THE COLUMN DENSITY DISTRIBUTION AND CONTINUUM OPACITY OF THE INTERGALACTIC AND CIRCUMGALACTIC MEDIUM AT REDSHIFT ($z = 2.4^*$)

Gwen C. Rudie1, Charles C. Steidel1, Alice E. Shapley2, and Max Pettini3

1 Cahill Center for Astronomy and Astrophysics, California Institute of Technology, MS 249-17, Pasadena, CA 91125, USA; gwen@astro.caltech.edu
2 Department of Astronomy, University of California, Los Angeles, 430 Portola Plaza, Los Angeles, CA 90024, USA
3 Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

Received 2012 August 24; accepted 2013 April 22; published 2013 May 15

ABSTRACT

We present new high-precision measurements of the opacity of the intergalactic and circumgalactic medium (IGM; CGM) at ($z = 2.4$). Using Voigt profile fits to the full Lyα and Lyβ 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 $14 \lesssim \log (N_{\text{HI}}/\text{cm}^{-2}) \lesssim 17.2$. 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., we find that there are fractionally more high-$N_{\text{HI}}$ absorbers than low-$N_{\text{HI}}$ 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 $\log (N_{\text{HI}}/\text{cm}^{-2}) < 17.2$. Reproducing the opacity measured in our data as well as the incidence of absorbers with $\log (N_{\text{HI}}/\text{cm}^{-2}) > 17.2$ requires a broken power law parameterization of the frequency distribution with a break near $N_{\text{HI}} \approx 10^{15}$ cm$^{-2}$. We compute new estimates of the mean free path ($\lambda_{\text{mfp}}$) to hydrogen-ionizing photons at $z_{\text{em}} = 2.4$, finding $\lambda_{\text{mfp}} = 147 \pm 15$ 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 $\lambda_{\text{mfp}} = 121 \pm 15$ Mpc. These $\lambda_{\text{mfp}}$ measurements are smaller than recent estimates and should inform future studies of the metagalactic UV background and of ionizing sources at $z \approx 2–3$.

Key words: intergalactic medium – quasars: absorption lines

Online-only material: color figures

1. INTRODUCTION

Observations of the Lyα 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 uncoupled 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, 2011; Ryan-Weber et al. 2009; Becker et al. 2009, 2011, 2012; Simcoe 2011) 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 ($N_{\text{HI}}$) 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 ($10^{12} \lesssim N_{\text{HI}} \lesssim 10^{22}$ cm$^{-2}$) which have necessitated surveys optimized for particular ranges in $N_{\text{HI}}$. This requirement is easily understood by considering the curve of growth of the Lyα line of neutral hydrogen.

The most common absorbers with $\log (N_{\text{HI}}/\text{cm}^{-2}) \lesssim 14$ are optically thin in the Lyα 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 ($\approx10$ km s$^{-1}$) high-signal-to-noise ratio ($S/N \sim 50$) spectra (Kim et al. 2002). At the highest column densities, $\log (N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 20.3$, 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 Lyα 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 $N_{\text{HI}}$, $14 \lesssim \log (N_{\text{HI}}/\text{cm}^{-2}) \lesssim 20$, lie on the flat part of the curve of growth thus requiring different techniques. Systems with $N_{\text{HI}} > 10^{17.2}$ cm$^{-2}$ have an optical depth to hydrogen-ionizing photons $\tau \gtrsim 1$, and so may be recognized by the strong breaks they produce in background QSO spectra at the Lyman limit (912 Å) in their rest frame. These so-called Lyman limit systems (LLSs) can thus be easily...
Finally there is the historically most-problematic range, $14 \lesssim \log (N_H / \text{cm}^{-2}) \lesssim 17.2$. These intermediate-$N_H$ 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-$N_H$ 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, $f(N, \chi)$, for the first time in a statistically robust sample. In our derivation of an analytic representation of $f(N, \chi)$ 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 $\log (N_H / \text{cm}^{-2}) < 17.2$ 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. (2012) 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. (2012) 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 interstellar medium (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 2 with a description of the unique data that enable these very precise $N_H$ measurements. In Section 3 we present the new $N_H$ frequency distribution measurements. Section 4 considers the frequency distribution within the CGM of galaxies and Section 5 quantifies the total opacity of the IGM and the environment of the source galaxy. Updated measurements of the mean free path ($\lambda_{\text{mfp}}$) to hydrogen-ionizing photons are presented in Section 6 along with their implications. The paper is summarized in Section 7.

Throughout this paper we assume a $\Lambda$-CDM cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. All distances are expressed in physical (proper) units unless stated otherwise. We use the 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.

2. DATA AND ANALYSIS

The data presented in this paper constitute a subset of the Keck Baryonic Structure Survey (KBSS). 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 hyperluminous ($m_V \sim 15.5-17$) 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 KBSS will be presented by C. C. Steidel et al. (in preparation), 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. (2012). 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 $R \approx 45,000$ (FWHM $\approx 7$ km s$^{-1}$), S/N $\sim 50-200$ pixel$^{-1}$, 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 $\lambda 1025.7$ down to at least $z = 2.2$ in all 15 of our sight lines, and to significantly lower redshift in many. The additional constraints provided by Lyβ and in many cases, additional Lyman series transitions allow for much more accurate measurements of H I for $N_H = 10^{14}-10^{17}$ 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. (2012). 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 1.

The line fitting was facilitated by a semi-automatic code that estimates input parameters (redshift $z$, H I column density $N_H$, and Doppler parameter $b_d$ of each absorber) using a cross-correlation technique as described in Rudie et al. (2012). These parameters were then input into the $\chi^2$ minimization code VPFIT$^5$ 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 $\log (N_H / \text{cm}^{-2})$ for these absorber measurements are listed in Table 2. Examples of fits to absorbers with $\log (N_H / \text{cm}^{-2}) > 15.5$ 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 $12.0 < \log (N_H / \text{cm}^{-2}) < 17.2$ and

$^4$ 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 1.

$^5$ http://www.ast.cam.ac.uk/~rfc/vpfit.html; ©2007 R. F. Carswell, J. K. Webb.
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 sight line, \( \Delta X_i \), is defined as

\[
dX = \frac{H_0}{H(z)} (1 + z)^2 dz
\]

\[
\Delta X_i = \int_{z_{\text{min}}}^{z_{\text{max}}} dX = \int_{z_{\text{min}}}^{z_{\text{max}}} \frac{(1 + z)^2}{\sqrt{\Omega_m + \Omega_\Lambda (1 + z)^3}} dz.
\]

The total pathlength covered by this survey is \( \Delta X = 26.9 \) and the \( \Delta X \) covered as a function of redshift summed over the 15 lines of sight is shown in Figure 1. The pathlength-weighted mean redshift of the sample is \( \langle z \rangle = 2.37 \). Measurements of \( f(N, X) \) derived from the KBSS sample of absorbers with 12.5 < log\((N_{\text{HI}}/\text{cm}^{-2})\) < 17.2 are presented in Figure 2. The data points are located at the mean log\((N_{\text{HI}})\) of the bin edges.

3. \( N_{\text{HI}} \) 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 \( N_{\text{HI}} \). We consider the differential column density distribution per unit path length, \( f(N_{\text{HI}}, X) \), defined as

\[
f(N_{\text{HI}}, X) dN_{\text{HI}} dX = \frac{m}{\Delta N_{\text{HI}}/\Delta X} dN_{\text{HI}},
\]

where \( m \) is the observed number of absorbers with column densities in the range \( \Delta N_{\text{HI}} \), \( \Delta X \) is the comoving pathlength of the survey summed over the sight lines. This comoving pathlength was introduced by Bahcall & Peebles (1969) in

2.02 < \( z < \) 2.84 over a total redshift pathlength of \( \Delta z = 8.27 \). This represents the largest Ly\( \alpha \) 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.

