



Metal Lines Associated with the Ly α Forest from eBOSS Data

Li Yang^{1,2,3} , Zheng Zheng³ , Hélión du Mas des Bourboux³ , Kyle Dawson³ , Matthew M. Pieri⁴, Graziano Rossi⁵,
Donald P. Schneider^{6,7} , Axel de la Macorra⁸, and Adrean Muñoz Gutiérrez⁸

¹ Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, People's Republic of China; liyang@shao.ac.cn

² School of Astronomy and Space Sciences, University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing 100049, People's Republic of China

³ Department of Physics and Astronomy, University of Utah, 115 S 1400 E, Salt Lake City, UT 84112, USA; zhengzheng@astro.utah.edu

⁴ Aix Marseille Université, CNRS, CNES, Laboratoire d'Astrophysique de Marseille, Marseille, France

⁵ Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Republic of Korea

⁶ Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

⁷ Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA

⁸ Instituto de Física, Universidad Nacional Autónoma de México, Apdo. Postal 20-364, México

Received 2022 January 6; revised 2022 June 16; accepted 2022 June 21; published 2022 August 22

Abstract

We investigate the metal species associated with the Ly α forest in eBOSS quasar spectra. Metal absorption lines are revealed in stacked spectra from cross-correlating the selected Ly α absorbers in the forest and the flux fluctuation field. Up to 13 metal species are identified as being associated with relatively strong Ly α absorbers (those with flux fluctuations $-1.0 < \delta_{\text{Ly}\alpha} < -0.6$ and with a neutral hydrogen column density of $\sim 10^{15-16} \text{ cm}^{-2}$) over the absorber redshift range of $2 < z_{\text{abs}} < 4$. The column densities of these species decrease toward higher redshift and for weaker Ly α absorbers. From modeling the column densities of various species, we find that the column density pattern suggests contributions from multiple gas components, both in the circumgalactic medium (CGM) and the intergalactic medium (IGM). While the low-ionization species (e.g., C II, Si II, and Mg II) can be explained by high-density, cool gas ($T \sim 10^4 \text{ K}$) from the CGM, the high-ionization species may reside in low-density or high-temperature gas in the IGM. The measurements provide inputs for modeling the metal contamination in the Ly α forest baryon acoustic oscillation measurements. Comparisons with metal absorptions in high-resolution quasar spectra and hydrodynamic galaxy formation simulations can further elucidate the physical conditions of these Ly α absorbers.

Unified Astronomy Thesaurus concepts: Intergalactic medium (813); Lyman alpha forest (980); Metal line absorbers (1032)

Supporting material: machine-readable table, data behind figure

1. Introduction

The Lyman- α (Ly α) forest, namely the ensemble of absorption lines in a quasar's continuum at wavelengths below that quasar's Ly α emission, caused by the absorption of intervening neutral hydrogen (e.g., Cen et al. 1994; Bi & Davidsen 1997; Rauch 1998), has become a powerful cosmological and astrophysical probe.

As a tracer of the underlying matter density field, the Ly α forest is a powerful cosmological probe that has been successfully used to constrain cosmological parameters. In particular, the Ly α forest data from the spectroscopic observations of quasars in the Sloan Digital Sky Survey (SDSS) have been used to measure the baryon acoustic oscillations (BAOs), adding unique high-redshift ($z \sim 3$) data points to constrain the cosmology (e.g., Slosar et al. 2013; Bautista et al. 2017; du Mas des Bourboux et al. 2020).

The Ly α forest also encodes valuable information about the circumgalactic and intergalactic media (CGM and IGM, respectively), including their density and temperature distributions. While the CGM and IGM are the sources of the material for star formation and galaxy formation, their properties are also affected by gas outflows from galaxies as a result of feedback processes. The outflows can bring heavy elements

into the CGM and IGM. The metal content is connected to the stellar population, abundance pattern, and star formation history. The metal species at various ionization levels can constrain the ionizing ultraviolet (UV) background. Ly α forest observations provide a means of revealing the metals. As metal absorption in the Ly α forest is an important systematic component to be modeled in Ly α forest BAO measurements (Bautista et al. 2015), studying the metals in the Ly α forest also aids the cosmological application of the Ly α forest. High-resolution spectroscopic observations of quasars with long exposure times are usually chosen to study the metal lines in the Ly α forest or that are associated with the Ly α forest (e.g., Ellison et al. 2000; Schaye et al. 2003; Simcoe et al. 2004; Lehner et al. 2019, 2021). For example, Lehner et al. (2021) study $2.2 \lesssim z \lesssim 3.6$ H I-selected absorbers with neutral hydrogen column density in the range of $14.6 \leq \log(N_{\text{HI}}/\text{cm}^{-2}) \leq 20$. The metallicity is found to have a broad distribution (e.g., an interquartile range from -3.6 to -1.8 dex, with a median of -2.4 dex for $\log N_{\text{HI}} = 14.5\text{--}16.2$ absorbers), and the median metallicity increases and the interquartile range decreases with increasing neutral hydrogen column density.

High-resolution spectroscopic observations of quasars usually probe only the Ly α forest along the sightline of a single quasar or those of a small number of quasars. In contrast, the large spectroscopic survey of quasars resulting from the various phases of the SDSS survey has been able to probe hundreds of thousands of quasar sightlines with moderate resolution and short exposure times. Although the signal-to-noise ratios (S/Ns) and spectral resolutions from the SDSS



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

observations typically do not allow the identification of the metal absorption lines associated with the Ly α forest in any individual quasar spectrum, the large number of quasar spectra can be used to perform a statistical study of metal lines.

Pieri et al. (2010) and Pieri et al. (2014) develop a technique to study metal absorption lines using a composite spectrum of Ly α forest absorbers identified in the spectra of quasars from SDSS. Ly α absorbers in the Ly α forest of each quasar are selected based on the transmitted flux fraction. For each selected absorber, the quasar’s spectrum is shifted to the rest frame of the absorber. The shifted quasar spectra for all the selected absorbers are then stacked to produce a composite spectrum. Features uncorrelated with the absorbers are highly suppressed in the composite spectrum, and the coherently stacked metal absorption lines associated with the selected absorbers can be detected with high significance. Applying the technique to the data from SDSS (Pieri et al. 2010) and from the Baryon Oscillation Spectroscopic Survey (BOSS) of SDSS-III (Pieri et al. 2014) has revealed a series of metal absorption lines in the composite spectrum for probing the physical conditions and metal enrichment of the CGM and IGM at $z \sim 2-3$.

In this paper, following the same basic idea in Pieri et al. (2010), we study the metal lines associated with the Ly α forest⁹ by applying a cross-correlation method to obtain stacked spectra using data from the extended BOSS (eBOSS; Dawson et al. 2016) survey of SDSS-IV. We study the dependence of the metal absorption lines on the Ly α fluctuation field and their redshift evolution over $2 < z < 4$.

In Section 2, we describe the data and the method for computing the stacked spectra of Ly α absorbers. We then present the main results in Section 3, including the column densities of 13 metal species as a function of Ly α absorber strength and redshift. A simple model is discussed to understand the results. Finally, we summarize and discuss our findings in Section 4. We provide some details of the column density measurements in Appendix A.

2. Data Samples and Reduction Methods

The Ly α forest data in this work are based on the quasar observations from SDSS (York et al. 2000) gathered during SDSS-III by BOSS (Eisenstein et al. 2011; Dawson et al. 2013) and during SDSS-IV by eBOSS (Dawson et al. 2016; Blanton et al. 2017), with a small fraction having been observed during SDSS-I and II (Schneider et al. 2010).

