 300$mÅ($N_{\rm HI} \gtrsim 10^{14}$ cm^{-2}) can be matched to our high-z absorbers with $N_{\rm HI} > 10^{15.5}$ cm^{-2}. Note that we chose the $N_{\rm HI}$ threshold which best reproduced $f_{\rm c}$ measured on $\sim$Mpc scales.

These two comparisons, shown in Figure 2.20, exhibit remarkably similar patterns in the covering fraction of $N_{\rm HI}$ as a function of $D_{\rm tran}$. It therefore appears that low-z absorbers with $\log(N_{\rm HI}) = {\rm N}_0$ *on average* trace the same physical regions around galaxies as gas with $\log(N_{\rm HI}) = {\rm N}_0 + 1.5$ at $\langle z \rangle = 2.3$. While we do not directly measure universal overdensity, these results lend observational support to the predictions by Davé et al. (1999) and Schaye (2001) that gas at the same universal overdensity at low- and high-z would have lower $N_{\rm HI}$ at low-z. These measurements also suggest that the physical size of the CGM around a typical galaxy is 300 pkpc at both $z \sim 0$ and $\langle z \rangle = 2.3$.

2.6 Mapping the Circumgalactic Medium

Another way to visualize the distribution of neutral hydrogen surrounding star-forming galaxies is to "map" the distribution along the transverse direction and the line of sight at the same time. As in §2.4.3, we use $|\Delta v|$ to compute a line-of-sight distance assuming it is entirely due to the Hubble flow. In Figures 2.21 - 2.24, the vertical axes represent the $D_{\rm Hubble}$ velocity scale line-of-sight distribution of gas. The horizontal axes correspond to the physical impact parameter between the galaxy and the QSO line of sight, $D_{\rm tran}$. Colors encode the incidence, $\eta_{\rm abs}$, of the gas in a given pixel in the map and the black and white shading represents the signal-to-noise ratio (S/N) with which $\eta_{\rm abs}$ is detected.

The S/N is determined using Poisson statistics. For a given bin in $D_{\rm tran}$ and $D_{\rm Hubble}$, we count the total number of absorption lines found in the random distribution, $n_{\rm abs,ran}$, and the number of random locations considered at that impact parameter, $n_{\rm gal,ran}$. We also count the number of real galaxies at that impact parameter, $n_{\rm gal}$. The "noise" level of the map is then taken to be:

$$\sigma_{\rm sig} = \frac{\sqrt{n_{\rm gal}\left(n_{\rm abs,ran}/n_{\rm gal,ran}\right)}}{n_{\rm gal}}, \tag{2.12}$$

where the quantity inside the square root is the number of absorption systems expected (based on the incidence for the random sample) given the number of galaxies in the real sample at a given $D_{\rm tran}$. The square root of this quantity represents the Poisson uncertainty in the number of absorbers that would be detected; the division by $n_{\rm gal}$ results in an expression for the error in $\eta_{\rm abs}$ *per galaxy.* This value is akin to the shaded error bars surrounding the median value of the random sample in Figure 2.9. The signal-to-noise ratio is then

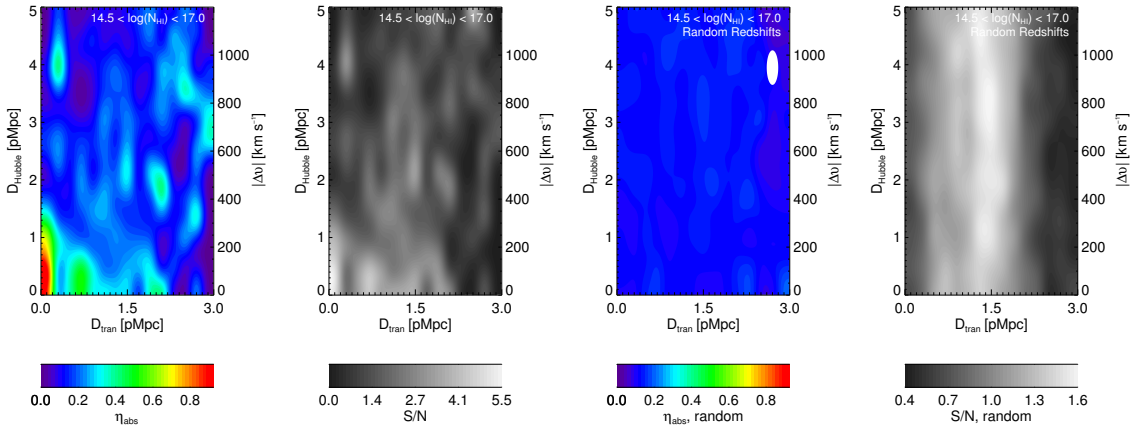


Figure 2.21 The incidence of $14.5 < \log(N_{\rm HI}) < 17$ absorbers. *Far Left*: Map of the $\eta_{\rm abs}$ with respect to the positions of galaxies (at the origin). *Second from Right*: The distribution at random locations in the IGM. The greyscale panels show the S/N ratios of the maps calculated as discussed in §2.6. The "bin" size for this map is 200 x 200 pkpc and the smoothing scale is 200 kpc in $D_{\rm tran}$ and 140 km s^{-1} (600 pkpc) along the line of sight. The FWHM of the gaussian smoothing beam is represented by the ellipse in the top right-hand corner of the 3rd panel from the left. Note that there is a strong peak in the incidence of high column density absorbers at close galacto-centric distances. The color bars in the $\eta_{\rm abs}$ maps have been re-normalized by χ_v = 6.3 in order to match the values of $\eta_{\rm abs}$ summed over $|\Delta v| <$ 300 km s^{-1}. The vertical structure in the S/N maps is caused by the variation in the number of galaxies in each bin of $D_{\rm tran}$.

taken to be:

$$\mathrm{S/N} = \frac{\eta_{\rm abs}}{\sigma_{\rm sig}}. \tag{2.13}$$

The maps are generated using three different bin sizes. We consider the intermediate bin size (200 x 200 pkpc) for two cuts in $N_{\rm HI}$ (Figures 2.21 and 2.23) which cover impact parameters out to 3 pMpc. For high-$N_{\rm HI}$ absorbers which exhibit a peak at small scales, we also consider maps with 100 x 100 pkpc bins (Figure 2.22). For low-$N_{\rm HI}$ absorbers, we provide a map with 400 x 400 pkpc bins to emphasize the large scale distribution (Figure 2.24). After the number of absorbers per bin is measured, the values are placed onto a 10 times finer grid and smoothed by a gaussian kernel with a full width half maximum equal to the bin size in impact parameter (x-axis) and a velocity scale of 140 km s^{-1} or 600 pkpc in Hubble distance (y-axis). Binning and smoothing is required to make maps of this type as the values of impact parameter and velocity are discrete. The scale of the smoothing, however, is motivated by the data. The smoothing in $D_{\rm tran}$ is selected to obtain a large enough sampling of the inner and outer bins such that the noise in the random distribution is reduced. The velocity scale smoothing corresponds roughly to the amplitude of the redshift uncertainties (see §2.2.2).

In Figures 2.21 – 2.24, the normalization of the color bar is re-scaled from the raw $\eta_{\rm abs}$ per 100 x 100, 200 x 200, or 400 x 400 pkpc bin by the multiplicative factor χ_v, where:

$$\chi_v = \frac{D_{\rm Hubble}\left(300 \text{ km s}^{-1}\right)}{D_{\rm Hubble}\,(\text{bin})} \approx \frac{1.26 \text{ pMpc}}{D_{\rm Hubble}\,(\text{bin})} \tag{2.14}$$

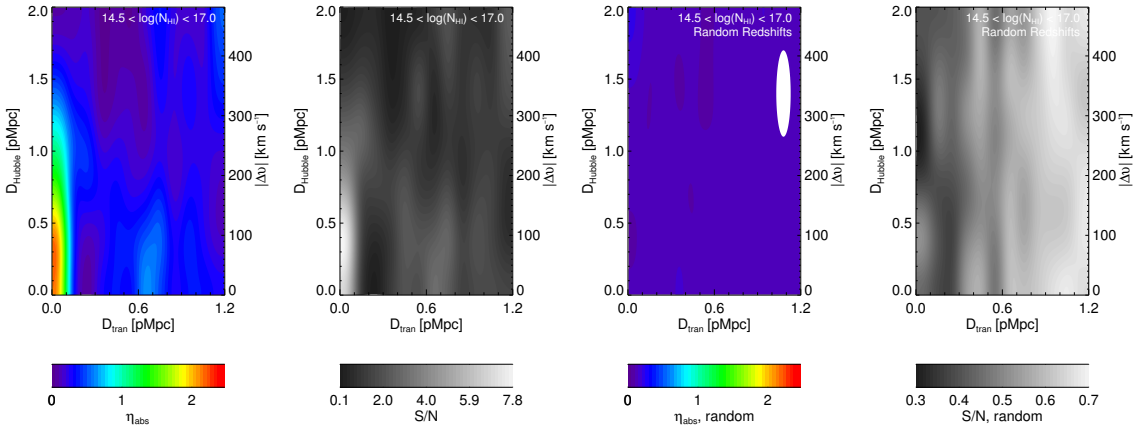


Figure 2.22 Same as Figure 2.21 but zoomed in to show the small scale distribution. The "bin" size for this map is 100 x 100 pkpc. The smoothing scale is 100 kpc in impact parameter and 140 km s^{-1} (600 pkpc) along the line of sight. The color bars for the incidence maps have been re-normalized χ_v = 12.6 in order to match the $\eta_{\rm abs}$ summed over $|\Delta v|<$ 300 km s^{-1}.

where $D_{\rm Hubble}$(bin) is either 100, 200, or 400 pkpc. For the 100 pkpc bins, χ_v = 12.6. This renormalization brings the scale of $\eta_{\rm abs}$ into agreement with that shown in Figures 2.16 and 2.17.

Recently, Rakic et al. (2011b) constructed similar maps to those presented here using pixel optical depth analysis (instead of Voigt profile fits) for a subset of the data presented here. Many of the conclusions discussed in the following section are corroborated by the independent analysis of Rakic et al. (2011b, §4).

2.6.1 Absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2}

Shown in Figure 2.21 is a map of $\eta_{\rm abs}$ for $10^{14.5} < N_{\rm HI} < 10^{17}$ cm^{-2}. Here we compare a map of the absorber distribution around galaxies (left-most panel) to a map which represents the general IGM as measured using the random distribution (third panel from left). Clearly, the incidence of $N_{\rm HI} > 10^{14.5}$ cm^{-2} absorbers is higher near galaxies than in the general IGM. $\eta_{\rm abs}$ remains significantly and consistently higher than that seen in the random map within 300 km s^{-1} along the line of sight and out to 2 pMpc in impact parameter.

The maps presented in Figure 2.22 which show the inner distance and velocity ranges at higher resolution, allow one to compare the scales of the absorption excess in Hubble distance and in $D_{\rm tran}$ in order to detect redshift anisotropies indicative of particular kinematic patterns. Galaxy redshift errors and peculiar velocities of gas with respect to galaxies will generally expand the line-of-sight velocity distribution.[17] Therefore, if the signal along the line of sight is elongated more than expected given the redshift errors and the scale of the transverse distribution, the remaining velocity structure may be attributed to peculiar velocities of the gas with respect to galaxies.

Based on Figure 2.22, adopting the green contours as representative of the most extreme portion of the

[17] One exception to this "rule" is the case of inflowing gas on large scales where the inflow velocities counter the relative Hubble flow causing a reduction in the line-of-sight velocity distribution.

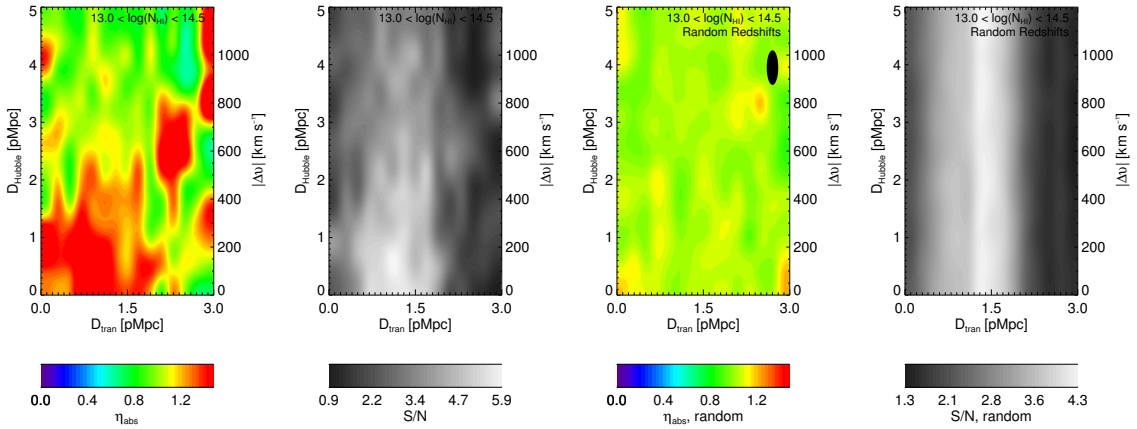


Figure 2.23 Same as Figure 2.21 but for low column density absorbers with $13 < \log(N_{\rm HI}) < 14.5$. Note, that the incidence of these absorption systems is higher in the large-scale regions surrounding galaxies, similar to that of high column density systems. However, unlike the high column density absorbers, there is no strong peak in the incidence rate of low-$N_{\rm HI}$ systems at small galacto-centric distances. The color bars for the incidence maps have been re-normalized $\chi_\nu = 6.3$ in order to match the $\eta_{\rm abs}$ summed over $|\Delta v| < 300$ km s^{-1}.

overdensity of gas (conservatively) corresponds to a velocity scale of $\sim$300 km s^{-1} and a physical impact parameter $\sim$125 pkpc[18]. Subtracting in quadrature the 140 km s^{-1} smoothing (which corresponds roughly to our redshift errors) leaves a residual velocity scale of $|\Delta v| \simeq 265$ km s^{-1}. This velocity, were it due to pure Hubble flow, would indicate a distance of ~ 1.1 pMpc. In order to consider the portion of the line-of-sight elongation likely due to peculiar velocities, we can subtract in quadrature the extent of the transverse distribution – in effect subtracting a possible Hubble flow broadening. Subtracting the $D_{\rm Hubble} = 125$ pkpc or ~ 35 km s^{-1} leaves $\simeq 260$ km s^{-1}; in other words, peculiar velocities of ± 260 km s^{-1} are indicated by the data.[19]

2.6.2 Absorbers with $N_{\rm HI} < 10^{14.5}$ cm^{-2}

Figure 2.23 displays a map of the incidence of lower-column density absorbers ($10^{13.0} < N_{\rm HI} < 10^{14.5}$ cm^{-2}). Again comparing the left-most panel (galaxy positions) to the third panel from the left (random IGM), one sees that the incidence of these absorbers is higher in the regions near galaxies than in the general IGM; however, the low-$N_{\rm HI}$ map exhibits no strong peak at small galactocentric distances.

Since there is little structure on small scales in Figure 2.23, the map has been more heavily smoothed to emphasize larger scales. Drawing attention to the edge of the red contours in Figure 2.24, it appears that the scale of the excess absorption is marginally *smaller* along the $D_{\rm Hubble}$ (line-of-sight) axis (1.5 pMpc) than along the transverse distance (2 pMpc). This suggests there may be coherent infall motion on $\gtrsim$ Mpc scales

[18] Maps produced using a symmetric smoothing kernel of 200 pkpc in both the line-of-sight and transverse distance (not shown) also yield excess absorption to 300 km s^{-1} along the line of sight. The extent of the line-of-sight distribution is robust to reductions in the line-of-sight smoothing kernel; however, the S/N associated with the excess is improved by the adopted smoothing kernel.

[19] Or, we have seriously underestimated the magnitude of our galaxy redshift errors.

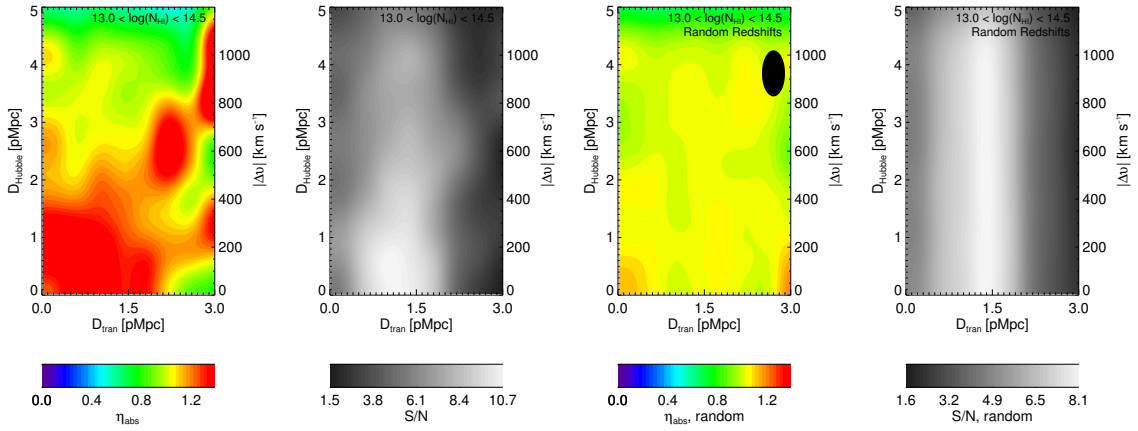


Figure 2.24 Same as Figure 2.23 but with 400 x 400 pkpc binning. Note that the line-of-sight distribution is compressed compared to the transverse distribution. The color bars for the incidence maps have been re-normalized χ_v = 3.15 in order to match the $\eta_{\rm abs}$ summed over $|\Delta v|< 300$ km s^{-1}.

compressing the distribution along the line of sight, the Kaiser (1987) effect. Rakic et al. (2011b) analyze the pixel statistics in the KBSS survey and discuss the kinematics of the circumgalactic gas in detail. They find a significant detection of compression of the signal along the line of sight which they interpret as large-scale infall.

2.6.3 Median Column Density Maps

An alternative visualization of the distribution of H I surrounding galaxies, avoiding the division of absorbers into low and high column density, is a map of the median $N_{\rm HI}$ in a given bin. Figure 2.25 shows maps where the color bar represents the median $N_{\rm HI}$ in a given bin of $D_{\rm tran}$ and $D_{\rm Hubble}$. On small scales, the behavior is similar to the high-$N_{\rm HI}$ absorbers, with significant elongation in the line of sight direction, due to a combination of peculiar motions of the gas and error in $z_{\rm gal}$.

In this section we have shown that the kinematics of absorbers with log($N_{\rm HI}$)>14.5 are consistent with a peculiar velocity component of $\sim \pm$260 km s^{-1} within $\sim$ 100 pkpc of galaxies, and separately that absorbers with log($N_{\rm HI}$)<14.5 may exhibit the kinematic signature of retarded Hubble flow on scales $\gtrsim$ 1 pMpc.

2.6.4 Explaining the Gas-Phase Kinematics

The kinematics of the gas surrounding galaxies may provide important clues about the physical processes occurring near galaxies. The two most plausible scenarios for the origin of the observed gas is that it traces the large scale structure (in which case it is likely falling onto the halo) or that it is the result of galactic winds, known to operate in these galaxies. Most likely, it is some combination of the two phenomena. Unfortunately, the sign of the velocity offset (red-shifted or blue-shifted) cannot tell us what the relative motion of the gas is with respect to the galaxy since its position along the line of sight is not known. For example, gas which

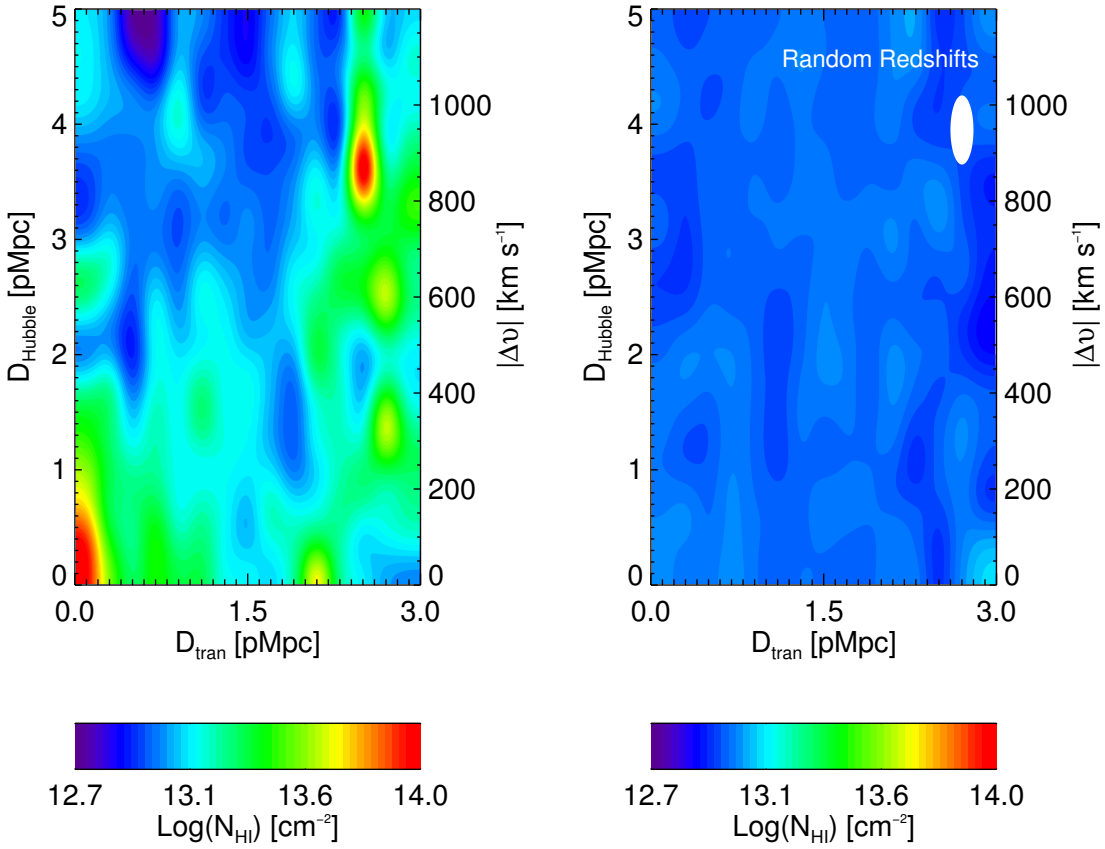


Figure 2.25 Maps of the median $N_{\rm HI}$ with respect to the positions of galaxies. The "bin" size for this map is 200 x 200 kpc. The smoothing scale is 200 kpc in impact parameter and 140 km s^{-1} (600 pkpc) along the line of sight. *Left:* The map of the true distribution of column densities. *Right:* The map of the column density distribution in the random sample. The significance with which we detect structures appearing in this map is quantified with the S/N maps presented in the preceding figures.

is redshifted with respect to the galaxy might be behind the galaxy, and thus be either outflowing or moving with the Hubble flow at larger distance. Conversely, redshifted absorbers could be foreground gas that is redshifted because it is falling onto the galaxy. There is no way to know *a priori* which scenario holds in which cases.

We recall that in §2.4.1 it was shown that the full extent in Δv spanned by absorbers comprising the net excess over random is $|\Delta v| \lesssim 700$ km s^{-1}, with the most significant portion of the excess within $|\Delta v| \lesssim 300$ km s^{-1}. Our map of high-$N_{\rm HI}$ systems confirms the 300 km s^{-1} excess and allows us to measure the extent of the peculiar velocities. After accounting for our estimated galaxy redshift errors (§ 2.6.1), the range of significant excess becomes $|\Delta v| \lesssim 260$ km s^{-1}.

In-falling cool gas is likely to be moving at relative velocities smaller than the galaxy circular velocity (Faucher-Giguère & Kereš 2011):

$$v_{circ} = \sqrt{\frac{GM_{\rm DM}}{r_{\rm vir}}} = 220 \text{ km s}^{-1}, \tag{2.15}$$

where we assume $M_{\rm DM} = 10^{12}$ M$_\odot$ and $r_{\rm vir} = 91$ pkpc (see §2.2.1 and §2.8.3). Since we measure only the line-of-sight component of this velocity, $|v_{\rm los}| \sim 200$km s^{-1} might serve as an approximate upper limit on the line-of-sight velocity component due to graviationally induced peculiar velocities. Unless we have significantly underestimated the measurement uncertainties in $z_{\rm gal}$, accreting gas would be expected to have a "quieter" velocity field than observed.

On the other hand, outflows observed in the galaxy spectra regularly show blue-shifted velocities as high as 800 km s^{-1}, and large velocity widths appear to be required to explain the strength of strongly saturated absorption at modest impact parameters ($D_{\rm tran} \lesssim 100$ pkpc), based on the galaxy-galaxy pairs analysis of Steidel et al. (2010). In the context of their simple model, the envelope of $|\Delta v|$ is simply the component of the outflow speed $v_{\rm out}$ projected along an observers line of sight; this maximum velocity also dictates how strong the resulting absorption features will be, with W_0 increasing proportionally to $v_{\rm out}$. To model the behavior of W_0 with impact parameter, $v_{\rm out}$ was the principle normalization factor and the inferred values were $650 \lesssim v_{\rm out} \lesssim 820$ km s^{-1} depending on the ion. In the present work, the first bin in $D_{\rm tran}$ extends from 50-100 pkpc, with a median $D_{\rm tran} = 80$ pkpc, with median $\log(N_{\rm HI}) \simeq 16.5$. Assuming the Steidel et al. (2010) geometric/kinematic model (with $R_{\rm eff} = 90$ pkpc) would predict that a line of sight with $D_{\rm tran}$= 80 kpc would have $|\Delta v| \lesssim 0.46 v_{\rm out}$ (equivalent to an observable velocity range between $\pm$325 and $\pm$400 km s^{-1}) and a threshold $N_{\rm HI} \sim 10^{17}$ cm^{-2} [analogous to our max($N_{\rm HI}$) statistic], both compatible with the observations from the QSO sightlines discussed above.

2.7 The Doppler Width

In addition to $N_{\rm HI}$ and $z_{\rm abs}$, the third component of a Voigt profile fit is the Doppler width, $b_{\rm d}$. In the case of an unsaturated line, the Voigt profile is well approximated by a Gaussian in optical depth, with a Doppler

parameter:

$$b_{\rm d} = \sqrt{2}\sigma = \frac{\rm FWHM}{2\sqrt{\ln 2}}. \tag{2.16}$$

In the case of purely thermal broadening, the gas temperature T can be inferred directly from the Doppler parameter:

$$b_{\rm d}(T) = \sqrt{\frac{2kT}{m}} \tag{2.17}$$

where m is the mass of the ion and k is the Boltzmann constant. For H I,

$$T = \frac{m_{\rm H}}{2k} {b_{\rm d}}^2 = 4 \times 10^4 \ {\rm K} \left(\frac{b_{\rm d}}{26 \ {\rm km \ s^{-1}}} \right)^2 ; \tag{2.18}$$

noting that 26 km s^{-1} is the median value of $b_{\rm d}$ in the full absorber catalog.

The physical state of the IGM is predicted to be governed by the balance of two principle processes: the adiabatic cooling caused by the expansion of the Universe and photoionization heating, generally from the UV background. In the regime where this holds and in the case that the baryonic overdensity roughly traces the dark matter overdensity, a natural consequence is a relationship between the temperature of the gas, T, and its density, ρ. Higher-density regions, having more gravitational resistance to the Hubble flow, cool less and thus have higher temperatures on average (see e.g., Hui & Gnedin 1997; Schaye et al. 1999). And indeed, observed distributions of thermally broadened absorption lines exhibit higher $b_{\rm d}$ for absorbers with higher $N_{\rm HI}$ (Pettini et al. 1990; Schaye et al. 2000; Bryan & Machacek 2000; Ricotti et al. 2000; McDonald et al. 2001).

Other processes can also broaden individual absorption components. For the most diffuse and physically extended absorbers, generally with lower $N_{\rm HI}$, the Hubble flow itself can contribute significantly to the line width. Turbulence (here meaning bulk motions of the gas) can also broaden absorption features such that:

$${b_{\rm d}}^2 = b_{\rm turb}^2 + \frac{2kT}{m} \tag{2.19}$$

where $b_{\rm turb}$ is the turbulent component of the line width.

In this section we discuss the distribution of Doppler widths ($b_{\rm d}$) observed in absorption systems as a function of their galactocentric distance. The observed trends are much more subtle than those in $N_{\rm HI}$, but their utility in discerning the physical state of the gas motivates a careful analysis.

2.7.1 The Dependence of $b_{\rm d}$ on Proximity to Galaxies, $D_{\rm tran}$

Figure 2.26 shows the Doppler width ($b_{\rm d}$) of absorbers with $N_{\rm HI} > 10^{13}$ cm^{-2} versus $D_{\rm tran}$. While considerably more noisy than the column density trends in the previous section, it appears that within the second bin (corresponding to $100 < D_{\rm tran} < 200$ pkpc) the median value of $b_{\rm d}$ is considerably higher than that of the random sample.

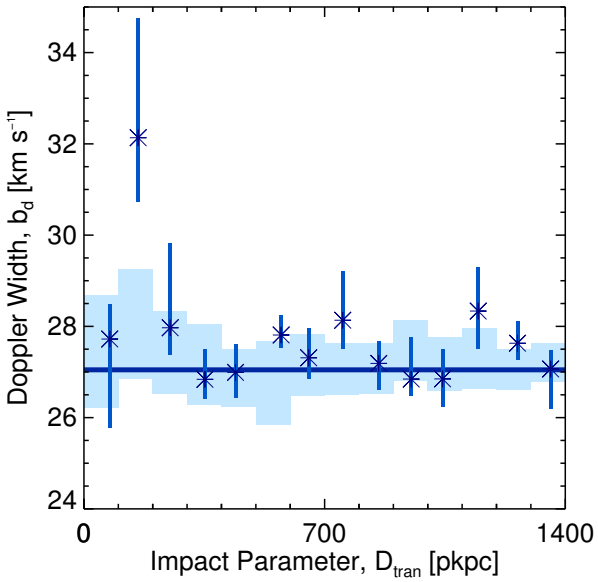


Figure 2.26 The Doppler width, $b_{\rm d}$, of absorbers with $\log(N_{\rm HI}) > 13$ and $|\Delta v| < 700$ km s^{-1} as a function of transverse distance from a galaxy. The symbols have the same meaning as those in Figure 2.9 except they refer to $b_{\rm d}$ instead of $N_{\rm HI}$. Note the peak in the second bin corresponding to absorbers with $100 < D_{\rm tran} < 200$ pkpc.

We consider the trend in $b_{\rm d}$ versus $N_{\rm HI}$ for both the sample of absorbers close to galaxies and that of the full absorber catalog. In this way, we can examine the $b_{\rm d}$ of absorbers at *fixed* $N_{\rm HI}$, as shown in Figure 2.27, in order to evaluate the effect of the $T-\rho$ relationship on this result. Here, each of the four closest bins in $D_{\rm tran}$ (Figure 2.26) is broken into individual panels. The asterisks represent the absorbers close to galaxies. The black horizontal lines and shading refer to the *full absorber catalog*.[20] Considering first the top right-hand panel with $100 < D_{\rm tran} < 200$ pkpc (the bin with the majority of the signal in Figure 2.26) for $N_{\rm HI} > 10^{14}$ cm^{-2}, the median $b_{\rm d}$ of absorbers close to galaxies is larger than the median of the full absorber sample at fixed $N_{\rm HI}$. This suggests that the larger $b_{\rm d}$ cannot be attributed solely to a dependence on $N_{\rm HI}$ as in the $T-\rho$ relation.

Figure 2.27 shows that the median $b_{\rm d}$ for the sample of absorbers within $D_{\rm tran} \sim 100-300$ pkpc of galaxies is systematically larger than that of absorbers from the full absorber catalogue across most bins in $N_{\rm HI}$. However, the absorbers closest to galaxies ($D_{\rm tran} < 100$ pkpc; top left panel of Figure 2.27) do not appear to exhibit the same systematic "excess" in median $b_{\rm d}$ (cf. Figure 2.26). There are too few galaxies in this inner bin to determine whether the result is statistically significant.

In order to quantify the statistical significance of the apparent excess $b_{\rm d}$ in the range $D_{\rm tran} < 300$ pkpc, we performed a Monte Carlo simulation. This was done by drawing random samples of absorbers from the full absorber catalogue with the same distribution in $N_{\rm HI}$ and of the same size as the absorber sample close to galaxies. The figure of merit was taken to be the average difference computed as follows: in each bin of $N_{\rm HI}$

[20] Note that within the full catalog for absorbers whose growth is expected to remain approximately linear ($N_{\rm HI} < 10^{15}$ cm^{-2}, $\rho/\overline{\rho} \lesssim 10$), the median values (black horizontal lines) exhibit increasing $b_{\rm d}$ with increasing $N_{\rm HI}$, as one would expect for the $T-\rho$ relation.

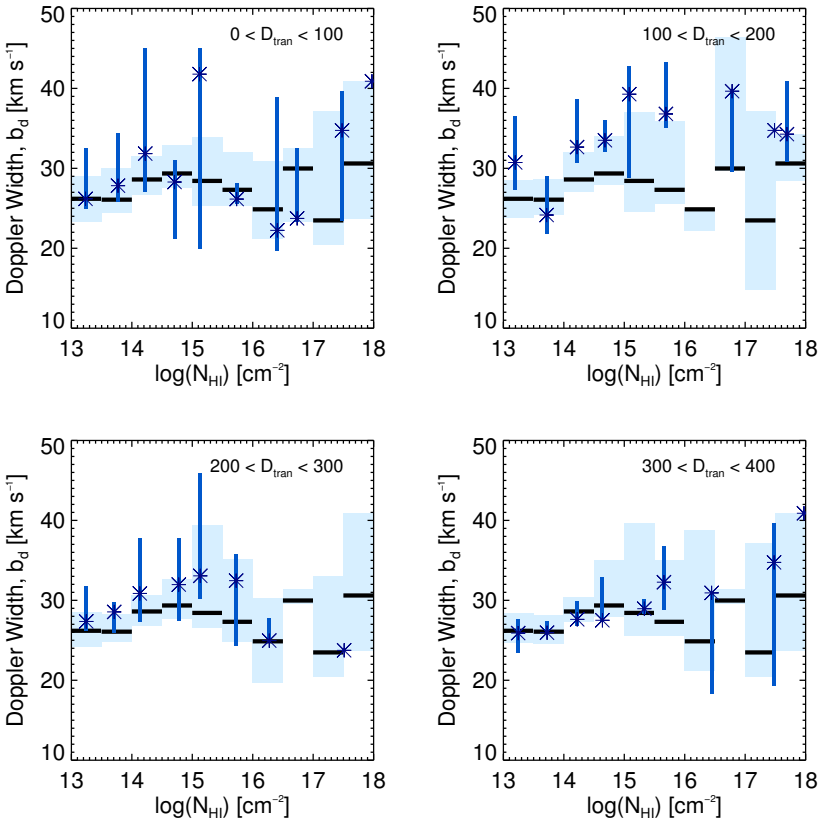


Figure 2.27 Comparison at fixed $N_{\rm HI}$ of absorbers close to galaxies with absorbers in the full H I catalog. The different panels show four bins of $D_{\rm tran}$. Asterisks represent the median value of $b_{\rm d}$ in a bin of $N_{\rm HI}$ for absorbers close to galaxies. The horizontal black lines show the median $b_{\rm d}$ of absorbers from the full H I catalog, and the light shaded boxes represent their dispersion calculated through bootstrap samples of the same size as those in the real sample. The velocity window is $|\Delta v| < 700$ km s^{-1}. In the bins without asterisks, there were no absorbers which satisfied both the $D_{\rm tran}$ and $N_{\rm HI}$ criterion; similarly, asterisks without error bars represent bins containing only a single absorber satisflying both the $D_{\rm tran}$ and $N_{\rm HI}$ criterion. Recall from Figure 2.26 that the strongest signal occurred at $100 < D_{\rm tran} < 200$ pkpc (shown in this Figure in the top right panel). Note that the median $b_{\rm d}$ of all absorbers close to galaxies with $N_{\rm HI} > 10^{14}$ cm^{-2} are larger than those in the full absorber catalog.

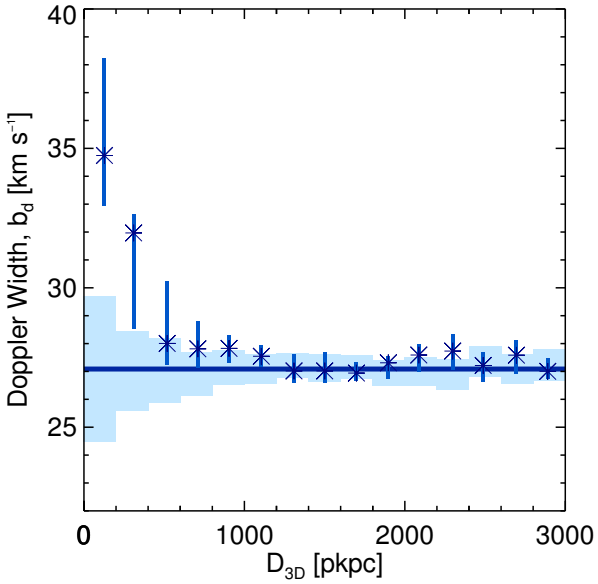


Figure 2.28 Same as Figure 2.26 but versus the 3D distance.

the difference between the median $b_{d,MC}$ from the Monte Carlo sample and the median b_d in the full absorber catalog was computed (asterisks minus black bars in Figure 2.27). The differences from the N_{HI} bins were then averaged. The fraction of Monte Carlo samples whose average b_d excess is greater than that of the D_{tran} $<$ 300 pkpc sightline sample is 0.005, meaning the excess is significant at approximately the 3σ level.

2.7.2 Doppler Widths vs. 3D Distance

In Figure 2.28 we consider the distribution of b_d as a function of D_{3D}, which allows one to consider trends as a function of distance without pre-selecting a velocity window. Here again, b_d is elevated for bins close to galaxies ($D_{3D} \lesssim 400$ pkpc). In the first bin, the median $b_d \approx 35$ km s^{-1} compared to b_d=27 km s^{-1} for the random sample.

We again perform a comparison at fixed N_{HI} and since the median b_d monotonically decrease with increasing D_{3D}, we consider cumulative distance bins. For absorbers with $D_{3D}<$ 600 pkpc, the median b_d for absorbers with $N_{HI}< 10^{16}$ cm^{-2} is larger than the median measured in the full absorber catalog. Notably, for $15 < \log(N_{HI})< 16$, absorbers close to galaxies ($D_{3D}<$ 400 pkpc) have a median $b_d > 40$ km s^{-1}. If interpreted as the result of an increase in gas temperature, the change in b_d is equivalent to more than doubling T. We also note that these intermediate N_{HI} systems that are characteristic of the CGM tend also to be associated with the strongest high ionization metals (O VI, C IV), often exhibiting evidence that the ionization has been produced by shocks in addition to photoionization (Simcoe et al. 2002).

As previously stated, the 3D distance is strongly influenced by the magnitude of $|\Delta v|$. Notably, the first three bins in D_{3D} (those with the most significant deviation in b_d) correspond to absorbers with $|\Delta v| < 150$ km s^{-1}.

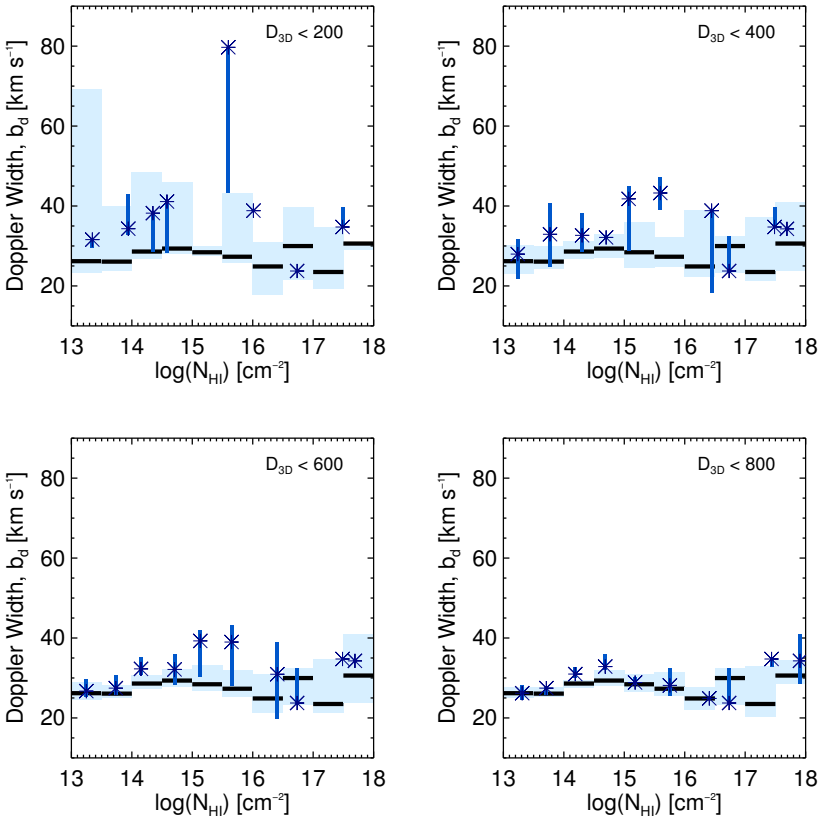


Figure 2.29 Same as Figure 2.27 but now divided into bins of 3D distance, D_{3D}, and with cumulative bins in D_{3D}. From Figure 2.28 one notes that the majority of the signal results from $D_{3D} <$ 400 pkpc (shown in this Figure in the upper two panels). Notably, absorbers near galaxies have higher b_d at nearly all values of N_{HI} at these impact parameters. Because the signal from all bins is combined, higher b_d at fixed N_{HI} is evident for absorbers with $D_{3D} <$ 600 pkpc.

In summary, considering the trends illustrated in Figures 2.26 – 2.29, it appears that absorbers of all $N_{\rm HI}$ with $100 < D_{\rm tran} < 300$ pkpc and $|\Delta v| < 150$ km s^{-1} have significantly larger Doppler widths than absorbers of the same $N_{\rm HI}$ in the full absorber sample.

2.7.3 Doppler Parameter: Possible Physical Explanations

The trends in Doppler widths, $b_{\rm d}$, could have a variety of physical origins. As previously discussed, elevated $b_{\rm d}$ could be related to the observed trend in $N_{\rm HI}$ through the $T-\rho$ relation in the IGM. However, we showed that at *fixed* $N_{\rm HI}$, the median $b_{\rm d}$ of absorbers close to galaxies is still significantly larger than in the full absorber catalog (Figures 2.27 and 2.29). This suggests that the $T-\rho$ relationship alone cannot account for the broader absorption lines close to galaxies.

Another possible contributor to increasing $b_{\rm d}$ near galaxies is a change in the *ionization* at a given $N_{\rm HI}$. If the galaxy itself produces a significant fraction of the local ionizing radiation field (i.e., in excess of that of the metagalactic UV background), then $N_{\rm HI}/N_{\rm H}$ would be smaller near galaxies. The fraction of the ionizing photons which escape the ISM of the galaxy is small (Nestor et al. 2011, Steidel et al. in prep), but their number, compared to the UV background is likely significant within $\sim$ 100 pkpc. In this case, the $T-\rho$ relation could hold, but the mapping of $N_{\rm HI}$ to ρ would be altered. If this were the case, we would expect the locations of the asterisks in Figures 2.27 and 2.29 to be consistent with a leftward shift of the black bars. However, this scenario is at odds with the observation that the median $b_{\rm d}$ of absorbers with $10^{14} < N_{\rm HI} < 10^{16}$ cm^{-2} and $100 < D_{\rm tran} < 200$ pkpc is higher than the median $b_{\rm d}$ for any bin in $N_{\rm HI}$ observed in the full absorber catalog (Figure 2.27). The same is true of absorbers with $D_{\rm 3D} < 400$ pkpc and $10^{14} < N_{\rm HI} < 10^{16}$ cm^{-2} (Figure 2.29).

Assuming that overdensity is not the main cause of the increased $b_{\rm d}$, we consider other plausible processes which could lead to an increase in the temperature or turbulence of the gas. In principle it is possible to disentangle the source of the broadening (turbulent vs. thermal) by comparing the widths of absorption lines arising from ions with different atomic weights. The thermal broadening component, $b_{\rm d}(T)$, depends on atomic mass, while the turbulent broadening component will remain constant so long as the ions reside in the same gas. While the HIRES QSO spectra in our sample include metal absorption lines which could be used for this purpose, the analysis of the metal lines is beyond the scope of this paper and will be presented elsewhere. At present, we consider the implications of both possible effects (temperature and turbulence).

Referring to Figure 2.28, there is an increase in the median $b_{\rm d}$ from 27 km s^{-1} to 35 km s^{-1}. Assuming purely thermal broadening, these values would indicate $T = 4.4 \times 10^4$ K and 7.4×10^4 K, respectively, or an increase in temperature of 68%. If we consider the absorbers with $b_{\rm d} \gtrsim$ 40 km s^{-1} as in Figure 2.29, these suggest a temperature of $\sim 10^5$ K. If instead the increase in line width is due to turbulence, an excess turbulent velocity of 20-30 km s^{-1} would account for the departure from the $b_{\rm d}$= 27 km s^{-1} global median. In either case, it is curious that the elevated Doppler widths occur in gas with $100 < D_{\rm tran} < 300$ pkpc but are not detected in the bin from 0 - 100 pkpc. While the non-detection at the smallest impact parameters is likely

not significant, clearly, whichever mechanism is responsible for this increase in width is acting significantly outside of the typical galaxy virial radius ($\sim$ 90 pkpc).

Unless AGN are involved (galaxies with detected AGN signatures were not used in our analysis), excess photo-heating for material closer to galaxies seems unlikely, since the UV radiation field from local stellar sources is likely to be (if anything) softer than the metagalactic background believed to dominate at an average IGM location.

There are at least two plausible processes which could increase $b_{\rm d}$ in the CGM of galaxies: galactic winds (outflows) and baryonic accretion. Accreting gas can be heated by shocks forming as the gas falls into galaxy halos. The galaxies in our sample have halo masses comparable to the theoretically expected transition mass between "cold mode" and "hot mode" accretion (van de Voort et al. 2011b; Faucher-Giguère et al. 2011; Kereš et al. 2009; Ocvirk et al. 2008); most of these authors predict that both hot and cold accretion occur in such halos at $z > 2$. The virial temperature for galaxies with halos of mass $\sim 10^{12}$ M$_\odot$ is expected to be $T_{\rm vir} \simeq 10^6$ K (van de Voort et al. 2011b), and Kereš et al. (2009) suggest that this hot gas may fill the volume surrounding galaxies to a few $r_{\rm vir}$. The H I absorbers in our sample are unlikely to trace gas this hot, however gas cooling from this temperature would be detectable.

The interaction of cool dense gas with either a hot halo or a rarified wind fluid is also expected to produce turbulent boundary layers via both Rayleigh-Taylor[21] and Kelvin Helmholtz[22] instabilities (Kereš & Hernquist 2009; Faucher-Giguère & Kereš 2011). These boundary layers are also where metallic ions such as O VI are expected to be most abundant.

Galactic super winds can naturally explain either thermal or turbulent broadening, potentially to large distances away from galaxies. The favored mechanism for large scale winds is shocks from supernovae which could easily increase the gas temperature of the ambient medium. Although the distance to which galactic winds propagate is not well constrained, there is evidence that they expand to at least $r_{\rm vir}$ (Steidel et al. 2010). Shen et al. (2011) recently considered high-resolution "zoom-in" simulations of a single Lyman Break Galaxy[23] at $z = 3$. In their model, the LBG itself (as opposed to its satellite galaxies) enriches the IGM with metals from supernovae-driven winds to $3r_{\rm vir}$ (see their Figure 9). Kollmeier et al. (2006) considered the effect of winds on the CGM of star-forming galaxies at $z = 3$ and found that winds could have a profound effect on the temperature of nearby gas: varying the energy released by supernovae in their simulations by a factor of 5 changed the temperature of gas close to galaxies by $\sim 250\%$ (see their Figure 10). More recently, van de Voort & Schaye (2011) considered the physical properties of gas surrounding galaxies with halo masses $M_{\rm DM} = 10^{12}$ M$_\odot$ at $z = 2$ as a function of the kinematics of the gas (i.e. inflowing vs. outflowing). They found a radial temperature profile of outflowing gas qualitatively consistent with the observed increase in $b_{\rm d}$. Outflows are also predicted to drive gas turbulence outside of virial halos through the conversion of

[21] Rayleigh-Taylor instabilities would form in the case of a lighter, less-dense fluid lying deeper in the gravitational potential of the galaxy than a denser cold-stream.

[22] Kelvin Helmholtz instabilities would form in the condition where the dense stream was moving with respect to the hot medium. This is especially plausible in the case of a fast moving galactic wind or a filament moving at the free-fall velocity.

[23] Lyman Break Galaxies at $z = 3$ have very similar properties to the galaxies studied in this paper.

kinetic energy from supernova driven bubbles into random motions through gas instabilities (see e.g., Evoli & Ferrara 2011).

While the exact cause of the enlarged line widths is not apparent, it is clear that the physical state of the gas within $\sim$ 300 pkpc is markedly different from that in "random" locations in the IGM - suggesting that galaxies and their potential wells deeply affect their surrounding CGM. This fact may introduce significant complications to studies of the general IGM which will be discussed further in §2.8.5.

2.8 General Discussion

In §2.4 – 2.7 we presented measurements and plausible interpretations of the properties of neutral hydrogen gas surrounding star-forming galaxies at high z. Here, we speculate on the possible implications of the observed trends.

2.8.1 Gaseous "Zones" around Galaxies

The distribution of strong absorbers (Figure 2.9) and of the incidence of absorbers (Figures 2.16 and 2.17) suggest three distinct "zones" in the gaseous envelopes surrounding galaxies.

1. The first zone corresponds to the volume within $D_{\rm tran} <$ 300 pkpc. For scale, the virial radius of a typical galaxy in our sample is $\sim$ 90 pkpc. Inside 300 pkpc, the values of the max($N_{\rm HI}$) statistic are significantly elevated and rise toward the position of the galaxy (Figure 2.9). Figure 2.16 shows rising $\eta_{\rm abs}$ with decreasing $D_{\rm tran}$ for the three bins within 300 pkpc. Similarly, at 300 pkpc, $\eta_{\rm abs}$ for absorbers with $N_{\rm HI} > 10^{14.5}$ drops sharply. In Figure 2.17, again at 300 pkpc one sees a rapid drop in the incidence of all absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2}.

2. The second zone lies between 300 pkpc $< D_{\rm tran} \lesssim$ 2 pMpc. In Figure 2.9, we observed that max($N_{\rm HI}$), while high compared to the random distribution, plateaus at a roughly constant value within this zone. $\eta_{\rm abs}$ also plateaus over the same range as shown in Figures 2.16 and 2.17.

3. The third zone corresponds to $D_{\rm tran} \gtrsim$ 2 pMpc. The transition between the second and third zone is not a sharp feature; rather between $\sim 2-3$ pMpc, the distribution of absorbers as seen in the max($N_{\rm HI}$) and $\eta_{\rm abs}$ statistics drops to become consistent with (or below that of) the IGM median.

Three distinct zones are also evident in the velocity distribution of absorbers.

1. The first zone, corresponding to the strongest peak in the velocity distribution, is within $|\Delta v| < 300$ km s^{-1} (Figures 2.6 and 2.7).

2. The second velocity zone includes gas to $|\Delta v| \approx 700$ km s^{-1}, which encompasses the full extent of the excess $N_{\rm HI}$ (Figures 2.5 and 2.6).

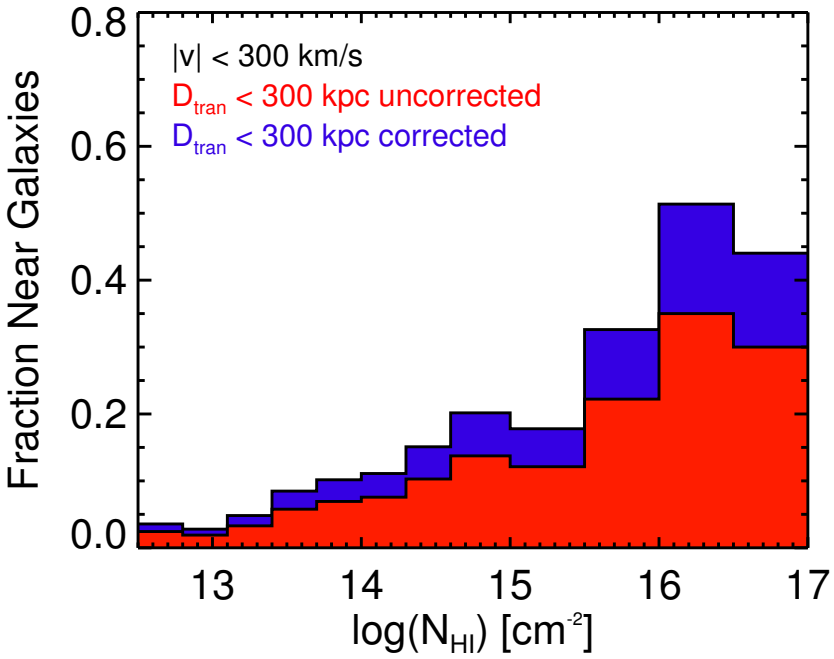


Figure 2.30 The fraction of all H I systems in our QSO sightlines that arise within $|\Delta v| < 300$ km s^{-1} and $D_{\rm tran} < 300$ pkpc of the position of a galaxy in our spectroscopic galaxy sample (lighter red histogram), as a function of $N_{\rm HI}$. The darker (blue) histogram indicates the fraction after correcting for the fact that not all galaxies in the photometric sample have been observed spectroscopically. These corrected points thus reflect the fraction of absorbers arising within the CGM of galaxies that meet our *photometric* selection criteria.

3. The third zone is $|\Delta v| > 700$ km s^{-1}, where the distribution with respect to galaxies is consistent with that of random locations (Figures 2.5 and 2.6).

Note that in Figure 2.7, the median $N_{\rm HI}$ of absorbers with $D_{\rm tran} < 300$ pkpc exhibits a sharp drop-off at $|\Delta v| \geq 300$ km s^{-1} suggesting that absorbers at $D_{\rm tran} > 300$ pkpc are likely responsible for the excess absorption at $300 < |\Delta v| < 700$ km s^{-1}.