Table 1

| Name                  | R.A.          | Decl.          | \( z_{\text{QSO}} \) | \( \lambda_{\text{min}} \) (Å) | \( z \) Range | \( \Delta z \) | \( \Delta X \) | \( N_{\text{gal}} \) | S/N Ly\( \alpha \) | S/N Ly\( \beta \) |
|-----------------------|---------------|----------------|------------------------|-----------------------------|---------------|----------------|----------------|-----------------|----------------|----------------|
| 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)     | 02:09:59.0    | −00:05:06.5    | 2.872                  | 3227                        | 2.1532–2.8339 | 0.68           | 2.26           | 1               | 82             | 55             |
| Q0499–1645            | 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\(^a\)            | 3186                        | 2.1132–2.6031 | 0.49           | 1.59           | 0               | 99             | 58             |
| SBS1217+499           | 12:19:30.9    | +49:50:12.4    | 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\(^a\)            | 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           | 1.74           | 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             |

Notes.
\(^a\) 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 (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 Ly\( \alpha \) and Ly\( \beta \) absorption.
\(^f\) The redshift of this QSO was revised after the fitting of the HIRES spectrum was completed. The maximum redshift assumed \( z_{\text{QSO}} = 2.640 \).
\(^g\) The redshift of this QSO was revised after the fitting of the HIRES spectrum was completed. The maximum redshift assumed \( z_{\text{QSO}} = 2.542 \).

Table 2

| \( \log(N_{\text{HI}}/\text{cm}^{-2}) \) | \( \sigma_{\log(N_{\text{HI}})} \) |
|-----------------|-----------------|
| 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            |
The Astrophysical Journal, 769:146 (24pp), 2013 June 1

Rudie et al.

Figure 1. Histogram of the pathlength sampled as a function of redshift. The pathlength-weighted average redshift of the sample is \( z = 2.37 \).

(A color version of this figure is available in the online journal.)

The most notable deviations from a single power law appear at the ends of the distribution. The low-\( N_{\text{HI}} \) 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 \( \log(\frac{N_{\text{HI}}}{\text{cm}^{-2}}) \gtrsim 19 \) 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 \( 19 < \log(\frac{N_{\text{HI}}}{\text{cm}^{-2}}) \leq 20.3 \) at \( z \approx 2 \) are by O’Meara et al. (2007); the frequency distribution of damped Ly\( \alpha \) systems with \( \log(\frac{N_{\text{HI}}}{\text{cm}^{-2}}) > 20.3 \) 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 \( \log(\frac{N_{\text{HI}}}{\text{cm}^{-2}}) < 17.2 \).

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

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

where \( n \) is the total number of absorbers in the sample and \( N_{\min} \) is 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

\[
\sigma_{\hat{\beta}} = \frac{\hat{\beta} - 1}{\sqrt{n}}.
\]

Similarly, the estimator of the normalization, \( C_{\text{HI}} \), is

\[
\hat{C}_{\text{HI}} = \frac{n (1 - \hat{\beta})}{\left( N_{\max}^{1-\hat{\beta}} - N_{\min}^{1-\hat{\beta}} \right) \Sigma \Delta X}
\]

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

4

Figure 2. The frequency distribution of absorbers from the KBSS H\textsc{i} sample (black data points). For bins with \( N_{\text{HI}} > 13.5 \text{ cm}^{-2} \), the number of absorbers per bin is displayed. The (red) line is a maximum likelihood power-law fit to the distribution of absorbers with \( 13.5 < \log(\frac{N_{\text{HI}}}{\text{cm}^{-2}}) < 17.2 \). The (green) triangle represents the frequency of LLS calculated from other studies as discussed in Section 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 5.1 discusses broken power law fits which better reproduce some aspects of the data, especially the incidence of LLSs.

(A color version of this figure is available in the online journal.)
In Figure 2 the MLE fit to the data is shown for absorbers with $13.5 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2$. 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 $N_{\text{HI}} = 13.5$ used in the fit is motivated as follows. As shown in Figure 2, $f(N_{\text{HI}}, X)$ for $\log(N_{\text{HI}}/\text{cm}^{-2}) \lesssim 13.5$ 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 $f(N_{\text{HI}}, X)$. Here we do not attempt to diagnose the cause of the low-$N_{\text{HI}}$ turnover, but simply note that above $\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 13.5$, $f(N, X)$ 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 $N_{\text{HI}}$ included in the fit as shown in Table 3. Note that a choice of $\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 13.3$ produces insignificant change in the maximum likelihood estimate of the slope, $\beta$.

Prior to the present work, the largest existing sample of Ly$\alpha$ forest absorbers at these redshifts was described by Kim et al. (2002). Their sample included 2130 H$\alpha$ absorbers with $12.5 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17$ observed along nine lines of sight. The most marked improvement over the previous measurements provided by our study is the use of higher-order transitions of H$\alpha$ which were generally not included in previous work. These allow for more precise measurements of $N_{\text{HI}}$ above the saturation point of Ly$\alpha$, $N_{\text{HI}} \approx 10^{14-14.5}$ cm$^{-2}$. Above $\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 15.5$ where the Ly$\beta$ transition saturates, the sample presented herein may suffer a degree of incompleteness at $z \lesssim 2.4$ as discussed in Appendix A.

In Figure 3, the KBSS data (black diamonds) are compared with the measurements of Kim et al. (2002) at $(z) = 2.1$ (blue squares). Notably, at $\log(N_{\text{HI}}/\text{cm}^{-2}) < 14$ there is good agreement between the two data sets, but for $\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 14$, the Kim et al. (2002) sample exhibits significant scatter not present in the KBSS measurements, with a systematic tendency to underestimate the number of intermediate-$N_{\text{HI}}$, ($14 \lesssim \log(N_{\text{HI}}/\text{cm}^{-2}) \lesssim 17$) 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 $\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 14$ and may explain why their measurements are systematically lower than those presented here.

### 3.3. Lyman Limit Systems

In this section, we quantify the incidence of absorbers with $N_{\text{HI}} > 10^{17.2}$ cm$^{-2}$. LLSs which are optically thick to hydrogen-ionizing photons. Because LLSs are relatively rare, and because their $N_{\text{HI}}$ 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 $\log(N_{\text{HI}}/\text{cm}^{-2}) > 17.2$ at $(z) \sim 2.4$ using data from two surveys 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 Stenger-Larrea et al. (1995). The subsamples are roughly equivalent in their statistical significance (comparable total pathlength and number of LLSs), with path-weighted mean redshifts of $(z) = 2.08$ and $(z) = 2.84$, respectively. The path-weighted mean redshift for the full sample is $(z_{\text{LLS}}) = 2.41$ based on 46 LLSs with $\log(N_{\text{HI}}/\text{cm}^{-2}) > 17.2$ in a total redshift path of $\Delta z = 25.95$ (AX = 88.31). These LLS samples provide a maximum likelihood estimate of the number of LLS per unit pathlength, $dN_{\text{LLS}}/dX = 0.52 \pm 0.08$.

A recent study of LLS near the redshift range of interest ($2 < z < 2.6$) using uniform HST data is described in O’Meara et al. (2013). They measure an integral constraint on the number of LLS absorbers with $\tau_{912} > 1$, $\log(N_{\text{HI}}/\text{cm}^{-2}) > 17.2$ and $\tau_{912} > 2$, $\log(N_{\text{HI}}/\text{cm}^{-2}) > 17.5$, where $\tau_{912}$ is the optical depth at the Lyman limit, 912 Å, in the absorber’s rest.

### Table 3

| $\log(N_{\text{HI},\text{min}})$ | $\beta$ | $\sigma_{\beta}$ | $\log(C_{\text{HI}})$ |
|-------------------------------|---------|-----------------|---------------------|
| 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           | 9.28                |
| 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               |

Note. * Maximum likelihood fits to the data with $\log(N_{\text{HI},\text{min}}) \leq \log(N_{\text{HI}}/\text{cm}^{-2}) \leq 17.2$. 