All quasar spectra used for this analysis were obtained using the eBOSS spectrographs mounted on the 2.5 m Sloan Foundation telescope (Gunn et al. 2006; Smee et al. 2013; Myers et al. 2015) at the Apache Point Observatory, which are publicly available in the sixteenth data release (DR16; Ahumada et al. 2020). The spectral resolution varies with wavelength (3600–10400 Å), from $R = 1800$ to 2200; we adopt $R = 2000$ in our analysis, and this approximation has no effect on our results.

The DR16 quasar catalog (DR16Q; Lyke et al. 2020) is adopted, to extract quasar spectra from `SpAll-v5_13_0.fits`¹⁰, processed with version `v5_13_0` of the eBOSS

spectroscopic pipeline. These are “coadded” spectra, typically constructed from four exposures of 15 minutes and resampled at wavelength pixels of width $\Delta \log \lambda \sim 10^{-4}$ ($c\Delta\lambda/\lambda \sim 69 \text{ km s}^{-1}$). Broad-absorption-line quasars with $\text{BALPROB} > 0.9$ (Lyke et al. 2020) are removed from our analysis. We end up with 725,635 quasars.

Following du Mas des Bourboux et al. (2020), we discard pixels flagged by the pipeline as problematic in the flux calibration or sky subtraction, and correct for small residual flux calibration errors from the eBOSS pipeline. We then mask out spectral intervals, in observed wavelength, where the variance increases sharply, owing to unmodeled emission lines from the sky, as well as the wavelength range of Ca II H and K absorption of the Milky Way. Damped Lyman- α systems (DLAs) are identified (Parks et al. 2018; Lyke et al. 2020; Chabanier et al. 2022), and all pixels where the absorption due to the DLA is higher than 20% are not used.¹¹ Finally, a spectral region is used only if it contains at least 50 pixels, to avoid overfitting the mean transmitted flux.

2.1. Ly α Absorbers and the Forests

In this work, Ly α absorption systems (Ly α absorbers) are selected based on the fluctuation field of the transmitted flux in the Ly α forest. For a selected Ly α absorber, the fluctuation fields in the Lyman series range and in the range longer than the Ly α wavelength are stacked for our study. We call these two ranges the Lyman-series forest and the metal absorption forest, respectively. To measure the flux fluctuation field of a given forest, we make use of the publicly available Python “Package for IGM Cosmological-Correlations Analyses” `picca`¹² (du Mas des Bourboux et al. 2021).

The measurement of the flux fluctuation field for the Ly α forest, the Lyman-series forest, and the corresponding metal absorption forest follows an approach similar to the one established for the Ly α forest BAO analysis, as presented in du Mas des Bourboux et al. (2020). In short, for each quasar q , the fluctuation field $\delta_q(\lambda)$ in the transmitted flux for each forest is computed as

$$\delta_q(\lambda) = \frac{f_q(\lambda)}{C_q(\lambda)\bar{F}(\lambda)} - 1, \quad (1)$$

where $f_q(\lambda)$ is the observed flux density and $C_q(\lambda)$ is the continuum (the flux density that would be observed in the absence of absorption). The quantity $\bar{F}(\lambda)$ is the mean transmission. The product $C_q(\lambda)\bar{F}(\lambda)$ is the mean expected flux for this quasar. It is modeled as a universal function in the quasar’s rest frame (with a uniform forest spectral template), corrected by a linear function of $\log \lambda$. The model parameters are determined by maximizing the likelihood between the model flux and the observed flux in the forest region for each quasar (see du Mas des Bourboux et al. 2020 for details).

As noted in Bautista et al. (2017) and du Mas des Bourboux et al. (2020), the above fitting method introduces a small

⁹ In our notation, the Ly α forest is the region blueward of a quasar’s Ly α emission line, and we do not specifically limit the column density of the absorption systems (absorbers). As we will see, the absorbers that we select can have contributions from high-column density systems. More accurately, we study metal lines associated with Ly α absorbers in the Ly α forest region.

¹⁰ https://data.sdss.org/datamodel/files/BOSS_SPECTRO_REDUX/RUN2D/spAll.html

¹¹ We perform a test of removing the pixels with DLA-caused absorption higher than 5%, and find that this more conservative cut has little effect on our final stacked results. Also, by default, we mask the Ly α and Ly β lines associated with DLAs. We do a further test of masking the metal lines from DLAs. We find that, except for weakening the shadow lines in the stacked spectra, it has no effect on the measurements of metals in the Ly α absorbers that we select, as the DLA metal lines do not contaminate them.

¹² <https://github.com/igmhub/picca>

Table 1
The Number of Ly α Absorbers in the Sample as a Function of Ly α Flux Fluctuation $\delta_{\text{Ly}\alpha}$ and Absorber Redshift z_{abs}

$\delta_{\text{Ly}\alpha}$	(2.0, 2.2)	(2.2, 2.4)	(2.4, 2.6)	(2.6, 2.8)	z_{abs} (2.8, 3.0)	(3.0, 3.2)	(3.2, 3.4)	(3.4, 3.6)	(3.6, 3.8)	(3.8, 4.0)
[−1.0, −0.9]	2420	4085	4628	4296	3880	2891	1935	925	424	137
[−0.9, −0.8]	2162	3864	4555	4133	3458	2510	1614	794	403	114
[−0.8, −0.7]	1886	3713	3997	3563	2988	2144	1372	607	265	83
[−0.7, −0.6]	1416	3027	3424	2966	2483	1820	977	474	192	63

Note. In total, there are 86,688 Ly α absorbers (16,621 quasars) at $2 < z_{\text{abs}} < 4$ and with $-1.0 < \delta_{\text{Ly}\alpha} < -0.6$.

systematic bias in $\delta_{\text{Ly}\alpha}$. The difference in the correlation functions between the true $\delta_{\text{Ly}\alpha}$ field and the one derived from the above fitting method, based on mock data (Figure 11 of Bautista et al. 2017), allows us to estimate this bias, and we find that the derived $\delta_{\text{Ly}\alpha}$ is about 0.02 lower than the true value, too small to have any significant effect on our results. Besides the systematic bias, the mean transmitted flux fitting for each individual quasar also introduces a statistical fluctuation in $\delta_{\text{Ly}\alpha}$, which is included in the uncertainty of $\delta_{\text{Ly}\alpha}$ (e.g., du Mas des Bourboux et al. 2020). Such a fluctuation would result in a spread in the true $\delta_{\text{Ly}\alpha}$ values for a given selected $\delta_{\text{Ly}\alpha}$. The high S/N (above ten; see below) that we impose to select Ly α absorbers limits such an effect so that it is small. In retrospect, this is supported by the clear trend that we find in the dependence of the metal column density on $\delta_{\text{Ly}\alpha}$ (to be presented later). Finally, it is possible that the mean transmitted flux fitting is not accurate at other wavelengths (e.g., around a metal line). As mentioned in Section 2.2, the stacked spectrum is normalized by a local linear fit around each metal line, which largely removes the effect caused by the inaccuracy in the above fitting.

With the measured Ly α fluctuation field $\delta_{\text{Ly}\alpha}$ in the Ly α forest, we construct a catalog of Ly α absorbers. We select Ly α absorbers in the redshift range $2.0 < z_{\text{abs}} < 4.0$, where z_{abs} is the redshift computed from the Ly α absorption. The lower redshift limit is a consequence of the Ly α forest exiting the observed wavelength coverage for quasars with $z \lesssim 2$. The upper redshift limit is adopted because there are only a small number of $z > 4.0$ quasars in DR16Q. For the purpose of our analysis, the Ly α forest is defined as being in the wavelength range of $\lambda_{\text{RF}} = 1041\text{--}1185 \text{ \AA}$ in the quasar’s rest frame (similar to that used by McDonald et al. 2006), where the proximity effects of the quasar’s Ly α and Ly β emission-line profiles are small (Iršič et al. 2013).