2.8.2 Defining the CGM

We adopt a working definition of the "circumgalactic medium" (CGM) as the region within $|\Delta v| < 300$ km s^{-1} and $D_{\rm tran} < 300$ pkpc of a galaxy. Figure 2.30 shows (in red) the fraction of absorbers from the full absorber catalog which fall within the CGM of a galaxy *in our spectroscopic sample* as a function of $N_{\rm HI}$.

Our spectroscopic galaxy sample is 70% complete with respect to the photometric parent sample within 300 pkpc of the line-of-sight to the QSOs.[24] The blue histogram in Figure 2.30 has been corrected for this incompleteness assuming that unobserved galaxies have the same overall redshift distribution and similar CGM to those that have been observed. Figure 2.30 shows that $\gtrsim$ 20% of all absorbers with $\log(N_{\rm HI}) >$

[24] Our absorption line catalogue is essentially 100% complete for absorbers with $\log(N_{\rm HI}) \gtrsim 13$.

14.5 and $\gtrsim 40\%$ of those with $\log(N_{\rm HI}) > 15.5$ arise in the CGM of galaxies that meet our selection criteria. Noting that the comoving number density of identically selected galaxies (with the same limiting magnitude of $\mathcal{R} = 25.5$) is $\Phi = 3.7 \times 10^{-3}$ cMpc^{-3} (Reddy et al. 2008), the CGM of LBGs at $\langle z \rangle = 2.3$ contains only $\simeq 1.5\%$ of the Universe's volume but accounts for nearly half of the total gas cross-section for $N_{\rm HI} > 10^{15.5}$ cm^{-2}. This is perhaps surprising given that our sample includes only galaxies with $L_{UV} \geq 0.25 L^*_{UV}$ and the faint-end slope of the UV luminosity function is very steep at $z \sim 2.3$; however, it is consistent with previous results based on strong metal-line absorption around similar galaxies at comparable redshifts (Adelberger et al. 2003, 2005a; Steidel et al. 2010).

The situation changes substantially for $N_{\rm HI} \lesssim 10^{14.5}$ cm^{-2}; we have seen that these lower $N_{\rm HI}$ absorbers are more loosely associated with galaxies in our sample, though their incidence is enhanced by $\sim 10-20\%$ relative to average locations in the IGM (e.g. see the top panels of Figure 2.16).

Stated another way, we have found that low and high- column density absorbers cluster differently with high-z, UV-bright galaxies. We showed (Figure 2.16) that absorbers with $\log(N_{\rm HI}) > 14.5$ exhibit a much stronger peak in incidence towards the positions of galaxies than do the absorbers with lower column densities. Similarly, in maps of the incidence (Figures 2.21 – 2.24) low-column density absorbers show only large scale correlation with the positions of galaxies while high-column density absorbers again showed a strong peak near galaxies both along the line of sight and in $D_{\rm tran}$.

Collectively, these lines of evidence point to a circumgalactic zone defined by the boundaries of $|\Delta v| <$ 300 km s^{-1} and $D_{\rm tran} <$ 300 pkpc and dominated by absorbers with $\log(N_{\rm HI}) > 14.5$. They also suggest that only systems with $\log(N_{\rm HI}) \lesssim 14$ cm^{-2} are confidently "IGM"– as $N_{\rm HI}$ increases beyond this limit, the likelihood that the gas lies within the CGM of a relatively bright galaxy increases very rapidly.

2.8.3 Small Scale Distribution of $N_{\rm HI}$

In an effort to understand the physics responsible for the distribution of $N_{\rm HI}$ as a function of transverse distance from galaxies, we compare the data to two theoretical models. First we consider the possibility that the H I column density traces the dark matter halo density profile.

The typical dark matter halo hosting a galaxy in our sample is theoretically expected to follow a NFW (Navarro et al. 1997) density profile:

$$\rho(r) = \frac{\rho_c(z)\ \delta_{\rm NFW}}{c\frac{r}{r_{\rm vir}}\left(1 + c\frac{r}{r_{\rm vir}}\right)^2} \tag{2.20}$$

where $\delta_{\rm NFW}$ is a normalization defined as

$$\delta_{\rm NFW} = \frac{200}{3}\frac{c^3}{\ln(1+c) - \frac{c}{1+c}} \tag{2.21}$$

and ρ_c is the critical density, c is the concentration parameter of the dark matter halo which we take to be $c = 4$ following Duffy et al. (2008), and $r_{\rm vir}$ is the virial radius for halos of average mass 10^{12} M$_\odot$ (91 pkpc;

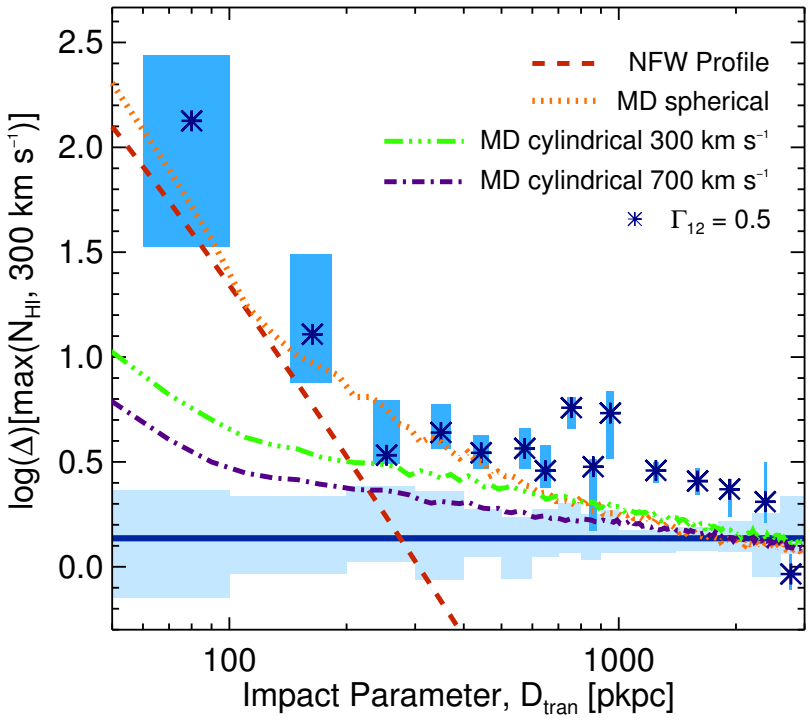


Figure 2.31 The max($N_{\rm HI}$, 300 km s^{-1}) statistic, converted into a universal overdensity using Schaye (2001) plotted against transverse distance to a galaxy on a logarithmic scale. The same data shown in Figure 2.10 with two models over-plotted for comparison. Here we use 100 pkpc bins within 1 pMpc and 400 pkpc bins at larger $D_{\rm tran}$ to simplify the plot. Shown in the (red) dashed line is the NFW radial density profile plotted as a universal overdensity. See text for more details. Shown in the light (orange) dotted , light (green) dash-dotted, and dark (purple) dash-dotted lines are the average dark matter profiles from the MultiDark simulation for halos with $M_{\rm DM} > 10^{11.8}$ $M_\odot$. The light (orange) dotted curve is the median value (across the halos considered) of the average density in concentric spherical shells of radius $D_{\rm tran}$. The light (green) dash-dotted curve is the median (across the halos considered) of the average dark matter density in concentric cylindrical shells of radius $D_{\rm tran}$ and length $\pm$300 km s^{-1}= $\pm$ 1200 pkpc. The dark (purple) dash-dotted curve is analogous to the green one, but the velocity window is $\pm$700 km s^{-1}.

Trainor & Steidel, in prep; Conroy et al. 2008; Adelberger et al. 2005b).

We define Δ, the matter overdensity, as:

$$\Delta = \frac{\rho}{\overline{\rho}} = \frac{\rho}{\Omega_{\rm M}(1+z)^3\ \rho_c(z=0)}. \tag{2.22}$$

Noting the definition for the critical density:

$$\rho_c(z) = \frac{3H^2(z)}{8\pi G} \tag{2.23}$$

we can then express the NFW profile as a function of overdensity, rather than density:

$$\Delta_{\rm NFW}(r) = \frac{1}{\Omega_{\rm M}(1+z)^3} \frac{\delta_{\rm NFW}}{c\frac{r}{r_{\rm vir}}\left(1+c\frac{r}{r_{\rm vir}}\right)^2} \frac{H^2(z)}{H_0^2}. \tag{2.24}$$

Using Jeans scale arguments, Schaye (2001) defined a scaling function which relates the gas overdensity ($\rho_b/\overline{\rho}_b$) to an expected value of $N_{\rm HI}$. While on the smallest scales where the growth is highly non-linear it is certainly not the case that $\rho_b/\overline{\rho}_b$ is equal to Δ, the matter overdensity (which is dominated by dark matter), the expectation is that they should mirror each other well at low overdensity ($\Delta \approx 1-10$) where the growth is yet to become significantly non-linear. We therefore assume $\rho_b/\overline{\rho}_b \approx \Delta$, in order to compare our measured max($N_{\rm HI}$) to theoretical density profiles. For the cosmology used in this paper, at the redshifts of our sample, the scaling function from Schaye (2001) is:

$$\Delta \approx \left(\frac{N_{\rm HI}}{10^{13.4}}\right)^{2/3} T_4^{0.17} \left(\frac{\Gamma_{12}}{0.5}\right)^{2/3} \left(\frac{1+z}{3.3}\right)^{-3}, \tag{2.25}$$

where T_4 is the gas temperature in units of 10^4 K and Γ_{12} is the hydrogen photoionization rate in units of 10^{-12} s^{-1} where the normalization is taken from Faucher-Giguère et al. (2008). Using this relation, we can convert our max($N_{\rm HI}$) statistic into Δ which can then be compared to the rescaled NFW profile.

The values after converting max($N_{\rm HI}$) to Δ are shown in Figure 2.31. The (red) dashed curve is the NFW density profile (Equation 2.24). The NFW halo has a radial distribution similar to that of the Δ inferred from our max($N_{\rm HI}$) statistic for small $D_{\rm tran}$, but falls well below the Jeans-scale-inferred overdensity for $D_{\rm tran} \gtrsim 300$ pkpc. This is expected since the NFW profile is measured in simulations for dark matter (DM) within the virial radius. In the following section, we compare to the DM density on larger scales using DM profiles drawn directly from simulations.

2.8.4 Large Scale Distribution of $N_{\rm HI}$

Considering the distribution of gas around galaxies on Mpc scales, it is interesting to note the similarity between the scale over which circumgalactic gas has higher density than an average location in the IGM ($D_{\rm tran} \lesssim 2$ pMpc) and the galaxy-galaxy autocorrelation scale length recently measured *from the same galaxies*, $r_0 = (6.5 \pm 0.5)h^{-1}$ cMpc ($\simeq 2.8$ pMpc at $\langle z \rangle = 2.3$; Trainor & Steidel, in prep). It would not be surprising if they were closely related.

Most of the recent work on the dark matter density profiles at several $r_{\rm vir}$ and beyond have focused on the local universe. Since we are interested in comparing the gas-inferred density profiles with an average dark matter profile at $z \sim 2-3$, we created median DM density profiles using the MultiDark simulation (Klypin et al. 2011). The radial density profile within 3 pMpc was computed from 100 halos with $\rm M_{DM} > 10^{11.8}$ $\rm M_\odot$[25]. Here we consider the dark matter density profile measured in two ways presented in terms of universal

[25]The halo mass was inferred from the number of particles found to "belong" to a halo using the Friends-of-Friends algorithm. The

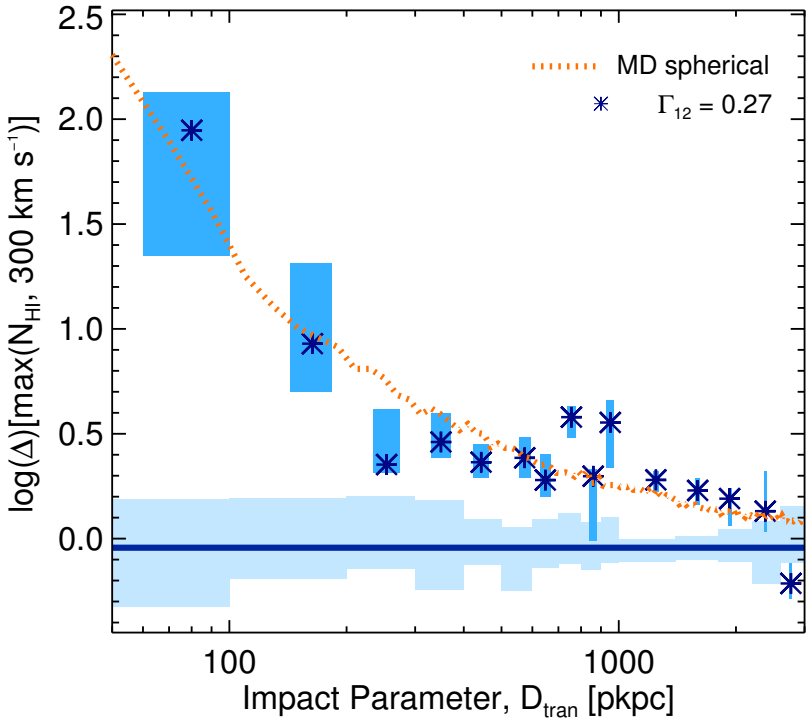


Figure 2.32 Same as Figure 2.31 but using a different normalization of the photoionization rate. In Figure 2.31 the data values were converted into Δ assuming $\Gamma_{12} = 0.5$ as suggested by Faucher-Giguère et al. (2008). Here, the data values are converted into Δ assuming $\Gamma_{12} = 0.27$ which bring them into the closest agreement with the MultiDark spherical density profile measurements.

overdensity as shown in Figure 2.31. (1) First, we consider the average density in concentric spherical shells of radius $D_{\rm tran}$ centered on the position of each halo. We take the median of this profile across the 100 halos considered (light orange dotted curve) and compare it to the max($N_{\rm HI}$, 300km s^{-1}) statistic converted into overdensity using equation 2.25. (2) We also consider the average dark matter density profile computed within concentric cylindrical shells of radius $D_{\rm tran}$ and length Δv= $\pm$300 km s^{-1} ($\pm$1200 pkpc) as shown in the light (green) dash-dotted curve. We consider the effect of the velocity window used as well - the dark (purple) dash-dotted curve is analogous to the green curve, but with a cylinder length of $\pm$700 km s^{-1}.

The spherically-averaged radial density profile is well-matched by the NFW profile on small scales as expected (Figure 2.31). After 200–300 pkpc (the same scale as the first "zone" of the CGM, §2.8.1), the spherically averaged radial density profile flattens compared with NFW. From $\sim$200–300 pkpc onwards, the inferred density declines smoothly, falling to the mean density of the universe [log(Δ)=0] at $D_{\rm tran}>$ 3 pMpc.

The cylindrically averaged density profile significantly dilutes the signal in the inner bins of $D_{\rm tran}$ suggesting that the max($N_{\rm HI}$) statistic traces more of a radial distribution than a "line-of-sight" distribution on small scales. On larger scales, the cylindrical and spherically averaged density profiles look quite similar. Between $D_{\rm tran} \approx$ 1.5 - 2 pMpc, the two cylindrically averaged curves become consistent with each other. This suggests that on scales of $\sim$ 2 pMpc the density along a sightline between $\pm$300 km s^{-1} and $\pm$300 to $\pm$700 km s^{-1} is similar – as one would expect far from galaxies as the density approaches the universal mean.

The density inferred from the H I measurements is markedly higher on $\sim$ pMpc scales than is predicted by either MultiDark DM density profile. There are three possible causes for this discrepancy: (1) It may be the case that the max($N_{\rm HI}$) statistic traces high-density substructure while the radial density profiles drawn from the simulations are (by construction) smoothed - averaging over these structures. The assumption of Jeans smoothing intrinsic to the arguments of Schaye (2001) should minimize the discrepancy; however, it is likely still that substructure may affect our max statistics. (2) One or more of the assumptions used in the Schaye (2001) conversion between $N_{\rm HI}$ and Δ could be invalid. For example one of the more uncertain quantities required in the conversion is the H I photoionization rate, Γ.[26] In Figure 2.32 we consider the change in the value of Γ needed to bring the measured Δ[max($N_{\rm HI}$, 300km s^{-1})] into agreement with the spherically averaged dark matter density profile. This plot suggests a normalization $\sim$40% lower than the estimates from Faucher-Giguère et al. (2008).[27] (3) Another possibility is that Jeans scale arguments do not hold in detail on these scales.[28]

---

halo-mass cut of $10^{11.8}$ M$_\odot$ is the same mass threshold which reproduces the clustering of galaxies in our spectroscopic sample, as described in Trainor & Steidel (in prep). The mean and median halo masses of galaxies with this halo-mass threshold are $10^{12.2}$ M$_\odot$ and 10^{12} M$_\odot$, respectively.

[26] Γ_{12} refers to the photoionization rate in units of 10^{-12} s^{-1}.

[27] The temperature of the gas is another quantity that could be adjusted. The temperature assumed currently is 10^4 K. To match our observations we would need to significantly lower the temperature assumed to $T < 10^3$ K which is unlikely to be the case. If anything, 10^4 K is likely a *lower* limit on the temperature of the IGM at $z \sim 2$ (see e.g., Becker et al. 2011a).

[28] We tested the agreement between the aforementioned models and our data converted using the Fluctuating Gunn-Peterson Approximation (FGPA; Rauch et al. 1997), finding that the disagreement is larger. Rakic et al. (2011b) also compare several methods for converting between H I optical depth (akin to $N_{\rm HI}$) and overdensity (see their §4.5).

2.8.5 The Implications of the Distinct Physics of the CGM

The final point is a note of caution regarding the effect the CGM of luminous galaxies may have on measurements which rely on "average" locations in the IGM. In addition to the rising $N_{\rm HI}$ and $\eta_{\rm abs}$ observed in the CGM of galaxies, we have shown that the physical state of the absorbing gas in this region is also distinct from the general IGM. The Doppler widths, $b_{\rm d}$, of individual absorbers are significantly larger in the CGM zone, even when compared at *fixed* $N_{\rm HI}$. Further, absorbers with $N_{\rm HI} > 10^{14.5}$ cluster tightly with galaxies (Figures 2.16, 2.17, and 2.22), and these absorbers are found in the CGM zone with very high frequency (Figure 2.30).

Taken together, these observations suggest that the physics of the "IGM" which has long been leveraged for measurements varying from the metagalactic background to the matter power spectrum, may be significantly affected by the baryonic physics of galaxy formation; particularly, low-resolution IGM studies dominated by H I absorbers $\sim$saturated in Lyα ($N_{\rm HI} > 10^{14.5}$ cm^{-2}) are likely to be affected by the presence of a nearby galaxy. Further observational and theoretical examination of this subject is crucial to a full understanding of the measurement power of the IGM.

In addition, studies of the distribution of metal absorption lines in the "IGM" often consider metallic species associated with HI absorbers of $\log(N_{\rm HI}) \gtrsim 14.5$ (Cowie et al. 1995; Songaila & Cowie 1996; Ellison et al. 2000; Carswell et al. 2002; Simcoe 2011). To a large degree, this is a practical choice as the associated metal absorbers occurring in sub-solar metallicity gas must be found and measured in spectra of finite S/N. However, the properties of the CGM discussed above suggest that, as many authors have speculated, these measurements more directly probe the metallicity of the CGM than that of the IGM. Observations at $z \gtrsim 5$ generally cannot measure the $N_{\rm HI}$ associated with metal systems due to the line density of the forest, but they are likely to suffer from the same effect, adding credence to the idea that the recent discovery of the high-z downturn of C IV (Ryan-Weber et al. 2006, 2009; Becker et al. 2009, 2011c; Simcoe et al. 2011) likely also traces CGM gas.

The observations of circumgalactic and intergalactic gas and its association with forming galaxies have now progressed to the point that more stringent comparison to theoretical expectations is possible. Even using the comparatively "blunt" tool of gas covering fraction, we have shown (§2.5.2) that the distribution of HI gas around galaxies exceeds the expectations of the models to which they have been compared, with the discrepancy increasing as $N_{\rm HI}$ decreases below 10^{17} cm^{-2}. It is clear that a combination of high-resolution (to capture the small-scale baryon physics) and cosmological (to provide sufficient statistical leverage) simulations will be required to make use of the more detailed physical measurements emerging from the observations: e.g., gas-phase kinematics, multi-phase velocity/density structure, redshift-space distortions, gas temperatures, turbulence, the distribution and mixing of metals, all as a function of galaxy properties. In order to match the real universe, simulations will need to realistically model the physics of feedback from star formation, supernovae, and AGN activity and its effect on gas-phase accretion from the CGM.

It will be important for future simulations to have sufficient dynamic range to make predictions about

the distribution of neutral hydrogen with $13 < \log(N_{\rm HI}) < 20$. The physical picture advocated by most modern theoretical treatments of the "IGM" is that the Lyα forest traces gas of modest overdensity within the filaments of the "cosmic web." The high covering fractions found in this study (e.g., the relatively large $f_{\rm c}$ of saturated HI extending $\gtrsim$2 pMpc from $\sim 10^{12}$ M$_\odot$ dark matter halos; §2.5) suggest that if this picture holds, the size of these filaments may remain quite large even as they approach the positions of high-bias halos. These initial comparisons suggest that the topology of the neutral hydrogen gas may be less "filamentary" (at all column densities) than current simulations of the "cosmic web" and "cold accretion" predict. If such disagreement between simulations and observations are common and persist they may become key to a deeper understanding of the baryon physics that ultimately controls the process of galaxy formation.

2.9 Summary

In this paper we have presented a detailed study of neutral hydrogen in the IGM surrounding 886 high-z star-forming galaxies using spectra of background QSOs from the Keck Baryonic Structure Survey. We draw conclusions from the analysis of 15 sightlines to hyper-luminous high-z QSOs using high-resolution (7 km s^{-1}), high-signal-to-noise ratio (50-200) spectra. This study constitutes the largest absorption line catalog ever created at these redshifts through a complete Voigt profile decomposition of the Lyα forest for all 15 sightlines. The high S/N and excellent UV/blue-wavelength coverage of the HIRES QSO spectra allow precise determination of the column densities of saturated H I absorption systems whose positions we have found to correlate strongly with those of the galaxies in our sample. This sample is also considered by Rakic et al. (2011b) who analyze the pixel statistics of the H I optical depth.

In this work, we study the absorption patterns of H I gas in the velocity and spatial locations surrounding each galaxy redshift. We have examined the correlations among column density ($N_{\rm HI}$), covering fraction ($f_{\rm c}$), incidence ($\eta_{\rm abs}$), absorption line width ($b_{\rm d}$), galactocentric transverse distance ($D_{\rm tran}$), velocity offset (Δv), and 3D distance ($D_{\rm 3D}$). Our principal results are as follows:

1. The redshifts of H I absorption systems in the QSO spectra correlate strongly with the redshifts of galaxies (§2.4.1). The distribution of relative velocities (Δv) is peaked at the systemic redshift of galaxies in our sample (Figure 2.5). The full excess absorption near galaxies falls within $\pm$700 km s^{-1} with the majority of the excess $N_{\rm HI}$ (i.e. the higher column density absorbers) falling within $\pm$300 km s^{-1} (Figures 2.6 and 2.7).

2. The column density ($N_{\rm HI}$) of absorbers is strongly correlated with the impact parameter to a galaxy ($D_{\rm tran}$, §2.4.2). We consider the highest-$N_{\rm HI}$ absorber within $\pm$300 km s^{-1} [max($N_{\rm HI}$)] of each of the 10 galaxies within 100 physical kpc and find that the median value is 3 orders of magnitude higher than at random places in the IGM (Figure 2.10). A trend of rising column densities as a function of decreasing $D_{\rm tran}$ is observed within the inner 300 physical kpc (1 cMpc) surrounding galaxies (Figures 2.9, 2.10,

and 2.11).

3. The column densities ($N_{\rm HI}$) of absorbers as measured through the max($N_{\rm HI}$) statistic plateau in the range 300 pkpc $< D_{\rm tran} \lesssim$ 2 pMpc (Figures 2.9 and 2.10, §2.4.2.1). Interestingly, the plateau is elevated in $N_{\rm HI}$ with respect to random places in the IGM by 0.3 − 0.6 dex. For $D_{\rm tran} \gtrsim$ 2 pMpc, the values of max($N_{\rm HI}$) decline and become consistent with random places in the IGM.

4. The median $N_{\rm HI}$ as a function of 3D distance[29] smoothly declines with increasing $D_{\rm 3D}$ and becomes consistent with random places in the IGM at $D_{\rm 3D} \approx$ 3 pMpc (Figure 2.14, §2.4.3).

5. The incidence[30] ($\eta_{\rm abs}$) of absorbers of all $N_{\rm HI}$ is higher within $D_{\rm tran} \lesssim$ 2 physical Mpc of the position of a galaxy than at random places in the IGM (Figure 2.16, §2.5). For $D_{\rm tran} \gtrsim$ 2 pMpc, $\eta_{\rm abs}$ is consistent with that of the random IGM.

6. For absorbers with log($N_{\rm HI}$) > 14.5 cm^{-2}, there is a strong peak in the incidence, $\eta_{\rm abs}$, close to galaxies at $D_{\rm tran} < 300$ pkpc (Figure 2.16, §2.5). No such peak is seen in the distribution of lower-column density absorbers. This suggests that absorbers with log($N_{\rm HI}$) > 14.5 cm^{-2} are more directly related to galaxies than absorbers of lower $N_{\rm HI}$.

7. The incidence, $\eta_{\rm abs}$, of absorbers with $N_{\rm HI} > 10^{14.5}$ cm^{-2} and $|\Delta v| <$ 300 km s^{-1} shows rising values within $D_{\rm tran} <$ 300pkpc. From 300 pkpc $< D_{\rm tran} \lesssim$ 2 pMpc, the $\eta_{\rm abs}$ roughly plateaus with $\eta_{\rm abs} \gtrsim$ 0.5. After $\sim$2 pMpc, the distribution becomes consistent with random places in the IGM (Figure 2.17, §2.5.1).

8. Within 300 pkpc and $\pm$350 km s^{-1} of a galaxy, the probability of intersecting an absorber with log($N_{\rm HI}$) > 14.5 is > 4 times higher than at a random place in the IGM. (Figure 2.18 and Table 2.3, §2.5.1).

9. Maps of the incidence, $\eta_{\rm abs}$, of low and high-$N_{\rm HI}$ absorbers as a function of $D_{\rm tran}$ and $D_{\rm Hubble}$ allow for the measurement of redshift anisotropies (§2.6). In low-$N_{\rm HI}$ gas [log($N_{\rm HI}$) < 14.5], on large scales (pMpc) the distribution along the line of sight is compressed compared to the distribution in the transverse direction. This is likely the signature of gas infall e.g. the Kaiser (1987) effect (Figure 2.24). In the high-$N_{\rm HI}$ gas [log($N_{\rm HI}$) > 14.5], we find the "finger of God" effect, the elongation of the distribution along the line of sight compared to the transverse distribution (Figure 2.22). Taking into account the effect of our redshift errors, the data are consistent with peculiar velocities of $\sim$ 250 km s^{-1}. These conclusions are in agreement with those presented in Rakic et al. (2011b) who used pixel optical depth analysis to demonstrate that the compression along the line of sight on large scales is highly significant ($> 3\sigma$).

[29] The 3D distance, $D_{\rm 3D}$, is computed using the quadratic sum of the physical impact parameter between the galaxy and the line of sight to the QSO ($D_{\rm tran}$) and the line-of-sight distance one would calculate assuming the Δv of each absorber was due only to the Hubble flow, $D_{\rm Hubble}$.

[30] The number of absorbers *per galaxy* with $|\Delta v| <$ 300 km s^{-1} and in a given bin of $D_{\rm tran}$.

10. The Doppler widths ($b_{\rm d}$) of individual absorption lines in the HIRES spectra rise as the galactocentric distance is decreased (§2.7). The majority of the increased width is found within $|\Delta v| < 150$ km s^{-1} and $100 < D_{\rm tran} < 300$ pkpc (Figures 2.26 and 2.28). The larger line widths could be caused by an increase in the gas temperature, the turbulence of the gas, or both. The median $b_{\rm d}$ close to galaxies are consistent with a gas temperature $\sim 10^5$ K, or a turbulent contribution of $\sim 20-30$ km s^{-1}.

11. Based on these measurements, we suggest a working definition of the circumgalactic medium (CGM) to be all locations within 300 pkpc and $\pm$300 km s^{-1} of a galaxy. We find that the CGM of our completeness corrected galaxy sample can account for nearly half of all absorbers with log(NHI) $>$ 15.5 (Figure 2.30, §2.8.2). Notably, the CGM of these galaxies comprises only 1.5% of the universe's volume at these redshifts.

These findings not only define relevant scales of the CGM of luminous galaxies at redshift $z \sim 2-3$, but also demonstrate that much of the moderately high-$N_{\rm HI}$ portion of the IGM likely originates in the regions surrounding luminous galaxies where the baryonic physics of galaxy formation appear to affect the physical state of the gas. This point may have important consequences for past and future studies of the "IGM" and should be considered in greater detail.

Our understanding of the physical properties of gas surrounding star-forming galaxies will greatly benefit from analysis of metallic absorption features in the HIRES spectra. These absorption lines probe the ionization state and metallicity of the gas and will provide direct constraints on the total mass in hydrogen and in metals that surround star-forming systems. This work is underway and will be published in the second paper in this series.

With this further analysis, we will have for the first time a quantitative snapshot of the baryonic interplay between the intergalactic medium and forming galaxies which is evidently central to the processes of galaxy formation and evolution.

Acknowledgements

The authors would like to thank Ryan Cooke who contributed the fits to the Damped Profiles in our QSO spectra. We are grateful to Bob Carswell for his assistance with the code VPFIT. Many thanks to Joop Schaye for his careful reading of the draft and insightful comments. We would also like to thank George Becker, Brian Siana, and Jean-René Gauthier for many helpful and interesting discussion. Thanks to Michele Fumagalli for providing the values listed in Table 2.5. We wish to acknowledge the staff of the the W.M. Keck Observatory whose efforts insure the telescopes and instruments perform reliably. Further, we extend our gratitude to those of Hawaiian ancestry on whose sacred mountain we are privileged to be guests.

This work has been supported by the US National Science Foundation through grants AST-0606912 and AST- 0908805. CCS acknowledges additional support from the John D. and Catherine T. MacArthur Foundation and the Peter and Patricia Gruber Foundation. Support for N.A.R. was provided by NASA through Hubble Fellowship grant HST-HF-01223.01 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. This research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration.

Chapter 3

The Temperature-Density Relation in the Intergalactic Medium at Redshift $\langle z \rangle = 2.4$

This chapter was previously published as Rudie et al. 2012, ApJL, 757, L30 with coauthors Charles C. Steidel and Max Pettini.

Abstract

We present new measurements of the temperature-density ($T-\rho$) relation for neutral hydrogen in the $2.0 < z < 2.8$ intergalactic medium (IGM) using a sample of $\sim$6000 individual H I absorbers fitted with Voigt profiles constrained in all cases by multiple Lyman series transitions. We find model independent evidence for a positive correlation between the column density of H I ($N_{\rm HI}$) and the minimum observed velocity width of absorbers ($b_{\rm min}$). With minimal interpretation, this implies that the $T-\rho$ relation in the IGM is not "inverted," contrary to many recent studies. Fitting $b_{\rm min}$ as a function of $N_{\rm HI}$ results in line width - column density dependence of the form $b_{\rm min} = b_0 \left(N_{\rm HI}/N_{\rm HI,0}\right)^{\Gamma-1}$ with a minimum line width at mean density ($\rho/\bar{\rho} = 1, N_{\rm HI,0} = 10^{13.6}$ cm^{-2}) of $b_0 = 17.9 \pm 0.2$ km s^{-1} and a power-law index of $(\Gamma-1) = 0.15 \pm 0.02$. Using analytic arguments, these measurements imply an "equation of state" for the IGM at $\langle z \rangle = 2.4$ of the form $T = T_0 \left(\rho/\bar{\rho}\right)^{\gamma-1}$ with a temperature at mean density of $T_0 = [1.94 \pm 0.05] \times 10^4$ K and a power-law index $(\gamma-1) = 0.46 \pm 0.05$.

3.1 Introduction

The "equation of state" of the low-density intergalactic medium (IGM) is believed to be controlled by two principal processes: photo-heating and adiabatic cooling. The cooling is most directly tied to the expansion of the Universe, while the heating is expected to be a complicated mixture of relic effects from the reionization of hydrogen and helium plus the current heating, predominantly from the UV background (Hui & Gnedin 1997; Schaye et al. 1999). This naturally imposes a relationship between the temperature and density of intergalactic gas. Denser gas is expected to trace larger over-densities, for which adiabatic cooling is suppressed because such regions are more bound against the expansion of the Universe. At the same time, denser gas has a much faster recombination time scale and thus presents a larger cross-section for photo-ionization which leads denser gas to experience greater heating.

Recently, this simple picture has been called into question on the basis of several statistical studies of the transmitted flux observed in individual pixels within the Lyman-α forest (Becker et al. 2007; Bolton et al. 2008; Viel et al. 2009; Lidz et al. 2010; Calura et al. 2012; Garzilli et al. 2012). These studies require comparison with large numerical simulations in order to interpret the physical implications of the detailed shape of these flux distributions. Further, the shape of the distribution may be sensitive to the full thermal history of the gas, not just the present temperature-density ($T-\rho$) relation (Peeples et al. 2010).

Becker et al. (2007) considered the transmitted flux probability distribution function (PDF) and compared it to the numerical models of Miralda-Escudé et al. (2000), finding that none of the models could explain the observations. They found that a $T-\rho$ relation in which lower density regions were *hotter* (an inverted $T-\rho$ relation) provided a better fit to the data. Since this study, several other authors have obtained similar results using transmitted flux PDF analysis (Bolton et al. 2008; Viel et al. 2009; Calura et al. 2012), wavelet analysis (Lidz et al. 2010), or a combination of both (Garzilli et al. 2012).

In contrast, early work on the $T-\rho$ relation and its evolution with redshift was performed via Voigt profile fitting of individual H I absorption lines. These early studies found a monotonic $T-\rho$ relation with a positive power-law index (Schaye et al. 2000; Ricotti et al. 2000; Bryan & Machacek 2000; McDonald et al. 2001).

In this letter, we return to the more direct test of the $T-\rho$ relation in the IGM using Voigt profile fits to individual H I absorbers. This method relies on the expected relationship between the column density of neutral hydrogen of an absorber, $N_{\rm HI}$, and its local overdensity, $\rho/\bar{\rho}$ (Davé et al. 1999; Schaye 2001). If such a relationship exists, then one expects to observe a correlation between $N_{\rm HI}$ and the velocity widths of individual thermally broadened absorbers, $b_{\rm d}$. Absorbers having only thermal broadening[1] are expected to have the smallest $b_{\rm d}$ at a given $N_{\rm HI}$, $b_{\rm min}$. Thus, by observing the behavior of the low-$b_{\rm d}$ "edge" of the distribution of $b_{\rm d}$ versus $N_{\rm HI}$, it is possible to infer the relationship between T and $N_{\rm HI}$ and through theory, the IGM $T-\rho$ relationship.

Here, we present analysis of individual absorbers fitted in 15 high-resolution, high-S/N spectra of lu-

[1]Other sources of broadening of absorbers include bulk motions of the gas (generally parameterized as a turbulent component of $b_{\rm d}$) and differential Hubble flow which broadens the absorption lines originating from the most diffuse and physically extended absorbers.

minous QSOs at $2.5 < z < 2.9$ from the Keck Baryonic Structure Survey (KBSS; Rudie et al. 2012b). We discuss the data set used in this study as well as the line-fitting procedure in Section 3.2. In Section 3.3, the fit to the minimum $b_{\rm d}$ ($b_{\rm min}$) as a function of $N_{\rm HI}$ is presented, followed by Section 3.4 in which the physical implications of the results in the context of the $T-\rho$ relation are discussed. The results and their implications are summarized in Section 3.5.

Throughout this paper we assume a Λ-CDM cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.3$, and $\Omega_\Lambda = 0.7$.

3.2 Data and Analysis

The data analyzed in this paper are taken from the Keck Baryonic Structure Survey (KBSS), a large spectroscopic survey designed to study the gaseous surroundings of high-z star-forming galaxies by combining redshift surveys for galaxies at $2 < z < 3$ in the fields of bright QSOs with absorption line data from those same QSOs. The data include high-resolution, high signal-to-noise echelle spectra of 15 luminous ($m_V \simeq 15.5-17$) QSOs located near the centers of the KBSS fields. The KBSS will be described fully by Steidel et al. (in prep); however, the data of relevance to this paper as well as the analysis of the absorption-line sample have been presented in Rudie et al. (2012b). Here, we give a brief summary.

The QSO spectra were obtained with the High Resolution Echelle Spectrometer (HIRES; Vogt et al. 1994) on the Keck I telescope. The HIRES spectra have $R \simeq 45,000$ (FWHM$\simeq 7$ km s^{-1}) and S/N $\sim 50-200$ per pixel. They cover the wavelength range 3100 – 6000 Å with no spectral gaps, allowing for the observation of Lyβ λ1025.7 down to at least $z = 2.2$ in all 15 of our sightlines. The additional constraints provided by Lyβ (and in many cases, additional Lyman series transitions) allow for highly accurate measurements of column densities and (most importantly for this paper) line widths for the H I absorbers in our sample.

The reduction of the QSO spectra and a detailed description of the process of fitting the forest are discussed in Rudie et al. (2012b). The final Voigt profile fits to the full Lyα and Lyβ forests of the 15 KBSS QSOs were completed using VPFIT[2] written by R.F. Carswell and J.K. Webb. The redshift range included in the fit is given in Table 3.1. In this paper, we use the path-length-weighted mean redshift of the sample $\langle z \rangle = 2.37$ as the fiducial redshift (Bahcall & Peebles 1969).[3]

The final H I absorber catalog includes 5758 absorbers with $12.0 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ and $2.02 < z < 2.84$ over a total redshift path length of $\Delta z = 8.27$. This H I sample is the largest ever compiled at these redshifts and is more than an order of magnitude larger than previous samples that included constraints from higher-order Lyman lines.

[2]http://www.ast.cam.ac.uk/∼rfc/vpfit.html; © 2007 R.F. Carswell, J.K. Webb

[3]The comoving pathlength, dX, was defined by Bahcall & Peebles (1969) such that absorbers with constant physical size and comoving number density would have a constant number per dX.

Table 3.1. KBSS Absorption Line Sample

Name	$z_{\rm QSO}$[a]	$z_{{\rm Ly}\alpha}$ range	S/N[b] Lyα	S/N[b] Lyβ
Q0100+130 (PHL957)	2.721	2.0617– 2.6838	77	50
HS0105+1619	2.652	2.1561– 2.6153	127	89
Q0142−09 (UM673a)	2.743	2.0260– 2.7060	71	45
Q0207−003 (UM402)	2.872	2.1532– 2.8339	82	55
Q0449−1645	2.684	2.0792– 2.6470	73	41
Q0821+3107	2.616	2.1650– 2.5794	50	33
Q1009+29 (CSO 38)	2.652	2.1132– 2.6031	99	58
SBS1217+499	2.704	2.0273– 2.6669	68	38
HS1442+2931	2.660	2.0798– 2.6237	99	47
HS1549+1919	2.843	2.0926– 2.8048	173	74
HS1603+3820	2.551	2.1087– 2.5066	108	58
Q1623+268 (KP77)	2.535	2.0544– 2.4999	48	28
HS1700+64	2.751	2.0668– 2.7138	98	42
Q2206−199	2.573	2.0133– 2.5373	88	46
Q2343+125	2.573	2.0884– 2.5373	71	45

[a]The redshift of the QSO (Trainor & Steidel 2012)

[b]The average signal to noise ratio per pixel of the QSO spectrum in the wavelength range pertaining to Lyα and Lyβ absorption.

3.3 The Temperature-Density Relation in the IGM

The equation of state of the IGM is expected to have the form

$$T = T_0 \left(\frac{\rho}{\bar{\rho}}\right)^{\gamma-1}, \tag{3.1}$$

where T_0 is the temperature at the mean mass density ($\overline{\rho}$) (Hui & Gnedin 1997).

Under the assumption of a relatively uniform ionizing radiation field, a power-law relationship between $N_{\rm HI}$ and ρ is also expected. Schaye (2001) presented a model for the low-density IGM in which "clouds" are in local hydrostatic equilibrium and therefore typically have sizes comparable to the local Jeans length. Employing this assumption, Schaye (2001) derived a relationship between $N_{\rm HI}$ and local overdensity, $\rho/\overline{\rho}$. Using updated cosmology and evaluating at the path-length-weighted mean redshift of the sample:

$$\rho_b/\overline{\rho}_b \approx \left(\frac{N_{\rm HI}}{10^{13.6}}\right)^{2/3} T_4^{0.17} \left(\frac{\Gamma_{12}}{0.5}\right)^{2/3} \left(\frac{1+z}{3.4}\right)^{-3}, \tag{3.2}$$

where T_4 is the gas temperature in units of 10^4 K and Γ_{12} is the hydrogen photoionization rate in units of 10^{-12} s^{-1} with the normalization suggested by Faucher-Giguère et al. (2008). Assuming this scaling[4],

[4]We note here that with $T_4 = 2$ as we find in this paper, the $N_{\rm HI}$ corresponding to $\bar{\rho}$ is $\log(N_{\rm HI}/{\rm cm}^{-2}) = 13.7$, a change of 0.1 dex.

absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \approx 13.6$ are expected to trace gas at the mean density of the universe at $z = 2.4$.

Thermally broadened absorbers are also expected to follow a power law relation between $b_{\rm d}$ and temperature: $b_{\rm d} \propto T^{1/2}$. Combining with the expected $T-\rho$ relation and the conversion between $N_{\rm HI}$ and ρ, the relationship between $b_{\rm min}$ and $N_{\rm HI}$ would be a power law of the form:

$$b_{\rm min} = b_0 \left(\frac{N_{\rm HI}}{N_{\rm HI,0}} \right)^{\Gamma-1} \tag{3.3}$$

where b_0 is the minimum line width of absorbers with $N_{\rm HI} = N_{\rm HI,0}$. With this formalism, $(\gamma-1)$ is proportional to $(\Gamma-1)$ (see, e.g., Schaye et al. 1999).

More explicitly, for pure thermal broadening:

$$b = (2k_B T/m_{\rm p})^{1/2} \tag{3.4}$$

where k_B is the Boltzmann constant and $m_{\rm p}$ is the mass of the proton. This suggests:

$$\log(T) = C + 2\log\left(\frac{b}{\rm km\ s^{-1}} \right) \tag{3.5}$$

where

$$C = \log\left(\frac{m_{\rm p}}{2k_B} \frac{({\rm km\ s^{-1}})^2}{\rm K} \right) = 1.78. \tag{3.6}$$

Rearranging the above equations, we expect the conversion between the index of $b_{\rm min}(N_{\rm HI})$ and the $T-\rho$ relationship to be roughly:[5]

$$\gamma - 1 \approx 3(\Gamma - 1). \tag{3.7}$$

In the sections that follow, we fit to the trend of the $b_{\rm min}$ as a function of $N_{\rm HI}$ for the data sample. We then use equations 3.5 and 3.7 to estimate the $T-\rho$ relation in the $\langle z \rangle = 2.4$ IGM.

3.3.1 A "Normal" $T-\rho$ Relation

The measured values of $b_{\rm d}$ and $N_{\rm HI}$ considered in this letter are presented in Figure 3.1. Examination of Figure 3.1 already suggests the main result of this paper; focusing on the low-$b_{\rm d}$ edge of the distribution, one can easily see that the lower envelope increases monotonically with $N_{\rm HI}$. Such behavior is expected for a *normal* temperature-density relationship in which denser regions are hotter.

3.3.2 Outlier Rejection

Before fitting to the "ridge-line" of $b_{\rm min}$ versus $N_{\rm HI}$, we exclude those absorbers that have large errors in their measured parameters. Schaye et al. (1999) suggest a simple algorithm for the outlier rejection in which

However, the uncertainty in other parameters (e.g. Γ_{12}) is large enough, that we do not consider this small effect in this paper.

[5] In the following equation, we have neglected the minor dependence of $\rho_b/\overline{\rho}_b$ on T from equation 3.2.

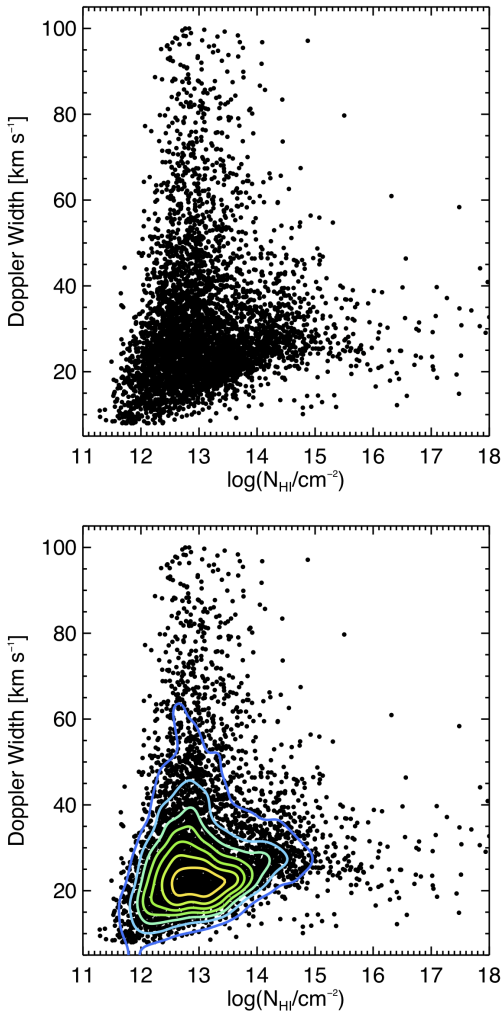


Figure 3.1 The Doppler widths of absorbers ($b_{\rm d}$) versus their column density ($N_{\rm HI}$) for all the absorbers in the HI sample with relative errors in $N_{\rm HI}$ and $b_{\rm d}$ less than 50%. Point-density contours are over plotted to guide the eye in the bottom panel. Note that the minimum value of $b_{\rm d}$ at each $N_{\rm HI}$ increases with increasing $N_{\rm HI}$ suggesting a normal $T-\rho$ relation.

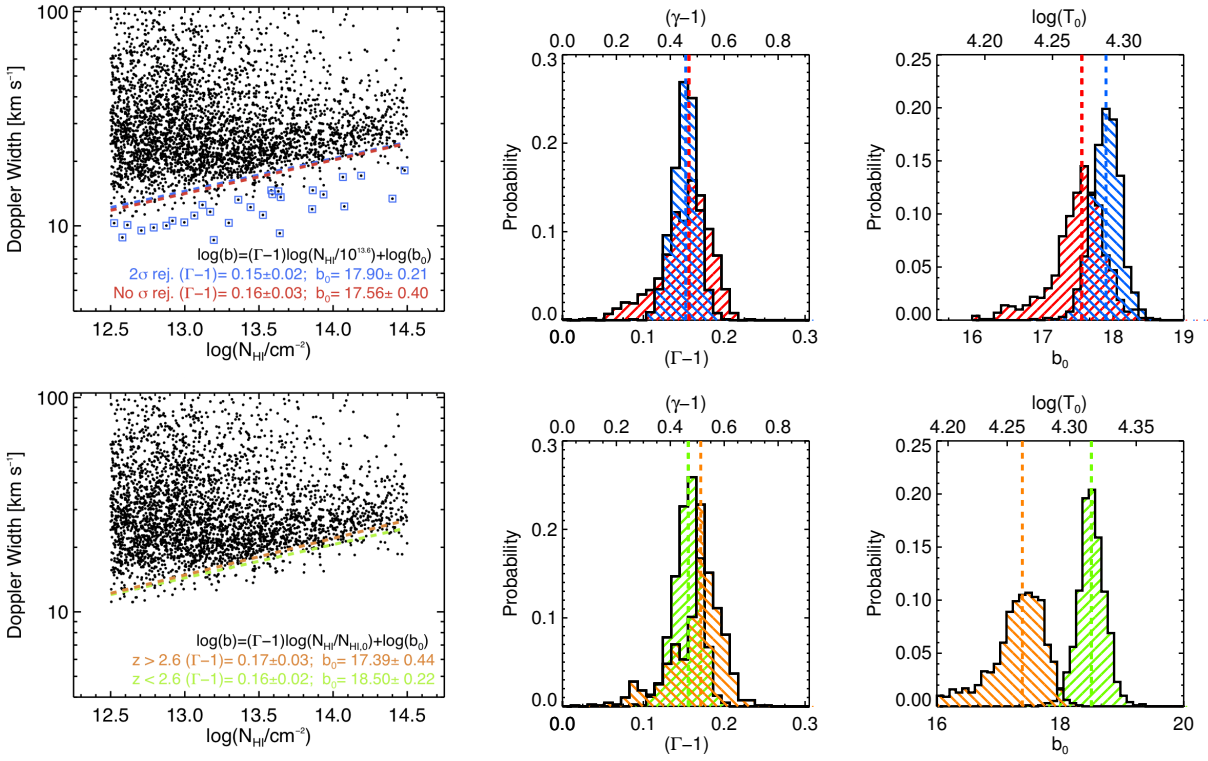


Figure 3.2 *Left panel:* $b_{\rm d}$ versus $N_{\rm HI}$ for all absorbers in the KBSS HI sample with $12.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 14.5$ and relative errors in $b_{\rm d}$ and $N_{\rm HI}$ less than 50%(black points). The curves show the resulting best power-law fits determined by applying the bootstrap resampling method to the sample followed by the iterative power-law fit method described by Schaye et al. (1999). *Top panels:* The blue and red lines pertain to the data set with the σ-rejected absorbers (blue boxes) excluded or included respectively. *Top Center and Right:* The probability distribution for the fitted parameters using the 50% relative-error rejection algorithm (red) and both the 50% relative-error rejection algorithm and the 2σ rejection algorithm (blue). *Bottom panels:* The orange and green curves pertain to the $z > 2.6$ and $z < 2.6$ subsamples of the σ-rejection data set. Here $N_{\rm HI,0}$ varies between the subsamples with $\log(N_{\rm HI,0}/{\rm cm}^{-2}) = 13.7$ for $z < 2.6$ and $\log(N_{\rm HI,0}/{\rm cm}^{-2}) = 13.4$ for $z > 2.6$ due to the redshift dependence in equation 3.2. Note that the $b_{\rm min}$-$N_{\rm HI}$ relation is roughly constant, but the expected $N_{\rm HI,0}$ evolution would require evolution in b_0 and T_0.

absorbers with relative error in $b_{\rm d}$ or $N_{\rm HI}$ larger than 50% are excluded.[6] This method is expected to primarily reject lines which originate in blends and are thus unlikely to have accurate line-width measurements. Additionally, we exclude those absorbers with $b_{\rm d}$= 8 km s^{-1}and $b_{\rm d}$= 100 km s^{-1}. These line widths correspond to the allowed line-width limits input to VPFIT and as such are artificial. The absorbers which remain after those exclusions are shown in Figure 3.1.

Considering Figure 3.1, we note a small set of points at very low $b_{\rm d}$ which seem to lie significantly below the main locus of points. These absorbers are likely to be unidentified metal line contaminants, and thus we consider a σ-rejection algorithm to exclude them.

The σ-rejection algorithm is applied as follows. Only absorbers with $b_{\rm d} < 40$ km s^{-1} are considered. The data are sorted by their $N_{\rm HI}$. The absorbers are placed into $N_{\rm HI}$ bins of width 0.25 dex and the mean and

[6]The relative error considered is computed from the formal error bars output by VPFIT for both $b_{\rm d}$ and $N_{\rm HI}$.

Table 3.2. Fits to the $b_{\rm min} - N_{\rm HI}$ and $T - \rho$ Relation in IGM

Outlier Rej.	z range	$\langle z \rangle$	$\log(N_{\rm HI,0}/{\rm cm}^{-2})$	$(\Gamma - 1)$	$(\gamma - 1)^{\rm a}$	b_0	$T_0/10^4$ K$^{\rm b}$
default	2.0 – 2.8	2.4	13.6	0.156 ± 0.032	0.47 ± 0.10	17.56 ± 0.40	1.87 ± 0.08
2σ-rej.	2.0 – 2.8	2.4	13.6	0.152 ± 0.015	0.46 ± 0.05	17.90 ± 0.21	1.94 ± 0.05
2σ-rej.	2.0 – 2.6	2.3	13.7	0.156 ± 0.016	0.47 ± 0.05	18.50 ± 0.22	2.07 ± 0.05
2σ-rej.	2.6 – 2.8	2.7	13.4	0.171 ± 0.032	0.51 ± 0.10	17.39 ± 0.44	1.83 ± 0.09

[a] $(\gamma - 1)$ is calculated from the measured value of $(\Gamma - 1)$ using equation 3.7.

[b] T_0 is calculated from the measured value of b_0 using equation 3.5.

standard deviation of $b_{\rm d}$ are computed for the bin. Absorbers with $b_{\rm d}$ 2σ or larger from the mean are flagged and excluded. This process is iterated; the set of absorbers excluded via this method with $b_{\rm d}$ lower than the mean are surrounded by (blue) boxes in Figure 3.2 and are excluded from the fitting procedure. We have checked the majority of σ-rejected absorbers by hand and find that they are either previously unidentified metal lines or parts of blended saturated H I systems whose individual properties are poorly constrained.

In the following section, we proceed with the measurement of the $b_{\rm min}- N_{\rm HI}$ relation. We show that the application of the σ-rejection algorithm does not produce significant changes to the results. In the conclusions of the paper, however, we prefer the combined error and σ-rejection method as it excludes the majority of the poorly measured absorbers as well as those which lie far below the $b_{\rm min} - N_{\rm HI}$ trend.

3.3.3 Fitting $b_{\rm min}$ versus $N_{\rm HI}$

In this section we measure the $b_{\rm min}$ – $N_{\rm HI}$ relation using the two data sets resulting from the outlier rejection methods described above. To fit the trend between $b_{\rm min}$ and $N_{\rm HI}$, we follow the iterative power-law method proposed by Schaye et al. (1999). A power-law relationship of the form shown in equation 3.3 with $\log(N_{\rm HI,0}/{\rm cm}^{-2})$= 13.6 is fit to the data. Data points which have $b_{\rm d}$ larger than one mean absolute deviation above the fit are discarded. The power law is refit to the remaining absorbers and the rejection and refitting are repeated until the power law converges. Then, absorbers more than one mean absolute deviation below the fit are removed *once*, and the power law is refit. This fit is taken as the final form of the minimum $b_{\rm min}$ – $N_{\rm HI}$ relation. Errors in the parameters of the fit are derived by applying the bootstrap resampling method to the data sample and following the above outlined procedure.