Figure 3. The frequency distribution of absorbers from the KBSS H$\alpha$ sample (black points) and the maximum likelihood fit to the data (red curve). For comparison, overplotted in (blue) squares are the data points from the $(z) = 2.1$ sample from Kim et al. (2002). Because the analysis of the KBSS sample includes higher-order H$\alpha$ transitions, we can measure the frequency of absorbers with $\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 14$ with much higher fidelity than has been previously possible. Notably, there is no evidence for a strong break in the power law distribution near $\log(N_{\text{HI}}/\text{cm}^{-2}) \approx 14$ as has been previously suggested. (A color version of this figure is available in the online journal.)
The Astrophysical Journal, 769:146 (24pp), 2013 June 1

Rudie et al.

4. FREQUENCY DISTRIBUTIONS WITH RESPECT TO THE POSITIONS OF GALAXIES

Rudie et al. (2012) considered the distribution of H1 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 pkpc of galaxies in the transverse direction and within 300 km s⁻¹ of the systemic redshift of galaxies; however, a tail of excess H1 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 ∼2.5 × 10¹¹ L☉ (Reddy et al. 2008, 2012). They have star formation rates of ∼30 M☉ yr⁻¹ (Erb et al. 2006) and halo masses of ∼10¹² M☉ (Adelberger et al. 2005; Conroy et al. 2008; Trainor & Steidel 2012; O. Rakic et al., in preparation).

Using the same KBSS galaxy sample as in Rudie et al. (2012), we measure the frequency distribution of H1 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 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.

The resulting frequency distributions for the 300 km s⁻¹ and 700 km s⁻¹ CGM are shown in Figure 4 in red and blue respectively. As found by Rudie et al. (2012), the frequency of all absorbers with log(N(H1)/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, X) within the CGM are listed in Table 4 and overplotted in Figure 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(H1) systems than low-N(H1) absorbers in the CGM. This result is qualitatively equivalent to results presented by Rudie et al. (2012) who found a strong correlation between N(H1) 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.

The characteristic N(H1) of the CGM is akin to the max(N(H1)) statistic presented in the left-hand panel of Figure 9 of Rudie et al. (2012). These authors typically found log(N(H1)/cm⁻²) > 14.5–15 within 300 pkpc of a galaxy. Misawa et al. (2007) considered the frequency distribution of absorbers surrounding H1 systems with log(N(H1)/cm⁻²) > 15, fitting to the absorber distribution within ±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

6. 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 H1 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 has twice the pathlength, ΔX = 1.9—and since f(N, X) is inversely proportional to the pathlength, the frequency of absorbers in the 700 km s⁻¹ CGM is approximately a factor of two smaller.

frame. These authors found dn_{LLS}/dX = 0.48^{+0.09}_{-0.12}, in excellent agreement with our compilation described above (dn_{LLS}/dX = 0.52 ± 0.08). For more optically thick systems, O’Meara et al. (2013) found dn_{LLS}/dX = 0.28^{+0.06}_{-0.06}. The consistency of these measurements with the lower-N(H1) frequency distribution will be discussed further in Section 5.

Unfortunately, surveys that rely on detecting breaks in the spectrum of QSOs at the redshift of the LLS can barely measure N(H1) accurately, especially for absorbers with τ_12 > 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(H1)/cm⁻²) ≲ 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(H1)/cm⁻²) < 17.2. First, we note that LLS surveys are sensitive to all absorbers with log(N(H1)/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(H1)/cm⁻²) ≳ 21, suggesting that there are very few absorbers with log(N(H1)/cm⁻²) > 22. Here, we assume that dn_{LLS}/dX is sensitive to absorbers with 17.2 < log(N(H1)/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(H1)/cm⁻²) < 17.2 held for the entire LLS regime. We then consider the incidence of LLSs given the fit parameters as follows:

\[ dn_{LLS}/dX = \int_{N_{\text{min}}}^{N_{\text{max}}} f(N, X) dN \]  

\[ dn_{LLS}/dX = \int_{N_{\text{min}}}^{N_{\text{max}}} C_{H1} N_{H1}^{-\beta} dN_{H1}, \]

where \( \log(N_{\text{min}}/\text{cm}^{-2}) = 17.2 \) and \( \log(N_{\text{max}}/\text{cm}^{-2}) = 22.0 \). We find that this extrapolation to log(N(H1)/cm⁻²) > 17.2 underestimates the number of LLSs significantly giving dn_{LLS}/dX = 0.21 (compared to dn_{LLS}/dX = 0.52 ± 0.08).

The green triangle in Figure 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 \( C_{H1,\text{LLS}} \) implied by the measured dn_{LLS}/dX by inverting Equation (9). Then the values of f(N, X) suggested by the data would be given by

\[ f(N_p, X) = C_{H1,\text{LLS}} N_p^{-\beta}, \]

for all values of \( N_p \). The height of the green triangle thus represents the values of f(N, X) if \( \beta_{\text{MLE}} \) 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_{LLS}/dX. Again, we can see clearly that the MLE fit to the frequency distribution (red dotted line) significantly underpredicts the measured LLS incidence.

We have seen that a single power law well approximates the intermediate-N(H1) absorbers; however, its extrapolation to log(N(H1)/cm⁻²) > 17.2 underestimates the number of LLSs significantly giving dn_{LLS}/dX = 0.21. In Section 5, we discuss other methods for parameterizing the distribution which better reproduce the high-N(H1) end of the frequency distribution.
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, \(C_{\text{HI}} = 7.262 \pm 0.431\), 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.

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 \(N_{\text{HI}}\) 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 \(N_{\text{HI}} = N_i\) is given by \(e^{-N_i\sigma}\) where \(\sigma\) is the photoionization cross section of hydrogen. In practice, \(\sigma\) depends sensitively on the energy of the photons considered:

\[
\sigma(v) \approx \sigma_{\text{LL}} \left( \frac{v}{v_{\text{LL}}} \right)^{-3},
\]

where \(\sigma_{\text{LL}} = 6.35 \times 10^{-18} \text{ cm}^2\) is the hydrogen ionization cross section to photons at the Lyman limit (912 Å) with energies of 1 Ry (13.6 eV), and \(v_{\text{LL}}\) 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 \(\sigma\) is removed.

In addition, for the time being, we calculate the opacity assuming that the universe is static. 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 Section 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 \(\log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2\) inferred directly from the data.

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

\[
\kappa = \frac{d\tau}{dr},
\]

where

\[
dr = \frac{1}{1+z} \frac{c}{H(z)} dz
\]

and \(c\) is the speed of light.

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

\[
\kappa_{\text{fit}} = \int_{N_{\text{min}}}^{N_{\text{max}}} f(N_{\text{HI}}, r)(1 - e^{-N_{\text{HI}}\sigma(v)}) dN_{\text{HI}}
\]

(Paresce et al. 1980).

\footnote{For the remainder of Section 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.}
Similarly, the opacity can be calculated directly from the data by summing over those absorbers \(i\) with \(N_{\text{min}} < N_{\text{HI}} < N_{\text{max}}\):
\[
\kappa_{\text{data}} = \frac{\sum (1 - e^{-N_{\text{HI}}/\sigma})}{\Delta r},
\]
where \(\Delta r\) is the sum of the proper distances probed by each sight line.

The uncertainty in the measurement of \(\kappa_{\text{data}}\) can be calculated using the bootstrap resampling method in which many sets of absorbers are drawn with replacement and \(\kappa_{\text{data}}\) is recalculated for each set. One can then estimate the probability distribution of \(\kappa_{\text{data}}\) by selecting the 68% confidence range. Values of \(\kappa_{\text{data}}\) for the IGM and CGM for \(13.5 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2\) is listed in the last column of Table 4. We emphasize that this measurement of the Ly\(\alpha\) opacity from absorbers with \(\log(N_{\text{HI}}/\text{cm}^{-2}) < 17\) is one of the principal results of this work.