In this work, we concentrate on overdense regions and define the Ly α absorber pixels as the local minima among three adjacent pixels in the Ly α forest. We select absorbers in four bins of flux fluctuation— $\delta_{\text{Ly}\alpha} = -0.95, -0.85, -0.75$, and -0.65 , respectively—with a bin width of 0.1. These are the ranges for absorbers with noticeable metal detection, which is the focus of this work. The analysis of hydrogen distribution properties for absorbers in other ranges of $\delta_{\text{Ly}\alpha}$, including those corresponding to underdense regions, will be presented in a separate paper. For each absorber pixel, we require that the S/N in $\delta_{\text{Ly}\alpha}$ is above 10.¹³ Such a cut preserves as many absorbers as possible for the study of metal absorption and, at the same time, it keeps meaningful divisions of the four absorber samples. To investigate the redshift evolution of the

metals in Ly α absorbers, we further divide the absorbers into 10 redshift bins, with $z_{\text{abs}} = 2.1 \pm 0.1, 2.3 \pm 0.1, \dots$, and 3.9 ± 0.1 . In Table 1, the numbers of the Ly α absorbers in different $\delta_{\text{Ly}\alpha}$ and redshift bins are provided.

Like the Ly α forest, we can define the Ly β forest region as being in the wavelength range of $\lambda_{\text{RF}} = 978\text{--}1014 \text{ \AA}$ in a quasar’s rest frame, and similarly for the higher-order Lyman series. In practice, we denote the wavelength range of $\lambda_{\text{RF}} = 787\text{--}1185 \text{ \AA}$ in a quasar’s rest frame as a single forest, namely the Lyman-series forest, even though there are metal lines in this range. For metal lines redward of a quasar’s Ly α emission, each can have its own forest region. For example, in du Mas des Bourboux et al. (2019), they adopt quasar rest-frame wavelength ranges of $1260\text{--}1375 \text{ \AA}$ as the Si IV (1394/1403 $\text{\AA}$ doublet) forest and $1420\text{--}1520 \text{ \AA}$ as the C IV (1548/1551 $\text{\AA}$ doublet) forest. In this work, we define a single metal forest as the range of $\lambda_{\text{RF}} = 1236\text{--}2776 \text{ \AA}$ in a quasar’s rest frame. The range corresponds to $\sim 5000 \text{ km s}^{-1}$ redward of the quasar’s Ly α emission and $\sim 2600 \text{ km s}^{-1}$ blueward of the quasar’s Mg II emission, and we also exclude the pixels within $\pm 15 \text{ \AA}$ (in the quasar’s rest frame) around the quasar’s C IV emission, reducing the effect of the quasar’s emission lines on the $C_q \bar{F}$ fit. Figure 1 illustrates the different forests and the $C_q \bar{F}$ fits in Equation (1) for each forest.

For quasars in the redshift range of $2.09 < z_q < 4.65$, after applying the absorber redshift cuts of $2 < z_{\text{abs}} < 4$ and strength cuts of $-1.0 < \delta_{\text{Ly}\alpha} < -0.6$, we end up with 16,621 Ly α forests ($\lambda_{\text{RF}} = 1041\text{--}1185 \text{ \AA}$), 16,363 Lyman-series forests ($\lambda_{\text{RF}} = 787\text{--}1185 \text{ \AA}$), and 16,607 metal absorption forests ($\lambda_{\text{RF}} = 1236\text{--}2776 \text{ \AA}$). Ly α absorbers selected from the Ly α forests will be correlated with flux fluctuations in the Lyman-series forests and metal absorption forests to study the metal content associated with those Ly α absorbers.

2.2. Stacked Spectra Based on Ly α Absorbers

The Ly α absorbers selected here are connected to neutral hydrogen in the CGM or IGM, and they can be affected by metal enrichment caused by star formation and feedback processes. While the corresponding metal absorption features may be too weak to be identified in the eBOSS quasar spectra for individual absorbers, they can be revealed by stacking the spectra associated with a large number of Ly α absorbers (e.g., Pieri et al. 2010, 2014).

With the catalog of Ly α absorber pixels and the fluctuation fields computed for the Lyman-series and metal absorption forests at different wavelengths, we perform a stacking analysis to reveal the metal absorption associated with Ly α absorbers of different strengths ($\delta_{\text{Ly}\alpha}$). We follow the same technique as in Font-Ribera et al. (2014) and du Mas des Bourboux et al. (2017) to perform the stacking. That is, we compute the cross-correlation function between the Ly α absorber pixels and the

¹³ The S/N in $\delta_{\text{Ly}\alpha}$ is calculated as $S/N = |\delta_q|/\bar{\sigma}_{\text{pip},q}$, where $\bar{\sigma}_{\text{pip},q} = \sigma_{\text{pip},q}/(C_q \bar{F})$, $\sigma_{\text{pip},q}$ is the pipeline estimate of the flux variance, and δ_q and $C_q \bar{F}$ are the same as in Equation (1).

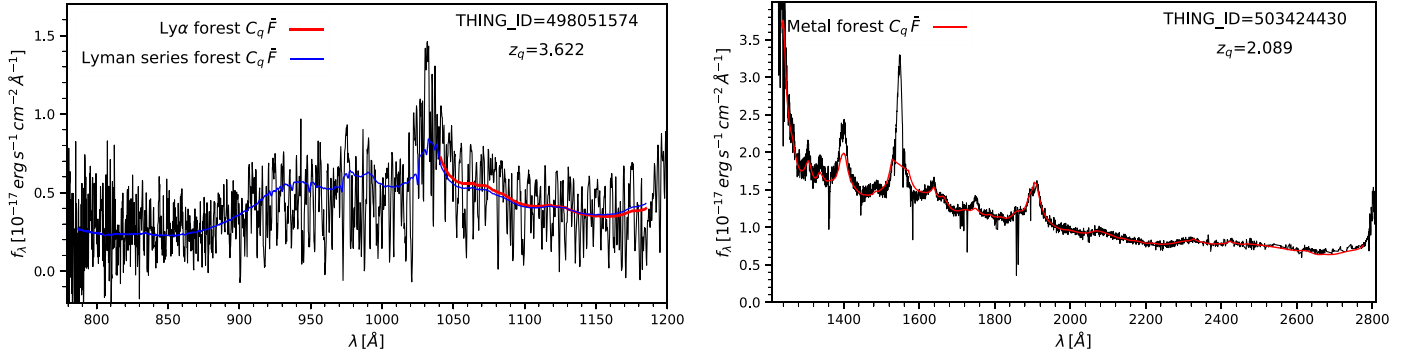


Figure 1. Illustration of the different forests and mean flux fits. The spectra are in a quasar’s rest frame. The colored curves are the fits for the mean expected flux ($C_q \bar{F}$ in Equation (1)) in the Ly α forest (red in the left panel), in the Lyman-series forest (blue in the left panel), and in the metal absorption forest (red in the right panel). See the text for further details. Note that in the right panel, the wavelength range within $\pm 15 \text{ \AA}$ of the quasar’s C IV emission is excluded for the $C_q \bar{F}$ fit, as the emission-line centers are too sharp to produce a good fit. The IDs of the two quasars, from DR16Q, are labeled in each panel.