Figure 3.2 shows the results of the iterative power-law fit for both data sets formed via the outlier rejection algorithms. The red curves correspond to the fit including the σ-rejected points, while the blue curves correspond to those without. The center and right panels show the distribution of fit parameters obtained through bootstrap resampling. As expected, those fits which exclude the low-$b_{\rm d}$ outliers via σ-rejection (blue histograms) result in tighter distributions for the fitted parameters; however, the best fit values are nearly indistinguishable.

The best fit parameters b_0 and $(\Gamma - 1)$ and their relative errors, as well as their physical counterparts

Table 3.3. Comparison with Previous Measurements of the $T-\rho$ Relationship

Source	Method	$\langle z \rangle$	$T_0/10^4$K	$(\gamma-1)$	Comment
Schaye et al. (2000)	Line Fitting	2.3	1.17	0.27	Their Figure 6
	Line Fitting	2.5	1.38	0.56	
Ricotti et al. (2000)	Line Fitting	2.75	2.52	0.22 ± 0.10	Their Table 6
	Line Fitting	1.90	1.77	0.32 ± 0.30	
Bryan & Machacek (2000)[b]	Line fitting	2.7	1.65[a]	0.29	
McDonald et al. (2001)	Line Fitting	2.41	1.74 ± 0.19	0.52 ± 0.14	Their Table 5, $z_{\rm sim}$ = 3
Zaldarriaga et al. (2001)	Transmission Power Spectra	2.4	$1-3$	0 to 0.6	Their Figure 3
Becker et al. (2007)	Flux PDF	$2-6$	$\cdots$	≈ -0.5	
Bolton et al. (2008)[c]	Flux PDF	2.07	$\cdots$	–0.33	
	Flux PDF	2.52	$\cdots$	–0.46	
	Flux PDF	2.94	$\cdots$	–0.56	
Viel et al. (2009)	Flux PDF	3	1.9 ± 0.6	-0.25 ± 0.21	
Lidz et al. (2010)	Wavelet Analysis	2.6	$1.6 \pm .6$[d]	$\cdots$	Their Figure 30
	Wavelet Analysis	2.2	$2.1 \pm .7$[d]	$\cdots$	
Becker et al. (2011b)	Curvature Analysis	2.4	1.11 ± 0.06[e]	$\cdots$	Their Table 3
Garzilli et al. (2012)[f]	Wavelet Analysis + Flux PDF	2.1	1.7 ± 0.2	0.11 ± 0.11	Their Table 2
	Wavelet Analysis	2.1	1.6 ± 0.5	> -0.14	
	Flux PDF	2.1	1.5 ± 0.3	-0.01 ± 0.14	
	Wavelet Analysis + Flux PDF	2.5	1.3 ± 0.4	> -0.05	
	Wavelet Analysis	2.5	1.6 ± 0.4	> -0.08	
	Flux PDF	2.5	1.4 ± 0.9	> -0.31	
Calura et al. (2012)	Flux PDF	3	1.93 ± 0.48	-0.10 ± 0.21	Their Table 5
	Flux PDF[g]	3	1.79 ± 0.35	-0.30 ± 0.12	
This Work	Line fitting	2.37	1.94 ± 0.05	0.46 ± 0.05	Results from data set using both relative error and σ-rejection

[a]The authors suggest this is an upper limit.

[b]Based on the spectrum of HS 1946+7658

[c]Measured by Kim et al. (2007), sample includes 18 QSOs (metals removed by hand to eliminate noise in flux PDF.)

[d]Marginalized over $(\gamma-1) = 0-0.6$

[e]Assuming $\gamma - 1 = 0.56$

[f]Find that flux PDF results in lower values of γ than wavelet analysis in general.

[g]Joint analysis with Kim et al. (2007) .

T_0 and $(\gamma - 1)$, are presented in Table 3.2. T_0 and $(\gamma - 1)$ are calculated assuming equations 3.5 and 3.7 respectively. Notably, both outlier rejection algorithms produce statistically consistent results suggesting that the measurements are robust to changes in the rejection algorithm; however, (as expected) slightly higher b_0 are preferred by the algorithms with σ-rejection.

Henceforth, we consider the results from the data set formed with both the error rejection and σ-rejection algorithms applied prior to the power-law fit to $b_{\rm min}(N_{\rm HI})$.

3.4 Results

As discussed above, the mapping of $b_{\rm min}(N_{\rm HI})$ to the equation of state for intergalactic gas requires a theoretical framework. Schaye et al. (2000) considered simulations of the IGM and found that the minimum values of $b_{\rm d}$ at each $N_{\rm HI}$ were larger than expected given purely thermal broadening. They suggested that all absorbers in the $N_{\rm HI}$ range used to measure the $T - \rho$ relationship experienced mild Hubble broadening (also expected to scale as $T^{1/2}$ for clouds having sizes similar to the local Jeans scale.) In this case, estimates made using the above equations would mildly over-predict the temperature at mean density. We prefer to report physical values tied closely to our measurements rather than rely directly on a specific set of simulations; however, our measurements of b_0 and $\Gamma - 1$ can be converted to physical parameters using a different set of assumptions. Our simplifying assumption of pure thermal line broadening for absorbers near $b_{\rm min}(N_{\rm HI})$ results in T_0 values slightly higher than some previous studies which used simulations as a reference (see Table 3.3).

In Table 3.3, we compare our results with those of previous studies. We find generally good agreement between our results and those of past studies which used line fitting. We further emphasize that our results differ significantly from those obtained using the transmitted flux PDF and similar statistical methods, calling into question the interpretation of such techniques.

3.4.1 He II Reionization

One of the expected results of He II reionization, believed to occur at $z \approx 3$, is a significant change in the $T - \rho$ relation due to excess photoionization heating (Hui & Gnedin 1997). These changes are not long lived, and the exact effects depend on the speed and patchiness with which reionization proceeds and the spectral hardness of the ionizing sources (see McQuinn et al. 2009). While detailed discussion is beyond the scope of this letter, we briefly mention that the absorbers in the KBSS H I sample show no strong evidence for a change in the slope of $b_{\rm min}(N_{\rm HI})$ for higher-redshift absorbers, which (presumably) are temporally closer to the He II reionization epoch. The bottom panels of Figure 3.2 show the iterative power law fit to absorbers with, respectively, $z < 2.6$ (green) and $z > 2.6$ (orange) from the sample with σ-rejection. Notably, the $b_{\rm min} - N_{\rm HI}$ relation appears to be independent of redshift (left panel); however, the expected evolution in the value of $N_{\rm HI}$ at the mean density (equation 3.2) results in differing values of $N_{\rm HI,0}$ for the two redshift bins. This in turn leads to expected variation in the values of b_0 and T_0 with z. The implied parameters from the fit are listed in

Table 3.2.

3.5 Conclusions

We have inferred the $T-\rho$ relationship in the $\langle z\rangle = 2.4$ IGM using Voigt profile fitting of individual Lyman lines. Fitting the trend of the minimum line width ($b_{\rm min}$) as a function of column density ($N_{\rm HI}$) with a power law, we find best fit values of $b_{\rm min}$ at mean density (which at $\langle z\rangle = 2.4$ corresponds to $N_{\rm HI,0} = 10^{13.6}$ cm^{-2}) $b_0 = 17.9 \pm 0.2$ km s^{-1} and a power-law index $(\Gamma - 1) = 0.15 \pm 0.02$. Assuming a monotonic relation between $N_{\rm HI}$ and $\rho/\bar{\rho}$, these data support the conclusion that the temperature-density ($T-\rho$) relationship in the IGM is not inverted at $\langle z\rangle = 2.4$ but instead has a power-law index $(\gamma - 1) = 0.46 \pm 0.05$. Further, our results suggest a temperature at mean density of $T_0 = [1.94 \pm 0.05] \times 10^4$K. Within our sample spanning the redshift range $2.0 \lesssim z \lesssim 2.8$, there is no evidence for significant evolution in the $b_{\rm min} - N_{\rm HI}$ relation or in the power-law index $(\gamma - 1)$ of the $T-\rho$ relation.

Acknowledgements

We thank George Becker for his early interest in the $T-\rho$ relation results from the KBSS sample, Joop Schaye who provided helpful advice and insightful comments, and Allison Strom and Ryan Trainor for their careful reading of early drafts. Thanks also to Olivera Rakic for her contributions to the reduction of the QSO data set and for her pertinent advice. The authors wish to acknowledge Ryan Cooke who contributed the fits to the damped profiles in our QSO spectra.

We wish to acknowledge the staff of the W.M. Keck Observatory whose efforts insure the telescopes and instruments perform reliably. We thank those of Hawaiian ancestry on whose sacred mountain we are privileged to be guests.

This work has been supported by the US National Science Foundation through grants AST-0606912 and AST- 0908805. CCS acknowledges additional support from the John D. and Catherine T. MacArthur Foundation and the Peter and Patricia Gruber Foundation. This research has made use of the Keck Observatory Archive (KOA).

Chapter 4

The Column Density Distribution and Continuum Opacity of the Intergalactic and Circumgalactic Medium at Redshift $\langle z \rangle = 2.4$

This chapter was previously published as Rudie et al. 2013, ApJ, 769, 146 with coauthors Charles C. Steidel, Alice E. Shapley, and Max Pettini.

Abstract

We present new high-precision measurements of the opacity of the intergalactic and circumgalactic medium (IGM, CGM) at $\langle z \rangle = 2.4$. Using Voigt profile fits to the full Lyman α and Lyman β forests in 15 high-resolution high-S/N spectra of hyperluminous QSOs, we make the first statistically robust measurement of the frequency of absorbers with H I column densities $14 \lesssim \log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 17.2$. We also present the first measurements of the frequency distribution of H I absorbers in the volume surrounding high-z galaxies (the CGM, 300 pkpc), finding that the incidence of absorbers in the CGM is much higher than in the IGM. In agreement with Rudie et al. (2012b), we find that there are fractionally more high-$N_{\rm HI}$ absorbers than low-$N_{\rm HI}$ absorbers in the CGM compared to the IGM, leading to a shallower power law fit to the CGM frequency distribution. We use these new measurements to calculate the total opacity of the IGM and CGM to hydrogen-ionizing photons, finding significantly higher opacity than most previous studies, especially from absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$. Reproducing the opacity measured in our data as well as the incidence of absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$ requires a broken power law parameterization of the frequency distribution with a break near $N_{\rm HI} \approx 10^{15}$ cm^{-2}. We compute new estimates of the mean free path ($\lambda_{\rm mfp}$) to hydrogen-ionizing photons at $z_{\rm em} = 2.4$, finding $\lambda_{\rm mfp} = 147 \pm 15$ Mpc when considering only IGM opacity. If instead, we consider photons emanating from a high-z star-forming galaxy and account for the local excess opacity due to the surrounding CGM of the galaxy itself, the mean free path is reduced to $\lambda_{\rm mfp} = 121 \pm 15$ Mpc. These $\lambda_{\rm mfp}$ measurements are smaller than recent estimates and should inform future studies of the metagalactic UV background and of ionizing sources at $z \approx 2-3$.

4.1 Introduction

Observations of the Lyman α forest imprinted upon the spectra of background QSOs provide a rare window into the nature and evolution of baryonic structures in the universe. With the advent of the theory of cold dark matter (CDM), it was recognized that self-gravitating halos could be the sites of uncollapsed Lyα clouds (Umemura & Ikeuchi 1985; Rees 1986). In the modern version of this model, the Lyα forest simply represents the array of density fluctuations present in the early universe and evolved under gravity to the present day (Tytler 1987; Cen et al. 1994; Hernquist et al. 1996), modulo the effect of the formation of stars, black holes, and galaxies which appear to have polluted a large fraction of the cosmic hydrogen with metals, even at early times (Songaila & Cowie 1996; Davé et al. 1998; Schaye et al. 2003; Simcoe et al. 2004; Ryan-Weber et al. 2009; Becker et al. 2009; Simcoe 2011; Simcoe et al. 2011; Becker et al. 2011c, 2012) and led to the ionization of the vast majority of the gas (Gunn & Peterson 1965).

There is an extensive literature concerned with the measurement of cosmological parameters using observations of the Lyα forest [for a review, see Rauch (1998) and Meiksin (2009)]. Considerable progress has been made in the last several decades in measuring the detailed properties of the intergalactic medium (IGM) and their implications for cosmology, structure formation, reionization, and the metagalactic background.

One of the most fundamental of these measurements is the number of absorbers as a function of column density ($N_{\rm HI}$) and redshift that can be seen along a given path through the Universe. This measurement has required significant observational effort, in a large part due to the very large dynamic range of the measurements ($10^{12} \lesssim N_{\rm HI} \lesssim 10^{22}$ cm^{-2}) which have necessitated surveys optimized for particular ranges in $N_{\rm HI}$. This requirement is easily understood by considering the curve of growth of the Lyα line of neutral hydrogen.

The most common absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 14$ are optically thin in the Lyman α transition and are thus on the linear part of the curve of growth. As such, their properties can be measured with a few high-resolution ($\sim$10 km s^{-1}) high-signal to noise ratio (S/N $\sim$ 50) spectra (Kim et al. 2002). At the highest column densities, $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 20.3$, the Lorentzian wings of the absorption line in Lyα become obvious against the continuum of the background source. Such rare pockets of predominantly neutral hydrogen, commonly referred to as Damped Lyman α absorbers (DLAs), can be identified and measured to high statistical precision with a large number of low to moderate resolution spectra of background QSOs as has been done using (e.g.) the QSO sample within the Sloan Digital Sky Survey (SDSS; Noterdaeme et al. 2009).

Absorbers with intermediate $N_{\rm HI}$, $14 \lesssim \log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 20$, lie on the flat part of the curve of growth thus requiring different techniques. Systems with $N_{\rm HI} > 10^{17.2}$ cm^{-2} have an optical depth to hydrogen-ionizing photons $\tau \geq 1$, and so may be recognized by the strong breaks they produce in background QSO spectra at the Lyman limit (912Å) in their rest frame. These so-called Lyman Limit Systems (LLS) can thus be easily discovered using low resolution spectra; however, a large sample is needed because they are comparatively rare. For studies at $z \lesssim 2.6$, surveys of LLSs require space-based observations as the Lyman limit shifts below the atmospheric cutoff at UV wavelengths. However, over the last 30 years with a combination of ground

and space based observations, the distribution of LLSs at $2 < z < 3$ has been well characterized (Tytler 1982; Sargent et al. 1989; Stengler-Larrea et al. 1995; Ribaudo et al. 2011; O'Meara et al. 2012).

Finally there is the historically most-problematic range, $14 \lesssim \log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 17.2$. These intermediate-$N_{\rm HI}$ absorbers are saturated in Lyα but have relatively low opacity to hydrogen-ionizing photons and therefore weak "Lyman breaks." For this reason, observational constraints on their statistical incidence have been approximate at best.

In this paper, we address these intermediate-$N_{\rm HI}$ absorbers using Voigt profile fits to all-available higher-order Lyman series transitions. We take advantage of the increased dynamic range of the measurements afforded by the decreased oscillator strength of the Lyβ, γ, etc. transitions to accurately measure the frequency distribution of these absorbers, $f(N,X)$, for the first time in a statistically robust sample. In our derivation of an analytic representation of $f(N,X)$ we place emphasis on finding a method which best reproduces the measured opacity of absorbers to hydrogen ionizing sources.

These new measurements are crucial to our understanding of the opacity of the universe to hydrogen-ionizing photons. While individually absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ have relatively low optical depths at their Lyman limit, they are much more common than LLSs, and therefore their ensemble contributes nearly half of the IGM's opacity. Here, we present the most precise measurements of the continuum opacity of absorbers that are optically thin at the Lyman limit.

In this paper, we also attempt to account for the fact that galaxies are located in regions of relatively high gas density. Rudie et al. (2012b) clearly demonstrated that typical galaxies at $z \approx 2.3$ have large quantities of H I surrounding them to 300 physical kpc (pkpc), and that there is excess H I compared to random places in the IGM to $\sim$2 physical Mpc (pMpc). In the case that galaxies such as those studied in Rudie et al. (2012b) are significant contributors to the ionizing background, the effect of their surrounding circumgalactic gas distribution is an important consideration. In this work, we attempt to separate the probability that an ionizing photon will escape the ISM of its galaxy from the probability that such a photon, after escaping to 50 pkpc, will also escape the enhanced opacity found in its circumgalactic medium (CGM).

We begin in Section 4.2 with a description of the unique data that enable these very precise $N_{\rm HI}$ measurements. In Section 4.3 we present the new $N_{\rm HI}$ frequency distribution measurements. Section 4.4 considers the frequency distribution within the CGM of galaxies and Section 4.5 quantifies the total opacity of the IGM and the environment of the source galaxy. Updated measurements of the mean free path ($\lambda_{\rm mfp}$) to hydrogen-ionizing photons are presented in Section 4.6 along with their implications. The paper is summarized in Section 4.7.

Throughout this paper we assume a Λ-CDM cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.3$, and $\Omega_\Lambda = 0.7$. All distances are expressed in physical (proper) units unless stated otherwise. We use the abbreviations pkpc and pMpc to indicate physical units. At the path-weighted mean redshift of the absorber sample ($\langle z \rangle = 2.37$), the age of the universe is 2.7 Gyr, the look-back time is 10.7 Gyr, and 8.2 pkpc subtends one arcsecond on the sky.

4.2 Data and Analysis

The data presented in this paper constitute a subset of the Keck Baryonic Structure Survey (KBSS). The KBSS was designed to detect and characterize the gaseous distribution surrounding star-forming galaxies at $2 < z < 3$ during the peak of cosmic star formation and black hole growth. The data include rest-frame UV spectra of 2188 star-forming galaxies located in 15 fields surrounding the lines of sight to hyper-luminous ($m_V \simeq 15.5-17$) QSOs. The present work focuses primarily on the H I absorption line statistics resulting from analysis of high-resolution, high signal-to-noise echelle spectra of the 15 KBSS QSOs. A full description of the KBSS will be presented by Steidel et al. (in prep), but the portions of the data most relevant to this work (as well as a detailed description of the analysis of the QSO data) can be found in Rudie et al. (2012b). Here, we provide a brief description.

The 15 KBSS QSOs were observed primarily with the High Resolution Echelle Spectrometer (HIRES; Vogt et al. 1994) on the Keck I telescope. The HIRES spectra have $R \simeq 45,000$ (FWHM$\simeq$ 7 km s^{-1}), S/N $\sim$ 50$-$200 per pixel, and cover at least the wavelength range 3100 – 6000 Å with no spectral gaps. The UV/blue wavelength coverage provides a significant advantage over other data sets as it enables the observation of Lyβ λ1025.7 down to at least $z = 2.2$ in all 15 of our sightlines, and to significantly lower redshift in many.[1] The additional constraints provided by Lyβ (and in many cases, additional Lyman series transitions) allow for much more accurate measurements of H I for $N_{\rm HI}= 10^{14}-10^{17}$ cm^{-2} than can be made with spectra including only the Lyα transition. A full discussion of the redshift sensitivity of our fitting procedure is given in Appendix A.

The reduction of the HIRES spectra was performed using T. Barlow's MAKEE package as described in detail by Rudie et al. (2012b). The analysis presented here includes a Voigt profile fit to the full Lyα and Lyβ forests in these spectra. The redshift range included in the fit is bounded at high redshift by the proximity zone of the QSO (taken to be 3000 km s^{-1}) in which absorbers may be ejected from or ionized by the QSO itself. The low-redshift cut was chosen to ensure that all H I absorbers included in the sample had been observed in both the Lyα and Lyβ transition. The final redshift windows used for each spectrum are detailed in Table 4.1.

The line fitting was facilitated by a semi-automatic code that estimates input parameters (redshift z, H I column density $N_{\rm HI}$, and Doppler parameter $b_{\rm d}$ of each absorber) using a cross-correlation technique as described in Rudie et al. (2012b). These parameters were then input into the χ^2 minimization code VPFIT[2] written by R.F. Carswell and J.K. Webb. The results were iteratively checked, altered, and re-run until a good fit was achieved. Median uncertainties in $\log(N_{\rm HI})$ for these absorber measurements are listed in Table 4.2. Examples of fits to absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 15.5$ are given in Appendix A along with an assessment of the completeness of the catalog as a function of redshift.

The result of this process is an H I absorber catalog including 5758 absorbers with $12.0 < \log(N_{\rm HI}/{\rm cm}^{-2}) <$

[1] The minimum wavelength that can be used for line fitting is determined by both the wavelength coverage of the spectrum (set by the spectrograph and the atmosphere) and by the location of the highest-redshift Lyman limit system along each line of sight. The minimum redshift for which we perform line fitting is shown in Table 4.1.

[2] http://www.ast.cam.ac.uk/$\sim$rfc/vpfit.html;

Table 4.1. KBSS Absorption Line Sample

Name	RA	Dec	z_{QSO}[a]	λ_{min}[b][Å]	z range	Δz	ΔX[c]	N_{gal}[d]	S/N Lyα[e]	S/N Lyβ[e]
Q0100+130 (PHL957)	01:03:11.3	+13:16:18.2	2.721	3133	2.0617– 2.6838	0.62	2.02	5	77	50
HS0105+1619	01:08:06.4	+16:35:50.0	2.652	3230	2.1561– 2.6153	0.46	1.50	0	127	89
Q0142−09 (UM673a)	01:45:16.6	−09:45:17.0	2.743	3097	2.0260– 2.7060	0.68	2.21	3	71	45
Q0207−003 (UM402)	02:09:50.7	−00:05:06.5	2.872	3227	2.1532– 2.8339	0.68	2.26	1	82	55
Q0449−1645	04:52:14.3	−16:40:16.2	2.684	3151	2.0792– 2.6470	0.57	1.84	2	73	41
Q0821+3107	08:21:07.6	+31:07:51.2	2.616	3239	2.1650– 2.5794	0.41	1.35	2	50	33
Q1009+29 (CSO 38)	10:11:55.6	+29:41:41.7	2.652[f]	3186	2.1132– 2.6031	0.49	1.59	0	99	58
SBS1217+499	12:19:30.9	+49:40:51.2	2.704	3098	2.0273– 2.6669	0.64	2.07	5	68	38
HS1442+2931	14:44:53.7	+29:19:05.6	2.660	3152	2.0798– 2.6237	0.54	1.76	6	99	47
HS1549+1919	15:51:52.5	+19:11:04.3	2.843	3165	2.0926– 2.8048	0.71	2.35	4	173	74
HS1603+3820	16:04:55.4	+38:12:01.8	2.551[g]	3181	2.1087– 2.5066	0.40	1.28	2	108	58
Q1623+268 (KP77)	16:25:48.8	+26:46:58.8	2.535	3126	2.0544– 2.4999	0.45	174	3	48	28
HS1700+64	17:01:00.6	+64:12:09.4	2.751	3138	2.0668– 2.7138	0.65	2.11	4	98	42
Q2206−199	22:08:52.1	−19:43:59.7	2.573	3084	2.0133– 2.5373	0.52	1.68	3	88	46
Q2343+125	23:46:28.3	+12:48:57.8	2.573	3160	2.0884– 2.5373	0.45	1.45	3	71	45

[a]The redshift of the QSO (Trainor & Steidel 2012)

[b]The minimum wavelength covered in the HIRES QSO spectrum

[c]The comoving pathlength in this line of sight (equation 4.3).

[d]The number of galaxies in our LRIS survey with impact parameters < 300 pkpc.

[e]The average signal to noise ratio per pixel of the QSO spectrum in the wavelength range pertaining to IGM Lyα and Lyβ absorption.

[f]The redshift of this QSO was revised after the fitting of the HIRES spectrum was completed. The maximum redshift assumed $z_{QSO} = 2.640$

[g]The redshift of this QSO was revised after the fitting of the HIRES spectrum was completed. The maximum redshift assumed $z_{QSO} = 2.542$

Table 4.2. Median log($N_{\rm HI}$) Uncertainty for Absorber Measurements

log($N_{\rm HI}$/cm^{-2})	$\sigma_{\log(N_{\rm HI})}$
12.0 – 12.5	0.10
12.5 – 13.0	0.07
13.0 – 13.5	0.07
13.5 – 14.0	0.06
14.0 – 14.5	0.04
14.5 – 15.0	0.03
15.0 – 15.5	0.03
15.5 – 16.0	0.04
16.0 – 16.5	0.07
16.5 – 17.0	0.06
17.0 – 17.5	0.09

17.2 and $2.02 < z < 2.84$ over a total redshift pathlength of $\Delta z = 8.27$. This represents the largest Lyα forest catalog to date at these redshifts and increases by an order of magnitude the number of absorbers measured using the additional constraints provided by higher-order Lyman series transitions.

4.3 $N_{\rm HI}$ Frequency Distribution

The differential column density distribution, first defined by Carswell et al. (1984), is a useful way of parameterizing the frequency of absorbers as a function of $N_{\rm HI}$. We consider the differential column density distribution per unit pathlength, $f(N_{\rm HI}, X)$, defined as

$$f(N_{\rm HI}, X) dN_{\rm HI} dX = \frac{m}{\Delta N_{\rm HI} \Delta X} dN_{\rm HI} dX \tag{4.1}$$

where m is the observed number of absorbers with column densities in the range $\Delta N_{\rm HI}$. ΔX is the comoving pathlength of the survey summed over the sightlines. This comoving pathlength was introduced by Bahcall & Peebles (1969) in part to understand if the statistics of absorbers were consistent with having been produced by galaxies. With this formalism, absorbers with a constant physical size and comoving number density will have constant $f(N, X)$. The comoving pathlength of a single sightline, ΔX_i, is defined as:

$$dX = \frac{H_0}{H(z)} (1+z)^2 dz \tag{4.2}$$

$$\Delta X_i = \int_{z_{min}}^{z_{max}} dX = \int_{z_{min}}^{z_{max}} \frac{(1+z)^2}{\sqrt{\Omega_\Lambda + \Omega_{\rm m}(1+z)^3}} dz. \tag{4.3}$$

The total pathlength covered by this survey is $\Delta X = 26.9$ and the ΔX covered as a function of redshift summed over the 15 lines of sight is shown in Figure 4.1. The pathlength-weighted mean redshift of the

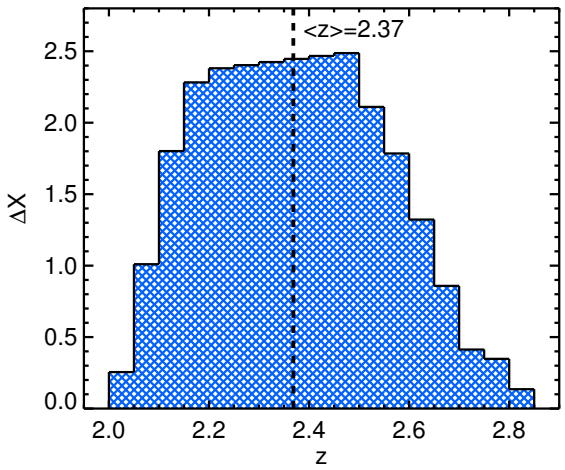


Figure 4.1 Histogram of the pathlength sampled as a function of redshift. The path-length-weighted average redshift of the sample is $z = 2.37$.

sample is $\langle z \rangle = 2.37$. Measurements of $f(N,X)$ derived from the KBSS sample of absorbers with $12.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ are presented in Figure 4.2. The data points are located at the mean $\log(N_{\rm HI})$ of the bin edges.

4.3.1 Parameterizing the Frequency Distribution

It has become standard to represent $f(N_{\rm HI},X)$ by a power law, or a series of broken power laws of the form:

$$f(N_{\rm HI},X)dN_{\rm HI}dX = C_{\rm HI}N_{\rm HI}^{-\beta}dN_{\rm HI}dX. \tag{4.4}$$

Tytler (1987) first noted that the full column density distribution from $12 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 21$ was reasonably well approximated by a single power law with $\beta = 1.5$. The detailed deviations from a single power law at various $N_{\rm HI}$ and the change in normalization and slope of $f(N_{\rm HI},X)$ as a function of redshift have been considered by numerous authors (e.g., Sargent et al. 1989; Petitjean et al. 1993; Davé & Tripp 2001; Kim et al. 2002; Penton et al. 2004; Janknecht et al. 2006; Lehner et al. 2007; Prochaska et al. 2010; Ribaudo et al. 2011; O'Meara et al. 2012).

The most notable deviations from a single power law appear at the ends of the distribution. The low-$N_{\rm HI}$ end appears to flatten, an effect generally attributed to a combination of incompleteness due to line blending and blanketing and a true physical turnover (Hu et al. 1995; Ellison et al. 1999; Schaye 2001; Kim et al. 2002). The deviation from a single power-law for $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 19$ is believed to be due to the transition of gas from optically thin to self-shielding regimes (Murakami & Ikeuchi 1990; Petitjean et al. 1992; Katz et al. 1996; Corbelli et al. 2001; Zheng & Miralda-Escudé 2002; Altay et al. 2011; McQuinn et al. 2011). The most recent estimates of the frequency distribution of absorbers with $19 < \log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 20.3$ at $z \approx 2$ are by O'Meara et al. (2007); the frequency distribution of damped Lyman α systems with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 20.3$

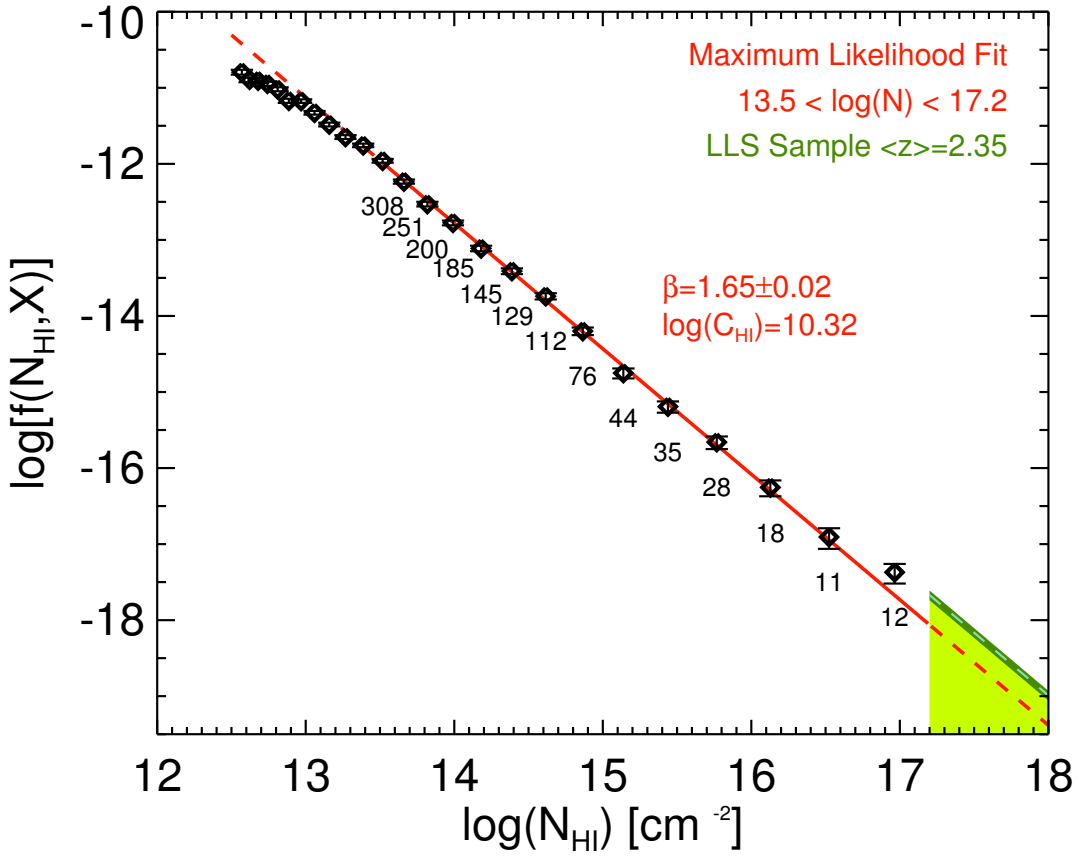


Figure 4.2 The frequency distribution of absorbers from the KBSS H I sample (black data points). For bins with $N_{\rm HI} > 13.5$ cm^{-2}, the number of absorbers per bin is displayed. The (red) line is a maximum likelihood power-law fit to the distribution of absorbers with $13.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$. The (green) triangle represents the frequency of LLS calculated from other studies as discussed in §4.3.3. The dark green band at the top of the triangle represents the error in this estimation. Note this is a single power law fit to the data. Section 4.5.1 discusses broken-power law fits which better reproduce some aspects of the data, especially the incidence of LLSs.

is measured to relatively high precision using the SDSS QSOs by Noterdaeme et al. (2009). In this paper, we leverage the KBSS QSO absorption line data set to measure $f(N,X)$ for absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$.

4.3.2 Maximum Likelihood Method

A statistically rigorous estimate of the power law exponent and normalization can be found via the Maximum Likelihood Method. The maximum likelihood estimator (MLE) for the power law exponent is:

$$\hat{\beta} = 1 + n \left[\sum_{i=1}^{n} \ln \frac{N_i}{N_{\rm min}} \right]^{-1} \tag{4.5}$$

where n is the total number of absorbers in the sample and $N_{\rm min}$ in the column density of the lowest-column absorber (see e.g., Clauset et al. 2007). The standard error on $\hat{\beta}$, derived from the width of the likelihood function, is:

$$\sigma_{\hat{\beta}} = \frac{\hat{\beta} - 1}{\sqrt{n}}. \tag{4.6}$$

Similarly, the estimator of the normalization, $C_{\rm HI}$, is

$$\hat{C}_{\rm HI} = \frac{n(1-\hat{\beta})}{\left(N_{\rm max}^{(1-\hat{\beta})} - N_{\rm min}^{(1-\hat{\beta})} \right) \Sigma \Delta X} \tag{4.7}$$

(Tytler 1987) where $\Sigma\Delta X$ is the sum of the pathlength over all 15 sightlines .

In Figure 4.2 the MLE fit to the data is shown for absorbers with $13.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$. Note that a single power law reproduces the distribution with reasonable fidelity. No strong breaks in the distribution are evident.

The choice of a minimum $N_{\rm HI}$= 13.5 used in the fit is motivated as follows. As shown in Figure 4.2, $f(N_{\rm HI},X)$ for $\log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 13.5$ appears to flatten. As mentioned above, this flattening has previously been observed (Hu et al. 1995; Ellison et al. 1999; Kim et al. 2002) and is generally attributed to a combination of incompleteness in the sample caused by line blending and a true physical turnover in $f(N_{\rm HI},X)$. Here we do not attempt to diagnose the cause of the low-$N_{\rm HI}$ turnover, but simply note that above $\log(N_{\rm HI}/{\rm cm}^{-2}) \geq 13.5$, $f(N,X)$ appears to be well fit by a single power law. More quantitatively, one can measure the change in power law slope introduced by varying the minimum value of $N_{\rm HI}$ included in the fit as shown in Table 4.3. Note that a choice of $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 13.3$ produces insignificant change in the maximum likelihood estimate of the slope, $\hat{\beta}$.

Prior to the present work, the largest existing sample of Lyα forest absorbers at these redshifts was described by Kim et al. (2002). Their sample included 2130 H I absorbers with $12.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17$ observed along 9 lines of sight. The most marked improvement over the previous measurements provided by our study is the use of higher-order transitions of H I which were generally not included in previous work.

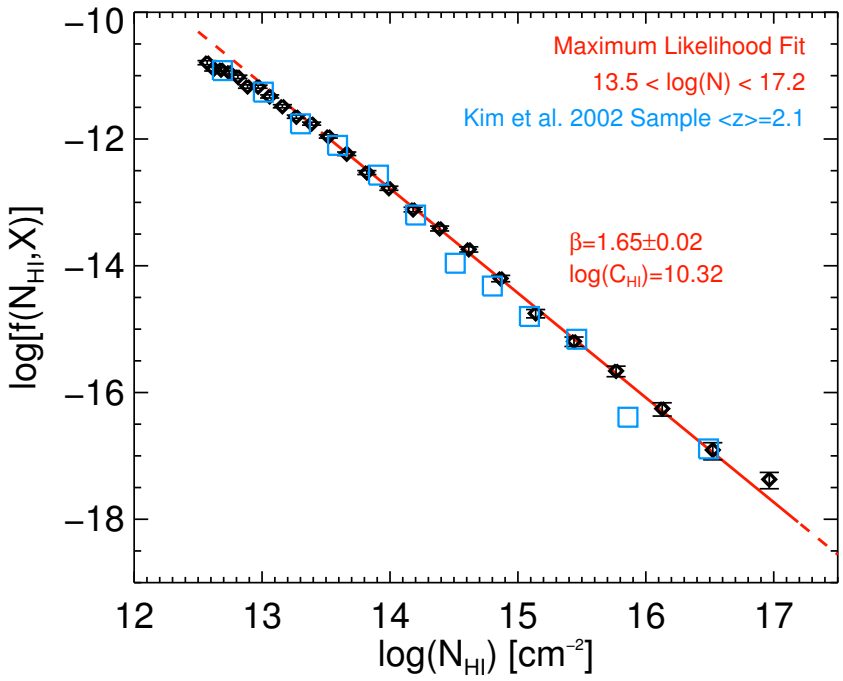


Figure 4.3 The frequency distribution of absorbers from the KBSS H I sample (black points) and the maximum likelihood fit to the data (red curve). For comparison, over plotted in (blue) squares are the data points from the $\langle z \rangle = 2.1$ sample from Kim et al. (2002). Because the analysis of the KBSS sample includes higher-order H I transitions, we can measure the frequency of absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 14$ with much higher fidelity than has been previously possible. Notably, there is no evidence for a strong break in the power law distribution near $\log(N_{\rm HI}/{\rm cm}^{-2}) \approx 14$ as has been previously suggested.

Table 4.3. Maximum Likelihood Fits[a]

$\log(N_{\rm HI,min})$	$\hat{\beta}$	$\sigma_{\hat{\beta}}$	$\log(\hat{C}_{\rm HI})$
12.5	1.518	0.008	8.43
12.6	1.544	0.008	8.80
12.7	1.567	0.009	9.13
12.8	1.585	0.010	9.38
12.9	1.605	0.011	9.67
13.0	1.616	0.012	9.82
13.1	1.627	0.013	9.99
13.2	1.636	0.014	10.11
13.3	1.651	0.015	10.33
13.4	1.658	0.016	10.44
13.5	1.650	0.017	10.32
13.6	1.636	0.018	10.11
13.7	1.639	0.020	10.15
13.8	1.645	0.021	10.25
13.9	1.649	0.023	10.31
14.0	1.654	0.025	10.39

[a]Maximum likelihood fits to the data with $\log(N_{\rm HI,min}) \leq \log(N_{\rm HI}/{\rm cm}^{-2}) \leq 17.2$.

These allow for more precise measurements of $N_{\rm HI}$ above the saturation point of Lyα, $N_{\rm HI} \approx 10^{14-14.5}$ cm^{-2}. Above $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 15.5$ where the Lyβ transition saturates, the sample presented herein may suffer a small degree of incompleteness at $z \lesssim 2.4$ as discussed in Appendix A.

In Figure 4.3, the KBSS data (black diamonds) are compared with the measurements of Kim et al. (2002) at $\langle z \rangle = 2.1$ (blue squares). Notably, at $\log(N_{\rm HI}/{\rm cm}^{-2}) < 14$ there is good agreement between the two data sets, but for $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 14$, the Kim et al. (2002) sample exhibits significant scatter not present in the KBSS measurements, with a systematic tendency to under-estimate the number of intermediate-$N_{\rm HI}$ ($14 \lesssim \log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 17$) absorbers compared to our measurements. Potential incompleteness found in the lower-z portion of the catalog presented herein (see Appendix A) may also affect the Kim et al. (2002) catalog for absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 14$ and may explain why their measurements are systematically lower than those presented here.

4.3.3 Lyman Limit Systems

In this section, we quantify the incidence of absorbers with $N_{\rm HI} > 10^{17.2}$ cm^{-2}, Lyman Limit Systems (LLSs) which are optically thick to hydrogen-ionizing photons. Because LLSs are relatively rare, and because their $N_{\rm HI}$ are challenging to determine accurately via line fitting, we use samples reported in the literature to measure the frequency of optically thick absorbers at the same redshifts as the KBSS sample.

We estimated the incidence of LLSs with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$ at $\langle z \rangle \sim 2.4$ using data from two sur-

veys that nicely bracket the relevant range of redshifts for our IGM sample. We evaluated the LLSs over a somewhat wider redshift range than $2 \lesssim z \lesssim 2.8$ (but having the same mean redshift) in order to mitigate the statistics of small numbers. Over the range $1.65 < z < 2.55$, we used data from Ribaudo et al. (2011) from their survey compilation of archival *Hubble Space Telescope* (*HST*) spectroscopy; results for $2.70 < z < 2.95$ are based on a subset of the ground-based surveys of Sargent et al. (1989) and Stengler-Larrea et al. (1995). The sub-samples are roughly equivalent in their statistical significance (comparable total pathlength and number of LLSs), with path-weighted mean redshifts of $\langle z \rangle = 2.08$ and $\langle z \rangle = 2.84$, respectively. The path-weighted mean redshift for the full sample is $\langle z_{\rm LLS} \rangle = 2.41$ based on 46 LLSs with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$ in a total redshift path of $\Delta z = 25.95$ ($\Delta X = 88.31$). These LLS samples provide a maximum likelihood estimate of the number of LLS per unit pathlength, $dn_{\rm LLS}/dX = 0.52 \pm 0.08$.

A recent study of LLS near the redshift range of interest ($2.0 < z < 2.6$) using uniform HST data is described in O'Meara et al. (2012). They measure an integral constraint on the number of LLS absorbers with $\tau_{912} > 1, \log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$ and $\tau_{912} > 2, \log(N_{\rm HI}/{\rm cm}^{-2}) > 17.5$, where τ_{912} is the optical depth at the Lyman Limit, 912Å, in the absorber's rest frame. These authors found $dn_{\tau>1}/dX = 0.48^{+0.09}_{-0.12}$, in excellent agreement with our compilation described above ($dn_{\rm LLS}/dX = 0.52 \pm 0.08$). For more optically thick systems, O'Meara et al. (2012) found $dn_{\tau>2}/dX = 0.28^{+0.06}_{-0.06}$. The consistency of these measurements with the lower-$N_{\rm HI}$ frequency distribution will be discussed further in §4.5.

Unfortunately, surveys that rely on detecting breaks in the spectrum of QSOs at the redshift of the LLS can rarely measure $N_{\rm HI}$ accurately, especially for absorbers with $\tau_{912} > 2$. As such, a measurement of the power law index of $f(N,X)$ in the LLS regime typically relies on the combination of integral constraints (dn/dX) for the LLS, subDLA, and DLA regimes, as well as the measured values of $f(N,X)$ for Lyα forest absorbers. As we now have much more statistically robust measurements of β for Lyα forest absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 17.2$, we re-examine the power law fit to $f(N,X)$ in the LLS regime.

Here, we test the consistency of the measured incidence of LLSs with the extrapolated slope inferred from the MLE method for absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$. First, we note that LLS surveys are sensitive to all absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$. Observations of DLAs (e.g., Noterdaeme et al. 2009) show that there is a strong turnover in the frequency of DLAs at $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 21$, suggesting there are very few absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 22$. Here, we assume that $dn_{\rm LLS}/dX$ is sensitive to absorbers with $17.2 < \log(N_{\rm HI}/{\rm cm}^{-2}) > 22.0$.

Next, we consider the incidence of LLSs that would be found if the MLE power law fit constrained by our data at $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ held for the entire LLS regime. We calculate the incidence of LLSs given the fit parameters as follows:

$$dn_{\rm LLS}/dX = \int_{N_{\rm min}}^{N_{\rm max}} f(N_{\rm HI},X) dN_{\rm HI} \quad (4.8)$$

$$dn_{\rm LLS,fit}/dX = \int_{N_{\rm min}}^{N_{\rm max}} C_{\rm HI} N_{\rm HI}^{-\beta} dN_{\rm HI}. \quad (4.9)$$

where $\log(N_{\rm min}/{\rm cm}^{-2}) = 17.2$ and $\log(N_{\rm max}/{\rm cm}^{-2}) = 22.0$. We find that this extrapolation to $\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$ underestimates the number of LLSs significantly giving $dn_{\rm LLS,fit}/dX = 0.21$ (compared to $dn_{\rm LLS,data}/dX = 0.52 \pm 0.08$).

The green triangle in Figure 4.2 represents a similar exercise graphically on the $f(N,X)$ plot. If we assume that the MLE determination of β for the forest holds in the LLS regime, we can calculate the value of $C_{\rm HI,LLS}$ implied by the measured $dn_{\rm LLS}/dX$ by inverting equation 4.9. Then the values of $f(N,X)$ suggested by the data would be given by:

$$f(N_{\rm p},X) = C_{\rm HI,LLS} N_{\rm p}^{-\beta}. \tag{4.10}$$

for all values of $N_{\rm p}$. The height of the green triangle thus represents the values of $f(N,X)$ if $\beta_{\rm MLE}$ holds of the LLS regime. The dark green band at the top of the triangle is an estimate in the uncertainty of this normalization due to the uncertainty in $dn_{\rm LLS}/dX$. Again, we can see clearly that the MLE fit to the frequency distribution (red dotted line) significantly under-predicts the measured LLS incidence.

We have seen that a single power law well approximates the intermediate-$N_{\rm HI}$ absorbers; however, its extrapolation to $\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$ underestimates the number of LLSs significantly giving $dn_{\rm LLS,fit}/dX = 0.21$. In §4.5, we discuss other methods for parameterizing the distribution which better reproduce the high-$N_{\rm HI}$ end of the frequency distribution.

4.4 Frequency Distributions with Respect to the Positions of Galaxies

Rudie et al. (2012b) considered the distribution of H I with respect to the positions of star-forming galaxies with $2.0 < z < 2.8$ using the absorption line data presented here. The study explored the kinematic and geometric structure of the CGM, concluding that the majority of the strong excess absorption is found within 300 physical kpc (pkpc) of galaxies in the transverse direction and within 300 km s^{-1} of the systemic redshift of galaxies; however, a tail of excess H I absorbers continues to 2 pMpc and 700 km s^{-1}.

The KBSS spectroscopic galaxy sample consist of UV-color selected galaxies with UV luminosities $0.25 < L/L^* < 3.0$ (Reddy & Steidel 2009) and a typical bolometric luminosity of $\sim 2.5 \times 10^{11} L_\odot$ (Reddy et al. 2008, 2012). They have star formation rates of ~ 30 M$_\odot$ yr^{-1} (Erb et al. 2006b) and halo masses of $\sim 10^{12}$ M$_\odot$ (Adelberger et al. 2005b; Conroy et al. 2008; Trainor & Steidel 2012; Rakic 2012).

Using the same KBSS galaxy sample as in Rudie et al. (2012b), we measure the frequency distribution of H I absorbers within the CGM as follows. First we locate those galaxies in our spectroscopic sample lying within 300 pkpc of the line of sight to the QSO and within the redshift range used for the Lyα forest decompositions (as listed in Table 4.1). Notably, we have no galaxies at impact parameters < 50 pkpc, and so the CGM distributions apply to the range $50-300$ pkpc. We then consider the column density distribution of absorbers within ± 300 km s^{-1} and ± 700 km s^{-1} of the redshifts of these galaxies. We note that all absorbers included in the 300 km s^{-1} CGM are also included in the 700 km s^{-1} CGM.

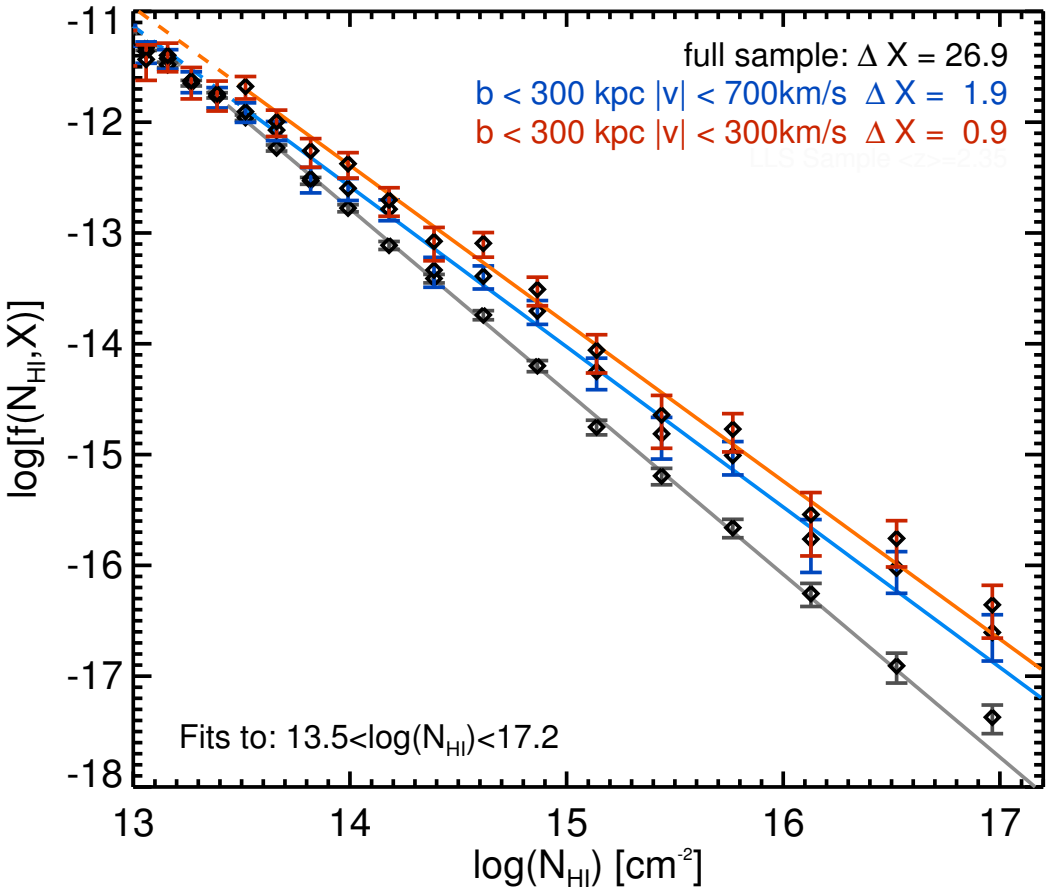


Figure 4.4 Same as Figure 4.2 but also considering the distribution of absorbers close to galaxies. The IGM distribution is shown in gray. Red points show the frequency of absorbers of a given $N_{\rm HI}$ within 300 km s^{-1} and 300 pkpc of a galaxy in our spectroscopic sample. Blue points show the frequency of those within 700 km s^{-1} and 300 pkpc. The curves are the maximum likelihood fits to the distributions over the range $13.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$. The parameters of the fit are given in Table 4.4 along with the opacities calculated from the data and the opacities implied by the fit. Notably this single power law fit to the data does not well reproduce the opacities calculated from the data. Section 4.5.1 discusses broken-power law fits which better reproduce these opacities.

Table 4.4. MLE Power Law fits to $f(N_{\rm HI}, X)$

Sample	β	$\log(C_{\rm HI})$	$\kappa_{\rm fit}$ [pMpc^{-1}]	$\kappa_{\rm data}$[a][pMpc^{-1}]
IGM	1.650 ± 0.017	10.322	0.0029	$0.0039^{+0.0005}_{-0.0005}$
CGM 700km s^{-1} 300 pkpc	1.447 ± 0.033	7.672	0.013	$0.017^{+0.004}_{-0.004}$
CGM 300km s^{-1} 300 pkpc	1.425 ± 0.037	7.562	0.022	$0.031^{+0.009}_{-0.009}$

[a]Opacity calculated from the the data using equation 4.15 for $13.5 \leq \log(N_{\rm HI}/{\rm cm}^{-2}) \leq 17.2$.

The resulting frequency distributions for the 300 km s^{-1} and 700 km s^{-1} CGM are shown in Figure 4.4 in red and blue respectively. As found by Rudie et al. (2012b), the frequency of all absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 13.5$ is higher within the CGM of galaxies than at random places in the IGM.

The maximum likelihood estimates of the power law approximation to $f(N_{\rm HI}, X)$ within the CGM are listed in Table 4.4 and overplotted in Figure 4.4. Notably, the power law slope of the CGM is significantly shallower than that of the IGM, in the sense that there are relatively more high-$N_{\rm HI}$ systems than low-$N_{\rm HI}$ absorbers in the CGM. This result is qualitatively equivalent to results presented by Rudie et al. (2012b) who found a strong correlation between $N_{\rm HI}$ and the fraction of absorbers found within the CGM of galaxies (their Figure 30). Further, the frequency of absorbers is higher within the 300 km s^{-1} CGM than within 700 km s^{-1}, as expected.[3]

The characteristic $N_{\rm HI}$ of the CGM is akin to the max($N_{\rm HI}$) statistic presented in the left-hand panel of Figure 9 of Rudie et al. (2012b). These authors typically found $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 14.5-15$ within 300 pkpc of a galaxy. Misawa et al. (2007) considered the frequency distribution of absorbers surrounding H I systems with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 15$, fitting to the absorber distribution within $\pm$1000 km s^{-1}of the redshifts of these tracer absorbers. Thus, the Misawa et al. (2007) sample traces a volume similar to that of the CGM. For their S2a sample which was selected to lie at least 5000 km s^{-1} lower than the redshifts of the QSOs, Misawa et al. (2007) find a power-law slope of $\beta = 1.390 \pm 0.027$, in reasonable agreement with the power-law exponents for our CGM measurements. The normalization, $C_{\rm HI} = 7.262 \pm 0.431$, is lower than our CGM results, consistent with the expectation that the Misawa et al. (2007) sample includes a larger fraction of "general" IGM compared to CGM.

[3]Since the majority of the excess absorption is found within 300 km s^{-1} of galaxies, the 700 km s^{-1} CGM is diluted by regions with H I properties more similar to the general IGM. As such, the difference between the 300 and 700 km s^{-1} CGM is mostly a result of the difference in pathlength considered. The pathlength for the 300 km s^{-1} CGM sample is $\Delta X = 0.9$ while the 700 km s^{-1} sample has twice the pathlength, $\Delta X = 1.9$ - and since $f(N_{\rm HI}, X)$ is inversely proportional to the pathlength, the frequency of absorbers in the 700 km s^{-1} CGM is approximately a factor of 2 smaller.