The left-hand panel of Figure 5 shows the opacity produced by various bins of \(\log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2\) within the IGM calculated directly from the data using Equation (15). These values are compared to the opacity found within the CGM in the right-hand panel of Figure 5. Notably, the opacity of the CGM of a galaxy (300 \(\text{km s}^{-1}\), 300 \(\text{pkpc}\)) 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 the data for absorbers with \(\log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2\) (Table 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 Ly\(\alpha\) opacity) is underpredicted by the MLE fit. In Section 5.1, we consider other parameterizations of the frequency distribution that better match \(dn_{\text{LLS}}/dX\) as well as the opacity from intermediate-\(N_{\text{HI}}\) systems [15 \(\lesssim \log(N_{\text{HI}}/\text{cm}^{-2}) \lesssim 17\)] calculated from the data.

5.1. Broken Power Law Parameterizations

While the MLE power law fits presented in Sections 3.2 and 4 provide a reasonable description of the overall column density distribution, because there are substantially more absorbers at low-\(N_{\text{HI}}\) the majority of the statistical power for the fit is derived from the low-\(N_{\text{HI}}\) absorbers. To estimate the Ly\(\alpha\) opacity of the intergalactic and circumgalactic gas, the absorbers of relevance are instead the much rarer, high-\(N_{\text{HI}}\) systems, particularly those near \(\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 17\).

In order to reproduce simultaneously the opacity from both absorbers with \(12.5 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2\) 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 binned8 data. Because the opacity is a strong function of \(N_{\text{HI}}\), it is most important that the number of absorbers with \(\log(N_{\text{HI}}/\text{cm}^{-2}) \approx 17\) is accurately reproduced. Therefore, to provide further flexibility, we allow for a discontinuous break in the power law parameterization. This allows for an independent fit to the high-\(N_{\text{HI}}\) absorbers [\(\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 15\)] 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 Ly\(\alpha\) 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-\(N_{\text{HI}}\) groups to minimize the discontinuity between the two power laws. We allow the break to vary in \(N_{\text{HI}}\) and select the break location to be that which best reproduces the opacity from absorbers with \(\log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2\) 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 6 and Table 5 along with their implied opacities, \(\kappa_{\text{fit}}\). The bin containing the break is indicated by the hatched box in Figure 6. Notably, the location of the break (\(N_{\text{HI}} \approx 10^{15} \text{ cm}^{-2}\)) is similar to the \(N_{\text{HI}}\) at which absorbers were found to significantly correlate with the positions of galaxies (Rudie et al. 2012).

This fitting technique reproduces the opacity from intermediate-\(N_{\text{HI}}\) absorbers measured from the data with high fidelity. Additionally, the IGM fit predicts a \(dn_{\text{LLS,fit}}/dX = 0.519\) for \(17.2 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 22.0\) in excellent agreement with \(dn_{\text{LLS, data}}/dX = 0.52 \pm 0.08\) determined from the compilation

8 We have verified that the fitted parameters are insensitive to the exact choice of binning.
of LLSs from the literature (Section 3.3). As in Section 3.3, the green triangle in Figure 6 represents the values of \( f(N, X) \) within the LLS regime implied by the measurement of \( dn_{LLS, data}/dX \) assuming a power law index, \( \beta = \beta_{high-N_{HI}} \). The agreement between the height on 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 (\( C_{N_{HI}} \)) is due to its covariance with the power law index, \( \beta \). In Table 6 we list the best fit determination of \( \beta_{high-N_{HI}} \) and \( C_{high-N_{HI}} \), and the values of \( \kappa_{fit} \) and \( dn_{LLS, fit}/dX \) implied by this analytic approximation. Also listed are the 1\( \sigma \) perturbed values of \( \beta_{high-N_{HI}} \), the corresponding normalization \( C_{high-N_{HI}} \), and their implied \( \kappa_{fit} \) and \( dn_{LLS, fit}/dX \). Notably, the perturbed analytic parameters reproduce the opacity due to Ly\( \alpha \) forest absorbers within the measurement errors. The incidence of LLSs is more sensitive to perturbed values of \( \beta \) 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 7 compares the IGM broken power law fit to the frequency of LLS with \( t_{912} > 1 \) and \( t_{912} > 2 \) from O’Meara et al. (2013) as described in Section 3.3. The blue triangle corresponds to \( t_{912} > 1 \) absorbers, and the red triangle represents the \( t_{912} > 2 \) absorbers. As in Section 3.3, the normalization of the triangles are set by the measured value of \( dn_{LLS, data}/dX \) under the assumption that an extrapolation of the broken power law fit to the Ly\( \alpha \) 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 distribution reproduces the measure LLS statistics from O’Meara et al. (2013). The green point in Figure 7 shows the measured frequency of absorbers with \( 1 < t_{912} < 2 \) [17.2 < \( \log(N_{HI}/cm^{-2}) \) < 17.5] calculated from the difference between the measured values of \( dn_{LLS, \tau>1}/dX \) and \( dn_{LLS, \tau>2}/dX \).

The power law fit to \( f(N, X) \) suggested by O’Meara et al. (2013) is overlaid as the red dotted line for comparison.
This curve would imply a value for the opacity from absorbers with $\log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2$ of $\kappa_{\text{data}} = 0.0022$, nearly a factor of two smaller than that measured from the KBSS sample ($\kappa_{\text{data}} = 0.0039 \pm 0.0005$).

There are likely some deviations in the detailed distribution of $f(N_{\text{HI}}, X)$ in the subDLA (19.3 < $\log(N_{\text{HI}}/\text{cm}^{-2}) < 20.3$) and DLA ($\log(N_{\text{HI}}/\text{cm}^{-2}) > 20.3$) regimes as discussed in Prochaska et al. (2009) and O’Meara et al. (2013). We show the current constraints on these regions of the frequency distribution in the right-hand panel of Figure 7. We chose to neglect the differences between the power law fit and observed constraints on $f(N, X)$. Holding the measured value of $dn_{\text{LLS}}/dX$ fixed, such an assumption makes our measurements of the opacity from absorbers with $N_{\text{HI}} > 10^{17.2} \text{ cm}^{-2}$ a lower limit.

However, since the attenuation resulting from any system with $\log(N_{\text{HI}}/\text{cm}^{-2}) > 18$ is similar, the detailed shape of the distribution for high-$N_{\text{HI}}$ 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 $dn_{\text{LLS}}/dX$ fixed at the observed value, this gives us a constraint on the number of systems with $N_{\text{HI}} > 10^{17.2} \text{ cm}^{-2}$. The assumption of a given value of the power-law index, $\beta$, allows us to infer the opacity from a measurement of $dn_{\text{LLS}}/dX$. Allowing $\beta$ to take a shallower value to match the plotted points of $f(N, X)$ for the subDLAs and DLAs only moves absorbers to higher $N_{\text{HI}}$, and thus higher opacity compared to a model with a steeper power-law index. Therefore, our suggested opacities are likely a slight underprediction 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 $\beta$ is significantly less than 10%.