fluctuation field δ of various forests, which is carried out with *picca*. In detail, the cross-correlation is computed as

$$\xi(\lambda) = \frac{\sum_{(i,k) \in A} w_i w_k \delta_i}{\sum_{(i,k) \in A} w_i w_k}, \quad (2)$$

where i denotes the forest pixel, k denotes the Ly α absorber pixel, δ_i is the flux fluctuation in one of the forests, and w_i and w_k are the weights (as defined in Equations (4) and (7) of du Mas des Bourboux et al. 2020).¹⁴ The wavelength bin A , represented by λ , is determined through $\lambda = \lambda_{\text{Ly}\alpha}(\lambda_i/\lambda_k)$, where $\lambda_{\text{Ly}\alpha} = 1215.67 \text{ \AA}$ is the rest-frame wavelength of the Ly α line and λ_i and λ_k are the observed wavelengths of the forest pixel and the Ly α absorber pixel, respectively. Clearly, λ is the wavelength in the absorber’s rest frame. Later, we also use $\delta(\lambda)$ to represent $\xi(\lambda)$. The computation is performed for absorbers in each $\delta_{\text{Ly}\alpha}$ bin and in each redshift bin. Because of the systematics in continuum fitting, the upper envelope of each stacked spectrum, $1 + \xi$, can slightly deviate from unity, and we apply a low-pass pseudo continuum to rectify this (e.g., Pieri et al. 2014). Specifically, for illustrating the stacked spectrum, we perform a spline fit to the $1 + \delta(\lambda)$ in regions excluding absorption lines and normalize the spectrum by the fit. To analyze each absorption line, we perform a local linear fit to the continuum and normalize the absorption line profile by the fit.

The covariance matrix of the cross-correlation is calculated by subsampling the data, similar to the approach in du Mas des Bourboux et al. (2017). We divide the sky into HEALPix pixels (Górski et al. 2005) and compute the cross-correlation using the data in each HEALPix pixel (subsample). Using a division of the sky with $n_{\text{side}} = 16$, we obtain a number of subsamples between 27 and 790, depending on the $\delta_{\text{Ly}\alpha}$ and z_{abs} bins. The covariance is then given by the following, neglecting the small correlations between subsamples:

$$C_{\alpha\beta} = \frac{1}{W_\alpha W_\beta} \sum_s W_\alpha^s W_\beta^s [\xi_\alpha^s \xi_\beta^s - \xi_\alpha \xi_\beta]. \quad (3)$$

¹⁴ The weights include contributions from both instrumental noise (readouts and photostatistics) and large-scale structure. Applying the weights helps to increase the S/N in the stacked spectrum. In the regime where the weights are dominated by photon noise, the stacking may potentially be biased toward selecting weaker absorbers. We perform tests with no weights applied and find that such a bias is negligible.

Here, $\xi_\alpha = \xi(\lambda_\alpha)$ is the mean measured correlation for wavelength λ_α , s denotes a subsample with summed weight W_α^s and measured correlation ξ_α^s , and $W_\alpha = \sum_s W_\alpha^s$.

The S/Ns of the stacked spectra depend on the number of Ly α absorber–forest pixel pairs. In Figure 2, we show the dependence of the number of such pairs on wavelength, $\delta_{\text{Ly}\alpha}$, and z_{abs} . Given our selection, the number has almost no dependence on $\delta_{\text{Ly}\alpha}$ and only a weak dependence on wavelength. The redshift dependence has two parts. First, the coverage shifts toward shorter wavelengths at higher z_{abs} , a reflection of the redshift effect with the fixed observed wavelength range. At the lowest absorber redshift, $z_{\text{abs}} \sim 2.1$, we are not able to see Lyman series beyond Ly α , while at the long-wavelength end it covers up to $\sim 2900 \text{ \AA}$. In contrast, at the highest absorber redshift, $z_{\text{abs}} \sim 3.9$, at the low-wavelength end the Lyman limit can be covered, but at the long-wavelength end we cannot go above $\sim 2150 \text{ \AA}$. Second, at a given wavelength, the redshift distribution of the quasars leads to a strong redshift dependence of absorber–forest pairs, varying from a few times 10^3 ($z_{\text{abs}} \sim 2.1$) to a few times 10^1 ($z_{\text{abs}} \sim 3.9$). Therefore, we expect to have fewer, less-constrained metal lines at higher absorber redshift. Since $\delta_{\text{Ly}\alpha}$ is already selected to have an S/N higher than 10, the number of pairs also indicates that the stacked spectra for each $\delta_{\text{Ly}\alpha}$ and at each z_{abs} have much higher S/Ns, making it possible to identify weak metal absorption lines. We note that there is a break around $1280 \text{ \AA}$, since for the absorber–forest pixel pairs we use the Ly α and Lyman-series forest (metal absorption forest) below (above) this wavelength for the forest pixels, which is also adopted in our stacked spectra.

As an illustration, in Figure 3, we show the stacked spectrum for Ly α absorbers in the whole redshift range of $2 < z_{\text{abs}} < 4$, with $-1.0 < \delta_{\text{Ly}\alpha} < -0.9$. The Lyman series, from Ly α to Ly11, can be easily identified. We leave the discussion of the Lyman series (including those in underdense regions) to a separate paper, and focus our discussion here on the metal lines that show up in relatively strong Ly α absorption systems.

The high S/N of such a stacked spectrum clearly helps to reveal an array of metal lines in the wavelength range of $900\text{--}2880 \text{ \AA}$, including low-ionization lines (e.g., Mg II and Fe II) and high-ionization lines (e.g., O VI and C IV). In particular, the lines with strong oscillator strengths are very prominent (e.g., the Si IV 1394/1403 $\text{\AA}$, C IV 1548/1551 $\text{\AA}$, and Mg II 2796/2804 $\text{\AA}$ doublets). There also appears to be a series of shadow lines. For example, the doublets at $\lambda \sim 1560 \text{ \AA}$