4.5 The Lyman Continuum Opacity

One of the principal uses of the column density distribution presented above is to estimate the transmissivity of the IGM to hydrogen-ionizing photons, hereafter referred to as Lyman continuum (LyC) photons. Understanding LyC emission and its transmission through the IGM is an important step in exploring the process of reionization at high-redshift, as well as estimating the average ionization level of the IGM and the metagalactic ionizing background as a function of redshift.

In order to understand and describe the process of transmission of LyC photons through the IGM, one must first quantify the sources of opacity in the IGM. Most notably, it is important to quantify the fractional contribution of absorbers of various $N_{\rm HI}$ to the full opacity of the forest. To do this, we begin with a simplified model of LyC emission and absorption.

The attenuation of LyC photons by an absorber with $N_{\rm HI}$= N_i is given by $e^{-N_i\sigma}$ where σ is the photoionization cross section of hydrogen. In practice, σ depends sensitively on the energy of the photons considered:

$$\sigma(\nu) \approx \sigma_{\rm LL}\left(\frac{\nu}{\nu_{\rm LL}}\right)^{-3} \tag{4.11}$$

where $\sigma_{\rm LL} = 6.35\times10^{-18}$ cm^2 is the hydrogen ionization cross section to photons at the Lyman Limit (912 Å) with energies of 1 Ry (13.6 eV), and $\nu_{\rm LL}$ is the frequency of such photons. For the moment, we will employ the simplifying assumption that all LyC photons have energies of 1 Ry so that the frequency dependence of σ is removed.

In addition, for the time being, we calculate the opacity assuming that the universe is static.[4] In practice, the expansion of the universe results in the redshifting of ionizing photons to non-ionizing energies; however since we are only considering 1 Ry photons, we will also temporarily ignore redshifting. In §4.6, we perform a more rigorous calculation to measure the true mean free path of LyC photons through the IGM and CGM at $z = 2.4$ including photons of all ionizing energies and including redshifting. The analytic calculation presented in this section is provided to build intuition regarding the sources of opacity in the universe, and also to quantify the opacity from absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ inferred directly from the data.

Using the above approximations, we calculate the opacity, κ, of the IGM and CGM to LyC photons and the contribution to this opacity from absorbers of various $N_{\rm HI}$. The opacity is defined as the change in optical depth per unit proper distance:

$$\kappa = \frac{d\tau}{dr} \tag{4.12}$$

where

$$dr = \frac{1}{1+z}\frac{c}{H(z)}dz \tag{4.13}$$

and c is the speed of light.

[4]For the remainder of Section 4.5 we use the redshift distribution of the sample to infer the physical distance over which absorbers are distributed. This is the only way in which redshifts enter into the calculation.

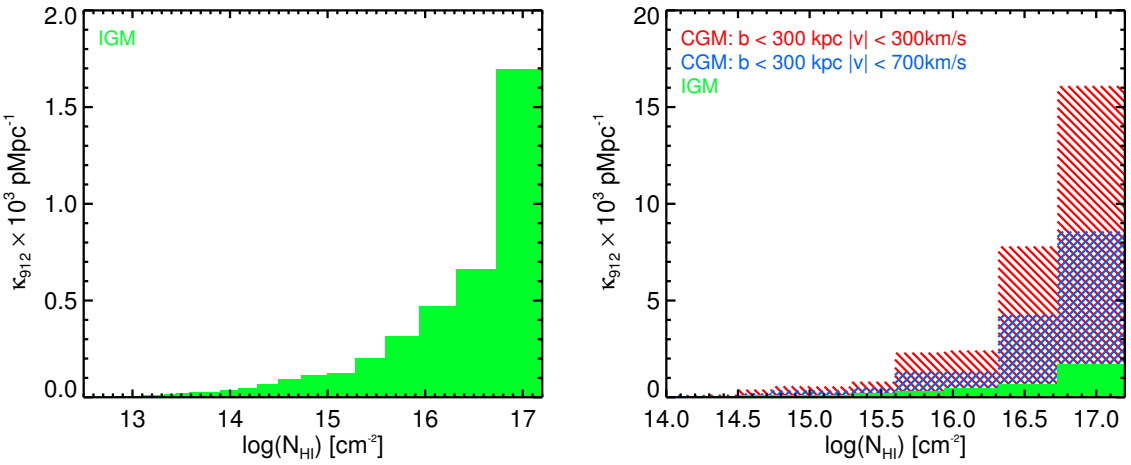


Figure 4.5 Contributions to the opacity from a given bin in $N_{\rm HI}$, measured directly from the data using equation 4.15. *Left:* The IGM opacity per bin in $N_{\rm HI}$. *Right:* The opacity per bin in $N_{\rm HI}$ within the IGM (green), the 700 km s^{-1} CGM (blue) and the 300 km s^{-1} CGM (red). Note that the scale of the left-hand panel is 10% of the scale of the right-hand panel. The opacity (optical depth per unit physical distance) of the CGM is an order of magnitude larger than the average opacity in the IGM.

Given an analytic fit to the frequency distribution per unit physical distance $f(N_{\rm HI}, r)$, the opacity due to Poisson-distributed absorbers with $N_{\rm min} \leq N_{\rm HI} \leq N_{\rm max}$ to photons of energy 13.6 eV is:

$$\kappa_{\rm fit} = \int_{N_{\rm min}}^{N_{\rm max}} f(N_{\rm HI}, r) \left(1 - e^{-N_{\rm HI}\sigma_{912}}\right) dN_{\rm HI} \tag{4.14}$$

(Paresce et al. 1980).

Similarly, the opacity can be calculated directly from the data by summing over those absorbers {i} with $N_{\rm min} < N_{\rm HI} < N_{\rm max}$:

$$\kappa_{\rm data} = \frac{\sum_i \left(1 - e^{-N_{{\rm HI},i}\sigma_{912}}\right)}{\Delta r} \tag{4.15}$$

where Δr is the sum of the proper distances probed by each sightline.

The uncertainty in the measurement of $\kappa_{\rm data}$ can be calculated using the bootstrap resampling method in which many sets of absorbers are drawn with replacement and $\kappa_{\rm data}$ is recalculated for each set. One can then estimate the probability distribution of $\kappa_{\rm data}$ by selecting the 68% confidence range. Values of $\kappa_{\rm data}$ for the IGM and CGM for $13.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ is listed in the last column of Table 4.4. *We emphasize that this measurement of the LyC opacity from absorbers with* $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17$ *is one of the principal results of this work.*

The left hand panel of Figure 4.5 shows the opacity produced by various bins of $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ within the IGM calculated directly from the data using equation 4.15. These values are compared to the opacity found within the CGM in the right-hand panel of Figure 4.5. Notably, the opacity of the CGM of a galaxy (300 km s^{-1}, 300 pkpc) is an order of magnitude larger than that of the average IGM.

Comparing the opacities calculated using the single power law MLE fit with those calculated directly from

Table 4.5. Broken Power Law Fits to $f(N_{\rm HI}, X)$

Sample	$\beta_{\rm low-N_{HI}}$	$\log(C_{\rm low-N_{HI}})$	break $N_{\rm HI}$	$\beta_{\rm high-N_{HI}}$	$\log(C_{\rm high-N_{HI}})$	$\kappa_{\rm fit}$ [pMpc^{-1}]	$\kappa_{\rm data}$[a][pMpc^{-1}]
IGM:	1.656± 0.030	10.398	15.14	1.479± 0.042	7.643	0.0038	$0.0039^{+0.0005}_{-0.0005}$
CGM 700km s^{-1} 300 pkpc	1.353± 0.088	6.354	14.87	1.385± 0.084	6.789	0.017	$0.017^{+0.004}_{-0.004}$
CGM 300km s^{-1} 300 pkpc	1.302± 0.084	5.823	14.87	1.345± 0.089	6.394	0.031	$0.031^{+0.009}_{-0.009}$

[a]Opacity calculated from the the data using equation 4.15 for $13.5 \leq \log(N_{\rm HI}/{\rm cm}^{-2}) \leq 17.2$.

Table 4.6. Effect of Uncertainty in β on κ and $dn_{\rm LLS}/dX$[a]

Sample	$\beta_{\rm high-N_{HI}}$	$\log(C_{\rm high-N_{HI}})$[b]	κ [pMpc^{-1}][c]	$dn_{\rm LLS}/dX$[d]
Data[e]	···	···	$0.0039^{+0.0005}_{-0.0005}$	0.52 ± 0.08
IGM Best Fit	1.479	7.643	0.0038	0.519
IGM Shallow Fit[f]	1.437	6.984	0.0040	0.656
IGM Steep Fit[g]	1.521	8.301	0.0035	0.413

[a]The effect of varying $\beta_{\rm high-N_{HI}}$ within its 1-σ uncertainty.

[b]The best fit normalization of $f(N,X)$ given the stated value of β. I.e., the covariance of $\log(C_{\rm high-N_{HI}})$ with $\beta_{\rm high-N_{HI}}$.

[c]The opacity from absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$.

[d]The incidence of LLSs

[e]Values computed directly from the data.

[f]The fit for the perturbed value of $\beta_{\rm high-N_{HI}}$ such that the distribution is more shallow.

[g]The fit for the perturbed value of $\beta_{\rm high-N_{HI}}$ such that the distribution is steeper.

the data for absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ (Table 4.4), one notes a statistically significant discrepancy in the sense that the MLE power law predicts far less opacity than is measured directly from the data. Similarly, the number of LLSs (which also contribute significantly to the LyC opacity) is under-predicted by the MLE fit. In Section 4.5.1, we consider other parameterizations of the frequency distribution that better match $dn_{\rm LLS}/dX$ as well as the opacity from intermediate-$N_{\rm HI}$ systems [$15 \lesssim \log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 17$] calculated from the data.

4.5.1 Broken Power Law Parameterizations

While the MLE power law fits presented in §4.3.2 and §4.4 provide a reasonable description of the overall column density distribution, because there are substantially more absorbers at low-$N_{\rm HI}$ the majority of the statistical power for the fit is derived from the low-$N_{\rm HI}$ absorbers. To estimate the LyC opacity of the intergalactic and circumgalactic gas, the absorbers of relevance are instead the much rarer, high-$N_{\rm HI}$ systems, particularly those near $\log(N_{\rm HI}/{\rm cm}^{-2}) \approx 17$.

In order to reproduce simultaneously the opacity from both absorbers with $12.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ as measured in our sample, as well as the number of LLSs, we adopt a different fitting method. Here we consider a least-squares fit to the binned[5] data. Because the opacity is a strong function of $N_{\rm HI}$, it is most important that the number of absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \approx 17.2$ is accurately reproduced. Therefore,

[5]We have verified that the fitted parameters are insensitive to the exact choice of binning.

to provide further flexibility, we allow for a discontinuous break in the power law parameterization. This allows for an independent fit to the high-$N_{\rm HI}$ absorbers [$\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 15$] which are the main source of opacity. Although no strong break is evident, any parameterization that does not include a break as well as all parameterizations with a continuous functional form at the position of the break cannot reproduce the opacity of Lyα forest absorbers as measured by the data.

We fit two power laws to the binned data. The bin that includes the break is included in the fit to the low and intermediate-$N_{\rm HI}$ groups to minimize the discontinuity between the two power laws. We allow the break to vary in $N_{\rm HI}$ and select the break location to be that which best reproduces the opacity from absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ and also the number of LLS in the case of the IGM sample. The break location for the two CGM distributions is forced to match. The best fits are presented in Figure 4.6 and Table 4.5 along with their implied opacities, $\kappa_{\rm fit}$. The bin containing the break is indicated by the hatched box in Figure 4.6. Notably, the location of the break ($N_{\rm HI} \approx 10^{15}$ cm^{-2}) is similar to the $N_{\rm HI}$ at which absorbers were found to significantly correlate with the positions of galaxies (Rudie et al. 2012b).

This fitting technique reproduces the opacity from intermediate-$N_{\rm HI}$ absorbers measured from the data with high fidelity. Additionally, the IGM fit predicts a $dn_{\rm LLS,fit}/dX = 0.519$ for $17.2 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 22.0$ in excellent agreement with $dn_{\rm LLS,data}/dX = 0.52 \pm 0.08$ determined from the compilation of LLSs from the literature (Section 4.3.3).[6] As in Section 4.3.3, the green triangle in Figure 4.6 represents the values of $f(N,X)$ within the LLS regime implied by the measurement of $dn_{\rm LLS,data}/dX$ *assuming a power law index,* $\beta = \beta_{\rm high-N_{HI}}$. The agreement between the height of the green triangle and the extrapolation of the power law fit (gray line) suggests that the broken power law parameterization well reproduces the LLS gross statistics.

Next we consider the uncertainty in the fitted parameters. The majority of the uncertainty in the normalization of the frequency distribution ($C_{\rm N_{HI}}$) is due to its covariance with the power law index, β. In Table 4.6 we list the best fit determination of $\beta_{\rm high-N_{HI}}$ and $C_{\rm high-N_{HI}}$ and the values of $\kappa_{\rm fit}$ and $dn_{\rm LLS,fit}/dX$ implied by this analytic approximation. Also listed are the 1-σ perturbed values of $\beta_{\rm high-N_{HI}}$, the corresponding normalization $C_{\rm high-N_{HI}}$, and their implied $\kappa_{\rm fit}$ and $dn_{\rm LLS,fit}/dX$. Notably, the perturbed analytic parameters reproduce the opacity due to Lyα forest absorbers within the measurement errors. The incidence of LLSs is more sensitive to perturbed values of β with offsets in the implied incidence of $\sim 1.5\sigma$ from the observed incidence. As the incidence of LLSs is not used to constrain the power-law fit, the disagreement between these values is not surprising.

Figure 4.7 compares the IGM broken power law fit to the frequency of LLS with $\tau_{912} > 1$ and $\tau_{912} > 2$ from O'Meara et al. (2012) as described in §4.3.3. The blue triangle corresponds to $\tau_{912} > 1$ absorbers, and the red triangle represents the $\tau_{912} > 2$ absorbers. As in section 4.3.3, the normalization of the triangles are set by the measured value of $dn_{\rm LLS,data}/dX$ *under the assumption that an extrapolation of the broken power law fit to the Ly*α *forest data holds within the LLS regime.* Again, the agreement between the extrapolated power-law fit (gray curve) and the blue and red triangles suggests that the analytic approximation to the frequency

[6] We remind the reader that the single power law MLE determination of the fit predicted a value of $dn_{\rm LLS,fit}/dX = 0.21$.

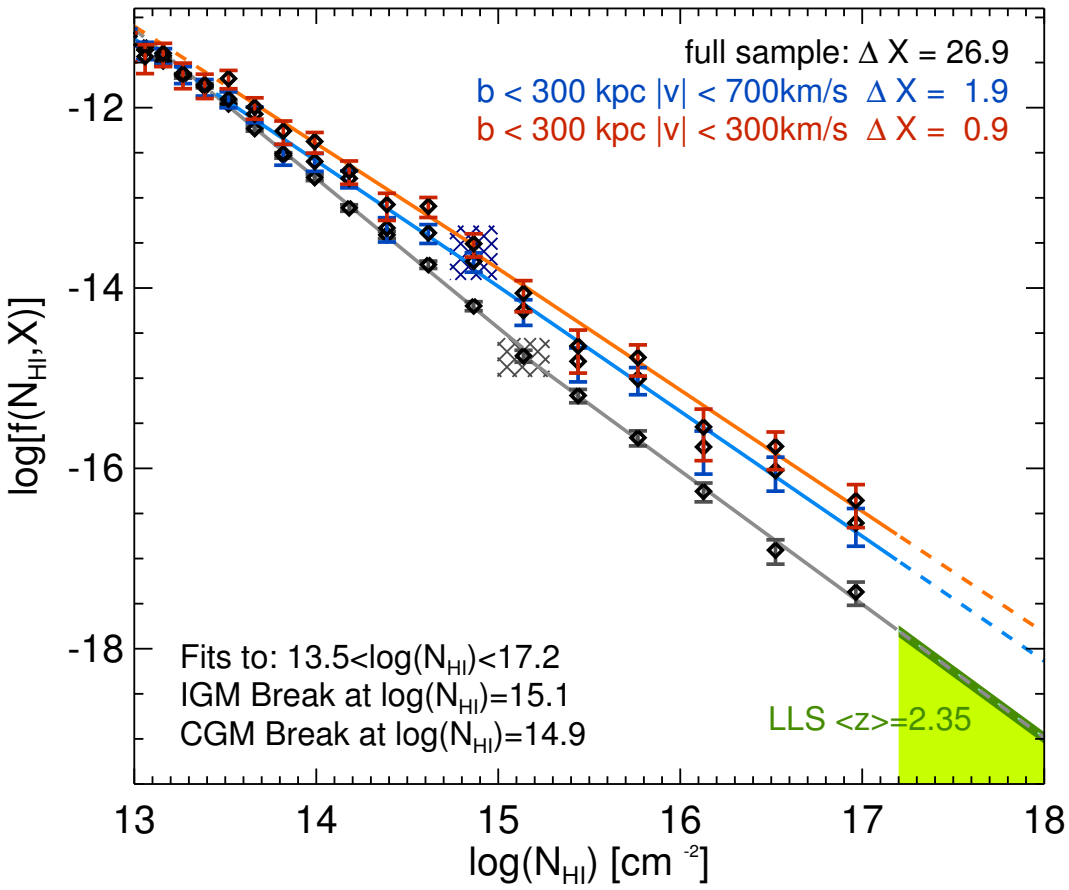


Figure 4.6 Broken power law least-squares fits to the binned IGM (gray) and CGM (blue and red) data. The slopes and normalizations for each data set (IGM or CGM) are allowed to vary independently; however, both ranges include the data point centered at the break (shaded regions). Note that the broken power-law fit is discontinuous at the break location. Including one data point in both ranges minimizes the discontinuity associated with the break and produces a better fit to the data. The hatched box shows the range of $N_{\rm HI}$ included in that point and thus in both fits. The green triangle represents the measured incidence of LLSs as discussed in §4.3.3. The parameters of the fit are given in Table 4.5 along with the opacity calculated from the data and the opacity produced by the fit to the distribution. Note that the opacity is reproduced to much higher fidelity using the broken power law fit.

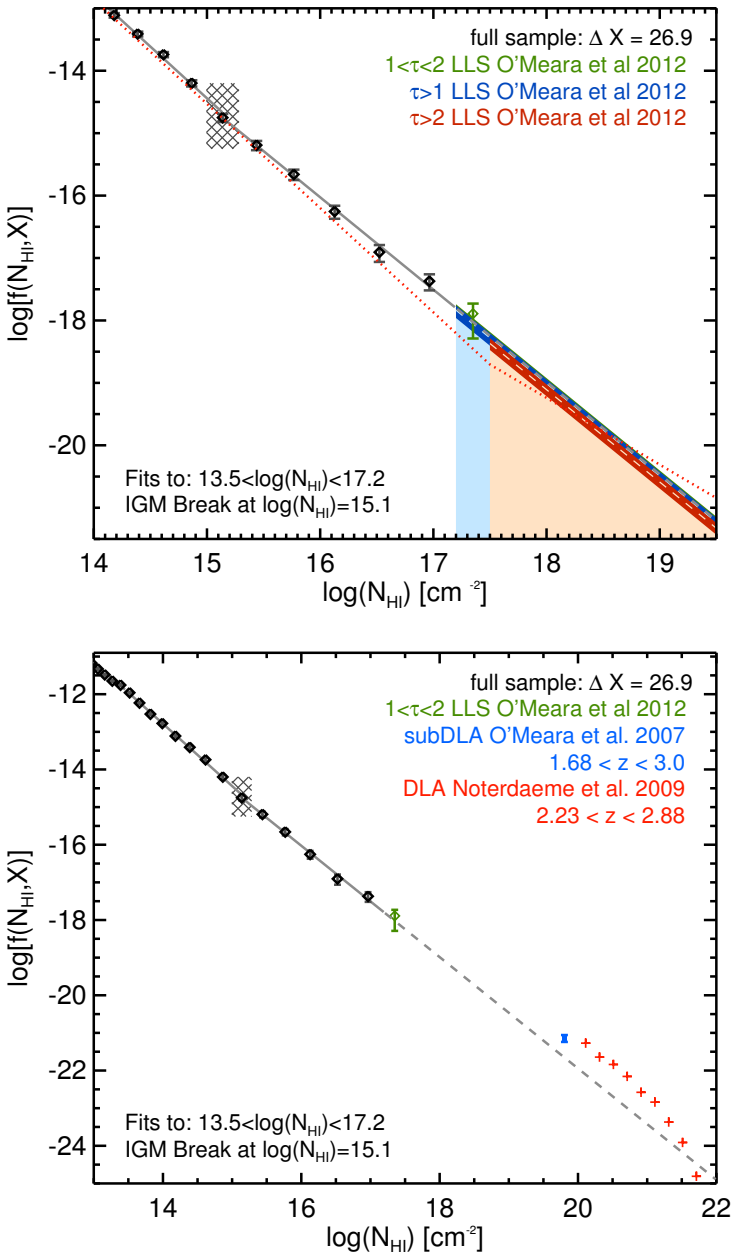


Figure 4.7 Same at Figure 4.6 but only considering the IGM distribution. *Top:* The results of a recent survey for Lyman limit absorbers by O'Meara et al. (2012). Note that these authors' constraints on $\tau_{912} > 1$ and $\tau_{912} > 2$ LLS are well reproduced by the extrapolation of the power law slope selected for the high-$N_{\rm HI}$ IGM absorbers. Computing the difference between these values allows for an assessment of the number of LLSs with $1 < \tau_{912} < 2$, here represented as the green data point. The (red) dotted curve shows the fit to the frequency distribution proposed by O'Meara et al. (2012). Note that this fit underpredicts the number of absorbers with $\log(N_{\rm HI})/{\rm cm}^{-2}) < 17.2$ as measured with the KBSS sample (black data points). *Bottom:* Including all currently known constraints at these redshifts for $f(N,X)$. In green the $1 < \tau_{912} < 2$ measurement from O'Meara et al. (2012). In blue are the subDLAs ($19.3 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 20.3$; $dn_{\rm subDLA}/dX = 0.110^{+0.026}_{-0.021}$) from O'Meara et al. (2007) and in red are the DLAs ($\log(N_{\rm HI}/{\rm cm}^{-2}) > 20.3$; $dn_{\rm DLA}/dX \approx 0.05$) from Noterdaeme et al. (2009). We note that the extrapolation of the intermediate-$N_{\rm HI}$ absorbers power law fit suggested by our Lyα forest data appears to provide a good fit to the data for $\log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 18$. Above this $N_{\rm HI}$, nature likely deviates significantly from this power law.

distribution reproduces the measure LLS statistics from O'Meara et al. (2012). The green point in Figure 4.7 shows the measured frequency of absorbers with $1 < \tau_{912} < 2$ [$17.2 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.5$] calculated from the difference between the measured values of $dn_{{\rm LLS},\tau>1}/dX$ and $dn_{{\rm LLS},\tau>2}/dX$.

The power law fit to $f(N,X)$ suggested by O'Meara et al. (2012) is overplotted as the red dotted line for comparison. This curve would imply a value for the opacity from absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ of $\kappa_{\rm fit} = 0.0022$, nearly a factor of two smaller than that measured from the KBSS sample ($\kappa_{\rm data} = 0.0039 \pm 0.0005$).

There are likely some deviations in the detailed distribution of $f(N_{\rm HI},X)$ in the subDLA ($19.3 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 20.3$) and DLA ($\log(N_{\rm HI}/{\rm cm}^{-2}) > 20.3$) regimes as discussed in Prochaska et al. (2009) and O'Meara et al. (2012). We show the current constraints on these regions of the frequency distribution in the bottom panel of Figure 4.7. We chose to neglect the differences between the power law fit and observed constraints on $f(N,X)$. Holding the measured value of $dn_{\rm LLS}/dX$ fixed, such an assumption makes our measurements of the opacity from absorbers with $N_{\rm HI} > 10^{17.2}$ cm^{-2} a *lower limit*. However, since the attenuation resulting from any system with $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 18$ is similar, the detailed shape of the distribution for high-$N_{\rm HI}$ will introduce relatively small changes in the opacity so long as the total number of such systems is reproduced.

The opacity inferred from the power law fit to $f(N,X)$ is a lower limit for the following reason. If we hold $dn_{\rm LLS}/dX$ fixed at the observed value, this gives us a constraint on the number of systems with $N_{\rm HI} > 10^{17.2}$ cm^{-2}. The assumption of a given value of the power-law index, β, allows us to infer the opacity from a measurement of $dn_{\rm LLS}/dX$. Allowing β to take a shallower value to match the plotted points of $f(N,X)$ for the subDLAs and DLAs only moves absorbers to higher $N_{\rm HI}$ and thus higher opacity compared to a model with a steeper power-law index. Therefore, our suggested opacities are likely a slight under-prediction of the true value. If we instead choose a flat value ($\beta = 0$) we would find 9.7% higher opacity from LLSs, suggesting the total uncertainty in the opacity from the assumed value of β is significantly less than 10%.

4.5.2 The Opacity of the IGM and CGM

Under the assumptions outlined above, and with a parameterization of the frequency distribution that reproduces both the incidence of LLS and the opacity from intermediate-$N_{\rm HI}$ absorbers, we can now measure the full opacity of the IGM and CGM. We calculate the opacity from LLS absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$ using equation 4.14[7] and the opacity from intermediate-$N_{\rm HI}$ absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ directly from the data using equation 4.15 (as described in the beginning of Section 4.5). The differential opacity from bins in $N_{\rm HI}$ as well as the cumulative opacity (from absorbers with $N_{\rm HI}$ smaller than a given value), are

[7] Note that the detailed distribution of κ per bin in $N_{\rm HI}$ would be affected by variation in $f(N_{\rm HI})$ at DLA and subDLA column densities. Because we adopt the extrapolation of the broken-power law fit to $f(N,X)$ to calculate the opacity, the values in individual bins of $N_{\rm HI}$ for $\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$ are not well constrained.

given in Table 4.7. The total opacity of the IGM is measured to be:

$$\kappa_{\rm IGM} = \kappa_{\rm data}\left(N_{\rm HI} < 10^{17.2}\right) + \kappa_{\rm fit}\left(N_{\rm HI} > 10^{17.2}\right) \quad (4.16)$$

$$\kappa_{\rm IGM} = 8.146 \times 10^{-3}\ {\rm pMpc}^{-1}. \quad (4.17)$$

Similarly,

$$\kappa_{\rm CGM\ 300\ km\ s^{-1}} = 94.4 \times 10^{-3}\ {\rm pMpc}^{-1} \quad (4.18)$$

$$\kappa_{\rm CGM\ 700\ km\ s^{-1}} = 46.5 \times 10^{-3}\ {\rm pMpc}^{-1}. \quad (4.19)$$

Table 4.7. Opacity and Optical Depth of Absorbers of Various $N_{\rm HI}$

$N_{\rm HI}$ min	$N_{\rm HI}$ max	method	$\kappa_{\rm IGM}$ $\times 10^3$ [pMpc^{-1}]	$\tau_{\rm IGM}$ $\times 100$	$\kappa_{\rm CGM\ 700}$ $\times 10^3$ [pMpc^{-1}]	$\tau_{\rm CGM\ 700}$ $\times 100$	$\kappa_{\rm IGM}$ $\times 10^3$ [pMpc^{-1}]	$\tau_{\rm IGM}$ $\times 100$	$\kappa_{\rm CGM\ 700}$ $\times 10^3$ [pMpc^{-1}]	$\tau_{\rm CGM\ 700}$ $\times 100$
			Differential: $N_{\rm HI}$ min $< \log(N_{\rm HI}/{\rm cm}^{-2}) < N_{\rm HI}$ max				Cumulative: $12.55 < \log(N_{\rm HI}/{\rm cm}^{-2}) < N_{\rm HI}$ max			
12.55	12.60	data	0.001	0.02	0.001	0.00	0.001	0.02	0.001	0.00
12.60	12.65	data	0.002	0.02	0.002	0.00	0.003	0.03	0.002	0.00
12.65	12.71	data	0.002	0.02	0.002	0.00	0.005	0.06	0.004	0.00
12.71	12.77	data	0.003	0.03	0.003	0.00	0.008	0.09	0.007	0.00
12.77	12.85	data	0.004	0.04	0.003	0.00	0.012	0.13	0.009	0.00
12.85	12.93	data	0.004	0.04	0.003	0.00	0.016	0.17	0.012	0.00
12.93	13.01	data	0.006	0.07	0.006	0.00	0.022	0.24	0.018	0.01
13.01	13.11	data	0.008	0.08	0.007	0.00	0.030	0.32	0.024	0.01
13.11	13.21	data	0.009	0.10	0.010	0.00	0.039	0.42	0.034	0.01
13.21	13.32	data	0.011	0.12	0.011	0.00	0.050	0.54	0.046	0.01
13.32	13.45	data	0.017	0.18	0.016	0.00	0.067	0.72	0.061	0.02
13.45	13.59	data	0.021	0.22	0.022	0.01	0.088	0.94	0.083	0.02
13.59	13.74	data	0.024	0.25	0.032	0.01	0.112	1.20	0.116	0.03
13.74	13.90	data	0.027	0.29	0.027	0.01	0.140	1.49	0.143	0.04
13.90	14.08	data	0.038	0.40	0.053	0.02	0.178	1.89	0.195	0.06
14.08	14.28	data	0.045	0.48	0.089	0.03	0.223	2.37	0.285	0.08
14.28	14.50	data	0.066	0.70	0.073	0.02	0.289	3.07	0.358	0.10
14.50	14.73	data	0.095	1.01	0.198	0.06	0.383	4.08	0.557	0.16
14.73	15.00	data	0.115	1.23	0.353	0.10	0.499	5.31	0.910	0.26
15.00	15.28	data	0.123	1.30	0.351	0.10	0.621	6.61	1.261	0.36
15.28	15.60	data	0.202	2.14	0.511	0.15	0.823	8.75	1.772	0.51
15.60	15.94	data	0.314	3.34	1.320	0.38	1.136	12.09	3.092	0.89
15.94	16.32	data	0.471	5.01	1.316	0.38	1.607	17.10	4.408	1.26
16.32	16.73	data	0.661	7.03	4.278	1.23	2.268	24.13	8.686	2.49
16.73	17.20	data	1.695	18.03	8.571	2.46	3.963	42.17	17.257	4.95
17.20	17.70	fit	1.570	16.70	9.249	2.65	5.534	58.87	26.506	7.60
17.70	18.20	fit	1.111	11.82	7.273	2.09	6.645	70.69	33.779	9.68

Another relevant quantity is the integrated opacity over a specific scale length, i.e., the optical depth τ. In particular, it is now possible to calculate the optical depth of the CGM of an individual galaxy. This quantity is particularly relevant for understanding the attenuation of ionizing flux emanating from a galaxies similar to those in the KBSS sample due to the CGM of the galaxy itself. Here we attempt to separate the probability that ionizing photons escape the ISM of a galaxy from the probability that photons, having already escaped the ISM, will also escape the enhanced absorption found within the CGM.

Specifically, the KBSS data set constrains the opacity and optical depth of the CGM of galaxies at impact parameters greater than 50 kpc[8]. Henceforth, we consider the classical "escape fraction" to be the probability that an ionizing photon reaches a distance of 50 pkpc, while the CGM is defined to begin at a distance of 50 pkpc from the galaxy. In order to calculate the optical depth, the scale length of interest must be assumed. For circumgalactic absorption, the velocity window associated with the CGM provides a natural scale length. As the relevant scale is along the line of sight, we use the distance, $\Delta r_{\rm CGM}$, associated with the line-of-sight velocity amplitude of the CGM:

$$\Delta r_{\rm CGM} = H^{-1}(z) d{\rm v} \tag{4.20}$$

where $H(z)$ is the Hubble constant evaluated at the mean redshift of the sample $\langle z \rangle = 2.4$ and dv is either 700 km s^{-1} or 300 km s^{-1}. We note that the KBSS data set constrains the distribution of gas at transverse distances > 50 pkpc from galaxies, not along the line of sight from galaxies. Here we assume that the gas distribution traced in the transverse plane well represents the gas distribution along the line of sight to a galaxy at distances > 50 kpc. Given the strong increase in $N_{\rm HI}$ at small galactocentric distances [as observed in Rudie et al. (2012b)], the optical depth associated with the CGM at distances > 50 kpc is likely significantly smaller than the optical depth associated with the closer-in CGM. Therefore the measurements presented here might be thought of as a lower limit on CGM absorption.

In Figure 4.8, we compare the optical depth of the CGM ($\tau_{\rm CGM} = \kappa_{\rm CGM} \Delta r_{\rm CGM}$) for the 300 km s^{-1}, 300 pkpc CGM (red) and the 700 km s^{-1}, 300 pkpc CGM (blue). Figure 4.8 also quantifies the optical depths from both forest absorbers ($\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$) and LLS in the CGM. The forest absorber opacities and optical depths are measured directly from the data, while the LLS opacity and optical depths are inferred from an extrapolation of the high-$N_{\rm HI}$ fit to the data using the broken power law parameterization. The values for $\tau_{\rm CGM}$ and $\kappa_{\rm CGM}$ in differential and cumulative $N_{\rm HI}$ bins are given in Table 4.7.

We find that the optical depth of the 700 km s^{-1} CGM is comparable to but slightly exceeds the optical depth in the 300 km s^{-1} CGM. This is as expected because the 300 km s^{-1} CGM includes the majority of the excess absorption; however, there is a small additional excess between 300 and 700 km s^{-1}, see Rudie et al. (2012b). For subsequent calculations, only the 700 km s^{-1} CGM will be considered since it captures the full optical depth of the CGM for $d > 50$ pkpc.

[8]This is because the closest galaxy to a QSO sightline in the KBSS sample is at a projected physical distance of 50 kpc.

Table 4.7—Continued

$N_{\rm HI}$ min	$N_{\rm HI}$ max	method	$\kappa_{\rm IGM}$ $\times 10^3$ [pMpc^{-1}]	$\tau_{\rm IGM}$ $\times 100$	$\kappa_{\rm CGM\ 700}$ $\times 10^3$ [pMpc^{-1}]	$\tau_{\rm CGM\ 700}$ $\times 100$	$\kappa_{\rm IGM}$ $\times 10^3$ [pMpc^{-1}]	$\tau_{\rm IGM}$ $\times 100$	$\kappa_{\rm CGM\ 700}$ $\times 10^3$ [pMpc^{-1}]	$\tau_{\rm CGM\ 700}$ $\times 100$
			Differential: $N_{\rm HI}$ min $< \log(N_{\rm HI}/{\rm cm}^{-2}) <$ $N_{\rm HI}$ max				Cumulative: $12.55 < \log(N_{\rm HI}/{\rm cm}^{-2}) <$ $N_{\rm HI}$ max			
18.20	18.70	fit	0.647	6.88	4.719	1.35	7.291	77.57	38.498	11.04
18.70	19.20	fit	0.372	3.96	3.030	0.87	7.664	81.53	41.527	11.91
19.20	19.70	fit	0.214	2.28	1.946	0.56	7.878	83.81	43.473	12.46
19.70	20.30	fit	0.141	1.50	1.439	0.41	8.019	85.31	44.912	12.88
20.30	22.00	fit	0.127	1.35	1.597	0.46	8.146	86.67	46.509	13.33

[a]Opacity calculated from the the data using equation 4.15 for $12.55 \leq \log(N_{\rm HI}/{\rm cm}^{-2}) \leq 17.2$.

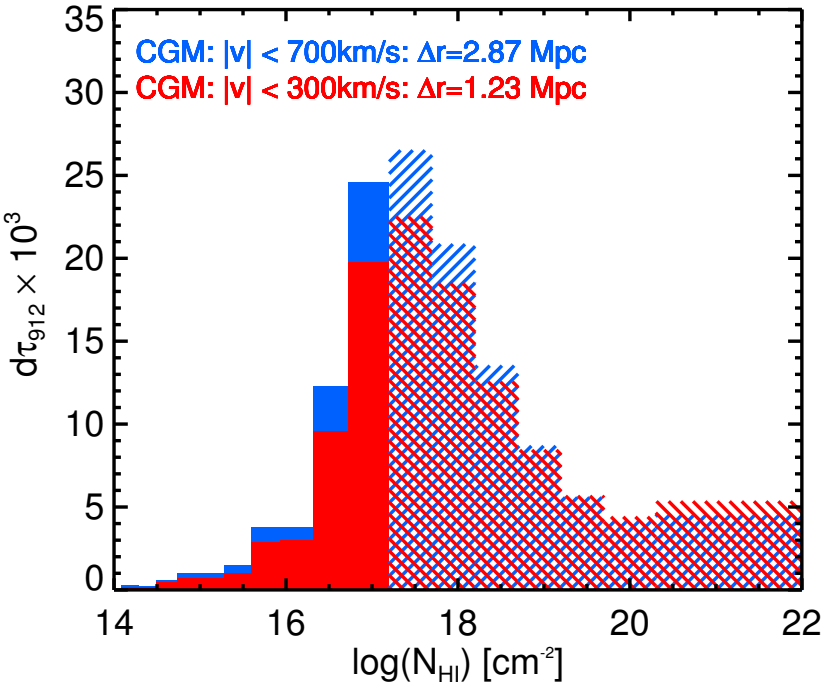


Figure 4.8 The optical depth of the CGM to13.6 eV photons. The bins with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ (solid shading) are measured directly from the data. The bins corresponding to LLSs ($\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$; hatched shadding) use the broken power law fit to the higher $N_{\rm HI}$ end of $f(N_{\rm HI}, X)$ as shown in Figure 4.6. The blue and red histograms show the total optical depth within one average CGM. This is equivalent to the opacity times the scale length $\Delta r_{\rm CGM}$. Note that the total optical depth in the 700 km s^{-1} CGM includes the opacity within the 300 km s^{-1} CGM but that there exists additional opacity due to absorbers with velocity offsets between $\pm$300 and $\pm$700 km s^{-1}.

4.5.3 The Distribution of Optical Depths in the IGM and CGM

With an estimate of the opacity from all $N_{\rm HI}$ absorbers, one can easily calculate the fractional contribution of the CGM vs the IGM, as well as the contribution from absorbers of different $N_{\rm HI}$, to the mean free path. For this simplified calculation, we define the mean free path, $\lambda_{\rm mfp}^{\rm static}$, to be the average distance an ionizing photon with energy 1 Ry travels in a static universe before having a probability of $1-e^{-1}$ of being absorbed. More quantitatively:

$$\lambda_{\rm mfp}^{\rm static} = \sum_i \Delta r_i \tag{4.21}$$

where

$$\sum_i \kappa_i \Delta r_i = 1. \tag{4.22}$$

Here we consider two possible mean free paths: (1) In the case of photons emanating from the ISM of a galaxy similar to those in our spectroscopic sample (see Section 4.4), they must first traverse the CGM of the galaxy in which they were formed and then the general IGM. Here we refer to this mean free path as $\lambda_{\rm mfp,\ IGM+CGM}^{\rm static}$.(2) The second case is the more general one and the mean free path that has been considered by previous authors - the mean free path of photons traveling in the IGM, a quantity relevant to photons once they have escaped the CGM of whichever structure in which they were formed. This distance is referred to as $\lambda_{\rm mfp,\ IGM}^{\rm static}$.

In the calculation that follows, we consider the contributions to the opacity from the 700 km s^{-1} CGM as well as the IGM. Under these assumptions:

$$\kappa_{\rm CGM}\Delta r_{\rm CGM} + \kappa_{\rm IGM}(\lambda_{\rm mfp,\ IGM+CGM}^{\rm static} - \Delta r_{\rm CGM}) = 1. \tag{4.23}$$

where $\Delta r_{\rm CGM} = 2.87$ pMpc as given in equation 4.20 using $dv = 700$ km s^{-1}, and $\kappa_{\rm IGM} = 0.0081$ pMpc^{-1} and $\kappa_{\rm CGM} = 0.047$ pMpc^{-1} are the total opacities over all $N_{\rm HI}$. Solving equation 4.23, we find $\lambda_{\rm mfp,\ IGM+CGM}^{\rm static}$= 109.3 pMpc. When we exclude the CGM component, we instead obtain $\lambda_{\rm mfp,\ IGM}^{\rm static}$= 122.8 pMpc, a 12% longer pathlength.

With $\lambda_{\rm mfp,\ IGM+CGM}^{\rm static}$ known, we can then calculate the fractional contribution to the optical depth from each bin of $N_{\rm HI}$ for both the IGM and CGM as shown in Figure 4.9 and Table 4.7 (middle columns). The cumulative version of this plot (i.e. the contribution to the optical depth for all $N_{\rm HI} < N_i$) is given in Figure 4.10 and in Table 4.7, right-hand columns. From these, we can see that the CGM accounts for 13% of the opacity within a $\lambda_{\rm mfp,\ IGM+CGM}^{\rm static}$. Further, absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ contribute 48% of the opacity to LyC photons.

In practice, the detailed distribution of opacity as a function of $N_{\rm HI}$ will be altered when photons of various energies are considered. The true $\lambda_{\rm mfp}$ of LyC photons will be non-negligibly affected by the redshifting of photons to lower energies, and by the presence of higher-energy radiation. For instance, 100 pMpc corresponds to $\Delta z \approx 0.25$ at these redshifts. Photons emitted by a source at $\langle z \rangle = 2.4$ with energies close to the

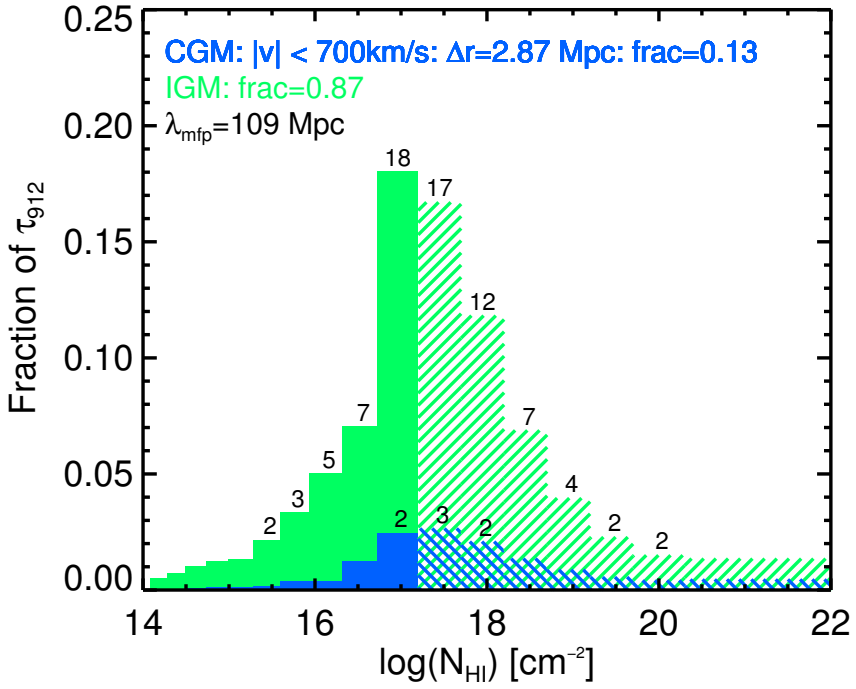


Figure 4.9 The fraction of the optical depth within 1 mean free path contributed by absorbers within a given bin of $N_{\rm HI}$ in the IGM (green) or 700 km s^{-1} CGM (blue). The solid histograms represent measurements from the data converted into an optical depth using equations 4.12 and 4.15. The hatched histograms represent values from the fit to the data extrapolated to higher values of $N_{\rm HI}$ and then converted into an optical depth using equations 4.12 and 4.14. The printed numbers give the percentage of the opacity due to that $N_{\rm HI}$ bin.

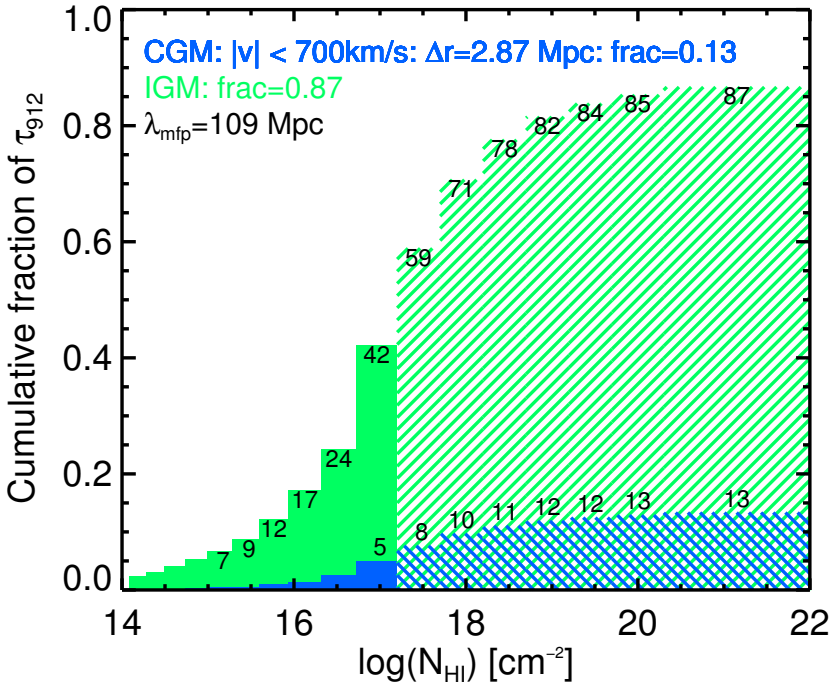


Figure 4.10 The cumulative fraction of the optical depth within 1 mean free path contributed by absorbers below a given $N_{\rm HI}$ in the IGM (green) or 700 km s^{-1} CGM (blue). The solid histograms represent measurements from the data converted into an optical depth using equations 4.12 and 4.15. The hatched histograms represent values from the fit to the data extrapolated to higher values of $N_{\rm HI}$ and then converted into an optical depth using equations 4.12 and 4.14. The printed numbers give the percentage of the opacity due to absorbers with $N_{\rm HI}$ less than or within the $N_{\rm HI}$ bin.

Lyman Limit therefore will redshift to non-ionizing energies long before they are attenuated by e.

In the next section, we consider a more rigorous calculation of $\lambda_{\rm mfp}$ including the aforementioned effects using Monte Carlo (MC) simulations of discrete lines of sight. As we will show, a more proper treatment of the physics does not substantially affect the value of $\lambda_{\rm mfp}$.

4.6 The Mean Free Path to Hydrogen-Ionizing Photons

As we have shown in Section 4.5, the mean free path of LyC photons is a cosmological distance at $z = 2.4$. As such, the consideration of photons with energies > 1 Ry as well as inclusion of redshifting of photons out of ionizing frequencies are necessary in order to measure the true $\lambda_{\rm mfp}$. It is also necessary to define very precisely what one means by the "mean free path," including the frame of reference in which the distance is measured. Here we define the mean free path, $\lambda_{\rm mfp}$, as the distance traveled from a source by a packet of photons before photons with energy of 1 Ry in the *absorber's*[9] frame are attenuated by an average factor of e. This distance is fundamentally related to a measurement of Δz, which we convert to the physical distance traveled by the photon through the expanding universe.

We perform a Monte Carlo simulation of the LyC absorption along many individual sightlines. For illustration, we begin with a discussion of MC runs with an emission redshift of $z = 2.4$.

4.6.1 Monte Carlo Simulations

To simulate absorption within intergalactic and circumgalactic space, we produced artificial absorption spectra whose absorption line distribution is matched to that observed in data at the same redshifts; such models are described in detail by Shapley et al. (2006). In this case, we produced 10,000 simulated lines of sight to sources with $z_{\rm em} = 2.4$, according to a frequency distribution of absorbers set using best-fit parameters from the broken power law fit to $f(N,X)$ (Table 4.5).

As was demonstrated in Section 4.5.1, various fitting methods produce largely discrepant values for the parameters of the analytic form of $f(N,X)$. The actual opacity represented by the data can be measured to significantly higher precision then the parameters of the analytic form of $f(N,X)$. For this reason, using perturbed values of the power law parameters will result in significantly larger variation in the opacity along a line of sight than occur in the actual universe as measured by the data. As such, we use only the best-fit values of the power law fit to $f(N,X)$ in the Monte Carlo.

The simulated absorbers are distributed throughout the redshift range of interest in such a way that they reproduce the overall redshift distribution of measured absorbers. The redshift distribution of absorbers is typically parameterized as:

$$\frac{dn}{dz} = n_0(1+z)^{\gamma}. \tag{4.24}$$

[9] We consider the energy of the photon at the absorber's redshift because only photons with energy sufficient to ionize hydrogen when they intersect an absorber are relevant to $\lambda_{\rm mfp}$.

For the MC simulations, we have assumed that the Lyman-α forest absorbers with low-$N_{\rm HI}$ (those absorbers fit by the low-$N_{\rm HI}$ power law in the broken power law parameterization: $\log(N_{\rm HI}/{\rm cm}^{-2}) < 15.1$ for the IGM and $\log(N_{\rm HI}/{\rm cm}^{-2}) < 14.9$ for the CGM) evolve with $\gamma = 2.5$ (Kim et al. 2002). We have verified that the line sample presented in this paper exhibits evolution consistent with these measurements. All high-$N_{\rm HI}$ absorbers are assumed to have $\gamma = 1.0$; the value for the LLS compilation reported in this paper (Section 4.3.3) measured through maximum likelihood estimation is $\gamma = 0.93 \pm 0.29$.

Briefly, we note that the redshift evolution for LLSs within the range $2 \lesssim z \lesssim 3$ is well constrained by the data presented by Ribaudo et al. (2011), Sargent et al. (1989), and Stengler-Larrea et al. (1995) as discussed in Section 4.3.3. Measurements of the evolution in the line density at $z \gtrsim 3$ (see e.g., Songaila & Cowie 2010; Prochaska et al. 2010) suggest much steeper evolution with redshift (high values of γ). Because our Monte Carlo simulations only include absorption at $2 \lesssim z \lesssim 3$, we use the value of $\gamma = 1$ as measured over the relevant redshift range and emphasize that the exact value of γ does not significantly affect the quoted $\lambda_{\rm mfp}$ within the redshift range $2 < z < 3$. [10] However, we caution that the extrapolations of the trend of our $\lambda_{\rm mfp}$ measurements to $z > 3$ is much more dependent on the chosen value of γ and is therefore unlikely to predict the true value at higher redshift.

An additional set of simulations was run to mimic lines of sight emanating from galaxies like those in the KBSS. In these simulations, H I absorbers were added within 700 km s^{-1} of $z_{\rm em}$ according to the CGM broken power law fit to $f(N,X)$. The remainder of the line of sight is drawn from the IGM distribution. These simulations produce the expected attenuation due to the CGM for sources within galaxies similar to those in our spectroscopic sample ($0.25 < L/L^* < 3$, $\langle z \rangle = 2.3$).

In addition to the simulated forest spectra that include both line and continuum opacity, we created another set where the opacity from individual Lyman lines is not included, leaving only Lyman continuum absorption. To measure the $\lambda_{\rm mfp}$, the principal concern is the fraction of photons absorbed while they are at ionizing energies. Once photons redshift out of the hydrogen-ionizing band, for the purpose of calculating the mean free path, their subsequent absorption or transmission is immaterial. For this reason, spectra with line and continuum opacity are useful for correcting observed objects for attenuation along the line of sight through the IGM (as will be discussed in Section 4.6.4), while the "continuum absorption only" spectra can be used to measure $\lambda_{\rm mfp}$ directly (as discussed in Section 4.6.2).

Figure 4.11 shows the average of 10,000 realizations of each Monte Carlo run plotted in the rest-frame of the emitter. The solid curves show the runs with Lyman line opacity included, and the dotted and dash-dotted lines show the MC runs that only consider the continuum opacity. The various colors pertain to either IGM ONLY (blue and purple) or IGM+CGM (red and orange) runs (note the presence of Lyα, β, γ, etc. absorption due to the CGM of the emitter in the red spectrum). Note that at the Lyman limit, the spectra that include CGM absorption show a sharper drop in transmission. This is due to LyC absorption from high-$N_{\rm HI}$ absorbers

[10] Considered in the redshift range $2.0 < z < 2.8$,varying γ by 1σ results in less than a 6% change in the incidence of LLSs and therefore less than a 3% change in $\lambda_{\rm mfp}$.

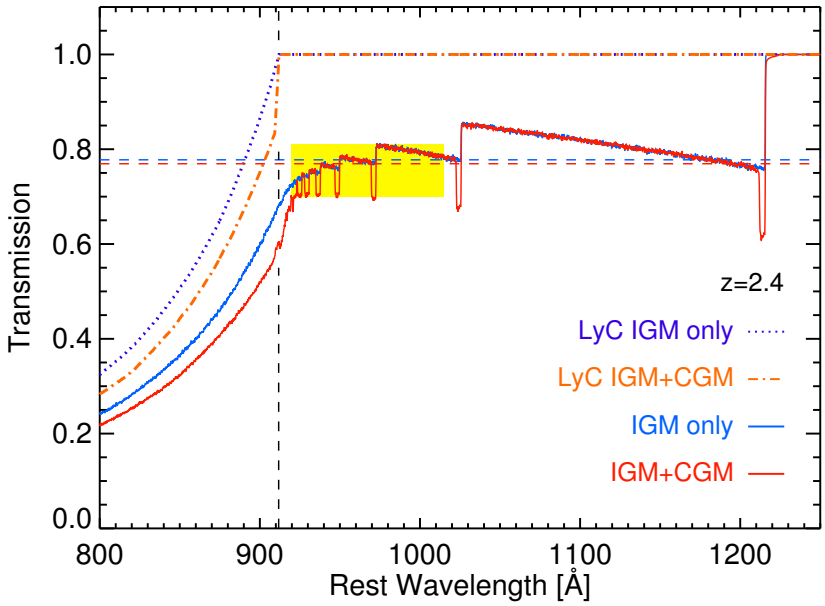


Figure 4.11 The average transmission spectra from 10,000 runs of the Monte Carlo simulations of IGM only (purple and blue) and IGM+CGM(red and orange) absorption in the foreground of a $z = 2.4$ emitter shown in the rest frame of the emitter. The solid (blue and red) curves show spectra using MC simulations that include line opacity. The dotted (purple) and dot-dashed (orange) curves show spectra from simulations only including continuum opacity. The highlighted yellow area marks the region between 920 -1015 Å in the frame of the emitter which is used to renormalize the line opacity spectra in Figure 4.12. The dashed horizontal lines mark the average value of the transmission in this region, $\langle 1 - D_B \rangle$, for the spectra containing line opacity. The vertical dashed line marks the position of the Lyman limit in the frame of the emitter. The line absorption seen in the red spectrum is due to the Lyman series line opacity from high-$N_{\rm HI}$ CGM absorbers.

found in the circumgalactic gas.