### 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-$N_{\text{HI}}$ absorbers, we can now measure the full opacity of the IGM and CGM. We calculate the opacity from LLS absorbers with $\log(N_{\text{HI}}/\text{cm}^{-2}) > 17.2$ using Equation (14) and the opacity from intermediate-$N_{\text{HI}}$ absorbers with $\log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2$

Note that the detailed distribution of $\kappa$ per bin in $N_{\text{HI}}$ would be affected by variation in $f(N_{\text{HI}})$ at DLA and subDLA column densities. Because we adopt the extrapolation of the broken power law fit to ($N_{\text{HI}}$) to calculate the opacity, the values in individual bins of $N_{\text{HI}}$, for $\log(N_{\text{HI}}/\text{cm}^{-2}) > 17.2$ are not well constrained.
directly from the data using Equation (15) (as described in the beginning of Section 5). The differential opacity from bins in \( N_{\text{HI}} \) as well as the cumulative opacity (from absorbers with \( N_{\text{HI}} \) smaller than a given value), are given in Table 7. The total opacity of the IGM is measured to be

\[
\kappa_{\text{IGM}} = \kappa_{\text{data}} \left( N_{\text{HI}} < 10^{17.2} \right) + \kappa_{\text{fit}} \left( N_{\text{HI}} > 10^{17.2} \right) \quad (16)
\]

\[
\kappa_{\text{IGM}} = 8.146 \times 10^{-3} \text{ pMpc}^{-1}. \quad (17)
\]

Similarly,

\[
\kappa_{\text{CGM}300 \text{ km s}^{-1}} = 94.4 \times 10^{-3} \text{ pMpc}^{-1} \quad (18)
\]

\[
\kappa_{\text{CGM}700 \text{ km s}^{-1}} = 46.5 \times 10^{-3} \text{ pMpc}^{-1}. \quad (19)
\]

Another relevant quantity is the integrated opacity over a specific scale length, i.e., the optical depth \( \tau \). 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.\(^{11}\) 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, \( \Delta r_{\text{CGM}} \), associated with the line-of-sight velocity amplitude of the CGM:

\[
\Delta r_{\text{CGM}} = H^{-1}(z) dv \quad (20)
\]

where \( H(z) \) is the Hubble constant evaluated at the mean redshift of the sample \( (\bar{z} = 2.4) \) 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 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 \( N_{\text{HI}} \) at small galactocentric distances (as observed in Rudie et al. 2012), 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 8, we compare the optical depth of the CGM (\( \tau_{\text{CGM}} = \kappa_{\text{CGM}} \Delta r_{\text{CGM}} \)) for the 300 km s\(^{-1}\), 300 pkpc CGM (red) and the 700 km s\(^{-1}\), 300 pkpc CGM (blue). Figure 8 also quantifies the optical depths from both forest absorbers (log\( (N_{\text{HI}}/\text{cm}^{-2}) < 17.2 \)) 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-\( N_{\text{HI}} \) fit to the data using the broken power law parameterization. The values for \( \tau_{\text{CGM}} \) and \( \kappa_{\text{CGM}} \) in differential and cumulative \( N_{\text{HI}} \) bins are given in Table 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. (2012). 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.

### 5.3. The Distribution of Optical Depths in the IGM and CGM

With an estimate of the opacity from all \( N_{\text{HI}} \) absorbers, one can easily calculate the fractional contribution of the CGM versus the IGM, as well as the contribution from absorbers of different \( N_{\text{HI}} \), to the mean free path. For this simplified calculation, we define the mean free path, \( \lambda_{\text{mfp}} \), 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,

\[
\lambda_{\text{mfp}} = \sum_{i} \Delta r_{i}, \quad (21)
\]

where

\[
\sum_{i} \kappa_{i} \Delta r_{i} = 1. \quad (22)
\]

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), 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

\(11\) This is because the closest galaxy to a QSO sight line in the KBSS sample is at a projected physical distance of 50 kpc.
as $\lambda_{\text{static, IGM+CGM}}$. (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 component, we instead obtain $\lambda_{\text{static, IGM+CGM}}$. Further, absorbers with $\log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2$ contribute 48% of the opacity to LyC photons.

In practice, the detailed distribution of opacity as a function of $N_{\text{HI}}$ will be altered when photons of various energies are considered. The true $\lambda_{\text{mfp}}$ 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 affected by a source at $\Delta_z \approx 0.25$ at these redshifts. Photons emitted by a source at $z = 2.4$ with energies close to the 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 $\lambda_{\text{mfp}}$ 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 $\lambda_{\text{mfp}}$.

### Table 7

| $N_{\text{HI}}$, Min | $N_{\text{HI}}$, Max | Method | $\kappa_{\text{ICM}} \times 10^3$ | $\kappa_{\text{CGM}} \times 10^3$ | $\kappa_{\text{ICM}} \times 10^3$ | $\kappa_{\text{CGM}} \times 10^3$ |
|----------------------|----------------------|--------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| 12.55                | 12.60                | data   | 0.001                       | 0.02                        | 0.001                       | 0.02                        |
| 12.60                | 12.65                | data   | 0.002                       | 0.02                        | 0.002                       | 0.02                        |
| 12.65                | 12.71                | data   | 0.002                       | 0.02                        | 0.002                       | 0.02                        |
| 12.71                | 12.77                | data   | 0.003                       | 0.02                        | 0.003                       | 0.02                        |
| 12.77                | 12.85                | data   | 0.004                       | 0.02                        | 0.004                       | 0.02                        |
| 12.85                | 12.93                | data   | 0.004                       | 0.02                        | 0.004                       | 0.02                        |
| 12.93                | 13.01                | data   | 0.006                       | 0.02                        | 0.007                       | 0.02                        |
| 13.01                | 13.11                | data   | 0.008                       | 0.02                        | 0.009                       | 0.02                        |
| 13.11                | 13.21                | data   | 0.009                       | 0.02                        | 0.010                       | 0.02                        |
| 13.21                | 13.32                | data   | 0.011                       | 0.02                        | 0.012                       | 0.02                        |
| 13.32                | 13.45                | data   | 0.017                       | 0.02                        | 0.018                       | 0.02                        |
| 13.45                | 13.59                | data   | 0.021                       | 0.02                        | 0.022                       | 0.02                        |
| 13.59                | 13.74                | data   | 0.024                       | 0.02                        | 0.025                       | 0.02                        |
| 13.74                | 13.90                | data   | 0.027                       | 0.02                        | 0.029                       | 0.02                        |
| 13.90                | 14.08                | data   | 0.038                       | 0.05                        | 0.040                       | 0.05                        |
| 14.08                | 14.28                | data   | 0.045                       | 0.08                        | 0.048                       | 0.08                        |
| 14.28                | 14.50                | data   | 0.059                       | 0.08                        | 0.066                       | 0.09                        |
| 14.50                | 14.73                | data   | 0.095                       | 0.09                        | 0.101                       | 0.10                        |
| 14.73                | 15.00                | data   | 0.115                       | 0.12                        | 0.123                       | 0.13                        |
| 15.00                | 15.28                | data   | 0.123                       | 0.13                        | 0.130                       | 0.15                        |
| 15.28                | 15.60                | data   | 0.202                       | 0.34                        | 0.213                       | 0.35                        |
| 15.60                | 15.94                | data   | 0.314                       | 0.51                        | 0.334                       | 0.53                        |
| 15.94                | 16.32                | data   | 0.471                       | 0.73                        | 0.510                       | 0.75                        |
| 16.32                | 16.73                | data   | 0.661                       | 0.92                        | 0.703                       | 0.94                        |
| 16.73                | 17.20                | data   | 1.695                       | 2.03                        | 1.803                       | 2.08                        |
| 17.20                | 17.70                | fit     | 1.570                       | 1.93                        | 1.670                       | 2.11                        |
| 17.70                | 18.20                | fit     | 1.111                       | 1.93                        | 1.182                       | 2.17                        |
| 18.20                | 18.70                | fit     | 0.657                       | 1.34                        | 0.688                       | 1.41                        |
| 18.70                | 19.20                | fit     | 0.372                       | 0.82                        | 0.396                       | 0.88                        |
| 19.20                | 19.70                | fit     | 0.214                       | 0.49                        | 0.228                       | 0.56                        |
| 19.70                | 20.30                | fit     | 0.141                       | 0.32                        | 0.150                       | 0.37                        |
| 20.30                | 22.00                | fit     | 0.127                       | 0.27                        | 0.135                       | 0.32                        |

Note. *Opacity calculated from the data using Equation (15) for 12.55 < log($N_{\text{HI}}/\text{cm}^{-2}$) < 17.2.

$\lambda_{\text{mfp, IGM+CGM}}$ as $\lambda_{\text{static}}$. Under these assumptions where $\Delta_{\text{CGM}} = 2.87$ pMpc as given in Equation (20) using $dv = 700$ km s$^{-1}$ and $\kappa_{\text{ICM}} = 0.0081$ pMpc$^{-1}$ and $\kappa_{\text{CGM}} = 0.047$ pMpc$^{-1}$ are the total opacities over all $N_{\text{HI}}$. Solving Equation (23), we find $\lambda_{\text{static, IGM+CGM}} = 109.3$ pMpc. When we exclude the CGM component, we instead obtain $\lambda_{\text{static, IGM}} = 122.8$ pMpc, a 12% longer pathlength.