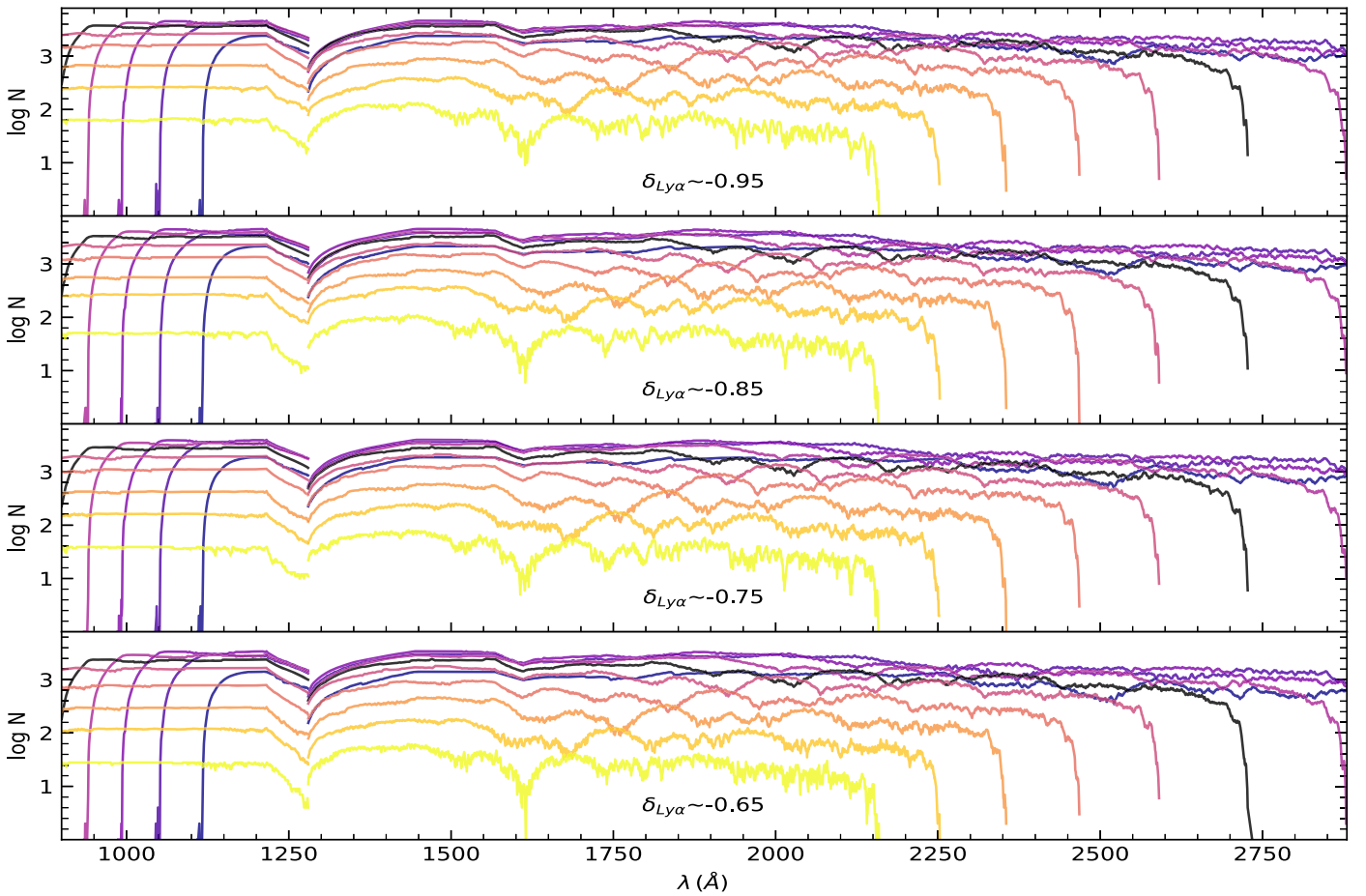


Figure 2. Numbers of Ly α absorber-spectral pixel pairs (see the tables in Appendix A for details). From top to bottom, the panels are for Ly α absorbers with $\delta_{\text{Ly}\alpha} \sim -0.95, -0.85, -0.75$, and -0.65 . In each panel, the color-coded curves (the same as in Figure 9) correspond to 10 absorber redshift bins, from $z_{\text{abs}} = 2.1$ to 3.9, with steps of 0.2, and roughly with a lower curve for higher redshift (at rest frame 1500 Å).

are shadows of the C IV 1548/1551 Å doublets caused by the contamination of Si III 1206 Å in Ly α . Similarly, the doublets at $\lambda \sim 2700$ Å are shadows of the Mg II 2796/2804 Å doublets from the contamination of Si II 1260 Å in Ly α . To avoid confusion, these shadow lines are not labeled in the figure, and a similar plot with the major shadow lines identified and labeled can be found in Appendix B (Figure 15). The lines contaminating Ly α are weak, and they play no significant role in our $\delta_{\text{Ly}\alpha}$ selection.

For each identified Lyman-series line and metal absorption line, we perform Voigt profile fitting to obtain the column density of each species. In performing the fitting, for a set of model parameters (column density N and Doppler/broadening parameter b), the Voigt model is convolved with a Gaussian kernel of $\text{FWHM} \sim 150 \text{ km s}^{-1}$, to account for the eBOSS spectral resolution.

For a given line, we typically choose more than five pixels on each side to determine the continuum, and we visually inspect the pixels to remove those belonging to shadow lines. For example, for Si IV 1394 Å, there is a shadow line of Si II 1527 Å at $\lambda \sim 1391$ Å caused by the contamination of C II 1335 Å in Ly α , and for Si IV 1403 Å, there is a shadow line of Si IV 1394 Å at $\lambda \sim 1404$ Å caused by the contamination of Si III 1207 Å in Ly α . In fitting the lines, we also exclude the pixels near the shadow lines associated with the contamination

of transitions around the Ly α line (such as Si II 1190 Å, Si II 1193 Å, Si III 1207 Å, NV 1239 Å, NV 1243 Å, Si II 1260 Å, O I 1302 Å, Si II 1304 Å, and C II 1335 Å). For example, the absorption feature around 1404.5 Å in the third panel of Figure 7 is the shadow of Si IV 1393 Å associated with the contamination of Si III 1207 Å in Ly α . The pixels around this feature are not included when we perform the Voigt fitting for the Si IV 1402 Å line. Following previous work (e.g., Pieri et al. 2014), for doublets, we choose to fit each line separately, and the difference in the constraints can provide an assessment of the possible systematics. For strongly blended lines that it is hard to measure independently, such as the blend of C II and O VI around 1037 Å and that of O I and Si II around 1303 Å, we perform a joint fit.

In general, we find that, consistent with previous work (e.g., Pieri et al. 2014), the broadening parameter b appears to be high, typically at the level of a couple of hundred km s^{-1} . This is related to the resolution of the eBOSS spectra. The relatively low resolution means that each identified Ly α absorber can consist of multiple absorbers, and the b parameter includes the contributions of thermal motion in each individual absorber and the relative motion among these absorbers. This work focuses on the column densities of various species. In Table 2 in Appendix A, the parameter constraints from the Voigt fitting for each metal line as a function of $\delta_{\text{Ly}\alpha}$ and z_{abs} are provided.