4.6.2 Measuring $\lambda_{\rm mfp}$

To measure $\lambda_{\rm mfp}$ from the Monte Carlo calculation, we use the spectra that include continuum absorption only.[11] We search the average spectrum for the location where the LyC only spectra reach a transmission equivalent to e^{-1}, as marked by the vertical dash-dotted lines in Figure 4.12. The rest wavelengths from Figure 4.12 can be converted into the physical distance traveled by a photon between $z_{\rm em} = 2.4$, the emission redshift of the spectra, and $z_{\rm mfp}$ where:

$$z_{\rm mfp} = \frac{\lambda_e - \lambda_{\rm LL}}{\lambda_{\rm LL}}(1 + z_{\rm em}) + z_{\rm em} \tag{4.25}$$

where $\lambda_{\rm LL} = 911.75$ Å is the wavelength of the Lyman limit and λ_e is the rest wavelength of the spectrum at which the transmission equals e^{-1}. To calculate $\lambda_{\rm mfp}$, we then compute the physical distance traveled by the

[11] Line opacity only reduces the transmission of non-ionizing photons which are unimportant for consideration of $\lambda_{\rm mfp}$.

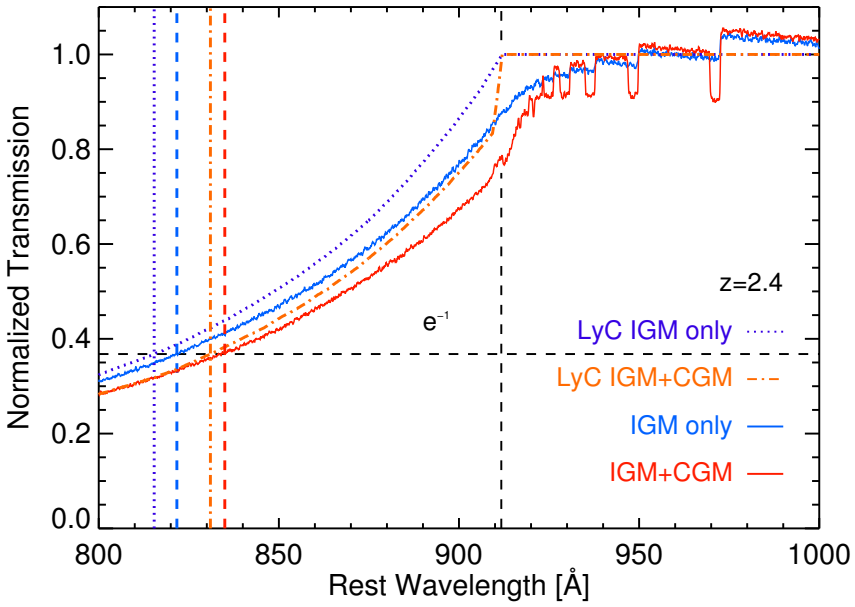


Figure 4.12 The average normalized transmission below the Lyman limit of a $z = 2.4$ emitter. The red and orange curves pertain to transmission including the CGM and the blue and purple curves pertain to IGM only transmissions. The dotted (purple) and dot-dashed (orange) lines show the same spectra as in Figure 4.11 from the continuum opacity only runs of the MC code. The solid (red and blue) curves show the spectra including line opacity normalized by $\langle 1-D_B \rangle$. The black vertical dashed line marks the position of the Lyman limit in the frame of the emitter. The dashed horizontal line shows where the normalized transmission is equal to e^{-1}. The colored vertical lines mark the position where each of the spectra crosses the e^{-1}, signifying the position of one $\lambda_{\rm mfp}$. These rest wavelengths are converted into physical distances corresponding to the $\lambda_{\rm mfp}$ as described in the text.

photon through the expanding universe, given by the integral over the proper line element:

$$\lambda_{\rm mfp} \equiv \int_{z_{\rm mfp}}^{z_{\rm em}} \frac{1}{1+z} \frac{c}{H(z)} dz \tag{4.26}$$

where c is the speed of light and

$$H(z) = H_0 \sqrt{\Omega_\Lambda + \Omega_{\rm m}(1+z)^3}. \tag{4.27}$$

Using the averaged LyC only spectra shown in Figure 4.12, we calculate the $\lambda_{\rm mfp}$ through the IGM assuming $z_{\rm em} = 2.4$ to be 147 pMpc (or $\Delta z = 0.359$). If instead the CGM is included (as would be the case for LyC photons emanating from a galaxy), the estimated $\lambda_{\rm mfp}$ falls to 120 pMpc ($\Delta z = 0.301$), $\sim$20% shorter. Note that the values of $\lambda_{\rm mfp}$ found in this sections using the MC simulations are similar to those estimated using the analytic approximations in Section 4.5 ($\lambda^{\rm static}_{\rm mfp,\ IGM}$= 122.8 pMpc, $\lambda^{\rm static}_{\rm mfp}$= 109.3 pMpc). This provides confidence that the conclusions of Section 4.5 hold even when a more rigorous treatment is considered.

4.6.3 Computing the Mean Free Path Uncertainties

To calculate the uncertainties in the $\lambda_{\rm mfp}$ measurement made in Sections 4.6.2, we must estimate the error in the opacity that originates from the Lyα forest absorbers ($\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$) and from LLSs ($\log(N_{\rm HI}/{\rm cm}^{-2}) > 17.2$). The errors associated with each must be estimated separately because the constraints on the incidence of these absorbers come from different surveys.

To estimate the uncertainty associated with LLSs, we return to the measured value of $dn_{\rm LLS}/dX = 0.52 \pm 0.08$ (Section 4.5). Using the method outlined in Section 4.5, we employ a fit to the frequency distribution to calculate the opacity, κ. Once again, we assume that an extrapolation of the broken power law fit to the slope of the high-$N_{\rm HI}$ absorbers holds in the LLS regime. We then calculate the normalization of the frequency distribution that reproduces $dn_{\rm LLS}/dX$ and its upper and lower limits. These normalizations and the assumed slope can then be used to measure $\kappa_{\rm LLS}$ and an approximate error. We find $\kappa_{\rm LLS} = 0.0043 \pm 0.0006$ Mpc^{-1}, a 15% uncertainty. If we allow the value of β to vary within the $1-\sigma$ confidence region (see Table 4.6), the uncertainty in our estimate of $\kappa_{\rm LLS}$ would be larger, $\sim 20-25\%$.

To estimate the $\lambda_{\rm mfp}$ uncertainties from absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ we employ the measurement of $\kappa_{\rm data}$ from Section 4.5. We thus assume that the uncertainty is dominated by sample variance rather than the uncertainties in the model. However, before using this value, we verify that the opacity (and its uncertainty) from $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ absorbers measured from the data are reproduced in the Monte Carlo simulations using the assumed power-law fits.

To estimate the opacity in the MC runs from forest absorbers, we ran another set of MC simulations of LyC only opacity. In this case, we included only absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$, excluding the contribution from LLSs. Using the ensemble of model spectra, we performed a jack-knife test with a sample size of 15 sightlines, designed to match the number of lines of sight in the KBSS QSO sample. Each set of 15 randomly-drawn spectra was averaged and the maximum attenuation with respect to the continuum was measured and recorded. The rest-frame wavelength at the position of maximum attenuation was converted into a proper distance at the emission redshift using equation 4.26. The average value of κ along those 15 lines of sight was taken to be the natural logarithm of the minimum flux divided by physical distance. The redshift range used in the simulations is $2.0 < z < 2.8$ in order to reproduce the range of the KBSS QSO sightline sample.

The jack-knife technique was applied 10,000 times to produce a distribution of the forest opacity, $\kappa_{\rm for}$, as shown in Figure 4.13. An estimate for the error in $\kappa_{\rm for}$ was then derived from the moments of the distribution. The resulting value of $\kappa_{\rm for} = 0.0037 \pm 0.0005$ is in good agreement with $\kappa_{\rm data} = 0.0039 \pm 0.0005$ calculated in Section 4.5, providing confidence that the MC simulations well reproduce both the mean value and the dispersion in $\kappa_{\rm for}$ for a sample of the same size as the KBSS.

With an estimate of $\kappa_{\rm LLS}$, $\sigma_{\kappa,{\rm LLS}}$ as well as $\kappa_{\rm for}$ and $\sigma_{\kappa,{\rm for}}$, it is straightforward to calculate the error in the

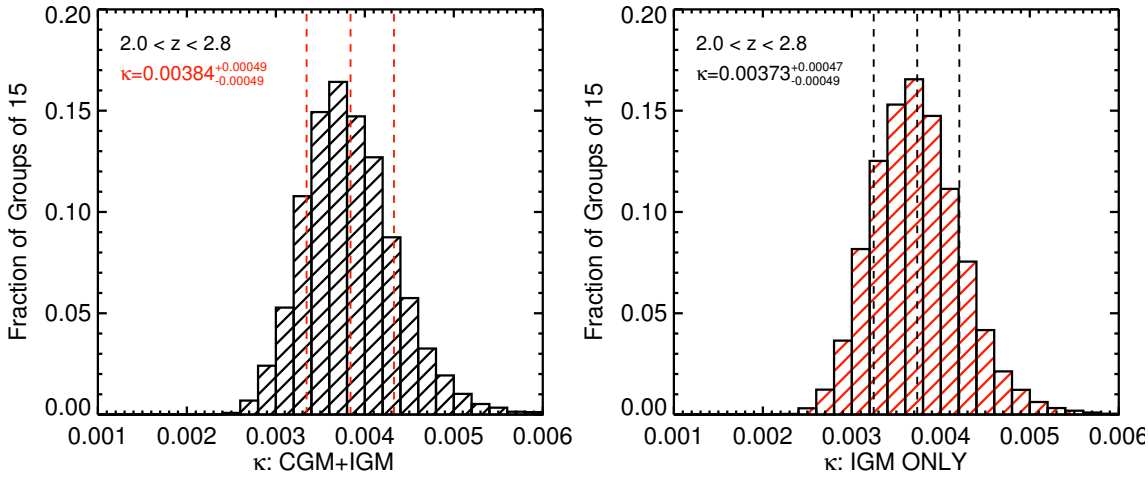


Figure 4.13 The variation in the measured opacity in groups of 15 sightlines due to all absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$. The simulations used in this plot are those with emission redshift $z = 2.8$. These show the results of 10,000 jack-knife tests in which 15 spectra were randomly drawn from the "LyC ONLY" Monte Carlo runs in which only absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ were included. See Section 4.6.3. The dotted vertical lines mark the $1-\sigma$ uncertainties in κ.

$\lambda_{\rm mfp}$. If

$$\lambda_{\rm mfp} = \frac{1}{\kappa_{\rm LLS} + \kappa_{\rm for}} \tag{4.28}$$

(see equation 4.23) then the uncertainty on $\lambda_{\rm mfp}$ is

$$\sigma_{\rm mfp} = \frac{\sqrt{\sigma^2_{\kappa,{\rm LLS}} + \sigma^2_{\kappa,{\rm for}}}}{(\kappa_{\rm LLS} + \kappa_{\rm for})^2} \tag{4.29}$$

where $\sigma_{\kappa,{\rm LLS}}$ and $\sigma_{\kappa,{\rm for}}$ are the uncertainties of $\kappa_{\rm LLS}$ and $\kappa_{\rm for}$ respectively. Employing these equations, we find a typical uncertainty in $\lambda_{\rm mfp}$ of $\sim$ 15 Mpc, which we adopt for both the IGM and IGM+CGM $\lambda_{\rm mfp}$ measurements. Since the contribution of LLSs to the opacity in the CGM is poorly constrained, the error on the CGM measurements is likely somewhat larger; however since the CGM contributes $< 20\%$ of the opacity, the effect on the net uncertainty is modest. We adopt this uncertainty for the measurement of $\lambda_{\rm mfp}$ at the mean redshift of the sample.

4.6.4 Simulated Sightlines with Line and Continuum Opacity

For some applications, it is more useful to know the expected attenuation along an ensemble of sight lines including line opacity. For instance, in studies of the escape fraction of LyC photons (Shapley et al. 2006; Iwata et al. 2009; Bridge et al. 2010; Siana et al. 2010; Nestor et al. 2011; Steidel et al. in prep) a correction must be made for the fraction of the attenuation at a given observed wavelength due to (non-ionizing) line blanketing. For these applications, our models including CGM attenuation are ideal. The sample of galaxies used in our CGM study have impact parameters of $50-300$ pkpc from the QSO line of sight, and so our MC

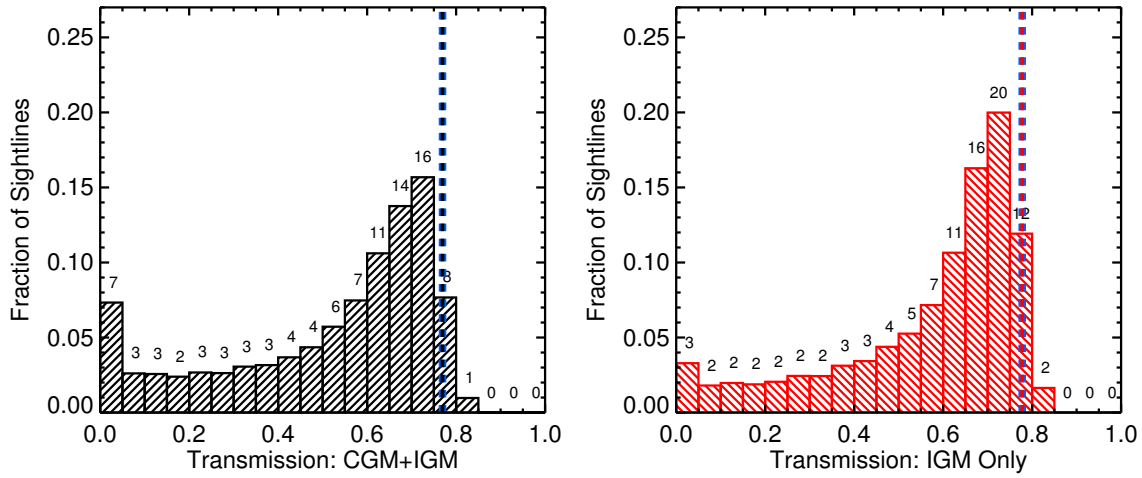


Figure 4.14 The distribution of transmissions from each sightline averaged over the band pass from 880-910 Å for the the IGM+CGM Monte Carlo simulations (left panel) and IGM only Monte Carlo simulations (right panel) that include line opacity. The simulations used for these figures are those with an emission redshift of $z = 2.4$. The (blue) dashed vertical line marks the average transmission in the Ly$\beta,\gamma,+$ forest $\langle 1-D_B \rangle$.

Table 4.8. Comparison of the $\lambda_{\rm mfp}$ Derived from the LyC and the Renormalized Forest Spectra

$z_{\rm em}$	$\lambda_{\rm mfp}$ LyC [pMpc] IGM ONLY	$\lambda_{\rm mfp}\langle 1-D_B \rangle$ [pMpc] IGM ONLY	$\lambda_{\rm mfp}$ LyC [pMpc] IGM+CGM	$\lambda_{\rm mfp}\langle 1-D_B \rangle$ [pMpc] IGM+CGM
2.2	191.8	185.9	160.2	151.4
2.4	146.5	135.7	120.0	113.4
2.6	115.7	106.8	93.7	85.6
2.8	91.9	85.1	73.5	66.0

simulations which include opacity from the CGM effectively model the typical gas distribution at radii larger than 50 kpc - and hence the attenuation suffered by photons after they leave the *ISM* of the galaxy, but before they leave the large gas densities associated with the CGM.

Using the sets of simulated spectra with $z_{\rm em} = 2.4$ that include Lyman line opacity for the IGM ONLY and the IGM+CGM, we measure the distribution of transmission of emitted 880-910 Å photons through the IGM and the IGM + CGM; see Figure 4.14. The transmission is generally lower along the IGM+CGM sight lines. The peak in the transmission PDF near ~ 0.75 corresponds to the average transmission within the Lyβ forest which is < 1 due to line blanketing from both higher Lyman series lines as well as Lyα absorption from lower redshift.

In Figure 4.12, the solid curves show the average spectrum of 10,000 MC runs that include line opacity. We use $\langle 1-D_B \rangle$, the average transmission in the region between 920 -1015 Å in the rest frame of the source (the region shown in the yellow box in Figure 4.11; Oke & Korycansky 1982) to "re-normalize" the continuum against which the LyC opacity will be measured. Using the renormalized spectrum as an approximation

to the "LyC only" MC spectra and measuring $\lambda_{\rm mfp}$ as described in Section 4.6.2, leads to $\lambda_{\rm mfp}$ values $5-10\%$ smaller than the "true" values. Table 4.8 compares the results of the MC simulations including line opacity (columns labeled $\langle 1-D_B \rangle$) with those based on the continuum only spectra (columns labeled LyC).

4.6.5 Comparison with Previous $\lambda_{\rm mfp}$ Measurements

In the previous sections, we derived the value of the mean free path $\lambda_{\rm mfp}$ to LyC photons traveling through the IGM to be:

$$\text{IGM ONLY: } \lambda_{\rm mfp}(z_{\rm em} = 2.4) = 147 \pm 15 \text{ Mpc}. \tag{4.30}$$

If instead, we assume such LyC photons emanate from galaxies similar to one of those in our spectroscopic sample, we find a value for $\lambda_{\rm mfp}$ including the CGM opacity to be:

$$\text{IGM+CGM: } \lambda_{\rm mfp}(z_{\rm em} = 2.4) = 121 \pm 15 \text{ Mpc}. \tag{4.31}$$

One important advance of the estimate of $\lambda_{\rm mfp}$ made in this paper is that all parts of the column density distribution have now been measured at the same mean redshift. This has not been possible previously due to the difficulty of collecting appropriate statistical samples for the various column density regions of the frequency distribution.

In this section, we compare these values to previous estimates from the literature. One caveat is that the measurements made in this paper typically differ in mean redshift from the samples on which previous $\lambda_{\rm mfp}$ measurements have relied. As such, an extrapolation of the literature measurements or our own measurements is generally necessary in order to compare the results, and therefore these comparisons are more dependent on the value of γ assumed. Nevertheless, in this section we compare to previous measurements for completeness. Figure 4.15 gives a graphical summary of the results of this comparison.

In general, the differences between our IGM ONLY $\lambda_{\rm mfp}$ results and those of other authors (especially those measured at similar z) underscore the degree to which $\lambda_{\rm mfp}$ is sensitive to changes in $f(N,X)$, highlighting the importance of the precise measurement of the intermediate $N_{\rm HI}$ systems made here.

One other caveat is the small degree of incompleteness which may be present in the $z \lesssim 2.4$, $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 15.5$ portion of our absorber sample. As described in Appendix A.1, this bias will only act to *shorten* the inferred value of $\lambda_{\rm mfp}$, pushing it to values farther below those inferred by other authors.

The mean free path and its evolution with redshift are generally parameterized as:

$$\lambda_{\rm mfp} = \lambda_{\rm mfp,0}\,(1+z)^{\xi} \tag{4.32}$$

where ξ captures the cosmology (as parameterized in equation 4.26) as well as the evolution in the number of absorbers as a function of redshift, dn/dz (described in Section 4.6.1). For the Monte Carlo simulations described in this section, the redshift evolution of $\lambda_{\rm mfp}$ is fixed by the assumed value of γ. We emphasize

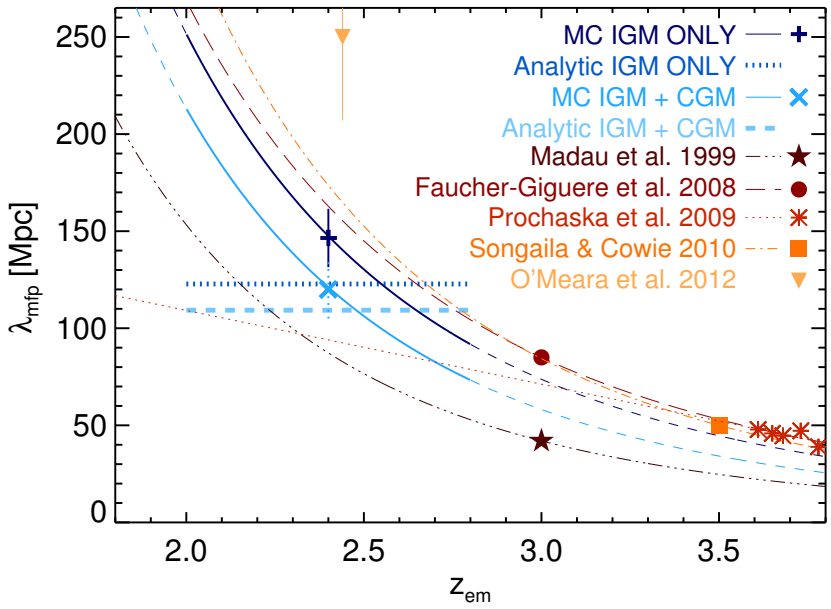


Figure 4.15 The $\lambda_{\rm mfp}$ as a function of $z_{\rm em}$ including previous estimates of the $\lambda_{\rm mfp}$ from the literature. The blue points and curves refer to values measured in this paper whereas estimates from the literature are plotted in shades of red. The blue cross and X and their corresponding curves show the values of $\lambda_{\rm mfp}$ calculated from the MC spectra for the IGM ONLY and IGM+CGM runs respectively. The dotted and dashed blue horizontal line show the results of the analytic calculation in Section 4.5 excluding and including the CGM respectively. The dark red filled star and the dash-dotted curve that intersects it are the suggested values of $\lambda_{\rm mfp}$ from Madau et al. (1999). The red filled circle and dashed curve are the results of Faucher-Giguère et al. (2008). Prochaska et al. (2009) suggested the values shown with the red/orange asterisks and the dotted line. The measurements of Songaila & Cowie (2010) are shown in the orange filled square and the dash-dotted curve that intersects it. The results of O'Meara et al. (2012) are shown in the orange filled triangle.

that the redshift evolution is only constrained by data with $2.0 \lesssim z \lesssim 3.0$, and as such the extrapolation of our results above to $z >> 3$ is ill-advised.

The redshift evolution of $\lambda_{\rm mfp}$, parameterized as a function of the emission redshift of the QSO, $z_{\rm em}$, returned by the MC simulations is represented by the solid blue curves in Figure 4.15 and is given by:

$$\text{IGM ONLY: } \lambda_{\rm mfp}(z_{\rm em}) = 147\left(\frac{1+z_{\rm em}}{3.4}\right)^{-4.3}\ \text{pMpc} \tag{4.33}$$

and

$$\text{IGM+CGM: } \lambda_{\rm mfp}(z_{\rm em}) = 121\left(\frac{1+z_{\rm em}}{3.4}\right)^{-4.5}\ \text{pMpc.} \tag{4.34}$$

The first estimate of $\lambda_{\rm mfp}$ was made by Madau et al. (1999) who found:

$$\lambda_{\rm mfp} = 42\left(\frac{1+z}{4.0}\right)^{-4.5}\ \text{Mpc} = 87\left(\frac{1+z}{3.4}\right)^{-4.5}\ \text{Mpc} \tag{4.35}$$

using a simple approximation for $f(N,X)$ with $\beta = 1.5$ and $\gamma = 2$ where their Einstein-de Sitter cosmology

has been converted to the cosmology used in this paper.

Faucher-Giguère et al. (2008) published an updated estimate to the $\lambda_{\rm mfp}$ assuming $\beta = 1.39$ as measured by Misawa et al. (2007). Misawa et al. (2007) measured the frequency distribution close to high-$N_{\rm HI}$ absorbers, and thus their sample is similar in slope to our CGM $f(N,X)$ distribution. However, the normalization chosen by Faucher-Giguère et al. (2008) was computed using LLS statistics for the full IGM, thereby closely matching our LLS points but somewhat under-predicting the incidence of lower-$N_{\rm HI}$ absorbers resulting in slightly larger values of $\lambda_{\rm mfp}$ at a given z:

$$\lambda_{\rm mfp} = 85\left(\frac{1+z}{4}\right)^{-4}\ {\rm Mpc} = 160\left(\frac{1+z}{3.4}\right)^{-4}\ {\rm Mpc}. \tag{4.36}$$

Prochaska et al. (2009) estimated $\lambda_{\rm mfp}$ using stacked QSO spectra to measure the LyC opacity. Notably, such a method would include any contribution from the CGM of their QSOs. They found

$$\lambda_{\rm mfp} = 48.4 - 38.0(z-3.6)\ {\rm Mpc} = 94 - 38.0(z-2.4)\ {\rm Mpc} \tag{4.37}$$

somewhat lower than our measurement at $z = 2.4$ and with very different evolution with redshift than indicated by our data.

Songaila & Cowie (2010) revisited the mean free path calculation after compiling a large sample of $z < 1$ and $z > 4$ LLS. They found

$$\lambda_{\rm mfp} = 50\left(\frac{1+z}{4.5}\right)^{-4.44^{+0.36}_{-0.32}}\ {\rm Mpc} = 170\left(\frac{1+z}{3.4}\right)^{-4.44^{+0.36}_{-0.32}}\ {\rm Mpc} \tag{4.38}$$

using $\beta = 1.3$. They also offer a variety of normalizations for different β values, including 130 Mpc for $\beta = 1.5$ at $z = 2.4$, in general agreement with our measured "IGM ONLY" value of $\lambda_{\rm mfp}$.

Recently O'Meara et al. (2012) presented a calculation of $\lambda_{\rm mfp}$ using their new measurements of the incidence of LLSs at $z \sim 2$. Their estimates of the incidence of absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ leads to a lower value for the opacity as described in Section 4.5.1. They found

$$\lambda_{\rm mfp} = 250 \pm 43\ {\rm Mpc} \tag{4.39}$$

at $z = 2.44$, nearly a factor of 2 larger than our estimate and significantly larger than the extrapolations of others' measurements at higher redshifts.

We note that for a steeper evolution in redshift as found in Songaila & Cowie (2010) and Prochaska et al. (2010), the general agreement between the extrapolation of the presented measurements at $\langle z \rangle = 2.4$ to $z > 3$ would result in a more discrepant value significantly below the measurements at $z > 3.5$ by Songaila & Cowie (2010) and Prochaska et al. (2009).

To summarize, the calculation described in Section 4.6 offers three significant improvements over previ-

ous measurements. First, all portions of $f(N,X)$ have been measured at the same mean redshift, avoiding the need to extrapolate any part of $f(N,X)$ to a different redshift using uncertain values of γ. A second advance is the accuracy of our Voigt profile fitting technique, achieved by including higher-order Lyman series transitions, yielding the best measurement of the power law slope, β, for absorbers with saturated Lyα profiles, $[14 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2]$. Because such absorbers account for $\sim$half of the opacity to LyC photons in the IGM, underestimating their contribution would result in significantly underestimating the total opacity of the IGM.

The third major improvement with respect to previous studies is the inclusion of absorption from the CGM which again decreases the measured value of $\lambda_{\rm mfp}$. If LyC photons are presumed to originate within galaxies similar to those in the KBSS sample, then every hydrogen-ionizing photon must first propagate through the CGM. As shown by Rudie et al. (2012b), regions of the IGM near galaxies have a much higher incidence of high-$N_{\rm HI}$ absorbers than the average IGM with the effect (again) of reducing $\lambda_{\rm mfp}$ relative to previous measurements.

4.6.6 Implications of the Higher Opacity of the IGM+CGM

A shorter mean free path has implications for the emissivity and demographics of ionizing sources at high redshift, as well as for calculations of the metagalactic UV radiation field. Briefly, if one considers the specific intensity of the UV background (J_ν, see e.g. Scott et al. 2000) or the photoionization rate (Γ, see e.g. Faucher-Giguère et al. 2008) to be well determined, a shorter mean free path directly implies a requirement for correspondingly larger ionizing emissivity of sources (ϵ_ν). Following Faucher-Giguère et al. (2008),

$$J_\nu(z) \approx \frac{1}{4\pi}\lambda_{\rm mfp}(\nu,z)\,\epsilon_\nu(z) \tag{4.40}$$

and

$$\Gamma \propto J_\nu. \tag{4.41}$$

Conversely, for a measured value of ϵ_ν from a given population of sources, a shorter $\lambda_{\rm mfp}$ implies a smaller contribution to both the specific intensity and the photoionization rate (Γ). Higher intergalactic opacity also requires that the background be produced by sources within a smaller volume, possibly leading to greater spatial variation in the radiation field intensity.

LyC escape fraction studies have often found very few detections of ionizing photons escaping from galaxies (Shapley et al. 2006; Bridge et al. 2010; Siana et al. 2010; however see Iwata et al. 2009, Nestor et al. 2011, and Steidel et al. in prep). For LyC studies of galaxies at high-z, the inclusion of the CGM is an important effect. Considering Figure 4.14, note that the number of galaxies with $\sim$zero transmission increases by a factor of 2 when the CGM is included. If the redshift evolution of the CGM is similar to that measured for the IGM, the effect of the CGM will become more pronounced at higher-z where the incidence

of high-$N_{\rm HI}$ systems in the CGM will be much higher.

Higher IGM and CGM opacities mean that there may be more sources contributing to the metagalactic UV radiation field than are implied by the number of actual LyC detections; on the other hand, the assumption of a $\lambda_{\rm mfp}$ that is too large may cause one to over-estimate the contribution to J_ν or Γ made by a population of detected LyC sources. The latter effect may reduce the tension between the apparent over-production of ionizing photons by observed LBGs at $z \sim 3$ (Nestor et al. 2011) relative to estimates of the total photoionization rate produced by all sources (Bolton et al. 2005; Faucher-Giguère et al. 2008).

4.7 Summary

Using a sample of 15 high-S/N, high-resolution QSO spectra drawn from the KBSS, we have produced the largest catalog of H I absorbers fit with Voigt profiles to Lyα and at least one higher order Lyman series transition. This analysis enables the first statistically rigorous measurements of the frequency distribution of H I absorbers with $14 \lesssim \log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 17$. In Section 4.3, we showed that the frequency of absorbers as a function of $N_{\rm HI}$ is well-parameterized by a single power-law from $13.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ with a maximum likelihood index of $\beta = 1.65 \pm 0.02$ and a normalization per unit pathlength of $\log(C_{\rm HI}) = 10.32$.

In Section 4.4, we measured the frequency distribution within 300 pkpc and both 300 km s^{-1} and 700 km s^{-1} of galaxies in the KBSS sample. We showed that the frequency of absorbers near galaxies is significantly higher and that the power-law index is shallower compared with that of the IGM, meaning there are disproportionately more high-$N_{\rm HI}$ systems than low. These findings are discussed further by Rudie et al. (2012b).

Section 4.5 examined statistically the total opacity of the IGM and CGM to hydrogen-ionizing Lyman continuum (LyC) photons. This section presented measurements of the opacity due to absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ with direct constraints from the data. We also found that a single MLE power-law fit over the full range of $N_{\rm HI}$ does not well reproduce this opacity, and also fails to reproduce the incidence of LLSs. We therefore adopted a broken power law fit that reproduces both the opacity measured within the data from absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ as well as the observed incidence of LLSs. The parameters of these fits can be found in Table 4.5.

We also measured the fractional LyC opacity in bins of $N_{\rm HI}$. We found that 48% of the opacity within one mean free path, $\lambda_{\rm mfp}$, is contributed by absorbers with $N_{\rm HI} < 10^{17.2}$ cm^{-2}, the vast majority of which results from absorbers with $14 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ whose frequency was poorly measured prior to this work.

In Section 4.6, we used Monte Carlo simulations to measure the value of the mean free path ($\lambda_{\rm mfp}$) of LyC photons at $\langle z \rangle = 2.4$. We measured both the $\lambda_{\rm mfp}$ of photons through the IGM (as has been considered many previous times) as well as the $\lambda_{\rm mfp}$ of photons formed in galaxies similar to those in the KBSS - photons that must first transverse the CGM of these galaxies before reaching the lower-opacity IGM. We found $\lambda_{\rm mfp}(z = 2.4) = 147 \pm 15$ Mpc for simulations including only IGM opacity and $\lambda_{\rm mfp}(z = 2.4) = 121 \pm 15$

Mpc when we include the CGM. These values of $\lambda_{\rm mfp}$ are lower than most previous estimates, which have not included the CGM and have generally underestimated the contribution of intermediate-$N_{\rm HI}$ absorbers. We also quantified the effect of the CGM on the distribution of transmission through random sight lines, relevant to measurements of the escape fractions of LyC photons from such galactic sources at high redshift.

We note that the $\lambda_{\rm mfp}$ value including attenuation from the CGM is relevant only if the dominant sources of LyC photons are galaxies similar to the UV color-selected galaxies in our sample ($0.25 < L/L^* < 3.0$ at $\langle z \rangle = 2.3$) - for other galaxies, these measurements are an approximation. However, in all cases the general statement holds that $\lambda_{\rm mfp}$ and related statistics are affected by the $\sim$Mpc-scale gaseous environments of the sources that contribute significantly to the metagalactic background.

The IGM and CGM opacities measured in this paper have significant implications for studies of ionizing sources at high redshift and for estimates of the metagalactic UV background at $z \approx 2-3$.

Acknowledgements

The authors wish to thank Claude-André Faucher-Giguère for his careful reading of the draft and pertinent comments. We also thank Olivera Rakic for her contributions to the reduction of the QSO data set and for her helpful advice. The authors wish to acknowledge Ryan Cooke who contributed the fits to the damped profiles in our QSO spectra. Our thanks to Bob Carswell for his assistance with VPFIT.

We wish to acknowledge the staff of the the W.M. Keck Observatory whose efforts insure the telescopes and instruments perform reliably. Further, we extend our gratitude to those of Hawaiian ancestry on whose sacred mountain we are privileged to be guests.

This work has been supported by the US National Science Foundation through grants AST-0606912 and AST- 0908805. CCS acknowledges additional support from the John D. and Catherine T. MacArthur Foundation and the Peter and Patricia Gruber Foundation. This research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration.

Chapter 5

Epilogue

The previous chapters of this thesis presented analysis of the distribution and physical properties of neutral hydrogen found within the $2 < z < 3$ IGM as well as the CGM of star-forming galaxies. Chapter 2 presented the first study of the CGM of high-z galaxies that (a) was based on a large ($N > 100$) sample of galaxies and (b) used line fitting techniques to measure the physical properties of the H I gas. This effort provides a significant step forward in our knowledge of the shape, scale, kinematics, and kinetic properties of gas within the CGM. With absorbers of $N_{\rm HI} \approx 10^{14.5}$ cm^{-2} preferentially residing in the space surrounding relatively massive galaxies, this study finds that a large fraction of what have been considered "intergalactic" absorbers are actually found within the CGM of relatively luminous star-forming galaxies. This implies that most of the *intermediate* and high-$N_{\rm HI}$ absorption systems within the forest are due to gas whose physical properties are likely to have been affected by their proximity to galaxies. The physics of the IGM may be less simple than previously believed.

Chapter 3 and 4 revisit previously measured statistics of the IGM. Chapter 3 focuses on the temperature-density relation in the low-density IGM, finding a "normal" relationship that suggests higher-density regions are somewhat warmer than those at lower density as expected in most theoretical treatments. Chapter 4 focuses on the statistics of the column density distribution of Lyα forest absorbers. New measurements of $f(N_{\rm HI}, X)$ are used to estimate the opacity of Lyα forest absorbers to hydrogen-ionizing photons. This is the first presentation of such a study based on direct measurements of absorbers with $14.0 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ using higher-order Lyman lines. We find that fully half of the opacity to ionizing photons within the IGM is due to optically-thin ($\tau_{\rm LL} > 1$) Lyα forest absorbers. Further, we quantify for the first time the column density distribution of H I within the CGM of star-forming galaxies and their corresponding continuum opacity to ionizing photons. Under the assumption that $L^*_{\rm UV}$ galaxies are significant sources of LyC photons in the high-z universe, we consider the effect of CGM absorbers on the mean free path. Collectively, these measurements imply that the ionizing UVB may be less spatially homogeneous than previous measurements would suggest.

In what remains of this thesis, I present my current work which focuses on metal absorbers within the CGM of the KBSS galaxies.

5.1 Metal Absorption in the CGM of $2 < z < 3$ Star-Forming Galaxies

Previous studies of metallic absorbers near galaxies at high-z have found evidence for solar metallicity gas $\gtrsim$ 100 kpc from a galaxy (Simcoe et al. 2006), O VI absorption likely associated with intergalactic shocks (Simcoe et al. 2006), and strong evidence of the direct association of C IV absorbers with high-z star-forming galaxies (Adelberger et al. 2003, 2005a). Collectively, this evidence suggests that high-z metal absorbers likely originate within metal-enriched galactic-scale outflows from forming galaxies; however, such conclusions remain tentative.

My current work is focused on significantly extending these previous studies using metal line Voigt-profile decomposition of the full KBSS sample. The UV metal absorption transitions detected in the HIRES QSO spectra commonly include C IV, O VI, and Si IV. Fits to these and other metallic species, combined with the H I measurements and photoionization models, will provide measurements of the metallicity and mass contained within the CGM. With preliminary analysis, I have found that roughly 40% of galaxies within 1400 kpc of the line of sight have detectable C IV absorption within $\pm$ 700 km s^{-1}, and that the frequency of C IV absorption is higher near galaxies within $\sim$ 500 kpc of the line of sight.

An important advance of this study will be the analysis of O VI across all 15 sight-lines. The presence of O VI, with an ionization potential of 114 eV, is a strong indication of highly-ionized, hot ($T \approx 10^5 - 10^{5.5}$ K) gas often from shock-heated regions. O VI is also a unique tracer of the boundary layer between the hot wind fluid observed locally in X-rays and the cool 10^4 K gas observed as lower-ionization absorbers believed to be entrained in the fluid. Because the ionization potential of O VI corresponds to the peak in the cooling curve, gas at temperatures characteristic of O VI is thermally unstable and therefore is typically cooling through that temperature. As such, O VI absorption likely traces *recent* activity in the IGM/CGM. Unfortunately, the wavelength of the O VI doublet falls only $\sim$ 10 Å from Lyβ. Thus at redshift $z \sim$2, where the forest is dense, the de-blending of O VI from contaminating Lyβ absorbers (and lower-redshift Lyα absorbers) is a painstaking but necessary process. I am currently searching for O VI near the redshifts of galaxies using my fit to the Lyβ forest to remove the majority of the contamination (see Figure 5.1).

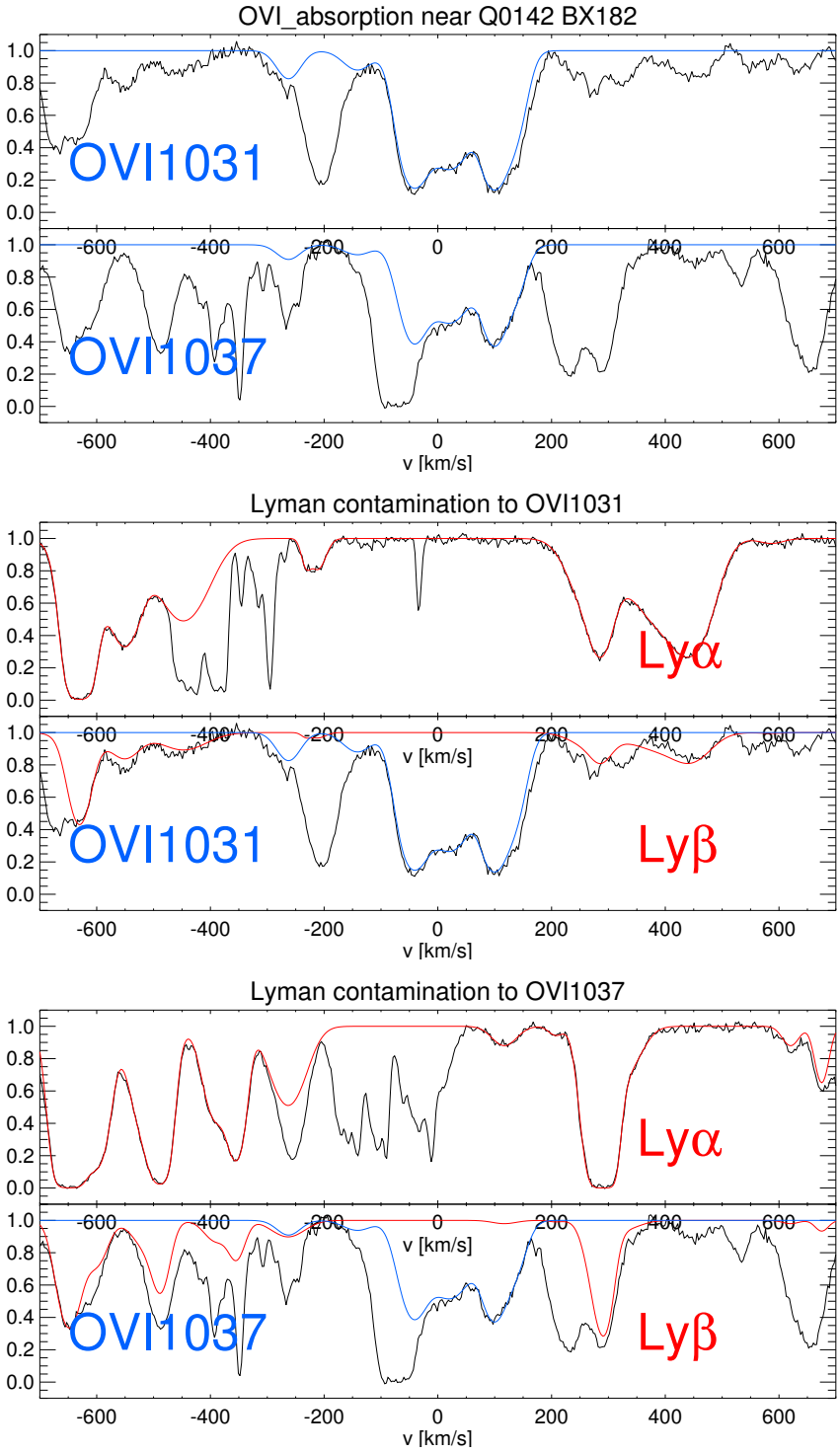


Figure 5.1 Constraints on contamination of the O VI ion due to Lyβ absorption. The top 2 panels show the the portion of the spectrum of Q0142 within $\pm$700 km s^{-1} of the redshift of BX182. The middle two panels show Lyβ contamination to the stronger transition (1031Å) in red along with the corresponding Lyα. The bottom 2 panels show contamination to the weaker O VI transition (1037Å). The fit to O VI at the redshift of the galaxy is shown in blue.

Table 5.1. Highest Column Density Component of Each Species

Galaxy	$D_{\rm tran}$	Log Column Density [cm^{-2}]								
	[pkpc]	H I	Si II	C II	Si III	C III	Si IV	C IV	N V	O VI
Q1442-BX333[a]	50	20.0	13.3	14.0	~ 13.8	$\sim$15.6	13.5	15.0	13.6	14.4
Q1549-D15	56	16.0	12.2	12.1	12.7	13.8	13.0	13.6	12.5	14.0
Q1442-MD50[a]	58	20.0	13.3	14.0	~ 13.8	$\sim$15.6	13.5	15.0	13.6	14.4
Q1549-BX79	62	16.8	12.5	13.1	13.2	> 15.0	12.8	13.5	13.3	14.3
Q0142-BX182	75	18.0	13.4	14.1	14.6	~ 15.7	13.3	14.1	13.6	14.5
Q0821-BX209	80	15.7	< 11.8	< 12.5	12.5	$\cdots$	11.5	12.7	13.0	14.0
Q2343-BX551	87	16.5	12.1	<13.0	13.1	$\cdots$	13.1	14.0	13.5	13.6
Q0100-BX210[b]	89	16.0	< 12.0	< 13.6	< 12.0	12.6	12.2	13.5	$\cdots$	$\cdots$
Q1623-BX432	97	15.6	< 11.5	< 12.0	12.0	$\cdots$	11.9	< 11.0	13.5	13.9

[a]Galaxies BX333 and MD50 both lie within 60 pkpc of the line of sight to Q1442 and within 225 km s^{-1} of each other as well as a sub-DLA ($N_{\rm HI} = 10^{20}$ cm^{-2}).

[b]Due to a large amount of contaminating absorption within the N V and O VI portions of the spectrum surrounding the redshift of Q0100-BX210, we do not attempt a measurement of these ions. See Figure 5.16.

5.1.1 Preliminary Metal-Line Fits

Figures 5.2 - 5.19 present preliminary fits to metal absorption lines within the CGM of galaxies in the KBSS that lie very near (< 100 pkpc) to the QSO sightline. For each galaxy, two sets of ions are considered: low-ionization lines such as C II, Si II, C III, Si III and high-ionization lines such as Si IV, C IV, N V, O VI.

Table 5.1 gives the column density of the strongest absorber of each species [analogous to the max($N_{\rm HI}$) absorber from Chapter 2]. Note that the strongest absorbers of each species do not necessarily result from co-spatial gas and therefore do not always have similar velocities.

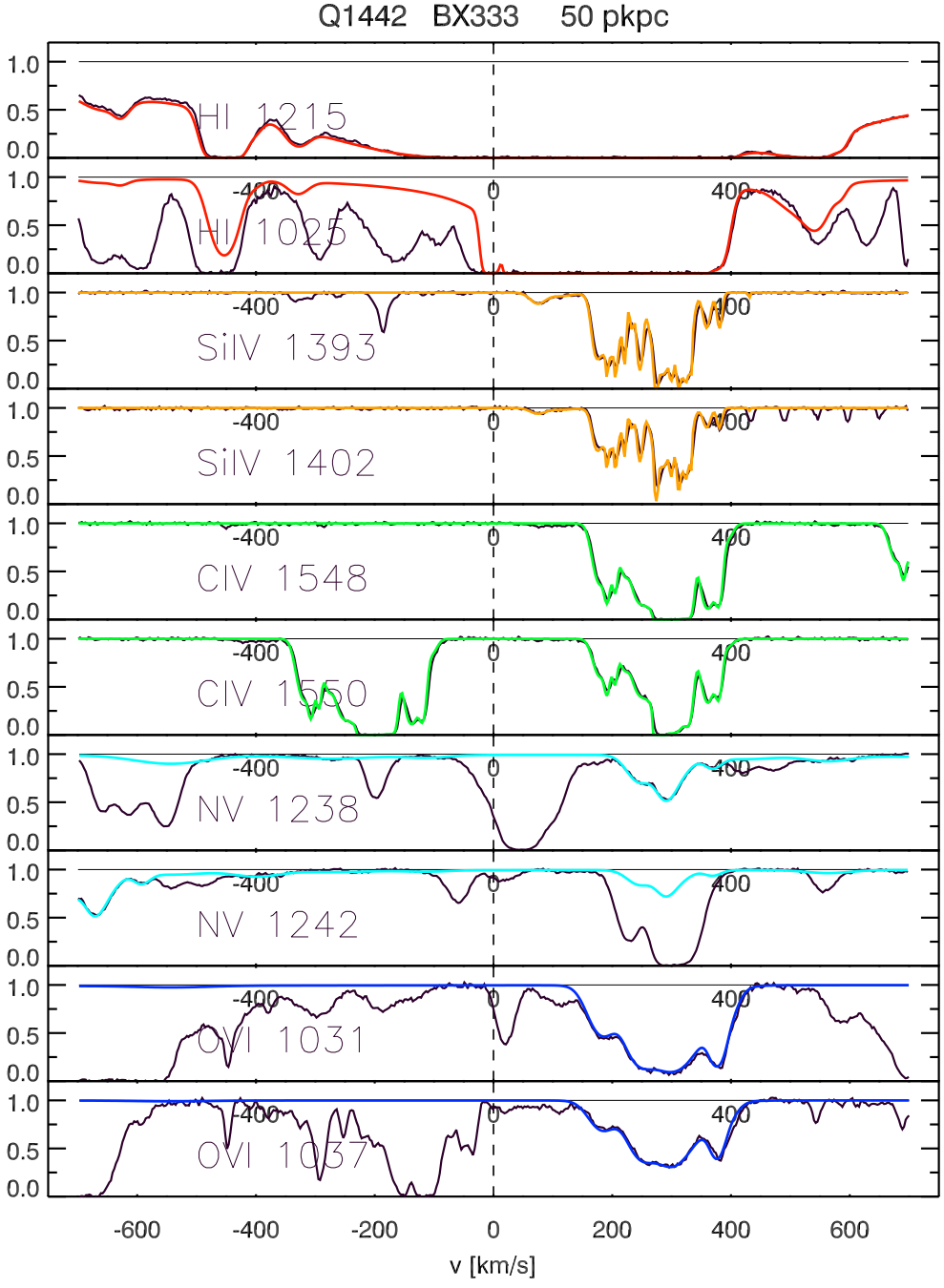


Figure 5.2 H I and high-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1442-BX333 which lies 50 pkpc from the line of the sight to the QSO. Note this galaxy's redshift lies +225 km s^{-1} from that of Q1442-MD50 (Figure 5.6).

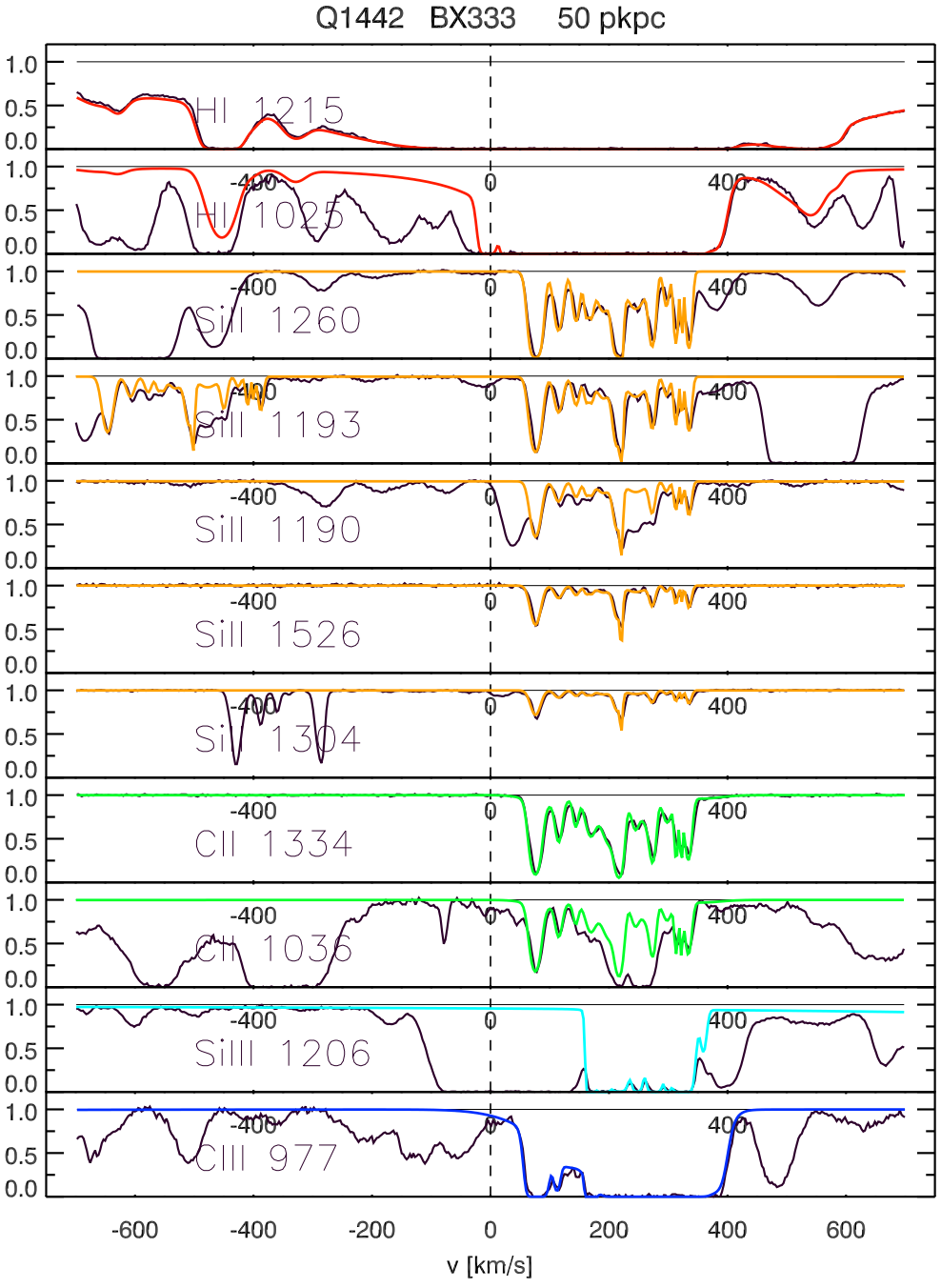


Figure 5.3 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1442-BX333 which lies 50 pkpc from the line of the sight to the QSO. Note this galaxy's redshift lies +225 km s^{-1} from that of Q1442-MD50 (Figure 5.6).