With $\lambda_{\text{static, IGM+CGM}}$ known, we can then calculate the fractional contribution to the optical depth from each bin of $N_{\text{HI}}$ for both the IGM and CGM as shown in Figure 9 and Table 7 (middle columns). The cumulative version of this plot (i.e., the contribution to the optical depth for all $N_{\text{HI}} < N$) is given in Figure 10 and in Table 7, right-hand columns. From these, we can see that the CGM accounts for 13% of the opacity within

### 6. The Mean Free Path to Hydrogen-Ionizing Photons

As we have shown in Section 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 $\lambda_{\text{mfp}}$. It is also
As such, we use only the best-fit values of the power law fit to sight than occur in the actual universe as measured by the data.

The printed numbers give the percentage of the opacity due to that absorber are relevant to the "mean free path," including the frame of reference in which the distance is measured. Here we define the mean free path, $\lambda_{\text{mfp}}$, as the distance traveled from a source by a packet of photons before photons with energy of 1 Ry in the absorber's frame are attenuated by an average factor of $e$. This distance is fundamentally related to a measurement of $\Delta z$, which we convert to the physical distance traveled by the photon through the expanding universe.

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

### 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 $\Delta z = 2.4$, according to a frequency distribution of absorbers set using best-fit parameters from the broken power law fit to $f(N, X)$ (Table 5).

As was demonstrated in Section 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 than 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 MC.

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

$$\frac{dn}{dz} = n_0(1 + z)^\gamma.$$  \hspace{1cm} (24)

For the MC simulations, we have assumed that the Ly$\alpha$ forest absorbers with low-$N_{\text{HI}}$ (those absorbers fit by the low-$N_{\text{HI}}$ power law) in the broken power law redshift parameterization: $\log(N_{\text{HI}}/\text{cm}^{-2}) < 15.1$ for the IGM and $\log(N_{\text{HI}}/\text{cm}^{-2}) < 14.9$ 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-$N_{\text{HI}}$ absorbers are assumed to have $\gamma = 1.0$; the value for the LLS compilation reported in this paper (Section 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 $2 < z < 3$ is well constrained by the data presented by Ribando et al. (2011), Sargent et al. (1989), and Stengler-Larrea et al. (1995) as discussed in Section 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 $\gamma$). Because our MC 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 $\gamma$ does not significantly affect the quoted $\lambda_{\text{mfp}}$ within the redshift range $2 < z < 3$.\footnote{Considered in the redshift range $2.0 < z < 2.8$, varying $\gamma$ by 1σ results in less than a 6% change in the incidence of LLSs and therefore less than a 3% change in $\lambda_{\text{mfp}}$.}

However, we caution that the extrapolations of the trend of our $\lambda_{\text{mfp}}$ measurements to $z > 3$ is much more dependent on the chosen value of $\gamma$ and is therefore unlikely to predict the true value at higher redshift.

\footnote{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 $\lambda_{\text{mfp}}$.}
An additional set of simulations was run to mimic lines of sight emanating from galaxies like those in the KBSS. In these simulations, H tidal absorbers were added within 700 km s−1 of $z_{em}$ 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 ($0.25 < L/L^* < 3; (z) = 2.3$).

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 LyC absorption. To measure the $\lambda_{mfp}$, 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 6.4), while the “continuum absorption only” spectra can be used to measure $\lambda_{mfp}$ directly (as discussed in Section 6.2).

Figure 11 shows the average of 10,000 realizations of each MC 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-$N_{H}$ absorbers found in the circumgalactic gas.

6.2. Measuring $\lambda_{mfp}$

To measure $\lambda_{mfp}$ from the MC calculation, we use the spectra that include continuum absorption only. 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 dashed-dotted lines in Figure 12. The rest wavelengths from Figure 12 can be converted into the physical distance traveled by a photon between $z_{em} = 2.4$, the emission redshift of the spectra, and $z_{mfp}$ where

$$z_{mfp} = \frac{\lambda_{e} - \lambda_{LL}}{\lambda_{LL}} (1 + z_{em}) + z_{em} \quad (25)$$

where $\lambda_{LL} = 911.75 \text{Å}$ is the wavelength of the Lyman limit and $\lambda_{e}$ is the rest wavelength of the spectrum at which the transmission equals $e^{-1}$. To calculate $\lambda_{mfp}$, we then compute the physical distance traveled by the photon through the expanding universe, given by the integral over the proper line element:

$$\lambda_{mfp} \equiv \int_{z_{em}}^{z_{mfp}} \frac{c}{1 + z} H(z) dz, \quad (26)$$

where $c$ is the speed of light and

$$H(z) = H_{0} \sqrt{\Omega_{\Lambda} + \Omega_{m}(1 + z)^3}. \quad (27)$$

Using the averaged LyC only spectra shown in Figure 12, we calculate the $\lambda_{mfp}$ through the IGM assuming $z_{em} = 2.4$ 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 $\lambda_{mfp}$ falls to 120 pMpc ($\Delta z = 0.301$).
\[ \lambda_{\text{mfp}} = \frac{1}{\kappa_{\text{LLS}} + \kappa_{\text{for}}} \] (see Equation (23)) then the uncertainty on \( \lambda_{\text{mfp}} \) is

\[ \sigma_{\text{mfp}} = \sqrt{\frac{\sigma_{\kappa_{\text{LLS}}}^2 + \sigma_{\kappa_{\text{for}}}^2}{(\kappa_{\text{LLS}} + \kappa_{\text{for}})^2}}, \] (29)

where \( \sigma_{\kappa_{\text{LLS}}} \) and \( \sigma_{\kappa_{\text{for}}} \) are the uncertainties of \( \kappa_{\text{LLS}} \) and \( \kappa_{\text{for}} \) respectively. Employing these equations, we find a typical uncertainty in \( \lambda_{\text{mfp}} \) of \( \sim 15 \) Mpc, which we adopt for both the IGM and IGM+CGM \( \lambda_{\text{mfp}} \) 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 \( \lambda_{\text{mfp}} \) at the mean redshift of the sample.
6.4. Simulated Sight Lines 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; C. C. Steidel et al., in preparation) 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 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 $z_{\text{em}} = 2.4$ 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 14. The transmission is generally lower along the IGM+CGM sight lines. The peak in the transmission distribution near $\lambda = 915$ Å is due to line blanketing from both higher Lyman series lines as well as Ly$\alpha$ absorption from lower redshift.

In Figure 12, the solid curves show the average spectrum of 10,000 MC runs that include line opacity. We use $(1 - D_N)$, 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 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 $\lambda_{\text{mfp}}$ as described in Section 6.2, leads to $\lambda_{\text{mfp}}$ values 5%–10% smaller than the “true” values. Table 8 compares the results of the MC simulations including line opacity (columns labeled $(1 - D_N)$) with those based on the continuum only spectra (columns labeled LyC).

6.5. Comparison with Previous $\lambda_{\text{mfp}}$ Measurements

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

$$\lambda_{\text{mfp}}(z_{\text{em}} = 2.4) = 147 \pm 15 \text{ Mpc}. \hspace{1cm} (30)$$

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

$$\lambda_{\text{mfp}}(z_{\text{em}} = 2.4) = 121 \pm 15 \text{ Mpc}. \hspace{1cm} (31)$$

One important advance of the estimate of $\lambda_{\text{mfp}}$ 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 $\lambda_{\text{mfp}}$ 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 $\gamma$ assumed. Nevertheless, in this section we compare to previous measurements for completeness. Figure 15 gives a graphical summary of the results of this comparison.