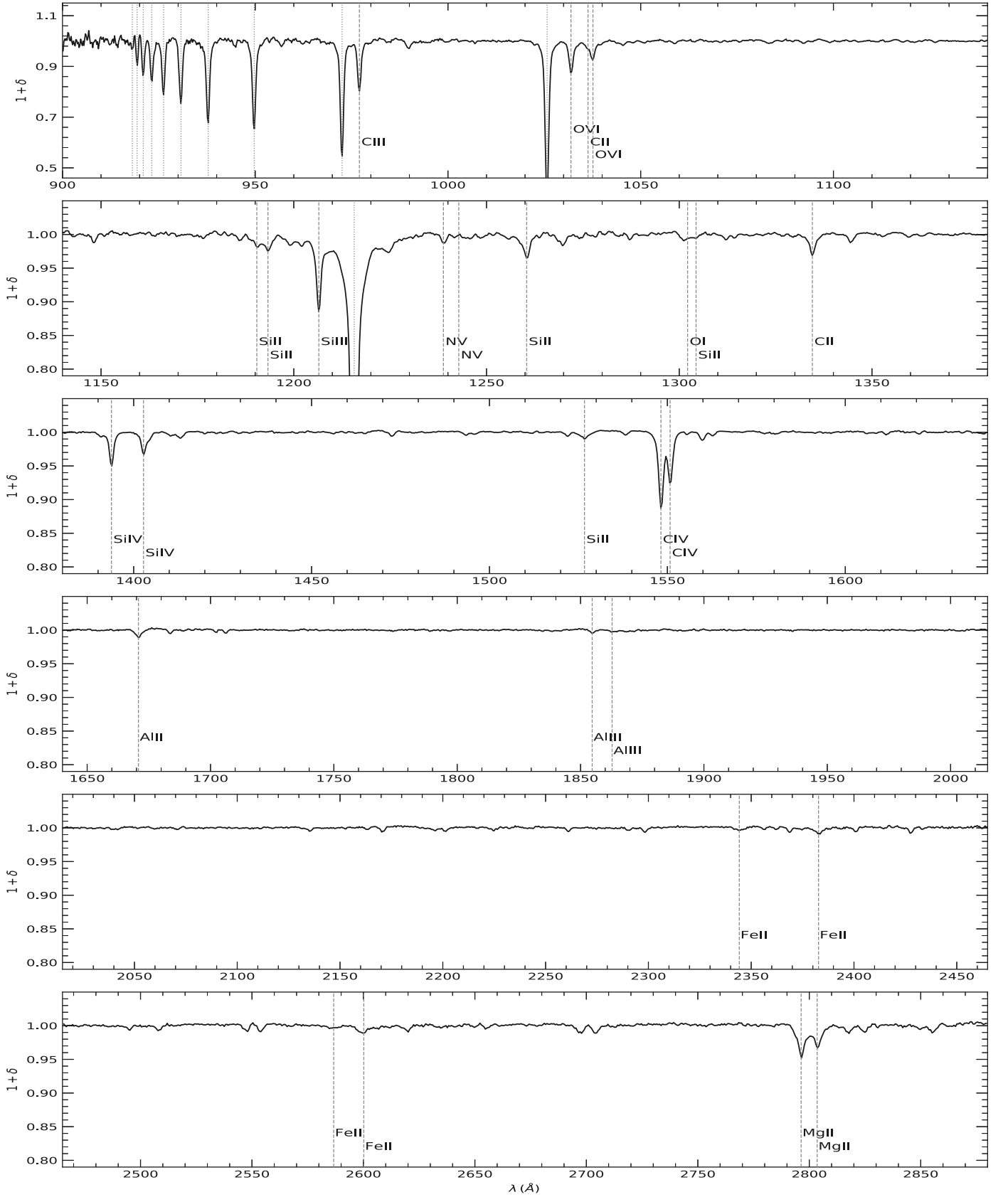


Figure 3. Illustration of the stacked spectrum by selecting $-1.0 < \delta_{\text{Ly}\alpha} < -0.9$ in the whole redshift range of $2 < z_{\text{obs}} < 4$. The vertical dashed lines indicate the metal lines identified, and the vertical dotted lines denote the locations of the Lyman series. The complete stacked spectrum is available as the data behind the figure. (The data used to create this figure are available.)

Table 2
The Measurements of the Metal Lines in the $2.0 < z_{\text{abs}} < 4.0$ Stacked Spectra

Species	λ (Å)	$-1.0 \leq \delta_{\text{Ly}\alpha} < -0.9$			$-0.9 \leq \delta_{\text{Ly}\alpha} < -0.8$			$-0.8 \leq \delta_{\text{Ly}\alpha} < -0.7$			$-0.7 \leq \delta_{\text{Ly}\alpha} < -0.6$		
		$\log N$	b	N_{pair}	$\log N$	b	N_{pair}	$\log N$	b	N_{pair}	$\log N$	b	N_{pair}
C III	977	...	...	...	...	...	...	...	...	...	...	...	...
O VI	1032	...	...	...	...	...	...	...	...	...	...	...	...
C II	1036	...	...	...	...	...	...	...	...	...	...	...	...
O VI	1038	...	...	...	...	...	...	...	...	...	...	...	...
Si II	1190	12.93 ± 0.18	204 ± 129	1, 203	...	...	...	...	...	...	...	...	...
Si II	1193	12.99 ± 0.08	244 ± 63	1, 183	...	...	...	...	...	...	...	...	...
Si III	1207	13.00 ± 0.02	203 ± 15	1, 084	12.44 ± 0.04	137 ± 21	946	12.06 ± 0.09	112 ± 47	817	...	...	...
N V	1239	13.01 ± 0.22	174 ± 118	828	12.66 ± 0.18	49 ± 88	718	12.86 ± 0.22	214 ± 148	618	12.98 ± 0.13	150 ± 75	482
N V	1243	12.87 ± 0.36	15 ± 243	804	13.26 ± 0.15	210 ± 103	696	...	...	...	...	...	...
Si II	1260	12.57 ± 0.07	201 ± 48	697	...	...	...	...	...	...	...	...	...
O I	1302	13.58 ± 0.16	287 ± 103	1, 617	...	...	...	...	...	...	...	...	...
Si II	1304	13.43 ± 0.14	251 ± 97	1, 619	...	...	...	...	...	...	...	...	...
C II	1335	13.77 ± 0.03	307 ± 28	1, 636	13.14 ± 0.07	194 ± 47	1, 462	...	...	...	...	...	...
Si IV	1394	13.16 ± 0.02	214 ± 14	1, 643	12.74 ± 0.03	189 ± 23	1, 467	12.14 ± 0.15	203 ± 109	1, 276	11.81 ± 0.22	99 ± 97	958
Si IV	1403	13.31 ± 0.03	230 ± 25	1, 641	12.81 ± 0.06	151 ± 35	1, 465	12.44 ± 0.12	94 ± 55	1, 274	...	...	...
Si II	1527	13.24 ± 0.04	330 ± 40	1, 564	11.86 ± 0.35	15 ± 457	1, 395	...	...	...	...	...	...
C IV	1548	13.93 ± 0.01	195 ± 5	1, 548	13.54 ± 0.01	139 ± 5	1, 380	13.22 ± 0.02	115 ± 8	1, 200	12.87 ± 0.04	88 ± 18	901
C IV	1551	14.09 ± 0.01	197 ± 6	1, 546	13.74 ± 0.01	166 ± 9	1, 379	13.31 ± 0.02	106 ± 12	1, 199	12.85 ± 0.08	63 ± 37	900
Al II	1671	12.16 ± 0.03	375 ± 38	1, 496	...	...	...	...	...	...	...	...	...
Al III	1855	11.83 ± 0.10	147 ± 55	1, 406	...	...	...	...	...	...	...	...	...
Al III	1863	12.02 ± 0.18	260 ± 143	1, 400	...	...	...	...	...	...	...	...	...
Fe II	2344	13.06 ± 0.04	485 ± 58	873	...	...	...	...	...	...	...	...	...
Fe II	2383	12.58 ± 0.06	367 ± 76	845	11.60 ± 0.17	46 ± 89	751	...	...	...	...	...	...
Fe II	2587	12.95 ± 0.07	417 ± 76	728	...	...	...	...	...	...	...	...	...
Fe II	2600	12.74 ± 0.04	363 ± 42	711	...	...	...	...	...	...	...	...	...
Mg II	2796	12.69 ± 0.04	266 ± 32	516	12.26 ± 0.11	347 ± 110	452	10.92 ± 0.49	15 ± 471	390	...	...	...
Mg II	2804	12.98 ± 0.10	409 ± 109	511	12.29 ± 0.13	176 ± 86	447	...	...	...	...	...	...

Note. The column density N and the Doppler parameter b from the Voigt profile fitting are in units of cm^{-2} and km s^{-1} , respectively. The complete set of metal-line measurements is available in the online Journal in a machine-readable format.

(This table is available in its entirety in machine-readable form.)

3. Results

With the stacked spectra, we first derive the H I column density of each set of Ly α absorbers, by analyzing the Lyman-series lines. Then we study how the metal lines depend on the strength of the Ly α absorbers and their redshift evolution. Finally, we present an illustrative model of the metal column densities, to provide insights into the physical properties of the Ly α absorbers selected in this work.