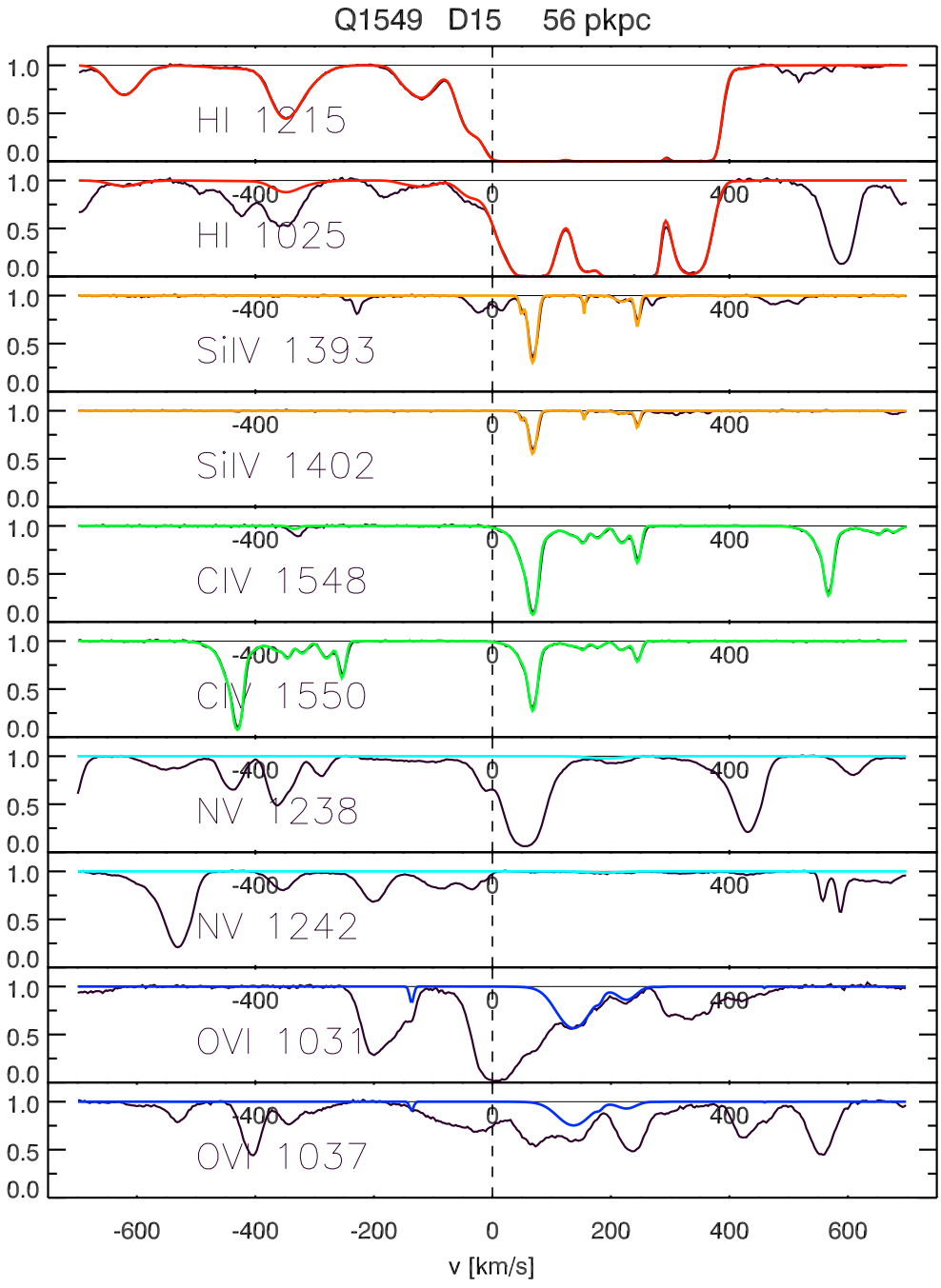


Figure 5.4 H I and high-ionization metal line absorption within $\pm$700 km s^{-1} of the galaxy Q1549-D15 which lies 56 pkpc from the line of the sight to the QSO.

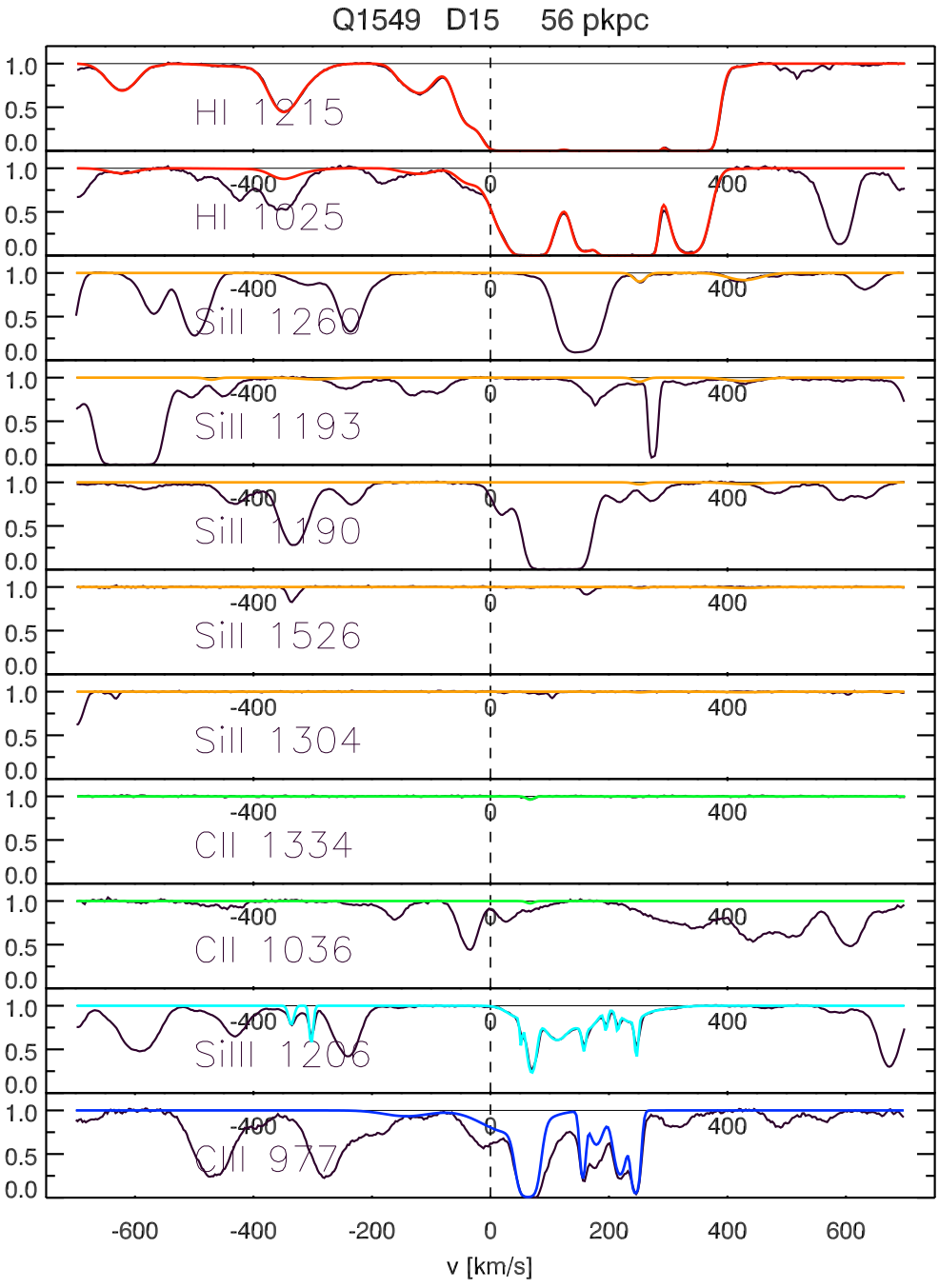


Figure 5.5 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1549-D15 which lies 56 pkpc from the line of the sight to the QSO.

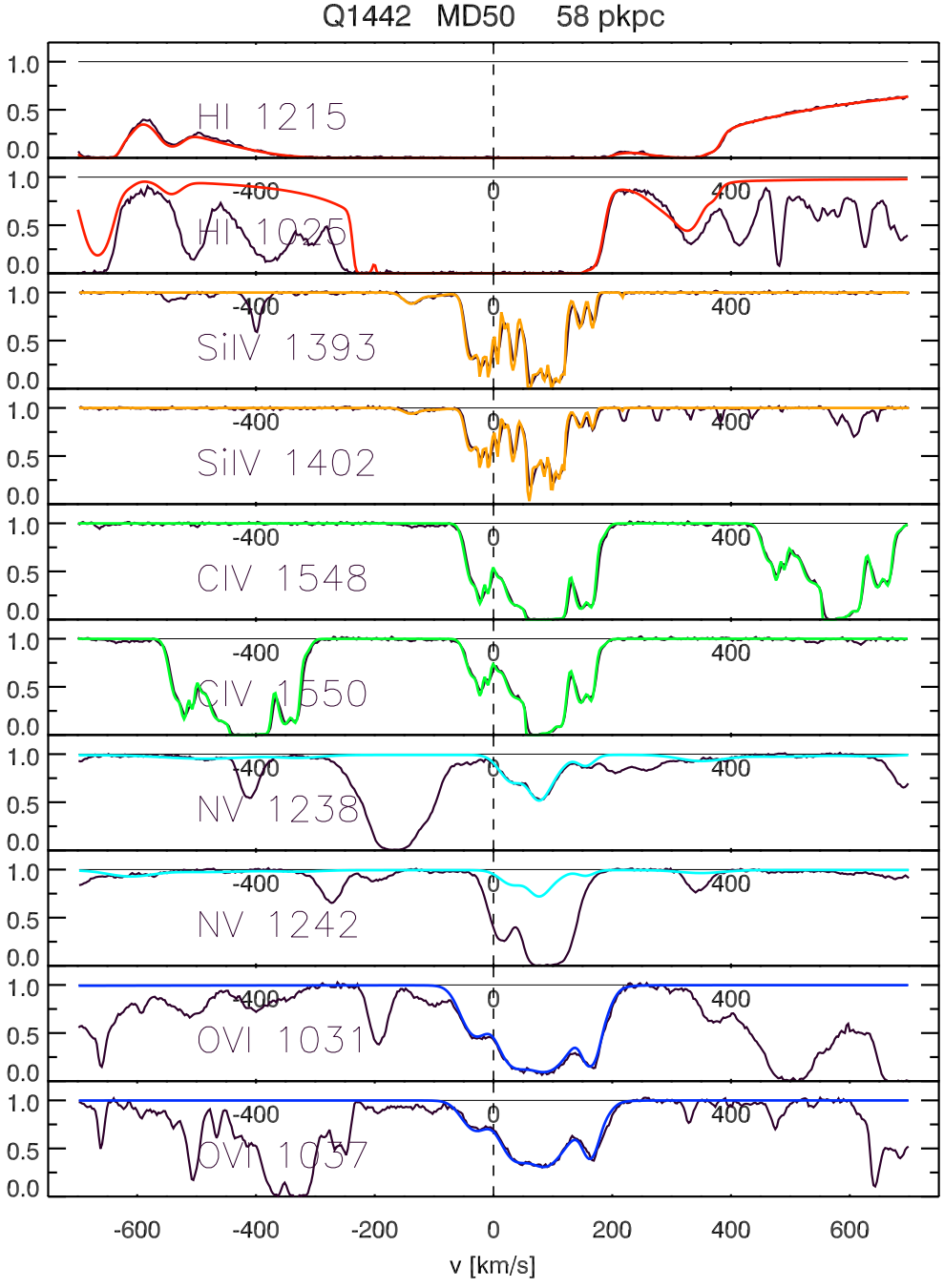


Figure 5.6 H I and high-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1442-MD50 which lies 58 pkpc from the line of the sight to the QSO. Note this galaxy's redshift is –225 km s^{-1} from that of Q1442-BX333 (Figure 5.2).

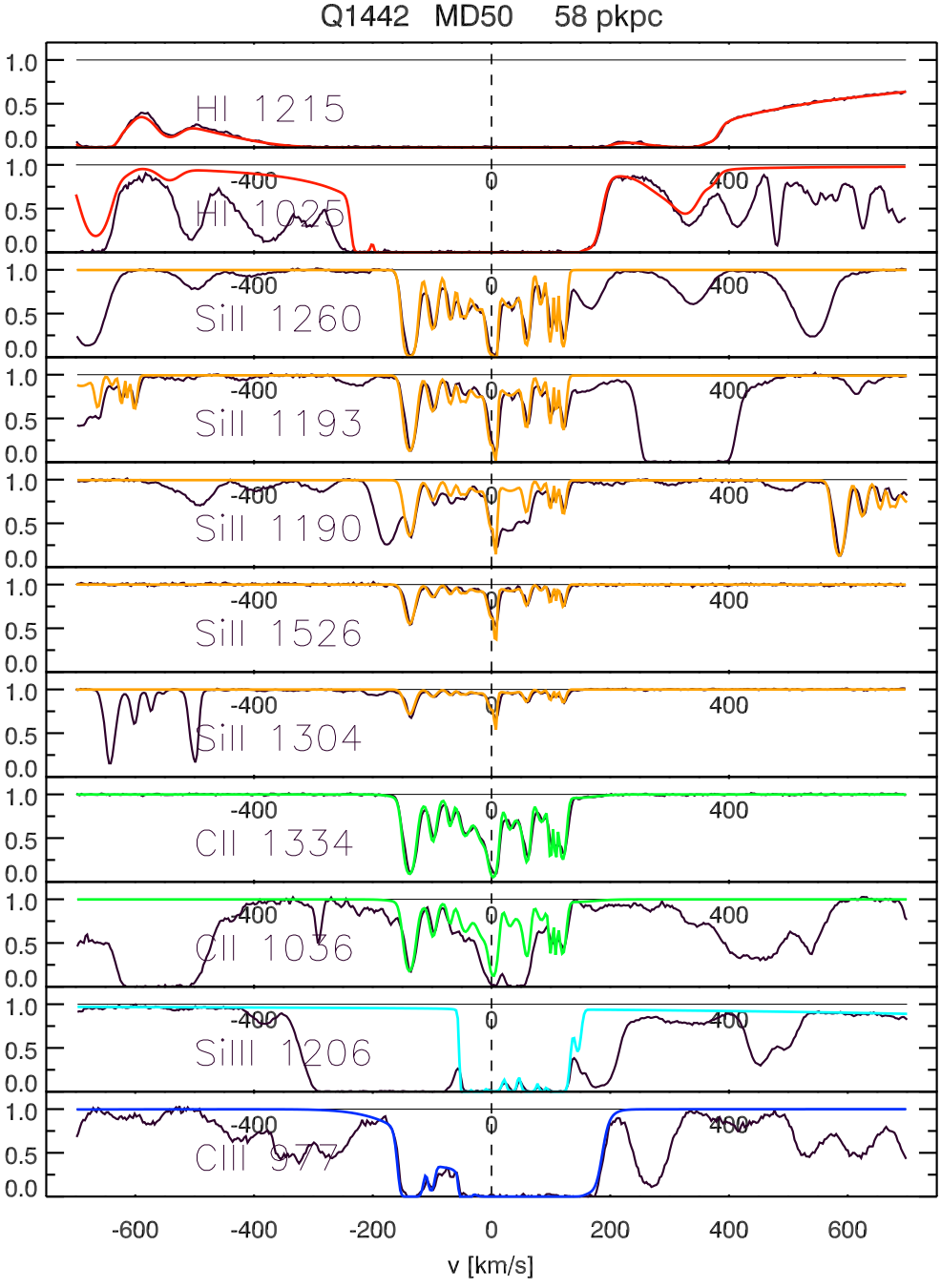


Figure 5.7 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1442-MD50 which lies 58 pkpc from the line of the sight to the QSO. Note this galaxy's redshift is −225 km s^{-1} from that of Q1442-BX333 (Figure 5.2).

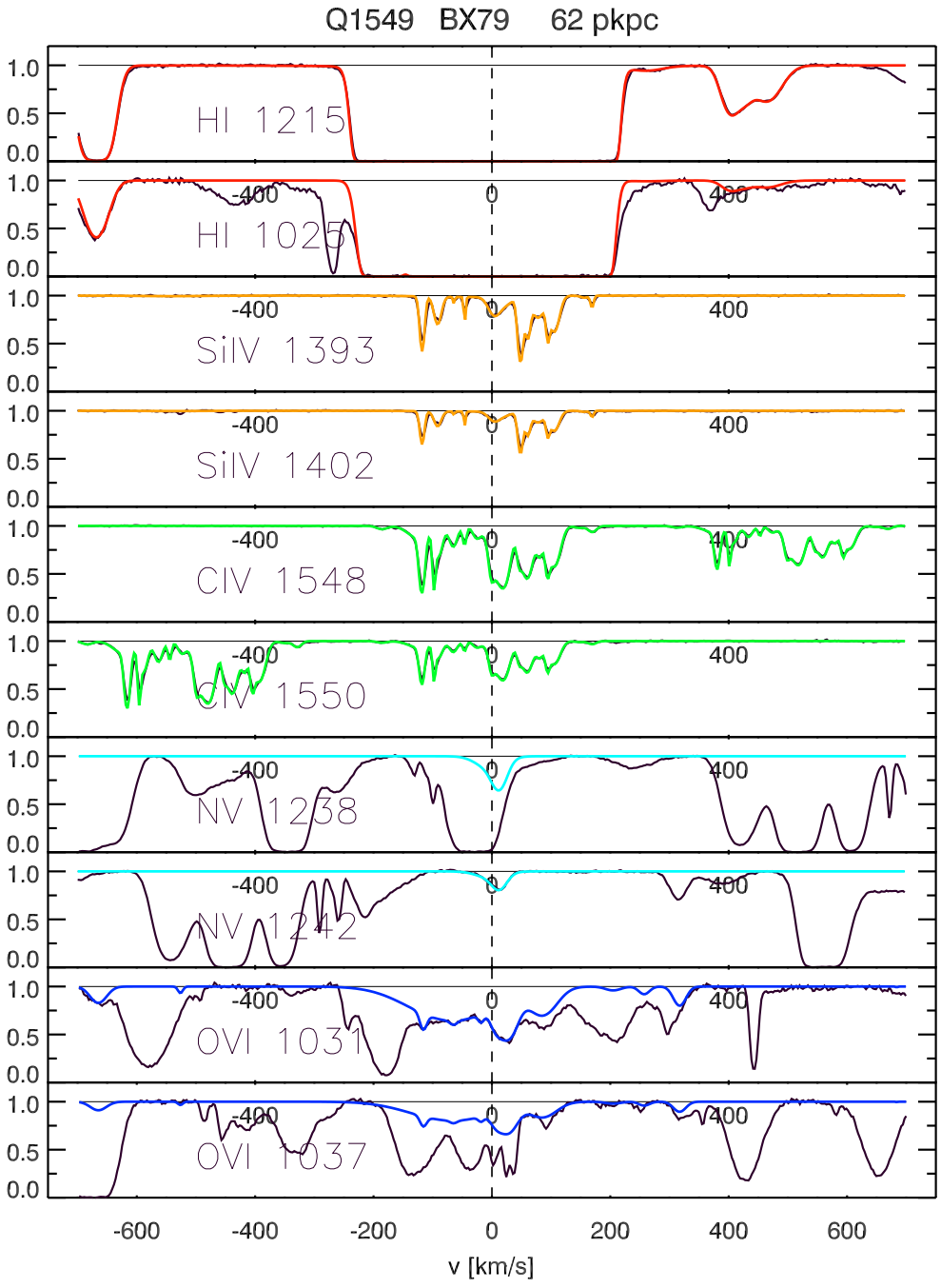


Figure 5.8 H I and high-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1549-BX79 which lies 62 pkpc from the line of the sight to the QSO.

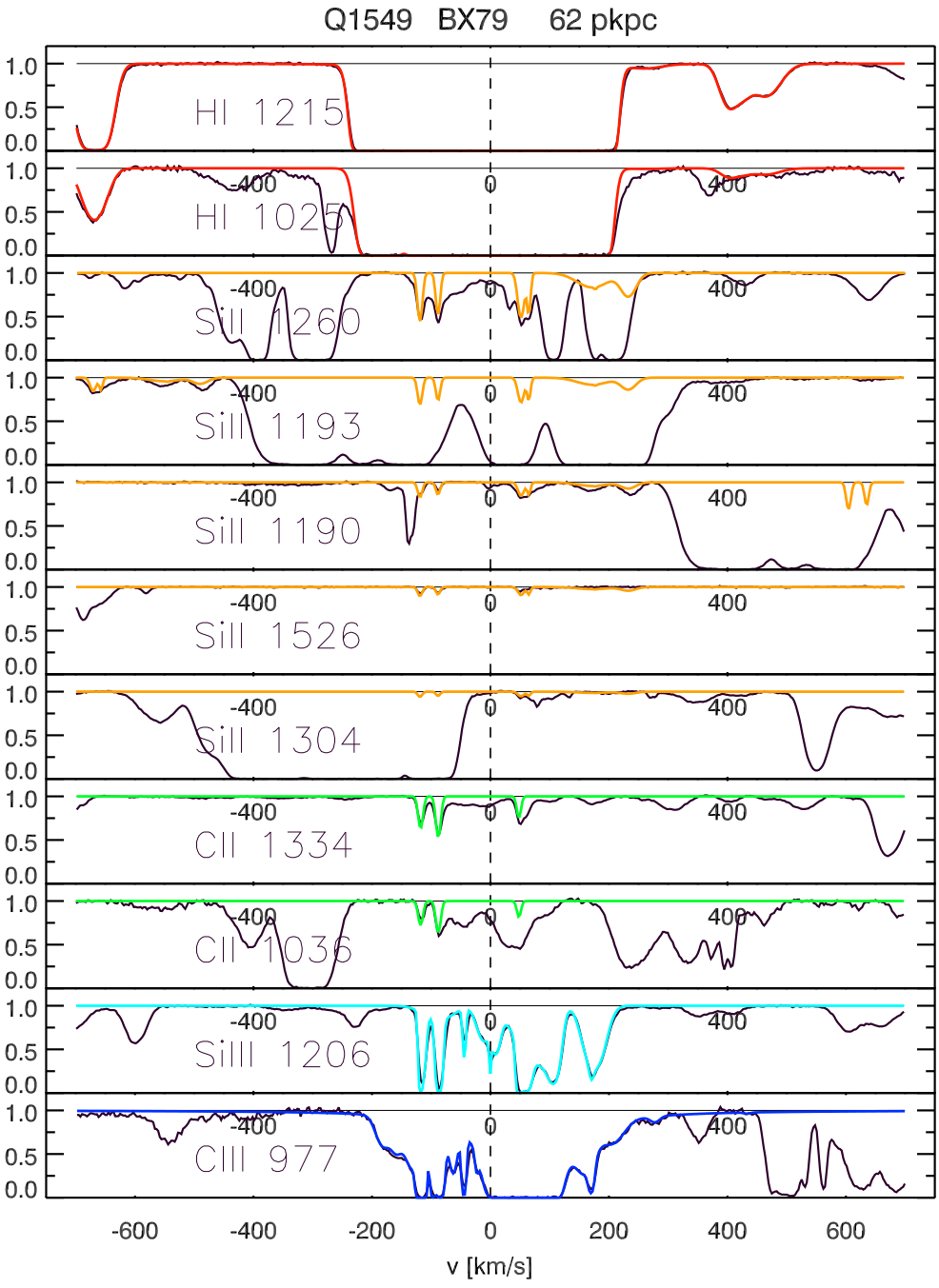


Figure 5.9 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1549-BX79 which lies 62 pkpc from the line of the sight to the QSO.

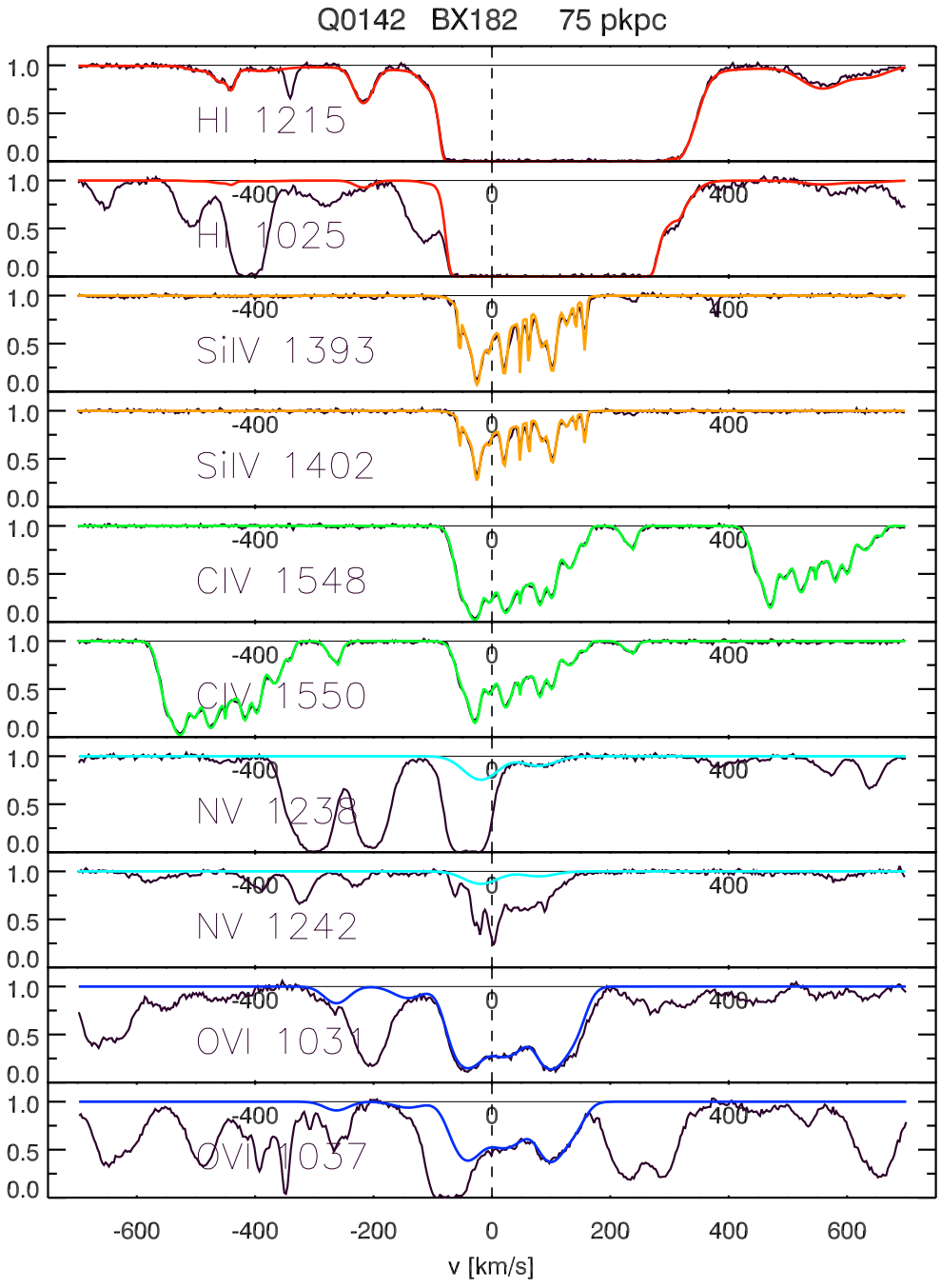


Figure 5.10 H I and high-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q0142-BX182 which lies 75 pkpc from the line of the sight to the QSO.

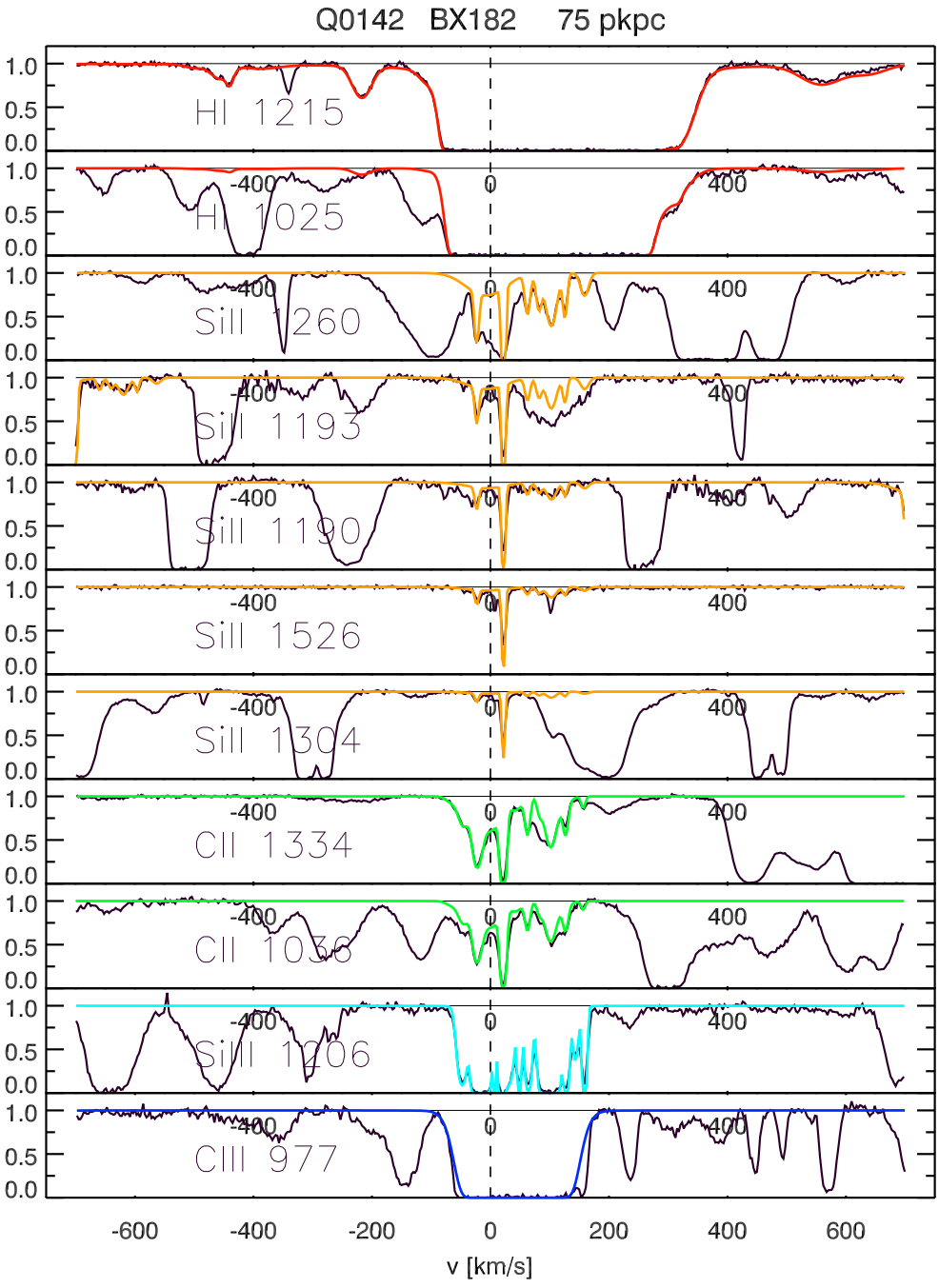


Figure 5.11 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q0142-BX182 which lies 75 pkpc from the line of the sight to the QSO.

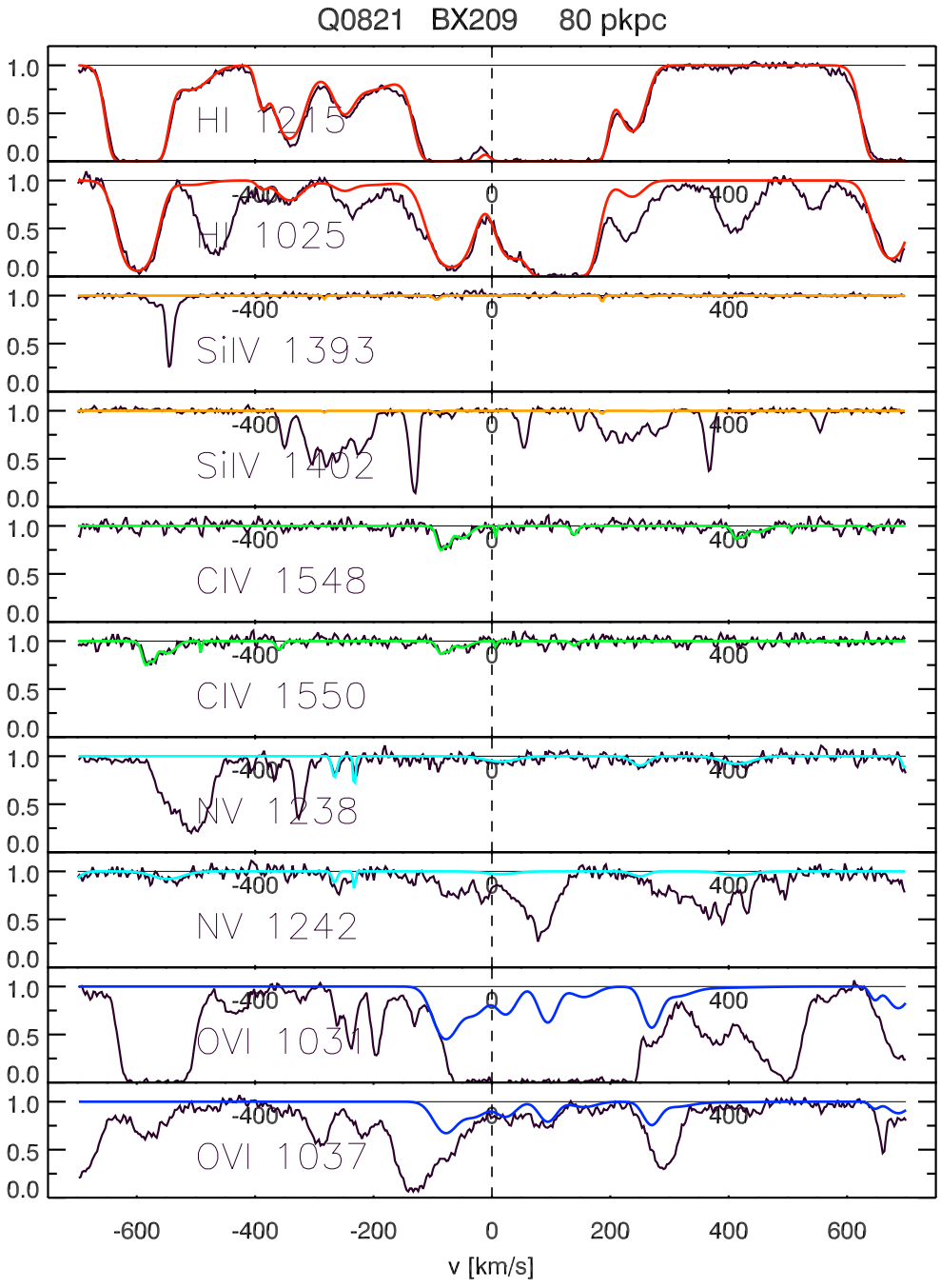


Figure 5.12 H I and high-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q0821-BX209 which lies 80 pkpc from the line of the sight to the QSO.

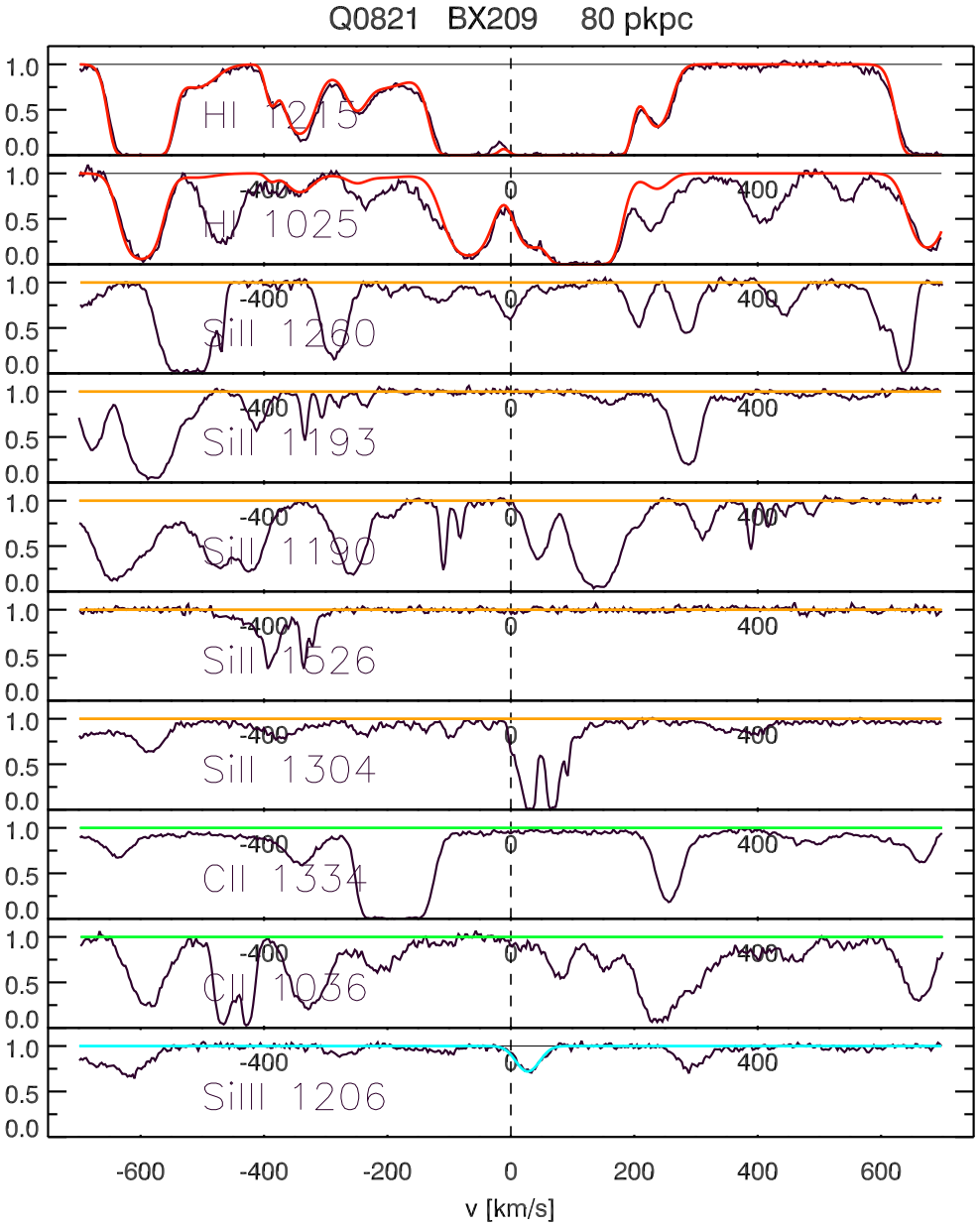


Figure 5.13 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q0821-BX209 which lies 80 pkpc from the line of the sight to the QSO. Note that we do not have sufficient wavelength coverage to see the C III transition. No Si II or C II absorption is detected in this region.

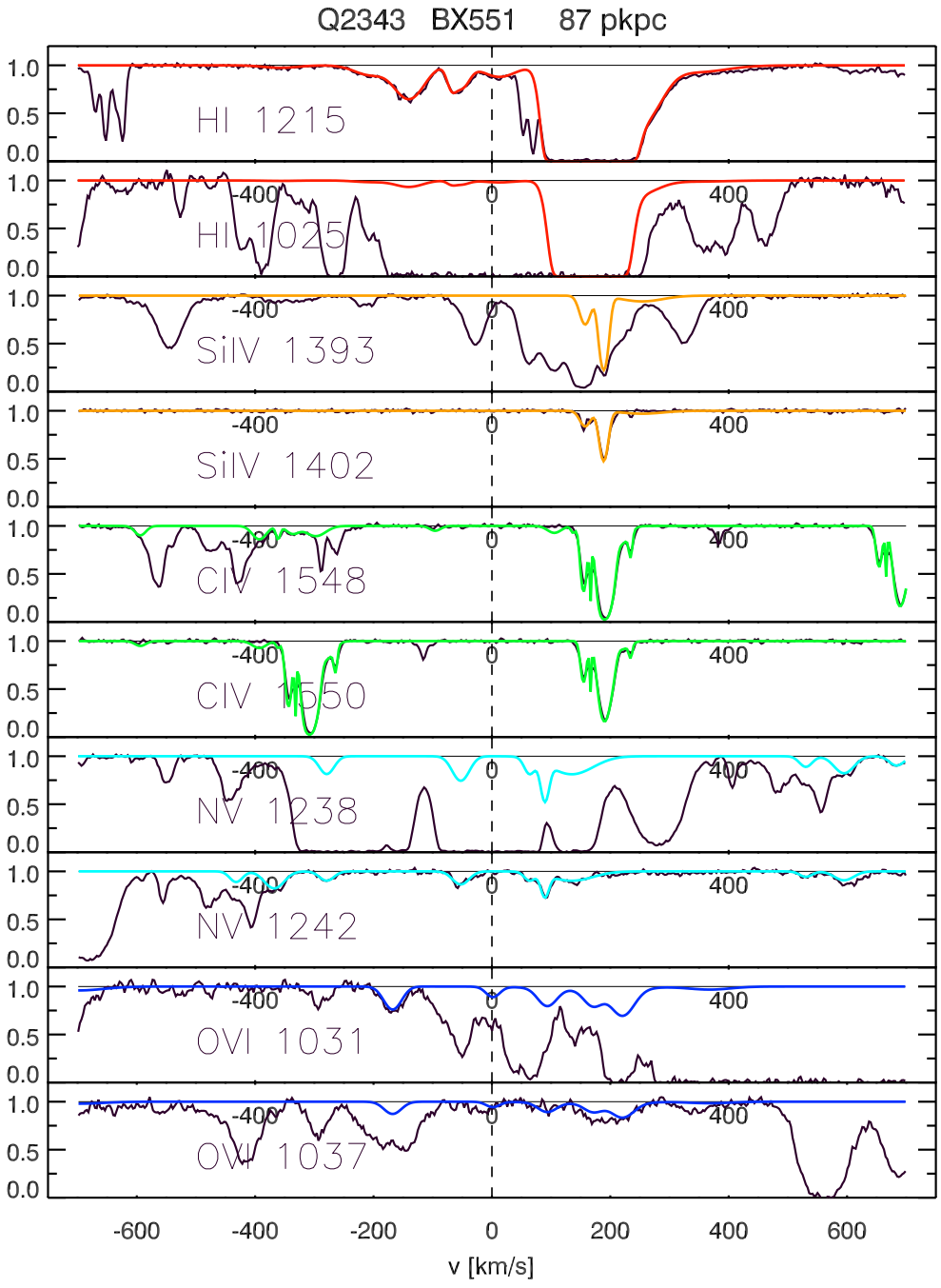


Figure 5.14 H I and high-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q2343-BX551 which lies 87 pkpc from the line of the sight to the QSO.

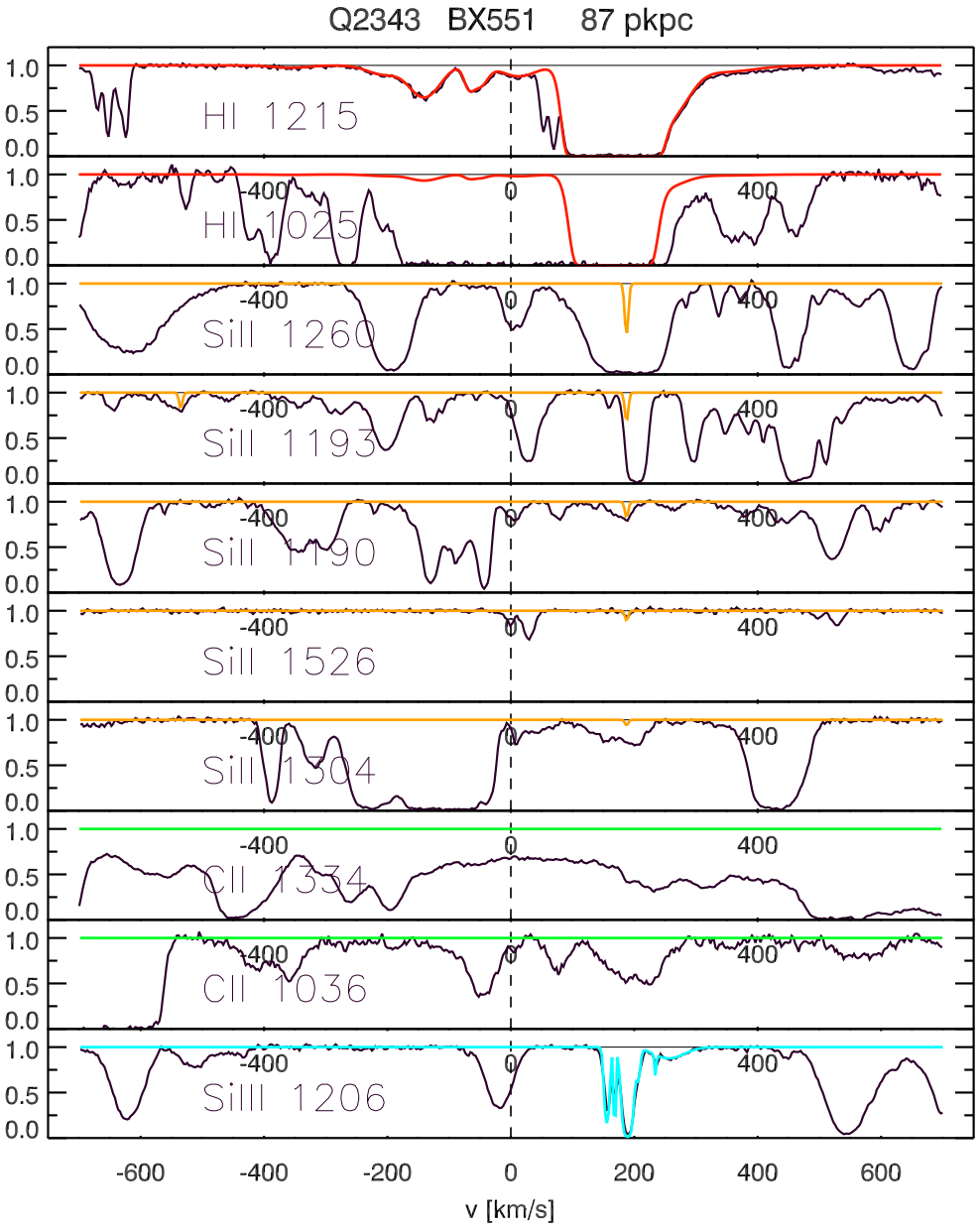


Figure 5.15 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q2343-BX551 which lies 87 pkpc from the line of the sight to the QSO. Note that we do not have sufficient wavelength coverage to see the C III transition and due to contamination in the C II region, we can only set a limit on the C II column density.

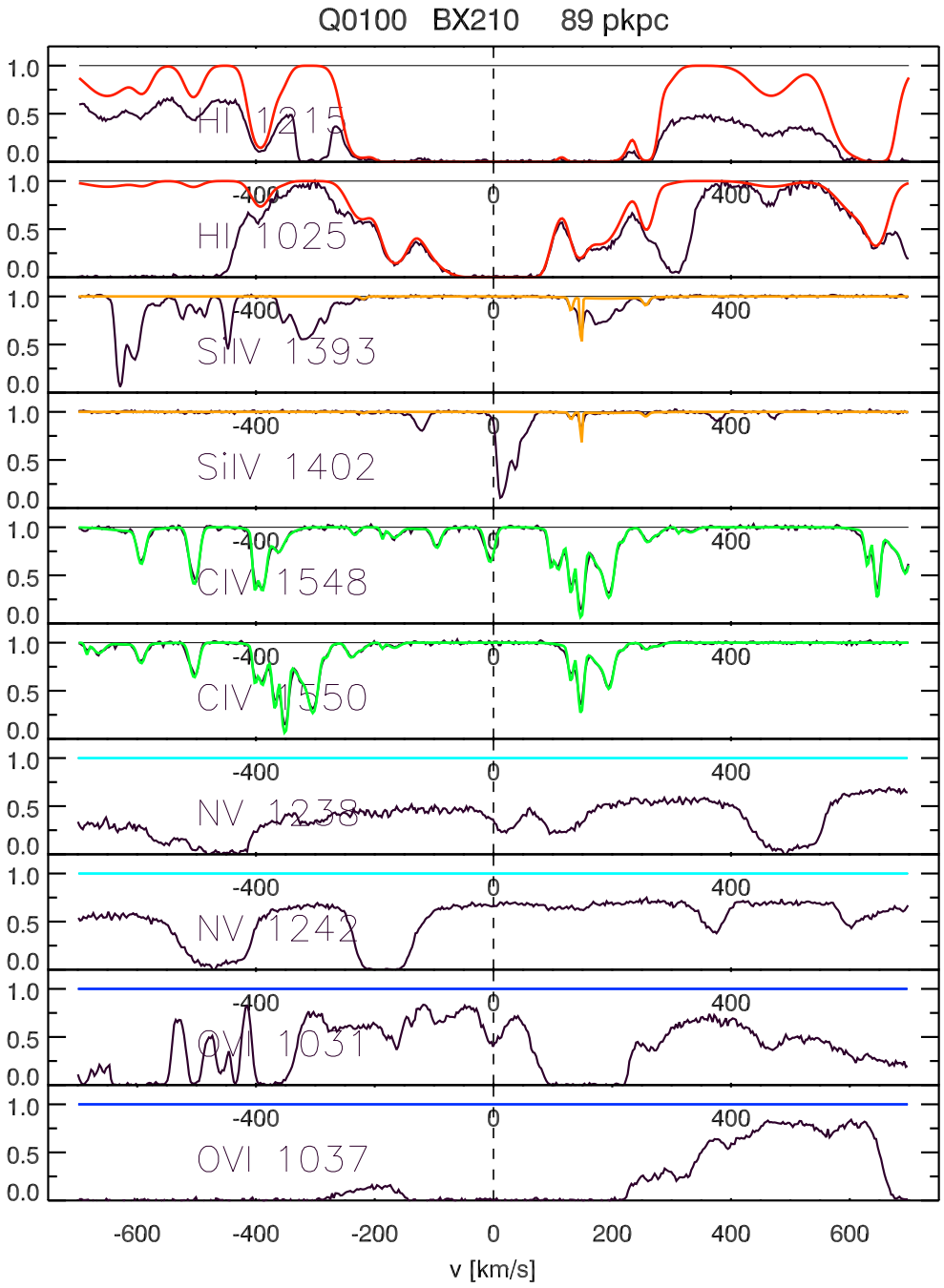


Figure 5.16 H I and high-ionization metal line absorption within $\pm$700 km s^{-1} of the galaxy Q0100-BX210 which lies 89 pkpc from the line of the sight to the QSO. The large amount of contaminating absorption in the N V and O VI portions of the spectrum are due Lyα and Lyβ respectively from a Damped Lyman α absorber (DLA) at a slightly higher redshift than that of the galaxy. The same DLA also causes the depression in the continuum in the Lyα portion of the spectrum.

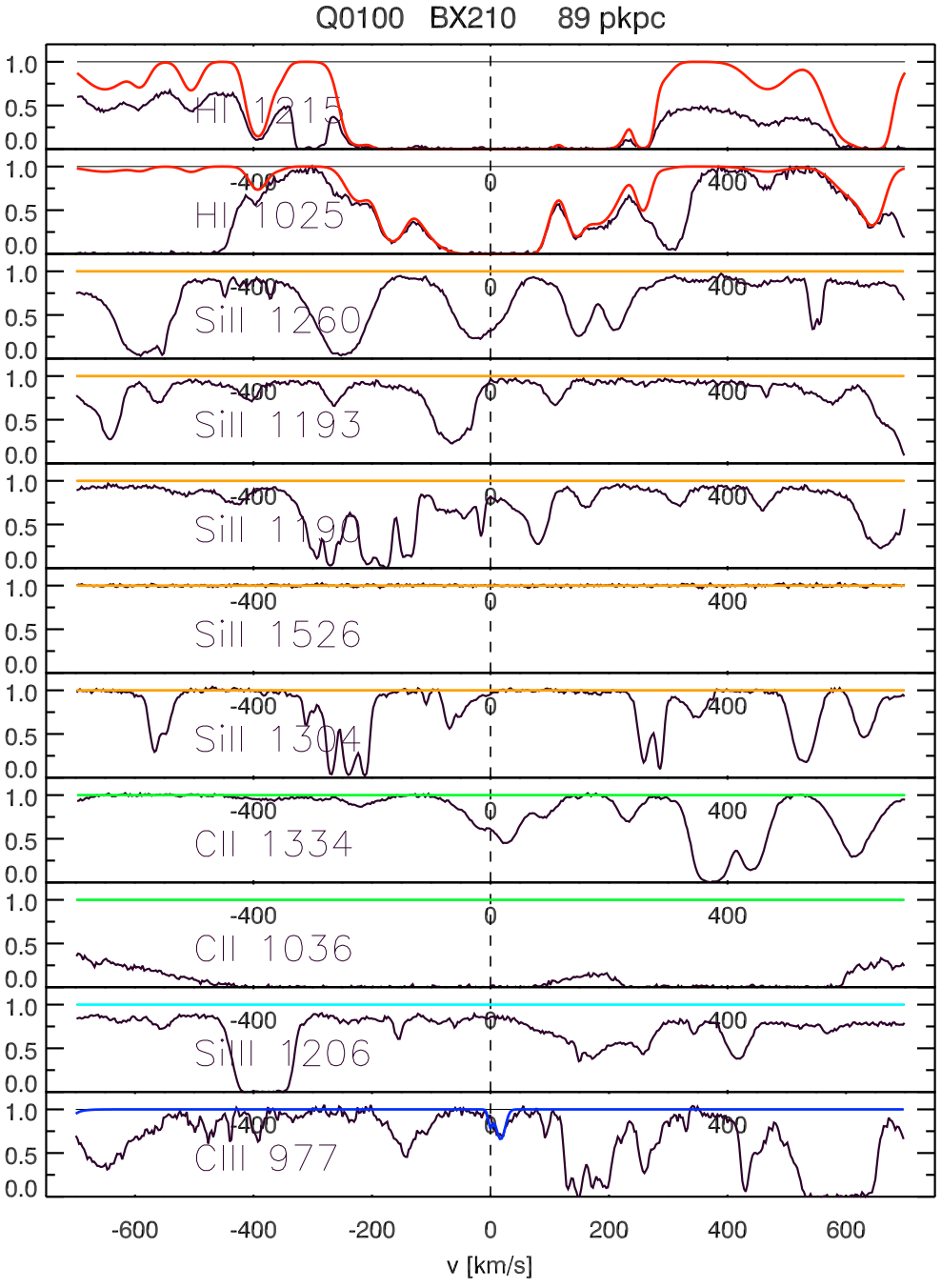


Figure 5.17 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q0100-BX210 which lies 89 pkpc from the line of the sight to the QSO. Do to the large amount of contamination in the C II region, only a weak constraint on the column density is possible. No Si II or Si III absorption is detected.