Table 8

| $z_{\text{em}}$ | $\lambda_{\text{mfp}}$ LyC (pMpc) | $\lambda_{\text{mfp}}(1 - D_N)$ (pMpc) | $\lambda_{\text{mfp}}$ LyC (pMpc) | $\lambda_{\text{mfp}}(1 - D_N)$ (pMpc) |
|---------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| IGM ONLY      | IGM ONLY                        | IGM+CGM                        | 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                            |

Figure 14. The distribution of transmissions from each sight line averaged over the band pass from 880–910 Å for 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 $(1 - D_N)$.

(A color version of this figure is available in the online journal.)
In general, the differences between our IGM ONLY $\lambda_{\text{mfp}}$ results and those of other authors (especially those measured at similar $z$) underscore the degree to which $\lambda_{\text{mfp}}$ is sensitive to changes in $f(N, X)$, highlighting the importance of the precise measurement of the intermediate $N_{\text{H}_1}$ systems made here.

One other caveat is the small degree of incompleteness which may be present in the $4 \lesssim z \lesssim 3$. We emphasize 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 $\lambda_{\text{mfp}}$, parameterized as a function of the emission redshift of the QSO, $z_{\text{em}}$, returned by the MC simulations is represented by the solid blue curves in Figure 15 and is given by

\[ \lambda_{\text{mfp}} = \lambda^{\text{em}} (1 + z)^{\xi}, \quad (32) \]

where $\xi$ captures the cosmology (as parameterized in Equation (26)) as well as the evolution in the number of absorbers as a function of redshift, $dn/dz$ (described in Section 6.1). For the MC simulations described in this section, the redshift evolution of $\lambda_{\text{mfp}}$ is fixed by the assumed value of $\gamma$. We emphasize 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 $\lambda_{\text{mfp}}$, parameterized as a function of the emission redshift of the QSO, $z_{\text{em}}$, returned by the MC simulations is represented by the solid blue curves in Figure 15 and is given by

\[ \lambda_{\text{mfp}}(z_{\text{em}}) = 147 \left( \frac{1 + z_{\text{em}}}{3.4} \right)^{-4.3} \text{pMpc} \quad (33) \]

and

\[ \lambda_{\text{mfp}}(z_{\text{em}}) = 121 \left( \frac{1 + z_{\text{em}}}{3.4} \right)^{-4.5} \text{pMpc} \quad (34) \]

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 $\lambda_{\text{mfp}}$ assuming $\beta = 1.39$ as measured by Misawa et al. (2007). Misawa et al. (2007) measured the frequency distribution close to high-$N_{\text{H}_1}$ 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 underpredicting the incidence of lower-$N_{\text{H}_1}$ absorbers resulting in slightly larger values of $\lambda_{\text{mfp}}$ at a given $z$:

\[ \lambda_{\text{mfp}} = 85 \left( \frac{1 + z}{4.0} \right)^{-4} \text{Mpc} = 160 \left( \frac{1 + z}{3.4} \right)^{-4} \text{Mpc} \quad (36) \]

Prochaska et al. (2009) estimated $\lambda_{\text{mfp}}$ 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

\[ \lambda_{\text{mfp}} = 48.4 - 38.0(z - 3.6) \text{Mpc} = 94 - 38.0(z - 2.4) \text{Mpc} \quad (37) \]

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 $z > 4$ LLSs. They found

\[ \lambda_{\text{mfp}} = 50 \left( \frac{1 + z}{4.5} \right)^{-4.44 \pm 0.16} \text{Mpc} = 170 \left( \frac{1 + z}{3.4} \right)^{-4.44 \pm 0.16} \text{Mpc} \quad (38) \]

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

Recently O’Meara et al. (2013) presented a calculation of $\lambda_{\text{mfp}}$ using their new measurements of the incidence of LLSs at $z \sim 2$. Their estimates of the incidence of absorbers with $\log(N_{\text{H}_1}/\text{cm}^{-2}) < 17.2$ lead to a lower value for the opacity as described in Section 5.1. They found

\[ \lambda_{\text{mfp}} = 250 \pm 43 \text{Mpc} \quad (39) \]

at $z = 2.44$, nearly a factor of two 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 $(z) = 2.4$ 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 6 offers three significant improvements over previous 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 \( \gamma \). 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, \( \beta \), for absorbers with saturated Ly\( \alpha \) profiles, \([14 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2]\). Because such absorbers account for \( \approx \)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 \( \lambda_{\text{mfp}} \). 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. (2012), regions of the IGM near galaxies have a much higher incidence of high-\( N_{\text{HI}} \) absorbers than the average IGM with the effect (again) of reducing \( \lambda_{\text{mfp}} \) relative to previous measurements.

### 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 \( (J_\nu; \text{see, e.g., Scott et al. 2000}) \) or the photoionization rate \( (\Gamma; \text{see, e.g., Faucher-Giguère et al. 2008}) \) to be well determined, a shorter mean free path directly implies a requirement for correspondingly larger ionizing emissivity of sources \( (\epsilon_\nu) \). Following Faucher-Giguère et al. (2008),

\[
J_\nu(z) \approx \frac{1}{4\pi} \lambda_{\text{mfp}}(\nu, z) \epsilon_\nu(z)
\]

and

\[
\Gamma \propto J_\nu.
\]

Conversely, for a measured value of \( \epsilon_\nu \) from a given population of sources, a shorter \( \lambda_{\text{mfp}} \) implies a smaller contribution to both the specific intensity and the photoionization rate \( (\Gamma) \). 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; C. C. Steidel et al., in preparation). For LyC studies of galaxies at high-\( z \), the inclusion of the CGM is an important effect. Considering Figure 14, note that the number of galaxies with \( \sim \)zero transmission increases by a factor of two 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-\( N_{\text{HI}} \) 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 \( \lambda_{\text{mfp}} \) that is too large may cause one to overestimate the contribution to \( f_\nu \) or \( \Gamma \) made by a population of detected LyC sources. The latter effect may reduce the tension between the apparent overproduction 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).

### 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\( \text{I} \) absorbers fit with Voigt profiles to Ly\( \alpha \) and at least one higher order Lyman series transition. This analysis enables the first statistically rigorous measurements of the frequency distribution of H\( \text{I} \) absorbers with \( 14 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17 \). In Section 3, we showed that the frequency of absorbers as a function of \( N_{\text{HI}} \) is well parameterized by a single power law from \( 13.5 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2 \) with a maximum likelihood index of \( \beta = 1.65 \pm 0.02 \) and a normalization per unit pathlength of \( \log(C_{\text{HI}}) = 10.32 \).

In Section 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-\( N_{\text{HI}} \) systems than low. These findings are discussed further by Rudie et al. (2012).

Section 5 examined statistically the total opacity of the IGM and CGM to hydrogen-ionizing LyC photons. This section presented measurements of the opacity due to absorbers with \( \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2 \) with direct constraints from the data. We also found that a single MLE power law fit over the full range of \( N_{\text{HI}} \) 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 \( \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.2 \) as well as the observed incidence of LLSs. The parameters of these fits can be found in Table 5.

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

In Section 6, we used MC simulations to measure the value of the mean free path (\( \lambda_{\text{mfp}} \)) of LyC photons at \( z = 2.4 \). We measured both the \( \lambda_{\text{mfp}} \) of photons through the IGM (as has been considered many previous times) as well as the \( \lambda_{\text{mfp}} \) 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 \( \lambda_{\text{mfp}}(z = 2.4) = 147 \pm 15 \text{ Mpc} \) for simulations including only IGM opacity and \( \lambda_{\text{mfp}}(z = 2.4) = 121 \pm 15 \text{ Mpc} \) when we include the CGM. These values of \( \lambda_{\text{mfp}} \) are lower than most previous estimates, which have not included the CGM and have generally underestimated the contribution of intermediate-N\( \text{HI} \) 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-redshifts.

We note that the \( \lambda_{\text{mfp}} \) 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 \(z = 2.3\)—for other galaxies, these measurements are an approximation. However, in all cases the general statement holds that \(\lambda_{\text{mfp}}\) 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 \(z \approx 2-3\).