3.1. Stacked Lyman Series and H I Column Density

Figure 4 shows the stacked H I Lyman-series absorption profiles as a function of $\delta_{\text{Ly}\alpha}$. Compared with previous work (e.g., Pieri et al. 2014), the substantial increase in the S/N allows us to reach Lyman series of much higher order. Four absorber redshifts are chosen for the illustration, which have clear detection up to Ly λ (i.e., Ly11). For each set of a given $\delta_{\text{Ly}\alpha}$, the absorption depth decreases with decreasing line wavelength, resulting from the lower oscillator strength f for higher-order Lyman lines.

We perform Voigt profile fitting for each Lyman-series line, to obtain the column density N_{HI} and the b parameter (and, hence, the equivalent width). An example is shown in Figure 5 for $z_{\text{abs}} \sim 2.9$. As pointed out in Pieri et al. (2014), it is challenging to use lower-order Lyman-series lines (e.g., Ly α and Ly β) to derive precise measurements of the equivalent width. Higher-order, unsaturated lines are more suited for such a task, and the equivalent width (normalized by $f\lambda$) is expected to approach a constant value. As we are able to reach Lyman-series lines of much higher order, we indeed find that the equivalent width reaches a plateau for each set of lines of a given $\delta_{\text{Ly}\alpha}$ (as shown in Figure 5). This can be seen more clearly in the middle panel, where the column densities from higher-order Lyman-series lines are consistent with being at a constant value within the error bars. We take the average column density from Ly5/Ly6/Ly6/Ly7 and above as the N_{HI} measurements for $\delta_{\text{Ly}\alpha} \sim -0.65/-0.75/-0.85/-0.95$, respectively. We find that N_{HI} drops from $\sim 10^{16} \text{cm}^{-2}$ to $\sim 10^{15} \text{cm}^{-2}$ as $\delta_{\text{Ly}\alpha}$ varies from ~ -0.95 to ~ -0.65 , and that there is almost no dependence on absorber redshift.

In what follows, the H I column density is presented together with the column densities of the metal lines and adopted in the modeling of the absorber systems.

3.2. Dependence of Metal Column Densities on $\delta_{\text{Ly}\alpha}$ and Absorber Redshift

Since our focus in this work is on metal lines, we obtain the stacked profiles for 13 major metal species. In Figures 6, 7, and 8, we show the profiles for each line at all four values of $\delta_{\text{Ly}\alpha}$ and at different absorber redshift z_{abs} . We choose to present all the stacked profiles, regardless of the S/Ns, to obtain a full view of the data. The two panels in each row represent those profiles with high and low S/Ns, respectively. The stacked spectra in the right panels typically do not lead to any robust results with Voigt profile fitting.

In bins of the lowest redshift ($z_{\text{abs}} \sim 2.1$), we miss the species with the shortest wavelengths, like C III 977 Å, O VI 1032 Å, C II 1036 Å, and O VI 1038 Å, as they do not enter into the observed wavelength range. Similarly, in bins of the higher redshifts ($z_{\text{abs}} > 2.9$), the species with the longest wavelengths (Fe II and Mg II) are not probed by the data.

At each redshift, for each metal line, there is a clear trend of the line strength increasing with increasing $|\delta_{\text{Ly}\alpha}|$, implying that more metals are present in denser regions. At fixed $\delta_{\text{Ly}\alpha}$, each metal line becomes weaker toward higher redshift.

At a given $\delta_{\text{Ly}\alpha}$ and z_{abs} , Voigt profile fitting is carried out for each metal line, as long as the S/N allows it.

Specifically, if the minimum pixel value of an absorption line lies below 1σ of the noise level of the “continuum,” we perform the profile fitting. The best fits are shown as the red solid curves. The derived column densities N and Doppler parameters b can be found in Table 2 in Appendix A. Overall, if the results are averaged over the redshift range, we find broad agreements with those in Pieri et al. (2010) and Pieri et al. (2014).

Figure 9 summarizes the dependence of the column density on $\delta_{\text{Ly}\alpha}$ and z_{abs} for various metal species, as well as neutral hydrogen. In each panel, up to four sets of data points are shown, corresponding to the four values of $\delta_{\text{Ly}\alpha}$ (-0.95 , -0.85 , -0.75 , and -0.65). Each set of data points is color-coded by the absorber redshift z_{abs} . The column density ranges of the panels are all chosen to be across 3 dex for an easy comparison between the strengths of the column density dependence on $\delta_{\text{Ly}\alpha}$ and z_{abs} of different species.

At each redshift, the column density of each metal species drops as $\delta_{\text{Ly}\alpha}$ increases from -0.95 to -0.65 . The column density of neutral hydrogen (top left panel) drops by about 1 dex, from $\log(N_{\text{HI}}/\text{cm}^{-2}) \sim 16$ to ~ 15 . The changes of the column densities of high-ionization metal species (such as O VI, N V, C III, and C IV) seem to track those of neutral hydrogen, all having a ~ 1 dex decrease over the $\delta_{\text{Ly}\alpha}$ range. In contrast, the column densities of low-ionization metal species show a stronger dependence on $\delta_{\text{Ly}\alpha}$. For example, that of Mg II drops by almost 2 dex over the $\delta_{\text{Ly}\alpha}$ range. Such a difference implies that high- and low-ionization species arise from different environments, as will be shown with a simple model in the next subsection.

At a given $\delta_{\text{Ly}\alpha}$, the column density of neutral hydrogen shows almost no dependence on redshift. In comparison, those for metal species at fixed $\delta_{\text{Ly}\alpha}$ show a clear dependence on redshift, typically more than 0.5 dex from $z_{\text{abs}} \sim 2.1$ to $z_{\text{abs}} \sim 3.5$, which may result from the evolution of temperature and ionizing background.

To help us understand these results, we turn to a simple model of Ly α absorbers.

3.3. Illustrative Model of Metals in the Ly α Forest of High Neutral Hydrogen Column Density

The column densities of various species encode information about the physical properties of the absorption system. Similar to Pieri et al. (2014), we apply a simple illustrative photoionization model to interpret the column density measurements.

3.3.1. Parameter Dependence

To guide our choice of models for comparing to the measurements, in Figure 10 we show the dependence of the column density of various species on the stop neutral hydrogen column density (N_{HI}), total hydrogen number density (n_{H}), and temperature (T), with the gas in a plane-parallel geometry photoionized by an external ionizing UV background. The nearly linear dependence of the column density of a species on