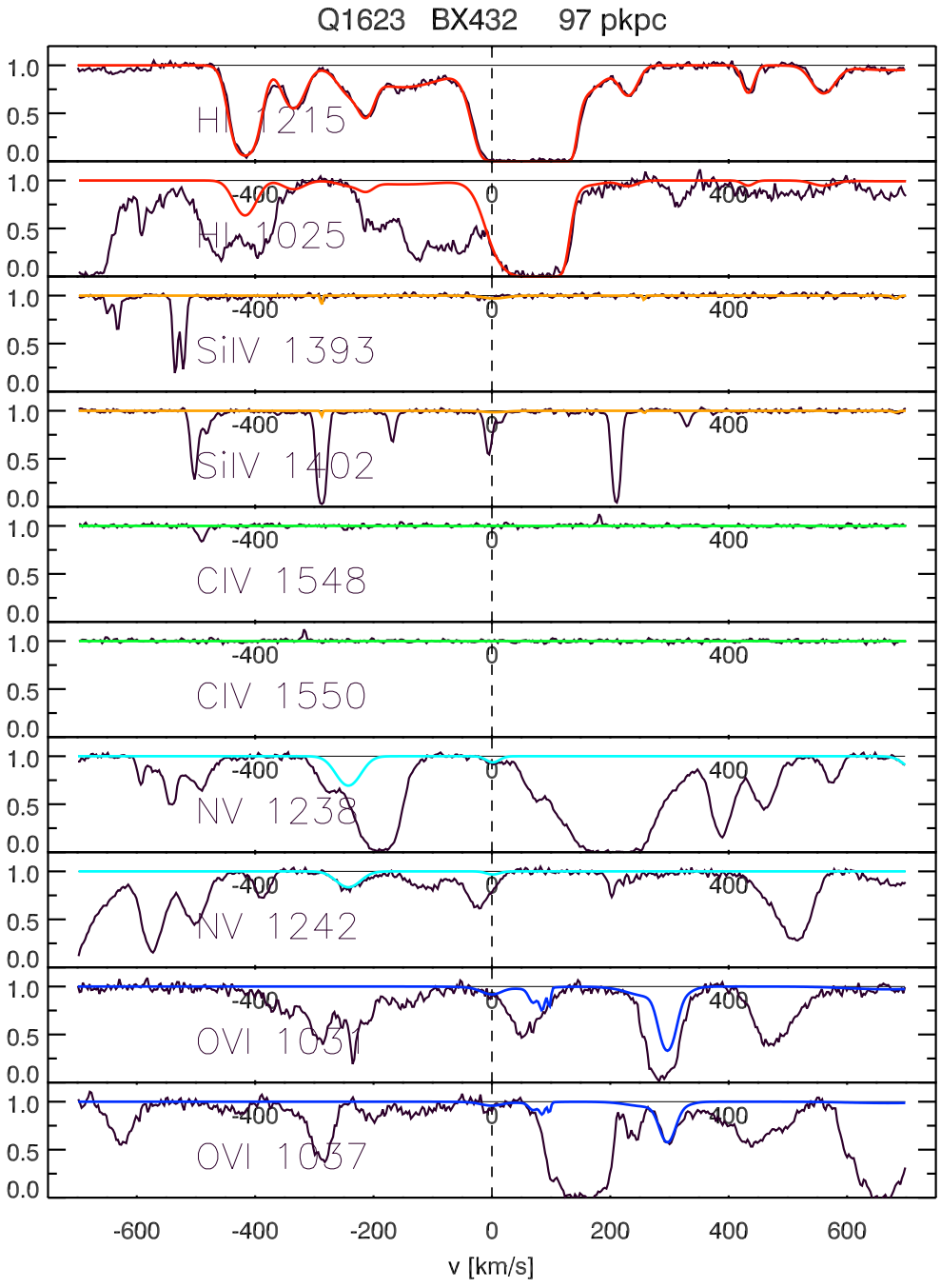


Figure 5.18 H I and high-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1623-BX432 which lies 97 pkpc from the line of the sight to the QSO.

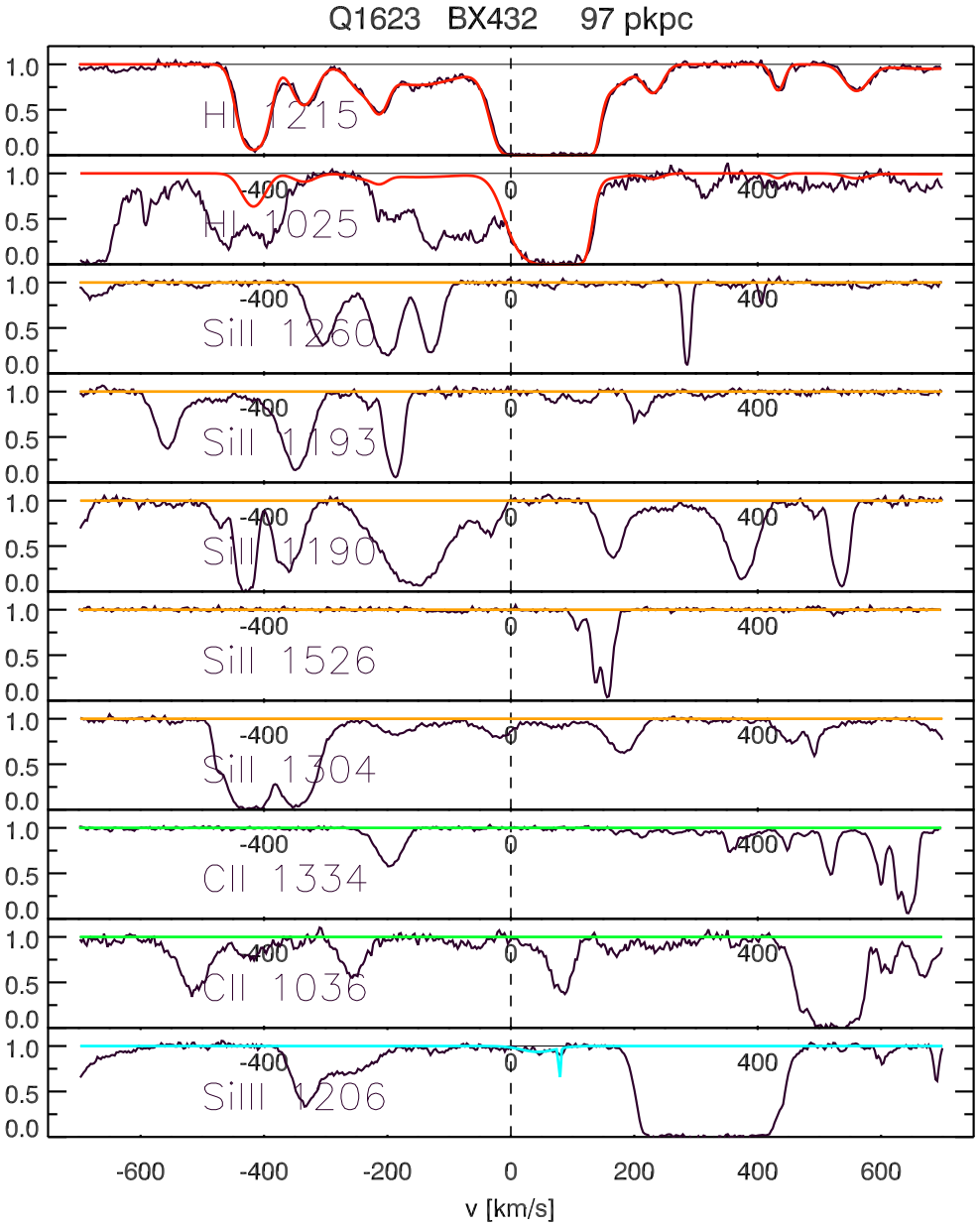


Figure 5.19 H I and low-ionization metal line absorption within ±700 km s^{-1} of the galaxy Q1623-BX432 which lies 97 pkpc from the line of the sight to the QSO. Note that we do not have sufficient wavelength coverage to see the C III transition. No Si II or C II absorption is detected in this region.

5.1.2 Preliminary Analysis

Comparison of different ionization states of the same element can be instructive. Figures 5.20 and 5.21 compare the fits to carbon- and silicon-enriched CGM gas 56 pkpc from the galaxy Q1549-D15. Here we see that the over-all velocity structure of the various ionization states are relatively well-aligned and are also aligned with the H I gas. The lowest-ionization state metal-line gas (C II and Si II) is nearly absent and the column densities of the intermediate (C III and Si III) and high (C IV and Si IV) ionization state gas are roughly equivalent (see Table 5.1).

CGM gas 50 pkpc from Q1442-BX333 and 56 pkpc from Q1442-MD50 shows somewhat different behavior. This metal complex is associated with an H I absorption systems with $N_{\rm HI}$= 10^{20} cm^{-2}, a much higher column density than the system near Q1549-D15 and a column density very close to the self-shielding limit for H I gas. Figures 5.22 and 5.23 compare the various ionization states of carbon and silicon for this gas. Considering first the low-ionization species, the component structure observed in C II and Si II is well matched by that of H I and its velocity center is typically offset from many of the high-ions, suggesting these lowest-ionization species are likely co-spatial with the majority of the H I. C III and Si III are much stronger (and are dominated by higher column density systems) than C II, Si II and C IV, Si IV, giving an indication of the dominant ionization state. The velocity center of the C III and Si III absorption complex seems to be intermediate between that of C II and Si II and C IV and Si IV.

Inspection of the highly-ionized metal line absorption near Q1549-D15 (Figure 5.4) and Q1442-MD50 (Figure 5.6) provides important insight into the physical nature of CGM absorbers. The shapes and velocity structure of N V and O VI are very different from C IV and Si IV as well as all of the low-ionization lines. This suggests that these ions trace different gaseous structures than C IV, Si IV, and other lower-ionization lines which is not surprising given the very different ionization potentials of these species (see Table 5.2.) Conversely, Si IV and C IV seem to have similar velocity structures, but the fits to these absorption systems result in much narrower $b_{\rm d}$ for Si IV ($b_{\rm d}\sim 2-9$ km s^{-1}), than for C IV ($b_{\rm d}\sim 5-15$ km s^{-1}). As discussed in 2.7 and 3.3, the widths of absorbers trace both thermal and turbulent sources of line broadening. In the case of thermal broadening, the line width is proportional to $\sqrt{T/m_{\rm ion}}$. The atomic mass of Si and C are 28 and 12 AMU respectively; therefore, in the case of thermal line broadening, we would expect:

$$\frac{b_{\rm CIV}}{b_{\rm SiIV}} = \sqrt{\frac{m_{\rm SiIV}}{m_{\rm CIV}}} \approx 1.5, \tag{5.1}$$

which is roughly the observed difference in the line widths of these ions. This implies much of the broadening of this gas is due to thermal effects. The relatively larger $b_{\rm d}$ of O VI and N V also suggest that they correspond to gas at comparatively higher temperatures (again unsurprising given their comparatively high ionization potentials).

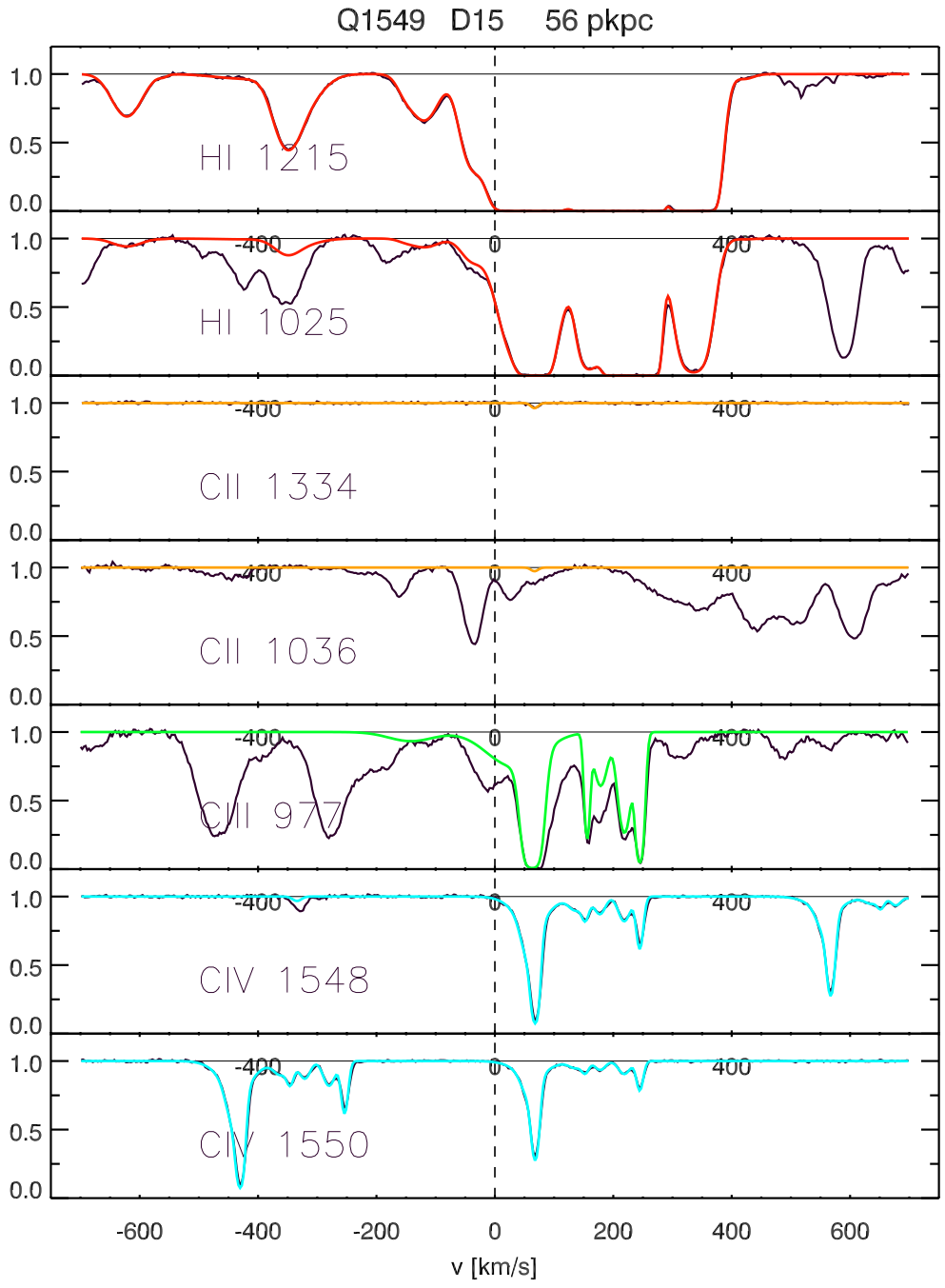


Figure 5.20 Comparison of the velocity structure in different ionization levels of carbon near the galaxy Q1549-D15.

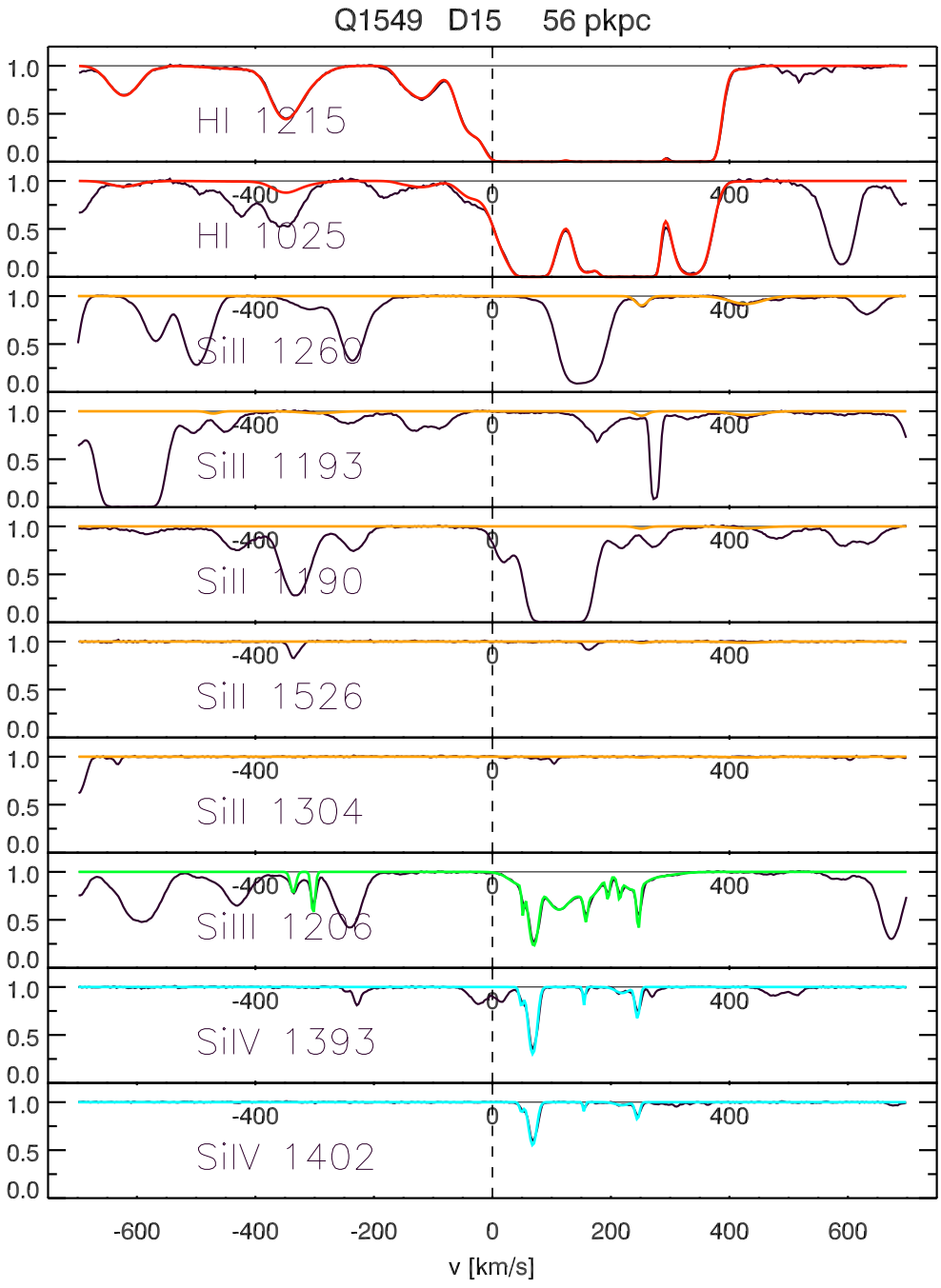


Figure 5.21 Comparison of the velocity structure in different ionization levels of silicon near the galaxy Q1549-D15.

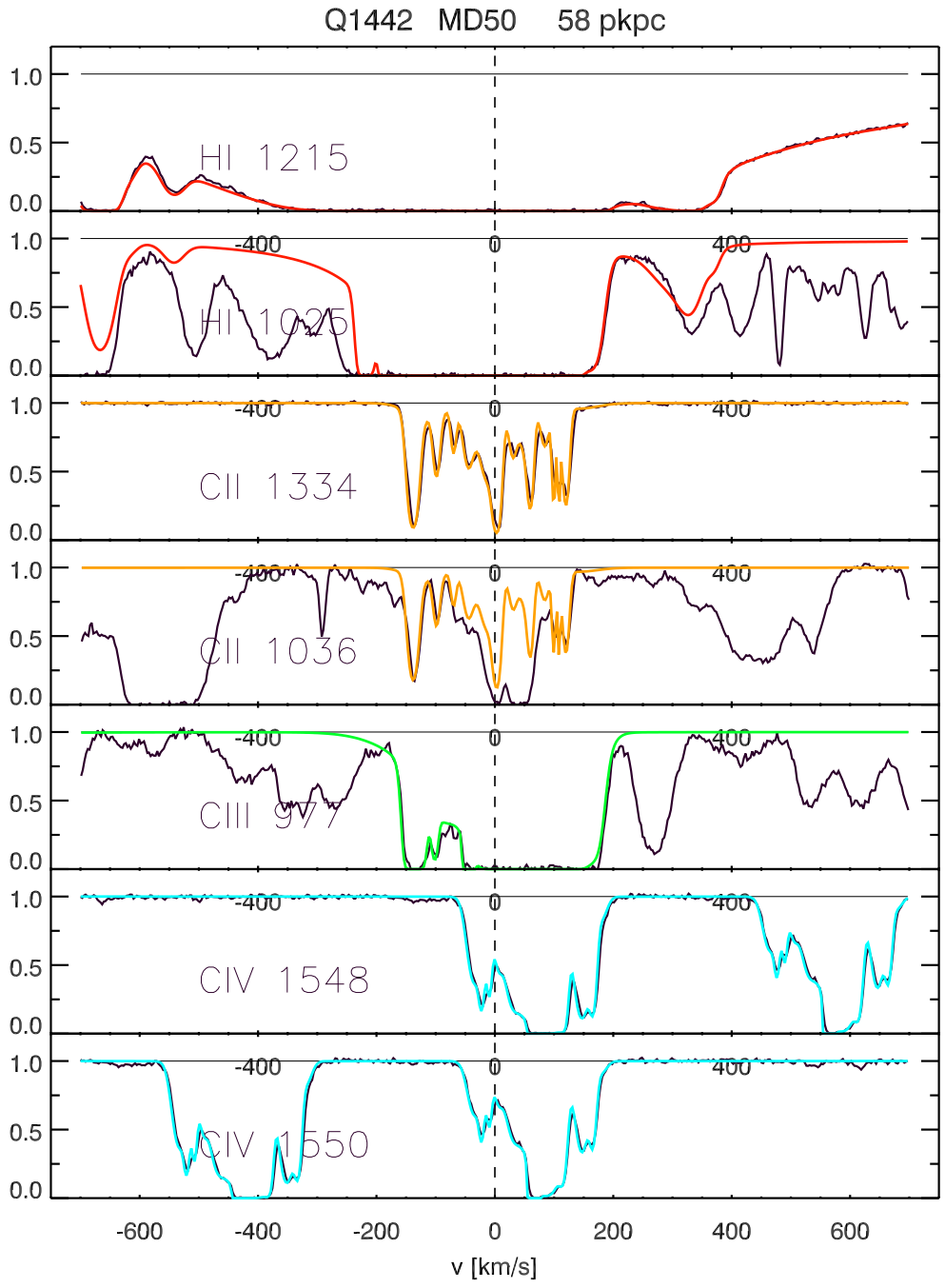


Figure 5.22 Comparison of the velocity structure in different ionization levels of carbon near the galaxy Q1442-MD50.

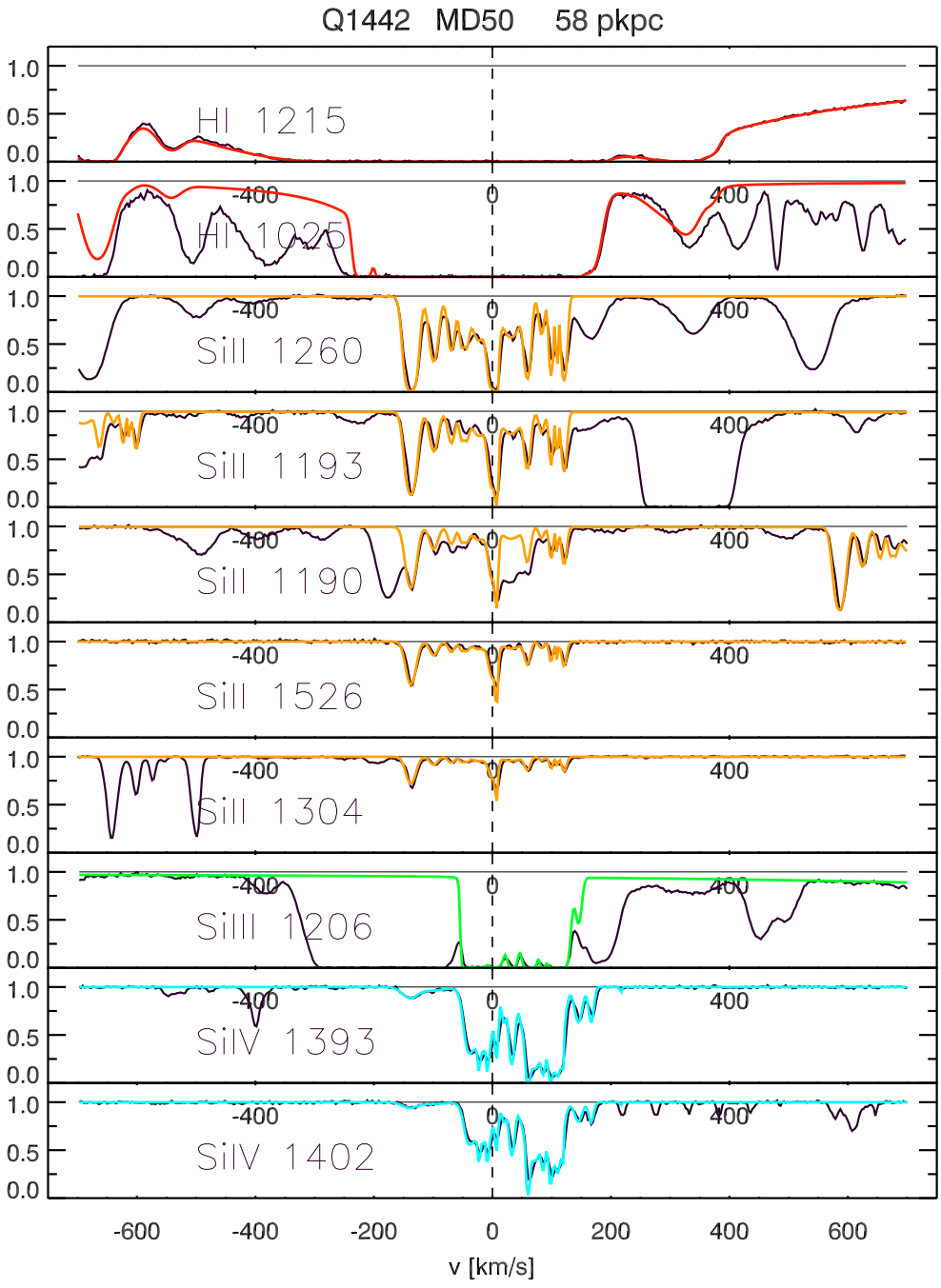


Figure 5.23 Comparison of the velocity structure in different ionization levels of silicon near the galaxy Q1442-MD50.

Table 5.2. The Ionization Potential (U) of Ions associated with Strong UV Transitions[a]

ion	U	ion	U
Si II	8.2 eV	Si IV	33.5 eV
C II	11.3 eV	C IV	47.9 eV
Si III	16.3 eV	N V	77.5 eV
C III	24.4 eV	O VI	113.9 eV

[a]From Cox (2000)

5.2 Looking Forward

The full analysis of the metal absorption in the entire KBSS sample in underway. Following the completion of metallic column density measurements, I will use the photoionization code CLOUDY (Ferland et al. 1998) to make ionization corrections and infer the physical properties of the gas. When complete, these will be the first measurements of the ionization state, metallicity, abundance patterns, and mass of gas within the CGM of a large sample of high-z galaxies. By combining these first measurements characterizing the CGM of a statistically representative sample of galaxies with measurements of the physical properties of the same galaxies (made possible by recent advances in NIR spectroscopy), we will gain new insight into the movement of baryons into and out of galaxies and the coevolution of galaxies with the IGM.

Appendix A

Fits to High-$N_{\rm HI}$ Absorbers

As discussed in §4.1, for $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 14.5$, the Ly$\alpha$ transition is saturated, complicating the measurement of absorbers with higher $N_{\rm HI}$ from Lyα only spectra. Similarly, for absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 15.5$, the Ly$\beta$ transition saturates. Thus, the precision and accuracy with which high-$N_{\rm HI}$ systems can be measured will depend to some degree on the number of accessible Lyman lines, thus imposing some redshift dependence.

For the sample presented in this paper, the QSO spectra cover the Lyα and Lyβ transition for all the absorbers. Given the typical spectral coverage of our sample (see Table 2.1), for absorbers with $z \gtrsim 2.3$, the Lyγ transition is also covered, and for those with $z \gtrsim 2.4$, Ly$\alpha, \beta, \gamma, \delta$ are generally observed. For absorbers with $z \gtrsim 2.5$, the spectral coverage of the QSOs typically allows for observation of transitions all the way to the Lyman Limit, thus allowing for accurate measures of $N_{\rm HI}$ to $\log(N_{\rm HI}/{\rm cm}^{-2}) \lesssim 17.2$.

An additional complication arises from the fact that high-$N_{\rm HI}$ absorbers are often found in blended systems of several high-$N_{\rm HI}$ absorbers. Such absorbers, when observed in only Lyα, β, often cannot be de-blended. The incidence of such blends increases with increasing redshift as the density of the forest increases. This blending effect is mitigated in our sample at $z > 2.4$ due to access to additional Lyman series lines, but may go unrecognized in the low end of the redshift interval ($z \lesssim 2.4$) which lacks optically thin Lyman lines. For such blended systems in the presented sample with $z \lesssim 2.4$, the number of subcomponents may be underestimated which may lead to some degree of incompleteness in the sample.

In this appendix we assess the accuracy of our fits to absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 15.5$.

A.0.1 Examples of Fits to High-$N_{\rm HI}$ Absorbers

Figures A.1 - A.3 show examples of the fits to the HIRES spectra for absorbers with $15.5 < (\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.0$. In each case, the HIRES data are shown in black, and the solid red curve shows the model with the best fit to the data. The dashed curves show the $\pm 3\sigma$ formal errors on $N_{\rm HI}$ as reported by VPFIT. As expected, the errors on $N_{\rm HI}$ are dependent on the number of transitions observed; however, the dependence is not trivial as the effect of (a) blending from other proximate absorbers, (b) the contamination of higher-order Lyman

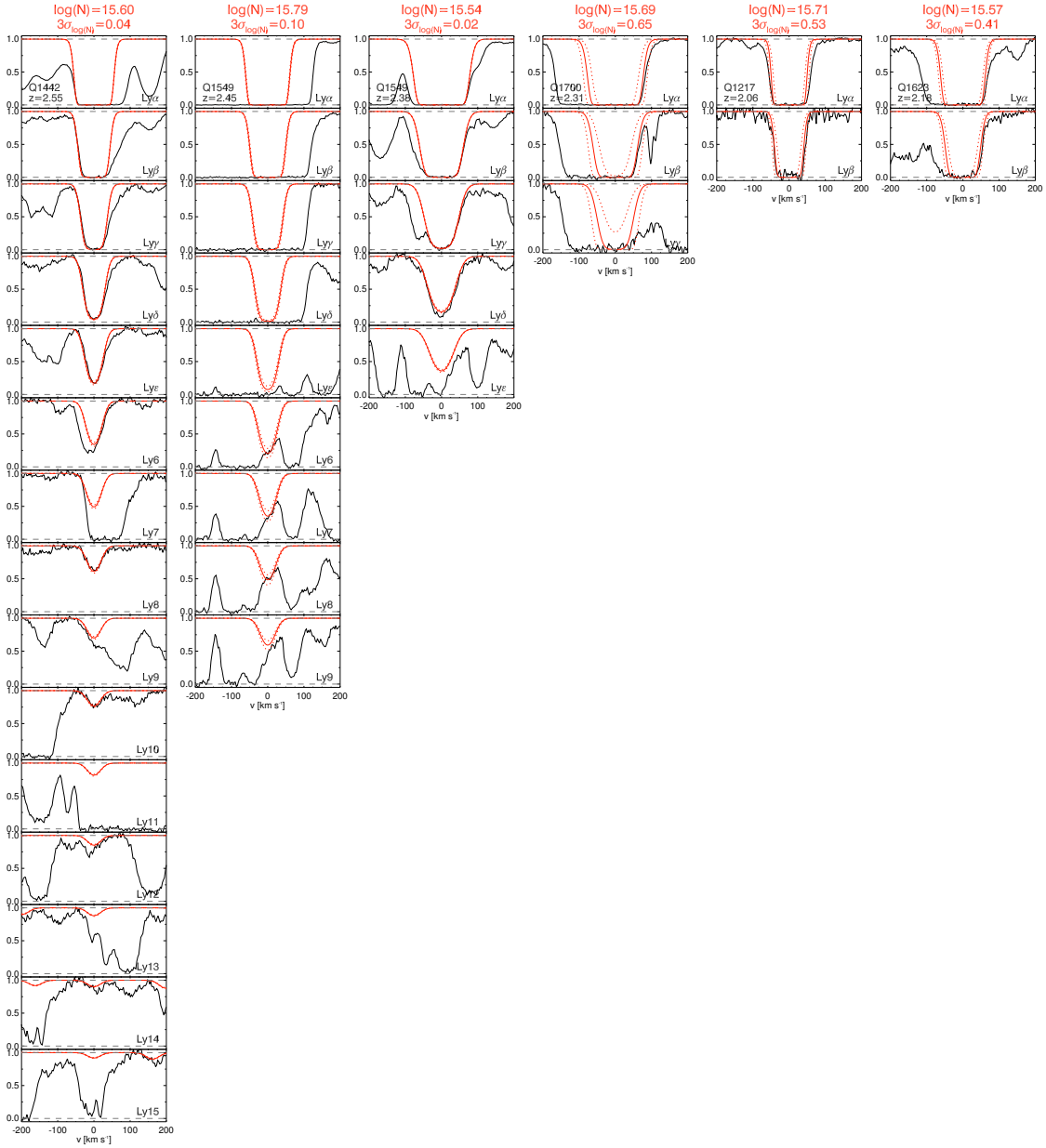


Figure A.1 Example fits to absorbers with $15.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 16.0$. The black curves show the HIRES data. Grey dashed lines mark the zero point and continuum of the QSO spectra. The solid red curve shows the best fit Voigt profile to each absorber. The dashed red curves show the formal $\pm 3\sigma$ $N_{\rm HI}$ error.

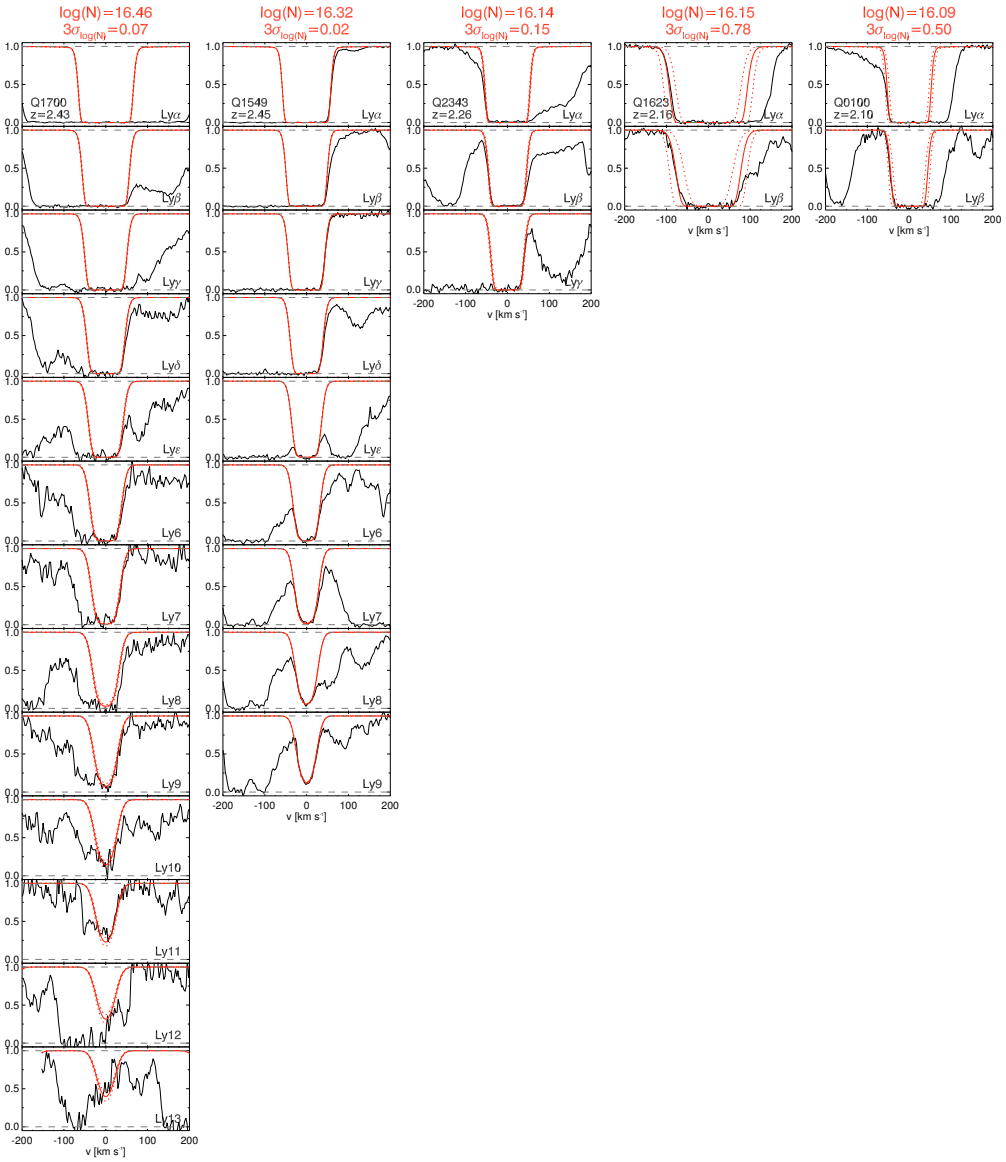


Figure A.2 Example fits to absorbers with $16.0 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 16.5$. The black curves show the HIRES data. Grey dashed lines mark the zero point and continuum of the QSO spectra. The solid red curve shows the best fit Voigt profile to each absorber. The dashed red curves show the formal $\pm 3\sigma$ $N_{\rm HI}$ error.

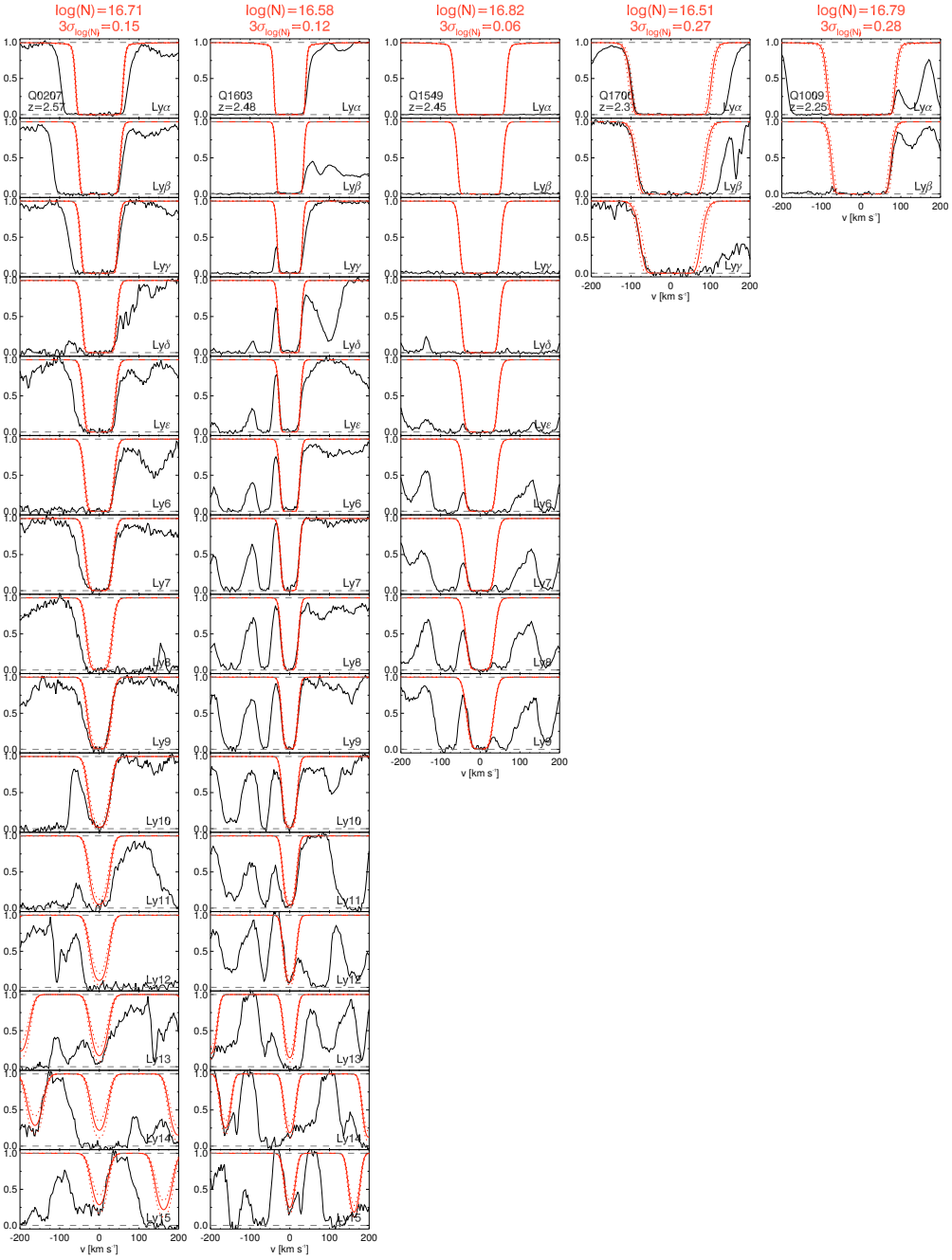


Figure A.3 Example fits to absorbers with $16.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.0$. The black curves show the HIRES data. Grey dashed lines mark the zero point and continuum of the QSO spectra. The solid red curve shows the best fit Voigt profile to each absorber. The dashed red curves show the formal $\pm 3\sigma$ $N_{\rm HI}$ error.

transitions due to unrelated absorbers, and (c) the S/N of the spectrum in the areas of interest all affect our ability to accurately determine the parameters of the absorber.

A.0.2 Testing Fits with Only Lyα and Lyβ

To assess the accuracy of fits to absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 15.5$ observed only in Lyα and Lyβ, we use the set of absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) \gtrsim 15.5$ for which $n \geq 5$ Lyman transitions were measured, generally drawn from the subsample with $z > 2.4$. We then attempt to refit these absorbers using VPFIT with constraints from the Lyα and Lyβ region only. The outcomes of these tests can be generalized into two principal categories: highly-blended and less-blended absorbers.

Absorbers for which a portion of one of both edges/wings of the absorption line can be observed in Lyα and/or Lyβ can be fit much more reliably. Such systems, when refit with VPFIT using only their Lyα and Lyβ transition result in weaker constraints on the line parameters as expected, but result in $N_{\rm HI}$ determinations consistent within their formal errors with those obtained via the higher-order transitions. The fits are not systematically offset to higher or lower values of $N_{\rm HI}$, and therefor should not *bias* the measurements of $f(N,X)$ in any large way.

For absorbers whose Lyα and Lyβ transition are fully saturated and both edges/wings of the absorption line cannot be observed due to other neighboring absorbers (e.g. see the left-most panel in Figure A.2 and the center panel in Figure A.3), essentially no information is available on the sub-component structure of the system. Such systems are often fit with a minimum of two components because it is rare that the wings/edges of the blended system are symmetric. In the absence of higher-order Lyman lines, the number of subcomponents is typically underestimated, likely leading to some degree of incompleteness in the lower-redshift portion of the catalog. There would also be a tendency to underestimate the total $N_{\rm HI}$ in such blends. If the $N_{\rm HI}$ of individual absorbers in the blends were systematically underestimated, this would *bias* the measurement of $f(N,X)$ to have steeper values of β than the true value.

In conclusion, the lack of higher-order constraints on the lower-z saturated absorbers have two principal effects on our measurements. (1) Relatively unblended absorbers yield somewhat weaker constraints on the parameters that are unlikely to bias the measurement of $f(N,X)$. (2) Highly-blended absorbers yield very weak constraints from Lyα and Lyβ only and therefore are typically fit with fewer subcomponents and with lower-$N_{\rm HI}$ values than would be assigned using additional Lyman series constraints. This results in some degree of incompleteness in the lower-z portion of the catalog.

A.0.3 Assessing Incompleteness in the $z \lesssim 2.4$, $\log(N_{\rm HI}/{\rm cm}^{-2}) > 15.5$ Catalog

To estimate the degree of potential incompleteness in the lower-z portion of the catalog due to line blending, we consider the fraction of the higher-z, high-$N_{\rm HI}$ absorbers observed in $n \geq 5$ Lyman series transitions that would be unconstrained or miscounted with observations of Lyα and Lyβ alone. This method likely over-

estimates the incompleteness as the degree of line blending and contamination is higher in the high-redshift forest than in the lower-z forest due to evolution in the line density as a function of redshift. Nevertheless, we report the results of this test, considering separately absorbers in half-dex bins of $N_{\rm HI}$. The results are summarized in Table A.1.

For absorbers with $15.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 16.0$, 23 are measured with 5 or more Lyman series transitions. Of these, only 2 would have essentially no constraint on the parameters of the absorber from Lyα and Lyβ alone (e.g. Figure A.1, second from the left). For 7/23 of these systems, the constraint on $N_{\rm HI}$ would be relatively weak, however, an individual component associated with that absorber can be clearly identified. All 7 of these systems are fit with much improved accuracy with the additional constraint from Lyγ. Among absorbers with $16.0 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 16.5$, we observed 12 absorbers with constraints from 5 or more Lyman series transitions. Of these, only 4 are strongly constrained from Lyα and Lyβ alone. With Lyγ, 8/12 could be fit, and with Ly$\alpha,\beta,\gamma,\delta$, 11/12 would be strongly constrained and could accurately be fit. For absorbers with $N_{\rm HI}$ just below that of LLSs ($16.5 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 17.0$), we observe 5 systems in 5 or more Lyman series transitions.[1] Of these, only 2 would have no constraint on $N_{\rm HI}$ from Lyα and Lyβ alone.

From this assessment based on higher-z measurements, we expect that the principal bias in our measurements of $N_{\rm HI}$ for high-$N_{\rm HI}$ systems for absorbers with only Lyα and Lyβ is to under-count the number of systems with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 15.5$. Within the present sample, systems with $16.0 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 16.5$ appear to be most-affected by blending and may therefore be the most incomplete. In the next appendix, we add further evidence to this claim by comparing the measurement of $f(N,X)$ for absorbers in two redshift subsamples.

A.1 The Effect of the Accuracy of $f(N,X)$ on the Measurement of the $\lambda_{\rm mfp}$

Here we assess the redshift dependence of the frequency distribution presented in §4.3 and §4.5. We split the absorber sample roughly in half and compare those absorbers with $z < 2.4$ and those with $z > 2.4$. For the higher-redshift subsample, the spectra typically cover at least 4 Lyman series transitions and most of this sample have all the Lyman series transitions observed. Figure A.4 shows the results of this comparison. Notably, the higher-redshift subsample (red points) is in reasonable agreement with the broken power law fit (gray curves) presented in §4.5.1 that were used in the Monte Carlo simulation in §4.6. Further, the higher-redshift sample, tends to lie slightly above the fit. Some portion of this is likely due to the redshift evolution of the line density of the forest; however, it is also possible that $f(N,X)$ of absorbers with $\log(N_{\rm HI}/{\rm cm}^{-2}) > 15.5$ is actually higher than that measured from the full sample. As discussed in Appendix A.0.3, we expect that the lower-z catalog likely suffers from incompleteness due to line blending, especially for absorbers with

[1] All five absorbers are observed in a minimum of 9 transitions. We do not observe any absorbers in this $N_{\rm HI}$ range with 5-8 Lyman series transitions observed.

Table A.1. Constraints on $N_{\rm HI}$ vs. the Number of Observed Lyman Series Transitions

$\log(N_{\rm HI}/{\rm cm}^{-2})$	Lines considered in the fit	# with 5+ Lyman series lines[a]	strongly constrained		weakly constrained[b]		unconstrained[c]	
			#	%	#	%	#	%
15.5 - 16.0	Lyα,β	23	14	61%	7	30%	2	9%
15.5 - 16.0	Lyα,β,γ	23	18	78%	3	13%	2	9%
16.0 - 16.5	Lyα,β	12	2	17%	2	17%	8	67%
16.0 - 16.5	Lyα,β,γ	12	2	17%	6	50%	4	33%
16.5 - 17.0	Lyα,β	5	1	20%	2	40%	2	40%
16.5 - 17.0	Lyα,β,γ	5	3	60%	0	0%	2	40%

[a]The number of absorbers in the catalog with $N_{\rm HI}$ in the range listed and at least 5 Lyman series transitions observed.

[b]For weakly constrained absorbers, the transitions listed give unambiguous evidence of the presence of an absorber, and typically have at minimum one edge of one transition well observed without a significant blend.

[c]For unconstrained absorbers, the transitions listed are fully saturated and blended with other high-$N_{\rm HI}$ lines (e.g. see middle panel Figure A.3) meaning this absorber would likely not be counted or be very poorly fit in data with only the transitions listed.

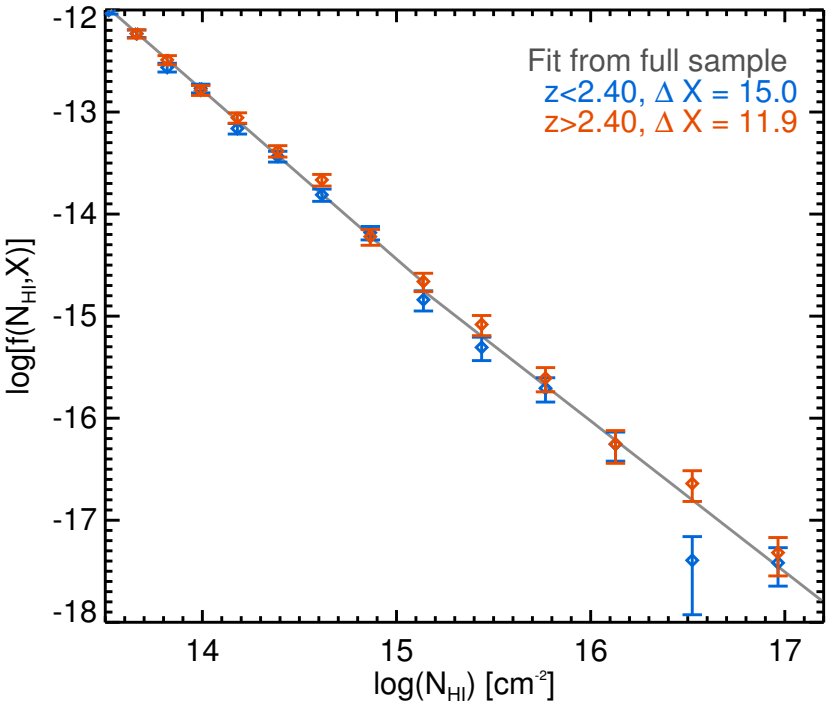


Figure A.4 The frequency distribution measured from the higher-redshift (red) and lower-redshift (blue) portions of the absorber sample (many of the points from the two subsamples lie on top of each other). The grey curve shows the broken power law fit to the full absorbers sample from §4.5.1.

$16.0 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 16.5$. In Figure A.4, we note the only bin for which the low-z and high-z data do not agree within 1σ contains absorbers with $16.0 < \log(N_{\rm HI}/{\rm cm}^{-2}) < 16.5$.

The principal result of this paper is the relatively high measured value of $f(N,X)$ for absorbers with $14.5 < (\log(N_{\rm HI}/{\rm cm}^{-2}) < 17.2$ which results in a value of $\lambda_{\rm mfp}$ comparable to or smaller than previous measurements. If we were to use the high-redshift sub-sample from Figure A.4 to measure the opacity and $\lambda_{\rm mfp}$, we would find higher IGM opacity and therefore a shorter (more discrepant) $\lambda_{\rm mfp}$. In any case, if there is a systematic bias in the measurements, it is one that would conspire to make $\lambda_{\rm mfp}$ appear *larger* than its true value. We therefore conclude that it is unlikely that a bias in the precision of our measurements of $f(N,X)$ as a function of z would change the principal result of the paper.

Appendix B

MOSFIRE

The Multi-Object Spectrometer for Infra-Red Exploration, MOSFIRE, is a multi-object NIR spectrograph designed for faint target spectroscopy at wavelengths of 0.97-2.45 μm over a 6.1×6.1 arcminute field. Observations cover one of the NIR atmospheric bands - $Y(0.97-1.12\mu\text{m})$, $J(1.15-1.35\mu\text{m})$, H $(1.46-1.81\mu\text{m})$, or $K(1.93-2.45\mu\text{m})$ - at a time. A mechanically-configurable cryogenic slit unit (the CSU) allows for the formation of up to 46 distinct slits within the 6' field, providing the significant multiplexing advantages now routine in optical spectrographs. With a low-noise state-of-the-art Teledyne Hawaii 2RG HgCdTe detector with 2K x 2K pixels, MOSFIRE also offers significant improvements in sensitivity over other facility NIR spectrographs. The optical design of MOSFIRE is shown in Figure B.1. MOSFIRE was successfully commissioned last Spring on the Keck I telescope, and began shared-risk science operations in 2012B.

This appendix briefly discusses my principal contributions to the MOSFIRE instrumentation project. Section B.1 focuses on MAGMA, the MOSFIRE slitmask design tool, for which I was co-author along with software engineer Jason Weiss and Caltech undergraduate Christopher Klein. Section B.2 discusses XTcalc, an GUI-based spectroscopic exposure time calculator for which I was author.

B.1 MAGMA: MOSFIRE Slitmask Design Software

The MOSFIRE Automatic GUI-based Mask Application, MAGMA, is an interactive slitmask design tool programed in Java. The tool is used both to design slitmasks as well as execute the mechanical repositioning of the 92 CSU bars which form the 46 slits for spectroscopic observations with MOSFIRE. The code is available as a stand-alone tool in order to prepare for observations and is integrated into the instrument control software suite.

The tool is designed to take in a set of coordinates of objects with assigned priorities. It then calculates the maximum priority mask that can be achieved at a series of center locations for the mask and a series of different position angles. The center position, PA, and the number of rotational and translational steps can be set by the user. Once the calculation is completed for the desired number of positions and PAs, the highest priority science mask is reported to the user and displayed in the GUI.