The authors 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 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 acknowledge the staff of 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 U.S. National Science Foundation through grants AST-0606912 and AST-0908805. C.C.S. 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.

APPENDIX A

FITS TO HIGH-\(N_{\text{HI}}\) ABSORBERS

As discussed in Section 1, for \(\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 14.5\), the Ly\(\alpha\) transition is saturated, complicating the measurement of absorbers with higher \(N_{\text{HI}}\) from Ly\(\alpha\) only spectra. Similarily, for absorbers with \(\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 15.5\), the Ly\(\beta\) transition saturates. Thus, the precision and accuracy with which high-\(N_{\text{HI}}\) 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 Ly\(\alpha\) and Ly\(\beta\) transition for all the absorbers. Given the typical spectral coverage of our sample (see Table 1), for absorbers with \(z \gtrsim 2.3\), the Ly\(\gamma\) 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 \(N_{\text{HI}}\) to \(\log(N_{\text{HI}}/\text{cm}^{-2}) \lesssim 17.2\).

An additional complication arises from the fact that high-\(N_{\text{HI}}\) absorbers are often found in blended systems of several high-\(N_{\text{HI}}\) absorbers. Such absorbers, when observed in only Ly\(\alpha\), \(\beta\), 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 \(\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 15.5\).

A.1. Examples of Fits to High-\(N_{\text{HI}}\) Absorbers

Figures 16–18 show examples of the fits to the HIRES spectra for absorbers with \(15.5 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.0\). 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 \(N_{\text{HI}}\) as reported by VPFIT. As expected, the errors on \(N_{\text{HI}}\) are dependent on the number of transitions observed; however, the dependence is not trivial as the effect of (1) blending from other proximate absorbers, (2) the contamination of higher-order Lyman transitions due to unrelated absorbers, and (3) the S/N of the spectrum in the areas of interest all affect our ability to accurately determine the parameters of the absorber.

A.2. Testing Fits with Only Ly\(\alpha\) and Ly\(\beta\)

To assess the accuracy of fits to absorbers with \(\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 15.5\) observed only in Ly\(\alpha\) and Ly\(\beta\), we use the set of absorbers with \(\log(N_{\text{HI}}/\text{cm}^{-2}) \gtrsim 15.5\) 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 Ly\(\alpha\) and Ly\(\beta\) 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 Ly\(\alpha\) and/or Ly\(\beta\) can be fit much more reliably. Such systems, when refit with VPFIT using only their Ly\(\alpha\) and Ly\(\beta\) transition result in weaker constraints on the line parameters as expected, but result in \(N_{\text{HI}}\) determinations consistent with those obtained via the higher-order transitions. The fits are not systematically offset to higher or lower values of \(N_{\text{HI}}\), and therefore should not bias the measurements of \(f(N, X)\) in any large way.

For absorbers whose Ly\(\alpha\) and Ly\(\beta\) transition are fully saturated and both edges/wings of the absorption line cannot be observed due to other neighboring absorbers (e.g., see the leftmost panel in Figure 17 and the center panel in Figure 18), 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 \(N_{\text{HI}}\) in such blends. If the \(N_{\text{HI}}\) of individual absorbers in the blends were systematically underestimated, this would bias the measurement of \(f(N, X)\) to have steeper values of \(\beta\) 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 Ly\(\alpha\) and Ly\(\beta\) only and therefore are typically fit with fewer subcomponents and with lower-\(N_{\text{HI}}\) 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.3. Assessing Incompleteness in the $z \lesssim 2.4$, $
HOffice{log}
HOffice{(N_{\text{HI}}/\text{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-$N_{\text{HI}}$ absorbers observed in $n \geq 5$ Lyman series transitions that would be unconstrained or miscounted with observations of Ly$\alpha$ and Ly$\beta$ alone. This method likely overestimates 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 $N_{\text{HI}}$. The results are summarized in Table 9.
Figure 17. Example fits to absorbers with $16.0 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 16.5$. The black curves show the HIRES data. Gray 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 ±3σ $N_{\text{HI}}$ error.

(A color version of this figure is available in the online journal.)

For absorbers with $15.5 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 16.0$, 23 are measured with five or more Lyman series transitions. Of these, only two would have essentially no constraint on the parameters of the absorber from Lyα and Lyβ alone (e.g., Figure 16, second from the left). For 7/23 of these systems, the constraint on $N_{\text{HI}}$ would be relatively weak; however, an individual component associated with that absorber can be clearly identified. All seven of these systems are fit with much improved accuracy with the additional constraint from Lyγ. Among absorbers with $16.0 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 16.5$, we observed 12 absorbers with
Figure 18. Example fits to absorbers with $16.5 < \log(N_{\text{HI}}/\text{cm}^{-2}) < 17.0$. The black curves show the HIRES data. Gray 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 $\pm 3\sigma N_{\text{HI}}$ error.

(A color version of this figure is available in the online journal.)
constraints from five or more Lyman series transitions. Of these, only four are strongly constrained from Lyα and Lyβ alone. With Lyγ, 8/12 could be fit, and with Lyα, β, γ, δ, 11/12 would be strongly constrained and could accurately be fit. For absorbers with \( N_{\text{HI}} \) just below that of LLSs (16.5 < \( \log(N_{\text{HI}}/\text{cm}^{-2}) \) < 17.0), we observe five systems in five or more Lyman series transitions.\(^{15}\) Of these, only two would have no constraint on \( N_{\text{HI}} \) from Lyα and Lyβ alone.

From this assessment based on higher-\( z \) measurements, we expect that the principal bias in our measurements of \( N_{\text{HI}} \) for high-\( N_{\text{HI}} \) systems for absorbers with only Lyα and Lyβ is to undercount the number of systems with \( \log(N_{\text{HI}}/\text{cm}^{-2}) \) > 15.5. Within the present sample, systems with 16.0 < \( \log(N_{\text{HI}}/\text{cm}^{-2}) \) < 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.

APPENDIX B

THE EFFECT OF THE ACCURACY OF \( f(N, X) \) ON THE MEASUREMENT OF THE \( \lambda_{\text{mfp}} \)

Here we assess the redshift dependence of the frequency distribution presented in Sections 3 and 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 four Lyman series transitions and most of this sample have all the Lyman series transitions observed. Figure 19 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 Section 5.1 that were used in the MC simulation in Section 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 \( \log(N_{\text{HI}}/\text{cm}^{-2}) \) > 15.5 is actually higher than that measured from the full sample. As discussed in Appendix A.3, we expect that the lower-\( z \) catalog likely suffers from incompleteness due to line blending, especially for absorbers with 16.0 < \( \log(N_{\text{HI}}/\text{cm}^{-2}) \) < 16.5. In Figure 19, we note the only bin for which the low-\( z \) and high-\( z \) data do not agree within 1\( \sigma \) contains absorbers with 16.0 < \( \log(N_{\text{HI}}/\text{cm}^{-2}) \) < 16.5.

The principal result of this paper is the relatively high measured value of \( f(N, X) \) for absorbers with 14.5 < \( \log(N_{\text{HI}}/\text{cm}^{-2}) \) < 17.2, which results in a value of \( \lambda_{\text{mfp}} \) comparable to or smaller than previous measurements. If we were to use the high-redshift subsample from Figure 19 to measure the opacity and \( \lambda_{\text{mfp}} \), we would find higher IGM opacity and therefore a shorter (more discrepant) \( \lambda_{\text{mfp}} \). In any case, if there is a systematic bias in the measurements, it is one that would conspire to make \( \lambda_{\text{mfp}} \) 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.

REFERENCES

Adelberger, K. L., Steidel, C. C., Pettini, M., et al. 2005, ApJ, 619, 697
Altay, G., Theuns, T., Schaye, J., Crighton, N. H. M., & Dalla Vecchia, C. 2011, ApJ, 737, L37

\(^{15}\) All five absorbers are observed in a minimum of nine transitions. We do not observe any absorbers in this \( N_{\text{HI}} \) range with 5–8 Lyman series transitions observed.