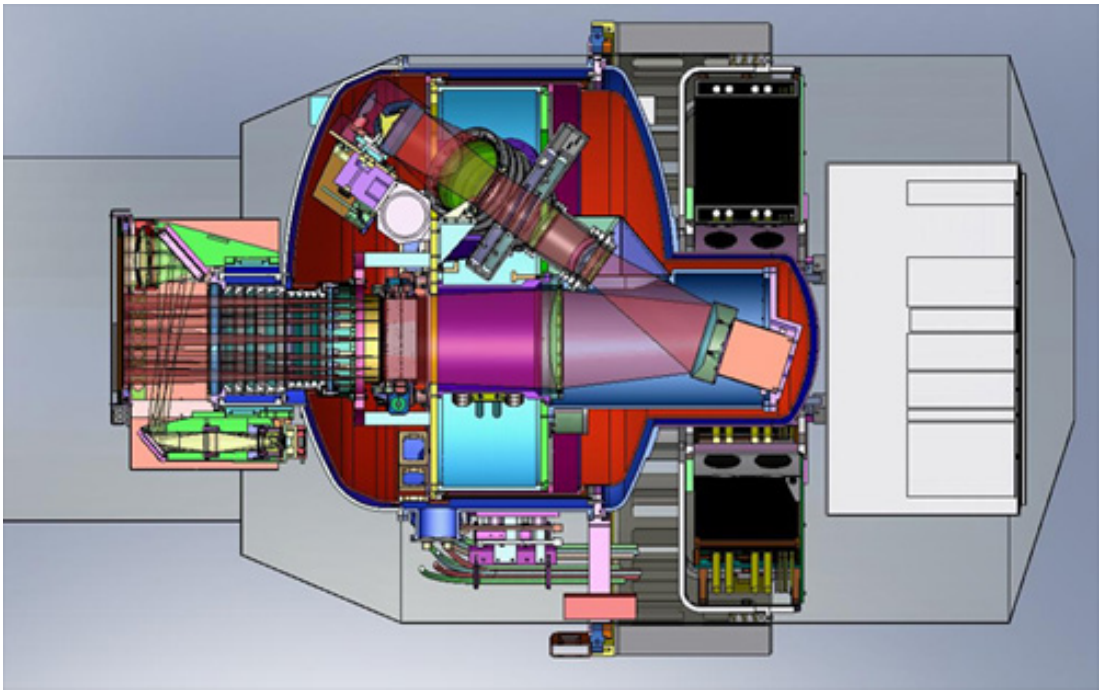
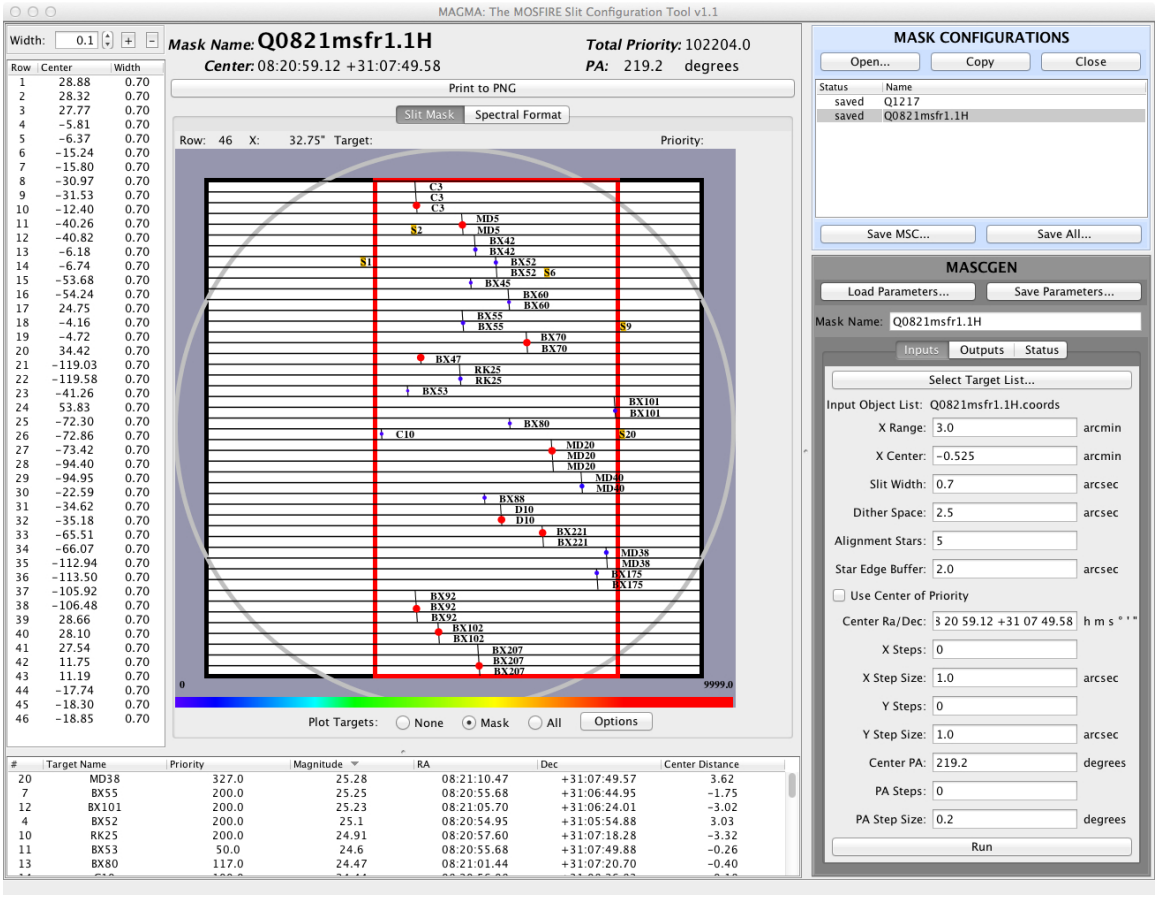


Figure B.1 The optical design of MOSFIRE.

An example of the stand alone code is shown in Figure B.2. In the bottom right-hand panel, the user specifies several input mask parameters. Users inputs include:

- a prioritized coordinate list of science targets and alignment stars. This list can be selected using the "Select Target List..." button.
- the X Range, the effective "field of view" in which a slit can be placed along the spectroscopic direction. This can be increased in order to include a larger field of view or decreased to constrain the specific spectral coverage of the observations of all slits, as the science requires.
- the X Center, the center position of the box with respect to the field of view. Again, this is adjusted to change the spectral coverage of the observations.
- the width of the slits
- the dither space: the minimum distance in arcseconds between the edge of the slit and a science target. This distance should be greater than the planned distance between dither locations.
- the number of alignment stars required. Generally, 3 stars are necessary to provide a good alignment; however, we recommend using 4-5 in case one or two must be rejected due to bad pixels, cosmic rays, etc. Note that the bars assigned to alignment stars will be repositioned to a science target once the mask alignment is complete.
- the star edge buffer: the minimum distance a star can be to the edge of a alignment box



#	Target Name	Priority	Magnitude	RA	Dec	Center Distance
20	MD38	327.0	25.28	08:21:10.47	+31:07:49.57	3.62
7	BX55	200.0	25.25	08:20:55.68	+31:06:44.95	-1.75
12	BX101	200.0	25.23	08:21:05.70	+31:06:24.01	-3.02
4	BX52	200.0	25.1	08:20:54.95	+31:05:54.88	3.03
10	RK25	200.0	24.91	08:20:57.60	+31:07:18.28	-3.32
11	BX53	50.0	24.6	08:20:55.68	+31:07:49.88	-0.26
13	BX80	117.0	24.47	08:21:01.44	+31:07:20.70	-0.40

Figure B.2 The MAGMA software. The box on the lower right allows for user input of the mask parameters. The box at the upper right allows users to save mask configuration files and load in pre-saved files. The center box shows the design of the mask produced by the software. The dots indicate the location of the object on the slit, the color of which designates its priority. The yellow boxes show the positions of the alignments boxes. The gray circle indicated the imaging field of view of MOSFIRE (slits placed outside the circle would not see light from the telescope) while the red box indicates the user-specified allowed positioning of the slits. The box is offset from the center of the circle in order to adjust the wavelength coverage of the science observations.

In addition to these, the user also specifies a number of parameters regarding how the search for the highest priority mask should be conducted. This includes:

- The center of the field at which to start the search. One can also choose to use the center coordinate from the object list.
- The number of steps in X and Y.
- The size of the steps in X and Y.
- The starting position angle of the mask from where to begin the search.
- The number and size of the PA steps.

Generally, the possible masks are very sensitive to changes in PA, and so a large number of PA and translational steps guarantee that the algorithm will find the highest priority mask possible; however, the algorithm is somewhat computational expensive. A typical run time for 60 one arcsecond steps in X and in Y along with 90 two degree steps in PA will generally take over an hour, depended upon processor and the number of targets in the priority list.

When all of the steps are calculated, a highest-priority mask is reported and its configuration is shown in the center panel of the GUI, see Figure B.2. The bottom panel shows various parameters of the targets that were selected to be put on the mask. The left hand panel shows the positions of the bars to be sent to the CSU as well as the selected slit width for each bar.

The spectral coverage of each slit given its placement within the field of view can be seen by clicking on the “Spectral Format” button near the top of the GUI above the graphic display as shown in Figure B.3. The numbers indicate the start and end wavelengths of each spectrum. Note that if the spectral coverage is not satisfactory, editing the X Range and X Center parameters will result in different spectral coverage variability between slits and different central wavelengths on each slit respectively.

When the mask is complete and ready to be saved, clicking the “Save All” button in the blue MASK CONFIGURATION section of the GUI will save all of the parameters of the mask so that they can be loaded at the observatory in order to configure the CSU. Such files can also be reloaded into MAGMA in order to alter the mask or make new masks with somewhat different parameters.

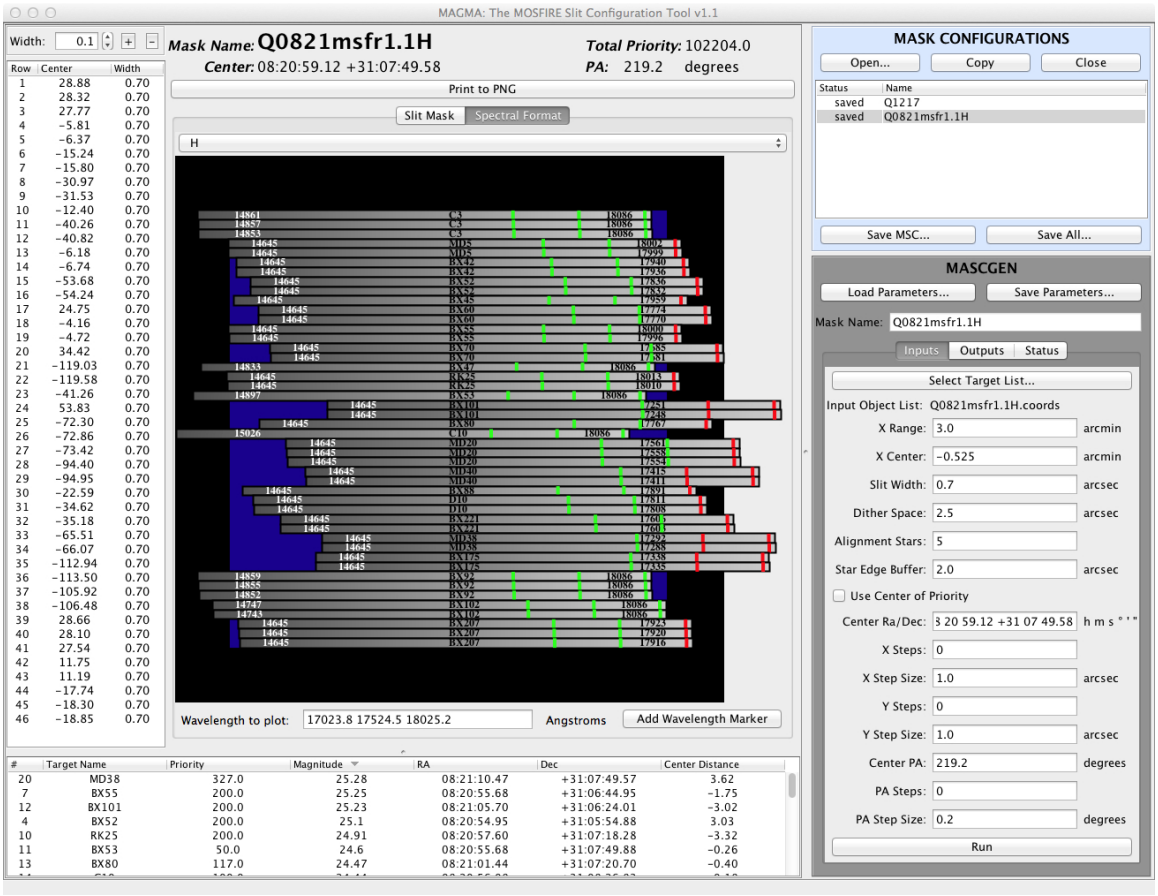


Figure B.3 Same as Figure B.2 only showing the Spectral Format viewing mode, in this case for the *H* band. The gray bars indicate the full wavelength span of the *H* band. The blue box indicated the detector footprint of the spectroscopic mode. The numbers on each gray bar indicate the star and end wavelengths covered by that slit. The green and red vertical bars mark the location a user specified wavelength along each slit, in this case corresponding to the location of [O III]λ5007 at a redshift of 2.4, 2.5, and 2.6. The line is green in the case that the wavelength falls on the detector footprint of the spectrograph and red if it falls outside the footprint of the detector.

B.2 XTcalc: MOSFIRE Exposure Time Calculator

XTcalc is a MOSFIRE-specific spectral exposure time calculator programed in IDL. The code can be run with the IDL Virtual Machine which does not require the user to have purchased an IDL license. Below is a users manual which explains the installation and usage of the code along with the mathematical basis of the calculation and the sources of noise included. Notably the code can be used to calculate the S/N both for a full user input spectrum or, more simply, for a spectral emission line of specified parameters. XTcalc can also generate simulated spectra given an input science spectrum which specifies the spectral shape of the science target.

B.2.1 Installation using IDL Virtual Machine

This is the default way to run the code. It does not require an IDL license.

B.2.1.1 Installation

First, the user should download and instal the IDL virtual machine software. The software can be downloaded from http://www.exelisvis.com/ProductsServices/IDL/IDLModules/IDLVirtualMachine.aspx

Note that XTcalc was developed to be launched with idl from the command line. To run IDL from the command line, you must source the 'idl_setup' file which is included in the 'bin' directory of the IDL installation.

Once you have IDL set up, unpack the tarball and then set the environment variables with the location of the unpacked tarball:

- 'MOSFIRE_XTCALC' with the path where you put the code: REQUIRED. Note the directory will be '[local_path]/XTcalc_dir'
- 'MOSFIRE_WD' optional path for input and output files for MOSFIRE. If not set, defaults to 'MOSFIRE_XTCALC'

B.2.1.2 Running the Code

From the command line, in the directory XTcalc_dir, run the following line:

idl -vm=run_XTcalc.sav

You can also supply the full path for run_XTcalc.sav.

B.2.2 Installation with Full IDL and License

If you already have a working IDL installation, the virtual machine is included. You can run the code as specified above. If you prefer, you can also compile the code and run it in the normal way with IDL as outlined below.

B.2.2.1 Installation

To run the code using a full IDL installation, XTcalc requires the astronomy idl library 'ASTROLIB'. Note that when running the code outside the virtual machine, any settings in your IDL setup file as well as any previous code run in your IDL session may affect the performance of the code. If you run into issues, please attempt to use the code with the virtual machine, as outlined above.

The following code and directories are required to run XTcalc. In the bin directory in the packaged version of XTcalc, there are three .pro files which must be compiled prior to running the code.

- XTcalc.pro: This is the main program which does the actual S/N and exposure time calculations.
- run_XTcalc.pro: This program creates and initializes the GUI.
- events_XTcalc.pro: These programs handle the GUI actions and events.

In addition to the code, the following directories are also required:

- mosfire: This directory contains the filter curves for MOSFIRE. The filter curves are named mosfire_[band].txt where [band] in Y, J, H, or K.
- Mauna_Kea_sky: This directory contains idl 'save' files which catalog the atmospheric transparency spectra. These files are called mktrans_zm_[water vapor]_[airmass].sav
- MosfireSkySpec: This directory contains the background sky spectra. They are formatted as idl 'save' files called [band]sky_cal_pA.sav
- MosfireSpecEff: This directory contains the throughput spectra for each band for MOSFIRE. The files are called [band]eff.sm.dat

The following environment variables should also be set:

- 'MOSFIRE_XTCALC' with the path where you put the directories: REQUIRED
- 'MOSFIRE_WD' optional path for input and output files for MOSFIRE. If not set, defaults to 'MOSFIRE_XTCALC'

B.2.2.2 Running the Code

- start IDL
- .r run_XTcalc
- .r XTcalc
- .r events_XTcalc
- run_XTcalc

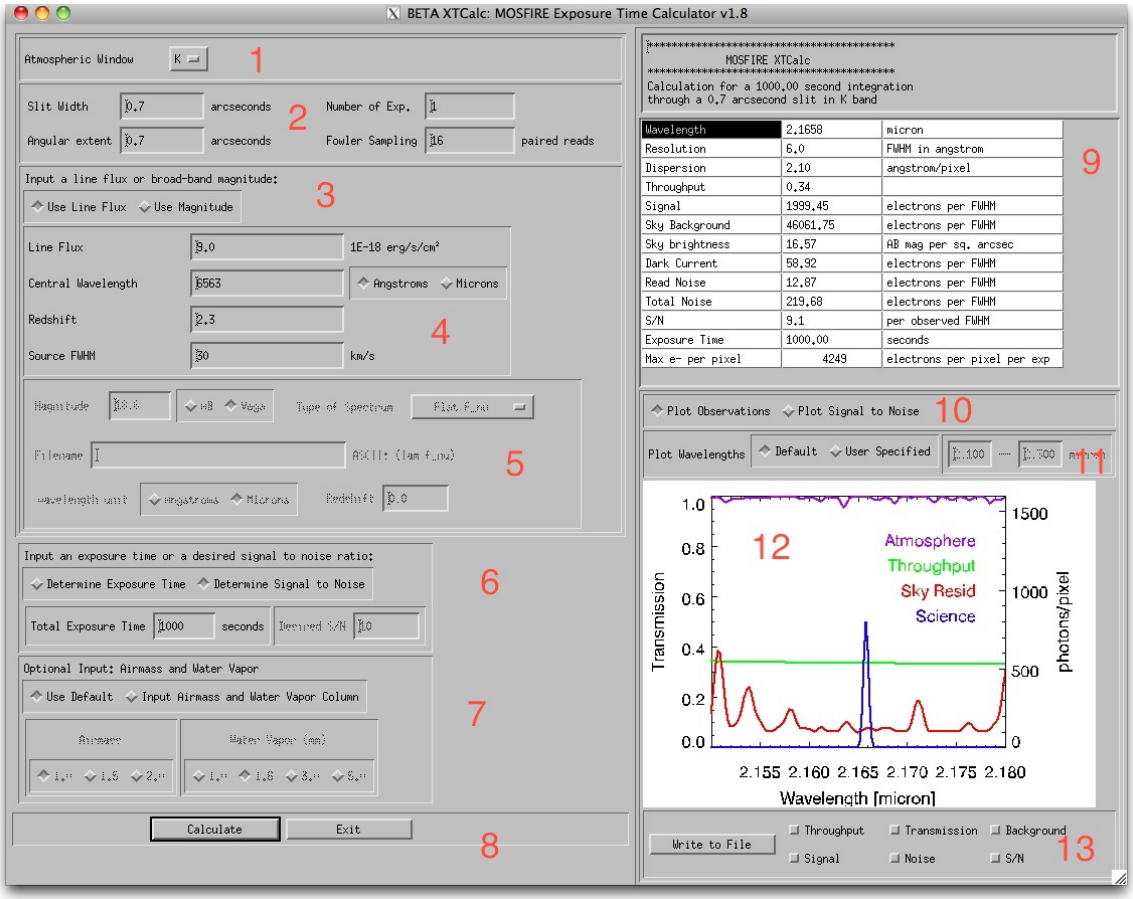


Figure B.4 The XTcalc GUI. The red numerals mark the positions of various features as outlined in the text.

B.2.3 Using XTcalc

XTcalc can be used in two principle modes: "Use Line Flux" or "Use Magnitude." The line flux mode is useful for understanding the S/N within the FWHM of an individual emission line. The magnitude mode calculates the S/N over the full band of observations. In this mode, the default spectral shape if a flat f_ν spectrum. However, users can also input their own ASCII spectrum to be considered. The program then scales the chosen spectral shape to match the input magnitude over the band pass considered.

B.2.3.1 Input

Please refer to Figure B.4 in the following section where the features of XTcalc are described. Sections 1-8 of the GUI allow users to input the properties of their objects. Sections 9-13 pertain to the output of the code.

- The atmospheric band (Y, J, H, or K) of the observations can be chosen with a pull down menu at position 1.

- At position 2, the slit width, angular extent of the object, number of exposures, and number of reads (fowler sampling) can be input.

- Position 3 allows users to choose either the line flux or continuum mode.

- Section 4 of the GUI pertains to the "Use line flux" mode. Within this section, the line flux, central wavelength, redshift, and FWHM of the line can be input.

- Section 5 pertains to the "Use Magnitude" function. Here the magnitude in AB or Vega can be specified. The spectral shape can be selected using the pull down menu at the top right-hand corner of the box. If the user would like to import their own spectra, choose "My Own Spectrum" from the pull down menu. The user will be prompted to select a file with the input ASCII spectrum with a format (lambda f_ν). The wavelengths can be provided in micron or Å. The spectrum can also be a rest wavelength spectrum, and the observed redshift can be entered.

- The box labeled "6" is where the user chooses to calculate the S/N given the exposure time or vice versa.

- Box 7 allows for adjustment in the atmospheric transparency model used in the calculation. Here the user can specify various airmasses and water vapor columns.

- Position 8 marks the location of the "Calculate" button which runs XTcalc with the specified parameters and the "Exit" button which closes the GUI.

B.2.3.2 Output

- Section 9 outputs the calculated values for the signal, various sources of noise, the S/N and the exposure time. The bottom line of the table is useful for checking the saturation level of the data. If the box turns yellow, the exposure time and number of exposures suggest that some of the pixels will pass the 1% linearity. Orange represents the 5% linearity. And red means some of the pixels are saturated.

- Section 12 provides graphical information regarding the science observations.

- Section 10 controls what is plotted in the graphical output at position 12. One can choose to plot the Signal and sky residual spectra. Alternatively, users can view the signal, noise and S/N of the spectra.

- Section 11 allows users to change the plotted range in section 12.

- Section 13 allows users to print out ASCII spectra of the throughput, transmission, background, signal, noise, and S/N for their proposed observations.

B.2.4 Throughput

The throughput of the instrument has been measured on sky and is implemented in XTcalc using these measurements. These throughput curves are data taken with MOSFIRE from which the atmospheric absorption has been removed.

B.2.5 Sky Background

The background sky spectrum assumed by XTcalc was taken with MOSFIRE during the engineering runs. The spectrum is integrated over the angular extent of the object θ and is scaled considering the slit width selected. **Note that when the calculator is used with a broad band magnitude, the angular size of the object is assumed to be the same as (or less than) the slit width. Slit losses are not currently calculated.**

We will refer to the spectrum of the sky background in units of photons $\sec^{-1}$ (spectral pixel)$^{-1}$ as $\beta_{\rm sky}$.

B.2.6 Atmospheric Transparency

The atmospheric transparency used in the calculator are the transparency spectra from Gemini Observatory: Lord, S. D., 1992, NASA Technical Memorandum 103957. The default parameters are 1.6 mm water vapor column and and airmass of 1, but they can be adjusted in the lower left hand panel of the GUI.

B.2.7 The Source Spectrum

B.2.7.1 Line Flux

The quantity calculated is the signal to noise ratio in the line. If the user inputs a S/N and asks for an exposure time, the program assumes the input value was the desired S/N *per FWHM* of the line. Note that this quantity depends on both the spectral resolution and the entered intrinsic line width.

B.2.7.2 Broad Band Magnitude

The quantity calculated is the signal to noise ratio per spectral pixel. If a S/N is input and the exposure time is calculated, the program assumes the input value is the desired S/N *per spectral pixel*.

Using the broad band magnitude calculator, the spectrum chosen is read in. The default is a spectrum that is flat in f_ν. The spectrum is convolved with the MOSFIRE resolution and resampled onto the MOSFIRE pixel scale. It is multiplied by the filter transmission spectrum, and then the spectrum is normalized such that the flux in the band pass matches the specified broad band magnitude. Multiplied by the atmospheric transparency spectrum and the instrument+telescope throughput spectrum. The resulting spectrum has units of photons $\sec^{-1}$ (spectral pixel)$^{-1}$ and will hence forth be refereed to as S.

Table B.1. Assumed Values for MOSFIRE

Name	Value	Units
Pixel Scale (spatial)	0.18	arcsec/pixel
Pixel Scale (dispersion)	0.24	arcsec/pixel
Detector Read Noise	21	electrons/pixel CDS
Dark Current	0.005	electrons/sec/pixel
K-band Dispersion	2.170	Å/pixel
H-band Dispersion	1.629	Å/pixel
J-band Dispersion	1.303	Å/pixel
Y-band Dispersion	1.086	Å/pixel
K-band Rθ Product	2534	arcsec
H-band Rθ Product	2562	arcsec
J-band Rθ Product	2317	arcsec
Y-band Rθ Product	2366	arcsec

B.2.8 Noise Calculation

The number of pixels over which the source is spread in the spatial and spectral dimension will factor into the calculation of the noise. The number of pixels along the spatial dimension, $N_{\rm Spatial}$, is taken to be the spatial extent of the object θ, divided by the spatial pixel scale, 0.18 arcseconds pixel^{-1}.

Here, we produce a spectrum of the noise. The units will be photons per spectral pixel. The increase in noise from spreading the light over the spatial direction will have already been factored in. The contributions to the noise, $\eta_{\rm tot}$, considered are the read noise and the Poisson noise from the background, dark current (assumed to be 0.005 electrons s^{-1} pixel^{-1}), and source.

B.2.8.1 Single Exposure

The read noise, $\eta_{\rm read}$, is the detector read noise, RN, in electron pixel^{-1} read^{-1} reduced by the Fowler sampling with $N_{\rm reads}$ and increased by the number of pixels.

$$\eta_{\rm read} = \frac{\rm RN}{\sqrt{N_{\rm reads}}} \sqrt{N_{\rm spatial}}. \tag{B.1}$$

The noise resulting from the dark current is Poisson in nature:

$$\eta_{\rm dark}(t) = \sqrt{N_{\rm Spatial} D t}. \tag{B.2}$$

Then the total noise per spectral pixel is:

$$\eta_{\rm tot} = \sqrt{St + \beta_{\rm sky} t + \eta^2_{\rm dark}(t) + \eta^2_{\rm read}} \tag{B.3}$$

where D is the dark current in units of electron sec^{-1} pix^{-1} (assumed to be 0.005). The source and sky Poisson noise is not multiplied by the number of spatial pixels in this expression because that has already been considered in the calculation of their spectrum in photon sec^{-1} (spectral pixel)$^{-1}$.

The signal to noise ratio per spectral pixel is then:

$$\mathrm{S/N} = \frac{S}{\eta}. \tag{B.4}$$

B.2.8.2 Multiple Exposures with a Two Point Dither

If the number of exposures, N_{exp}, is > 1, then a two point dither pattern is assumed in the S/N calculation. The exposure time input is considered to be the *total* exposure time, and the exposure time of the individual frames is taken to be t/N_{exp}. The total noise per spectral pixel after the frames are differences and co-added is:

$$\eta_{\mathrm{tot}} = \sqrt{N_{\mathrm{exp}}}\sqrt{S\frac{t}{N_{\mathrm{exp}}} + 2\left(\beta_{\mathrm{sky}}\frac{t}{N_{\mathrm{exp}}} + \eta^2_{\mathrm{dark}}\left(t/N_{\mathrm{exp}}\right) + \eta^2_{\mathrm{read}}\right)}. \tag{B.5}$$

The $\sqrt{N_{\mathrm{exp}}}$ results from the fact that the error in individual exposures is added in quadrature. And the 2 results from the additional contribution of read noise and Poisson noise from the background and dark current which occurs when using difference frames due to the noise in the estimate in the background.

This expression simplifies to:

$$\eta_{\mathrm{tot}} = \sqrt{St + 2\left(\beta_{\mathrm{sky}}t + \eta^2_{\mathrm{dark}}(t) + N_{\mathrm{exp}}\eta^2_{\mathrm{read}}\right)}. \tag{B.6}$$

The S/N per spectral pixel is then simply:

$$\mathrm{S/N} = \frac{St}{\sqrt{St + 2\left(\beta_{\mathrm{sky}}t + \eta^2_{\mathrm{dark}}(t) + N_{\mathrm{exp}}\eta^2_{\mathrm{read}}\right)}}. \tag{B.7}$$

B.2.8.3 S/N per Resolution Element or per FWHM

To instead have the S/N per resolution element or the S/N in a line, we multiply by the square root of the number of spectral pixels per resolution element or per FWHM, N_{Spectral}.

$$N_{\mathrm{Spectral}} = \frac{\mathrm{FWHM}}{\mathrm{dispersion}}, \tag{B.8}$$

for the S/N in a line, or

$$N_{\mathrm{Spectral}} = \frac{\mathrm{Resolution\ Element}}{\mathrm{dispersion}}, \tag{B.9}$$

for a continuum source.

The signal to noise reported for the "Use line flux" option is the signal to noise per FWHM within the line. The signal to noise reported for the "Use Magnitude" option in the median S/N per spectral pixel where

the median is taken over the full band pass of the selected filter.

B.2.9 GUI Table

The max e- per pixel will turn yellow if the observations will have a pixel past the 1% linearity point of the detector. If the observations surpass the 5% linearity level, the row will turn orange. If the pixel would saturate, the row will turn red.

Bibliography

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

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

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

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

Agertz, O., Teyssier, R., & Moore, B. 2009, MNRAS, 397, L64

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

Bahcall, J. N. & Peebles, P. J. E. 1969, ApJ, 156, L7

Barnes, L. A., Haehnelt, M. G., Tescari, E., & Viel, M. 2011, MNRAS, 416, 1723

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

—. 2011b, MNRAS, 410, 1096

Becker, G. D., Rauch, M., & Sargent, W. L. W. 2007, ApJ, 662, 72

—. 2009, ApJ, 698, 1010

Becker, G. D., Sargent, W. L. W., Rauch, M., & Calverley, A. P. 2011c, ApJ, 735, 93

Becker, G. D., Sargent, W. L. W., Rauch, M., & Carswell, R. F. 2012, ApJ, 744, 91

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2012, ArXiv: 1209.3013

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

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

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

Boksenberg, A. & Sargent, W. L. W. 1978, ApJ, 220, 42

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

Bolton, J. S., Viel, M., Kim, T.-S., Haehnelt, M. G., & Carswell, R. F. 2008, MNRAS, 386, 1131

Booth, C. M., Schaye, J., Delgado, J. D., & Dalla Vecchia, C. 2010, ArXiv e-prints

Bordoloi, R., Lilly, S. J., Knobel, C., Bolzonella, M., Kampczyk, P., Carollo, C. M., Iovino, A., Zucca, E., Contini, T., Kneib, J.-P., Le Fevre, O., Mainieri, V., Renzini, A., Scodeggio, M., Zamorani, G., Balestra, I., Bardelli, S., Bongiorno, A., Caputi, K., Cucciati, O., de la Torre, S., de Ravel, L., Garilli, B., Kovač, K., Lamareille, F., Le Borgne, J.-F., Le Brun, V., Maier, C., Mignoli, M., Pello, R., Peng, Y., Perez Montero, E., Presotto, V., Scarlata, C., Silverman, J., Tanaka, M., Tasca, L., Tresse, L., Vergani, D., Barnes, L., Cappi, A., Cimatti, A., Coppa, G., Diener, C., Franzetti, P., Koekemoer, A., López-Sanjuan, C., McCracken, H. J., Moresco, M., Nair, P., Oesch, P., Pozzetti, L., & Welikala, N. 2011, ApJ, 743, 10

Bouché, N., Murphy, M. T., Péroux, C., Davies, R., Eisenhauer, F., Förster Schreiber, N. M., & Tacconi, L. 2007, ApJ, 669, L5

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

Bridge, C. R., Teplitz, H. I., Siana, B., Scarlata, C., Conselice, C. J., Ferguson, H. C., Brown, T. M., Salvato, M., Rudie, G. C., de Mello, D. F., Colbert, J., Gardner, J. P., Giavalisco, M., & Armus, L. 2010, ApJ, 720, 465

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

Bryan, G. L. & Machacek, M. E. 2000, ApJ, 534, 57

Calura, F., Tescari, E., D'Odorico, V., Viel, M., Cristiani, S., Kim, T.-S., & Bolton, J. S. 2012, arXiv: 1201.5121

Carswell, B., Schaye, J., & Kim, T. 2002, ApJ, 578, 43

Carswell, R. F., Morton, D. C., Smith, M. G., Stockton, A. N., Turnshek, D. A., & Weymann, R. J. 1984, ApJ, 278, 486

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

Cen, R. & Ostriker, J. P. 2006, ApJ, 650, 560

Chelouche, D. & Bowen, D. V. 2010, ApJ, 722, 1821

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

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

—. 2001b, ApJ, 559, 654

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

Chen, H., Wild, V., Tinker, J. L., Gauthier, J., Helsby, J. E., Shectman, S. A., & Thompson, I. B. 2010, ApJ, 724, L176

Chen, H.-W., Lanzetta, K. M., Webb, J. K., & Barcons, X. 2001c, ApJ, 559, 654

Churchill, C. W., Kacprzak, G. G., & Steidel, C. C. 2005, in IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines, ed. P. Williams, C.-G. Shu, & B. Menard, 24–41

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

Clauset, A., Rohilla Shalizi, C., & Newman, M. E. J. 2007, arXiv: 0706.1062

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

Corbelli, E., Salpeter, E. E., & Bandiera, R. 2001, ApJ, 550, 26

Cowie, L. L., Songaila, A., Kim, T.-S., & Hu, E. M. 1995, AJ, 109, 1522

Cox, A. N. 2000, Allen's Astrophysical Quantities

Crighton, N. H. M., Bielby, R., Shanks, T., Infante, L., Bornancini, C. G., Bouché, N., Lambas, D. G., Lowenthal, J. D., Minniti, D., Morris, S. L., Padilla, N., Péroux, C., Petitjean, P., Theuns, T., Tummuangpak, P., Weilbacher, P. M., Wisotzki, L., & Worseck, G. 2011, MNRAS, 414, 28

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

Davé, R., Hellsten, U., Hernquist, L., Katz, N., & Weinberg, D. H. 1998, ApJ, 509, 661

Davé, R., Hernquist, L., Katz, N., & Weinberg, D. H. 1999, ApJ, 511, 521

Davé, R., Oppenheimer, B. D., & Finlator, K. 2011, MNRAS, 415, 11

Davé, R. & Tripp, T. M. 2001, ApJ, 553, 528

Dekel, A. & Birnboim, Y. 2006, MNRAS, 368, 2

—. 2008, MNRAS, 383, 119

Dekker, H., D'Odorico, S., Kaufer, A., Delabre, B., & Kotzlowski, H. 2000, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4008, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. M. Iye & A. F. Moorwood, 534–545

Devine, D. & Bally, J. 1999, ApJ, 510, 197

Duffy, A. R., Schaye, J., Kay, S. T., & Dalla Vecchia, C. 2008, MNRAS, 390, L64

Ellison, S. L., Lewis, G. F., Pettini, M., Sargent, W. L. W., Chaffee, F. H., Foltz, C. B., Rauch, M., & Irwin, M. J. 1999, Publications of the Astronomical Society of the Pacific, 111, pp. 946

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

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

Erb, D. K., Quider, A. M., Henry, A. L., & Martin, C. L. 2012, ApJ, 759, 26

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

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

—. 2006c, ApJ, 646, 107

Evoli, C. & Ferrara, A. 2011, MNRAS, 413, 2721

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

Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, 417, 2982

Faucher-Giguère, C.-A., Lidz, A., Hernquist, L., & Zaldarriaga, M. 2008, ApJ, 688, 85

Faucher-Giguère, C.-A. & Quataert, E. 2012, MNRAS, 425, 605

Ferland, G. J., Korista, K. T., Verner, D. A., Ferguson, J. W., Kingdon, J. B., & Verner, E. M. 1998, PASP, 110, 761

Finlator, K. & Davé, R. 2008, MNRAS, 385, 2181

Fumagalli, M., Prochaska, J. X., Kasen, D., Dekel, A., Ceverino, D., & Primack, J. R. 2011, MNRAS, 418, 1796

Fynbo, J. P. U., Jakobsson, P., Prochaska, J. X., Malesani, D., Ledoux, C., de Ugarte Postigo, A., Nardini, M., Vreeswijk, P. M., Wiersema, K., Hjorth, J., Sollerman, J., Chen, H.-W., Thöne, C. C., Björnsson, G., Bloom, J. S., Castro-Tirado, A. J., Christensen, L., De Cia, A., Fruchter, A. S., Gorosabel, J., Graham, J. F., Jaunsen, A. O., Jensen, B. L., Kann, D. A., Kouveliotou, C., Levan, A. J., Maund, J., Masetti, N., Milvang-Jensen, B., Palazzi, E., Perley, D. A., Pian, E., Rol, E., Schady, P., Starling, R. L. C., Tanvir, N. R., Watson, D. J., Xu, D., Augusteijn, T., Grundahl, F., Telting, J., & Quirion, P.-O. 2009, ApJS, 185, 526

Garzilli, A., Bolton, J. S., Kim, T.-S., Leach, S., & Viel, M. 2012, ArXiv e-prints

Gauthier, J., Chen, H., & Tinker, J. L. 2010, ApJ, 716, 1263

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

Heckman, T. M. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 254, Extragalactic Gas at Low Redshift, ed. J. S. Mulchaey & J. T. Stocke, 292

Heckman, T. M., Armus, L., & Miley, G. K. 1990, ApJS, 74, 833

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

Hopkins, A. M. & Beacom, J. F. 2006, ApJ, 651, 142

Hopkins, P. F., Quataert, E., & Murray, N. 2012, MNRAS, 421, 3522

Hu, E. M., Kim, T.-S., Cowie, L. L., Songaila, A., & Rauch, M. 1995, AJ, 110, 1526

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

Iwata, I., Inoue, A. K., Matsuda, Y., Furusawa, H., Hayashino, T., Kousai, K., Akiyama, M., Yamada, T., Burgarella, D., & Deharveng, J.-M. 2009, ApJ, 692, 1287

Janknecht, E., Reimers, D., Lopez, S., & Tytler, D. 2006, A&A, 458, 427

Jensen, B. L., Fynbo, J. U., Gorosabel, J., Hjorth, J., Holland, S., Möller, P., Thomsen, B., Björnsson, G., Pedersen, H., Burud, I., Henden, A., Tanvir, N. R., Davis, C. J., Vreeswijk, P., Rol, E., Hurley, K., Cline, T., Trombka, J., McClanahan, T., Starr, R., Goldsten, J., Castro-Tirado, A. J., Greiner, J., Bailer-Jones, C. A. L., Kümmel, M., & Mundt, R. 2001, A&A, 370, 909

Jones, T., Stark, D. P., & Ellis, R. S. 2012, ApJ, 751, 51

Kacprzak, G. G. & Churchill, C. W. 2011, ApJ, 743, L34

Kacprzak, G. G., Churchill, C. W., Evans, J. L., Murphy, M. T., & Steidel, C. C. 2011, MNRAS, 416, 3118

Kacprzak, G. G., Churchill, C. W., & Nielsen, N. M. 2012, ApJ, 760, L7

Kaiser, N. 1987, MNRAS, 227, 1

Katz, N., Weinberg, D. H., Hernquist, L., & Miralda-Escude, J. 1996, ApJ, 457, L57

Kennicutt, Jr., R. C. 1998, ARA&A, 36, 189

Kereš, D. & Hernquist, L. 2009, ApJ, 700, L1

Kereš, D., Katz, N., Davé, R., Fardal, M., & Weinberg, D. H. 2009, MNRAS, 396, 2332

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

Kim, T.-S., Bolton, J. S., Viel, M., Haehnelt, M. G., & Carswell, R. F. 2007, MNRAS, 382, 1657

Kim, T.-S., Carswell, R. F., Cristiani, S., D'Odorico, S., & Giallongo, E. 2002, MNRAS, 335, 555

Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, ApJ, 740, 102

Kollmeier, J. A., Miralda-Escudé, J., Cen, R., & Ostriker, J. P. 2006, ApJ, 638, 52

Kornei, K. A., Shapley, A. E., Martin, C. L., Coil, A. L., Lotz, J. M., Schiminovich, D., Bundy, K., & Noeske, K. G. 2012, ApJ, 758, 135

Kronberg, P. P., Biermann, P., & Schwab, F. R. 1985, ApJ, 291, 693

Lacki, B. C., Thompson, T. A., Quataert, E., Loeb, A., & Waxman, E. 2011, ApJ, 734, 107

Lanzetta, K. M. & Bowen, D. 1990, ApJ, 357, 321

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

Law, D. R., Steidel, C. C., Erb, D. K., Larkin, J. E., Pettini, M., Shapley, A. E., & Wright, S. A. 2009, ApJ, 697, 2057

Law, D. R., Steidel, C. C., Shapley, A. E., Nagy, S. R., Reddy, N. A., & Erb, D. K. 2011, ArXiv e-prints

Lehner, N., Howk, J. C., Tripp, T. M., Tumlinson, J., Prochaska, J. X., O'Meara, J. M., Thom, C., Werk, J. K., Fox, A. J., & Ribaudo, J. 2013, ArXiv e-prints

Lehner, N., Savage, B. D., Richter, P., Sembach, K. R., Tripp, T. M., & Wakker, B. P. 2007, ApJ, 658, 680

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

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

Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., & Fruchter, A. 1996, MNRAS, 283, 1388

Madau, P., Haardt, F., & Rees, M. J. 1999, ApJ, 514, 648

Martin, C. L. 2005, ApJ, 621, 227

McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., & Cen, R. 2001, ApJ, 562, 52

McLean, I. S., Becklin, E. E., Bendiksen, O., Brims, G., Canfield, J., Figer, D. F., Graham, J. R., Hare, J., Lacayanga, F., Larkin, J. E., Larson, S. B., Levenson, N., Magnone, N., Teplitz, H., & Wong, W. 1998, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 3354, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. A. M. Fowler, 566–578

McQuinn, M., Lidz, A., Zaldarriaga, M., Hernquist, L., Hopkins, P. F., Dutta, S., & Faucher-Giguère, C.-A. 2009, ApJ, 694, 842

McQuinn, M., Oh, S. P., & Faucher-Giguère, C.-A. 2011, ApJ, 743, 82

Meiksin, A. A. 2009, Reviews of Modern Physics, 81, 1405

Ménard, B., Wild, V., Nestor, D., Quider, A., & Zibetti, S. 2009, ArXiv e-prints

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

Miralda-Escudé, J., Haehnelt, M., & Rees, M. J. 2000, ApJ, 530, 1

Misawa, T., Tytler, D., Iye, M., Kirkman, D., Suzuki, N., Lubin, D., & Kashikawa, N. 2007, AJ, 134, 1634

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

Moster, B. P., Somerville, R. S., Maulbetsch, C., van den Bosch, F. C., Macciò, A. V., Naab, T., & Oser, L. 2010, ApJ, 710, 903

Murakami, I. & Ikeuchi, S. 1990, PASJ, 42, L11

Murray, N., Quataert, E., & Thompson, T. A. 2010, ApJ, 709, 191

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Nelson, D., Vogelsberger, M., Genel, S., Sijacki, D., Kereš, D., Springel, V., & Hernquist, L. 2013, MNRAS, 429, 3353

Nestor, D. B., Johnson, B. D., Wild, V., Ménard, B., Turnshek, D. A., Rao, S., & Pettini, M. 2010, ArXiv e-prints

Nestor, D. B., Shapley, A. E., Steidel, C. C., & Siana, B. 2011, ApJ, 736, 18

Nielsen, N. M., Churchill, C. W., & Kacprzak, G. G. 2012, ArXiv e-prints

Noterdaeme, P., Petitjean, P., Ledoux, C., & Srianand, R. 2009, A&A, 505, 1087

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

Oke, J. B., Cohen, J. G., Carr, M., Cromer, J., Dingizian, A., Harris, F. H., Labrecque, S., Lucinio, R., Schaal, W., Epps, H., & Miller, J. 1995, PASP, 107, 375

Oke, J. B. & Korycansky, D. G. 1982, ApJ, 255, 11

O'Meara, J. M., Prochaska, J. X., Burles, S., Prochter, G., Bernstein, R. A., & Burgess, K. M. 2007, ApJ, 656, 666

O'Meara, J. M., Prochaska, J. X., Worseck, G., Chen, H.-W., & Madau, P. 2012, arXiv: 1204.3093

Oppenheimer, B. D. & Davé, R. 2008, MNRAS, 387, 577

Oppenheimer, B. D., Davé, R., Kereš, D., Fardal, M., Katz, N., Kollmeier, J. A., & Weinberg, D. H. 2010, MNRAS, 406, 2325

Panter, B., Jimenez, R., Heavens, A. F., & Charlot, S. 2007, MNRAS, 378, 1550

Paresce, F., McKee, C. F., & Bowyer, S. 1980, ApJ, 240, 387

Peeples, M. S., Weinberg, D. H., Davé, R., Fardal, M. A., & Katz, N. 2010, MNRAS, 404, 1281

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

—. 2004, ApJS, 152, 29

Petitjean, P., Bergeron, J., & Puget, J. L. 1992, A&A, 265, 375

Petitjean, P., Webb, J. K., Rauch, M., Carswell, R. F., & Lanzetta, K. 1993, MNRAS, 262, 499

Pettini, M., Hunstead, R. W., Smith, L. J., & Mar, D. P. 1990, MNRAS, 246, 545

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

Pettini, M., Steidel, C. C., Adelberger, K. L., Dickinson, M., & Giavalisco, M. 2000, ApJ, 528, 96

Prochaska, J. X., Chen, H.-W., Dessauges-Zavadsky, M., & Bloom, J. S. 2007, ApJ, 666, 267

Prochaska, J. X., O'Meara, J. M., & Worseck, G. 2010, ApJ, 718, 392

Prochaska, J. X., Weiner, B., Chen, H.-W., Mulchaey, J., & Cooksey, K. 2011, ApJ, 740, 91

Prochaska, J. X., Worseck, G., & O'Meara, J. M. 2009, ApJ, 705, L113

Rakic, O. 2012, PhD thesis, Leiden Observatory, The Netherlands

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

—. 2011b, arXiv:1109.4944

Rauch, M. 1998, ARA&A, 36, 267

Rauch, M., Miralda-Escude, J., Sargent, W. L. W., Barlow, T. A., Weinberg, D. H., Hernquist, L., Katz, N., Cen, R., & Ostriker, J. P. 1997, ApJ, 489, 7

Reddy, N., Dickinson, M., Elbaz, D., Morrison, G., Giavalisco, M., Ivison, R., Papovich, C., Scott, D., Buat, V., Burgarella, D., Charmandaris, V., Daddi, E., Magdis, G., Murphy, E., Altieri, B., Aussel, H., Dannerbauer, H., Dasyra, K., Hwang, H. S., Kartaltepe, J., Leiton, R., Magnelli, B., & Popesso, P. 2011, ArXiv e-prints

—. 2012, ApJ, 744, 154

Reddy, N. A. & Steidel, C. C. 2009, ApJ, 692, 778

Reddy, N. A., Steidel, C. C., Pettini, M., Adelberger, K. L., Shapley, A. E., Erb, D. K., & Dickinson, M. 2008, ApJS, 175, 48

Rees, M. J. 1986, MNRAS, 218, 25P

Rees, M. J. & Ostriker, J. P. 1977, MNRAS, 179, 541

Ribaudo, J., Lehner, N., & Howk, J. C. 2011, ApJ, 736, 42

Richards, G. T., Strauss, M. A., Fan, X., Hall, P. B., Jester, S., Schneider, D. P., Vanden Berk, D. E., Stoughton, C., Anderson, S. F., Brunner, R. J., Gray, J., Gunn, J. E., Ivezić, Ž., Kirkland, M. K., Knapp, G. R., Loveday, J., Meiksin, A., Pope, A., Szalay, A. S., Thakar, A. R., Yanny, B., York, D. G., Barentine, J. C., Brewington, H. J., Brinkmann, J., Fukugita, M., Harvanek, M., Kent, S. M., Kleinman, S. J., Krzesiński, J., Long, D. C., Lupton, R. H., Nash, T., Neilsen, Jr., E. H., Nitta, A., Schlegel, D. J., & Snedden, S. A. 2006, AJ, 131, 2766

Ricotti, M., Gnedin, N. Y., & Shull, J. M. 2000, ApJ, 534, 41

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

Rudie, G. C., Steidel, C. C., & Pettini, M. 2012a, ApJ, 757, L30

Rudie, G. C., Steidel, C. C., Shapley, A. E., & Pettini, M. 2013, ApJ, 769, 146

Rudie, G. C., Steidel, C. C., Trainor, R. F., Rakic, O., Bogosavljević, M., Pettini, M., Reddy, N., Shapley, A. E., Erb, D. K., & Law, D. R. 2012b, ApJ, 750, 67

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

Ryan-Weber, E. V., Pettini, M., & Madau, P. 2006, MNRAS, 371, L78

Ryan-Weber, E. V., Pettini, M., Madau, P., & Zych, B. J. 2009, MNRAS, 395, 1476

Sargent, W. L. W., Steidel, C. C., & Boksenberg, A. 1989, ApJS, 69, 703

Sargent, W. L. W., Young, P. J., Boksenberg, A., & Tytler, D. 1980, ApJS, 42, 41

Schaye, J. 2001, ApJ, 559, 507

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

Schaye, J., Theuns, T., Leonard, A., & Efstathiou, G. 1999, MNRAS, 310, 57

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

Schechter, P. 1976, ApJ, 203, 297

Schmidt, M. 1959, ApJ, 129, 243

—. 1962, ApJ, 136, 684

—. 1963, Nature, 197, 1040

Scott, J., Bechtold, J., Dobrzycki, A., & Kulkarni, V. P. 2000, ApJS, 130, 67

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

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

Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., & Erb, D. K. 2006, ApJ, 651, 688

Shen, S., Madau, P., Aguirre, A., Guedes, J., Mayer, L., & Wadsley, J. 2011, arXiv: 1109.3713

Siana, B., Teplitz, H. I., Ferguson, H. C., Brown, T. M., Giavalisco, M., Dickinson, M., Chary, R.-R., de Mello, D. F., Conselice, C. J., Bridge, C. R., Gardner, J. P., Colbert, J. W., & Scarlata, C. 2010, ApJ, 723, 241

Simcoe, R. A. 2011, ApJ, 738, 159

Simcoe, R. A., Cooksey, K. L., Matejek, M., Burgasser, A. J., Bochanski, J., Lovegrove, E., Bernstein, R. A., Pipher, J. L., Forrest, W. J., McMurtry, C., Fan, X., & O'Meara, J. 2011, ApJ, 743, 21

Simcoe, R. A., Sargent, W. L. W., & Rauch, M. 2002, ApJ, 578, 737

—. 2004, ApJ, 606, 92

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

Songaila, A. & Cowie, L. L. 1996, AJ, 112, 335

—. 2010, ApJ, 721, 1448

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

Steidel, C. C. 1992, PASP, 104, 843

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

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

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

Steidel, C. C., Dickinson, M., & Persson, S. E. 1994, ApJ, 437, L75

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

Steidel, C. C., Kollmeier, J. A., Shapley, A. E., Churchill, C. W., Dickinson, M., & Pettini, M. 2002, ApJ, 570, 526

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

Stengler-Larrea, E. A., Boksenberg, A., Steidel, C. C., Sargent, W. L. W., Bahcall, J. N., Bergeron, J., Hartig, G. F., Jannuzi, B. T., Kirhakos, S., Savage, B. D., Schneider, D. P., Turnshek, D. A., & Weymann, R. J. 1995, ApJ, 444, 64

Stewart, K. R., Brooks, A. M., Bullock, J. S., Maller, A. H., Diemand, J., Wadsley, J., & Moustakas, L. A. 2013, ArXiv e-prints

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

—. 2011b, ApJ, 738, 39

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

Thom, C., Tumlinson, J., Werk, J. K., Prochaska, J. X., Oppenheimer, B. D., Peeples, M. S., Tripp, T. M., Katz, N. S., O'Meara, J. M., Brady Ford, A., Davé, R., Sembach, K. R., & Weinberg, D. H. 2012, ApJ, 758, L41

Trainor, R. F. & Steidel, C. C. 2012, ApJ, 752, 39

Tumlinson, J., Thom, C., Werk, J. K., Prochaska, J. X., Tripp, T. M., Weinberg, D. H., Peeples, M. S., O'Meara, J. M., Oppenheimer, B. D., Meiring, J. D., Katz, N. S., Davé, R., Ford, A. B., & Sembach, K. R. 2011, Science, 334, 948

Tytler, D. 1982, Nature, 298, 427

—. 1987, ApJ, 321, 49

Umemura, M. & Ikeuchi, S. 1985, ApJ, 299, 583

van de Voort, F. & Schaye, J. 2011, arXiv: 1111.5039

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

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

Veilleux, S., Cecil, G., & Bland-Hawthorn, J. 2005, ARA&A, 43, 769

Viel, M., Bolton, J. S., & Haehnelt, M. G. 2009, MNRAS, 399, L39

Vogt, S. S., Allen, S. L., Bigelow, B. C., Bresee, L., Brown, B., Cantrall, T., Conrad, A., Couture, M., Delaney, C., Epps, H. W., Hilyard, D., Hilyard, D. F., Horn, E., Jern, N., Kanto, D., Keane, M. J., Kibrick, R. I., Lewis, J. W., Osborne, J., Pardeilhan, G. H., Pfister, T., Ricketts, T., Robinson, L. B., Stover, R. J., Tucker, D., Ward, J., & Wei, M. Z. 1994, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 2198, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. D. L. Crawford & E. R. Craine, 362–+

Wakker, B. P. & Savage, B. D. 2009, ApJS, 182, 378

Weiner, B. J., Coil, A. L., Prochaska, J. X., Newman, J. A., Cooper, M. C., Bundy, K., Conselice, C. J., Dutton, A. A., Faber, S. M., Koo, D. C., Lotz, J. M., Rieke, G. H., & Rubin, K. H. R. 2009, ApJ, 692, 187

Werk, J. K., Prochaska, J. X., Thom, C., Tumlinson, J., Tripp, T. M., O'Meara, J. M., & Meiring, J. D. 2012, ApJS, 198, 3

Wiersma, R. P. C., Schaye, J., Dalla Vecchia, C., Booth, C. M., Theuns, T., & Aguirre, A. 2010, MNRAS, 409, 132

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

Young, P. J., Sargent, W. L. W., Boksenberg, A., Carswell, R. F., & Whelan, J. A. J. 1979, ApJ, 229, 891

Zaldarriaga, M., Hui, L., & Tegmark, M. 2001, ApJ, 557, 519

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

Zheng, Z. & Miralda-Escudé, J. 2002, ApJ, 568, L71

Zibetti, S., Ménard, B., Nestor, D. B., Quider, A. M., Rao, S. M., & Turnshek, D. A. 2007, ApJ, 658, 161

...