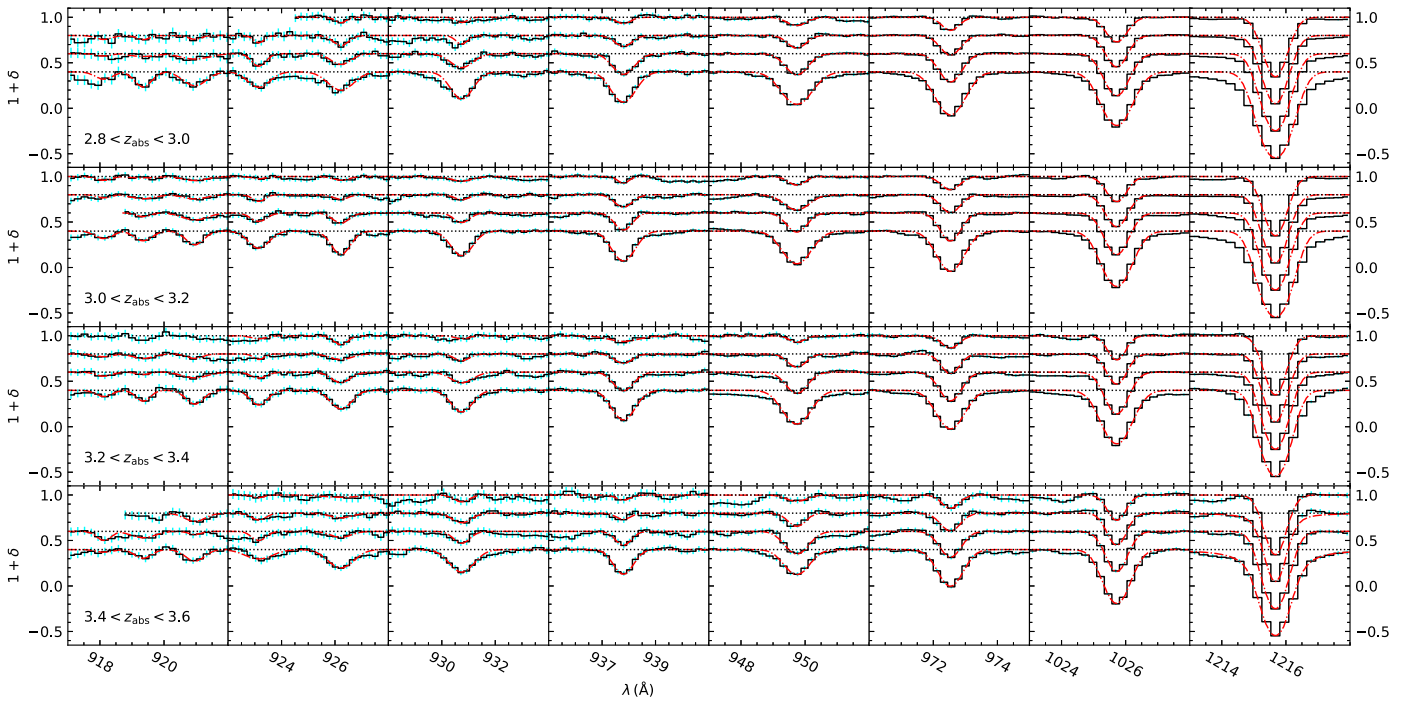


Figure 4. Stacked H I Lyman-series absorption-line profiles as a function of $\delta_{\text{Ly}\alpha}$ and absorber redshift. The absorber redshift is labeled in each panel, in the first column. At each redshift, the four sets of profiles are for $\delta_{\text{Ly}\alpha} \sim -0.65, -0.75, -0.85$, and -0.95 (from top to bottom), with offsets of 0, -0.2 , -0.4 , and -0.6 , respectively, for clarity. The dotted lines show the continuum levels and the dotted-dashed curves are the Voigt profile fits. Seen from right to left, the panels in each row are the line profiles of Ly α , Ly β , Ly γ , ..., and Ly λ (i.e., Ly11). For this figure, the absorption redshifts are chosen so that the Ly11 line can be detected for $\delta_{\text{Ly}\alpha} \sim -0.95$.

metallicity (with other parameters fixed) is not shown here. Figure 14 in Appendix A further explores the parameter dependence.

In Figure 10, the species are ordered such that the ionization potential decreases from left to right. The calculations are performed with version 17.02 of the photoionization code Cloudy, last described by Ferland et al. (2017). The gas was assumed to be a uniform slab of constant density in thermal and ionization equilibrium and exposed to the external UV background. The models explored are centered on a model with total hydrogen density $\log(n_{\text{H}}/\text{cm}^{-3}) = -3$ and temperature $\log(T/\text{K}) = 4.5$. The metallicity is fixed at $[X/\text{H}] = -1$. Solar abundance is assumed and the Haardt & Madau (2012) UV background at $z=2.75$ is adopted in the calculations. With the central model, we vary the three parameters, N_{H} , n_{H} , and T , one at a time. Each panel shows one set of changes, and the thicker lines are for higher values of the corresponding varied parameter.

As with the metallicity dependence, there is a clear trend that the column density of a species increases with increasing stop neutral hydrogen column density (top panel). The dependence is stronger for low-ionization species. For example, for a 3 dex change in N_{H} , the column density of Mg II increases by about 4.5 dex, while the change is only about 1.5 dex for that of O VI. This is expected in the photoionization model. At higher hydrogen column density, the ionizing background is more attenuated within the absorption system. The subsequent lower photoionization rate causes a given species to stay at a low-ionization state rather than a high-ionization state.

As for the dependence on the total hydrogen density n_{H} , a higher n_{H} results in a higher recombination rate, making it hard for a species to maintain a state of high-level ionization. As a consequence, the column densities of high-ionization species

drop by large factors (middle panel). For example, as n_{H} increases from $\log(n_{\text{H}}/\text{cm}^{-3}) = -3$ by 1 dex, the column densities of O VI, N V, and C IV decrease by about 3 dex, 2.5 dex, and 2 dex, respectively. The recombination from higher-level ionization, on the other hand, increases the column densities of low-ionization species by mild amounts, e.g., ~ 1 dex for Mg II. In contrast, lowering n_{H} by 1 dex leads to substantial drops in the column densities of low-ionization species (e.g., by ~ 2.5 dex for Mg II) and clear increases in the column densities of high-ionization species (e.g., by ~ 2.5 dex for O VI). As we will see, such an increase for high-ionization lines at low density provides a possible mechanism for explaining the measurements.

Increasing the temperature by 0.5 dex from $\log(T/\text{K}) = 4.5$ (and hence reducing the recombination rate) has the effect of enhancing the high-ionization species relative to the low-ionization ones (bottom panel). Compared to the lower-density case (middle panel), the relative abundances of high-ionization lines do not change much. For example, the column density of O VI still appears to be lower than that of C IV. The column density changes in the low-ionization species are not as drastic as for the lower-density case. Compared with the case of increasing temperature by 0.5 dex, decreasing the temperature by 0.5 dex only has a small effect, shifting the column densities by at most 0.5 dex.

3.3.2. A Two-component Model

To see how the measured column densities of various species compare with the Cloudy models, we use the absorption system selected by $\delta_{\text{Ly}\alpha} = -0.95$ and $2.6 < z_{\text{abs}} < 2.8$ as an example. The data points in the top panel of Figure 11 show the inferred column densities (see

Table 2 for the $2.6 < z_{\text{abs}} < 2.8$ absorbers). For those lines with column density measurements barely missed at this redshift, but with measurements at the neighboring redshifts, such as C II 1036 Å and O VI 1038 Å, we apply interpolations to obtain the $z_{\text{abs}} \sim 2.7$ values. For a species with multiple lines, we find that the (logarithmic) column density measurements from different lines can have a systematic difference up to 0.10–0.25 dex (e.g., C IV or Fe II), and we take the maximum column density as the measurement.¹⁵

Overall, the column densities of high- and low-ionization species are within a relatively narrow range of about 2 dex. In contrast, the single-component models presented in Figure 10 all show a large range in column density. In particular, these models are unable to give the column densities of high-ionization species at a comparable level to those of low-ionization species, as seen in the data. Motivated by Pieri et al. (2014), where the high-ionization and low-ionization species are compared to different models, we explore the scenario that the inferred column density pattern results from a superposition of the contributions from a high-metallicity and a low-metallicity component.

We build two sets of Cloudy models with the gas in a plane-parallel geometry, one with metallicity $[X/H] = -2$ and the other with $[X/H] = -0.3$. For each set, we compute the models on a grid, $\log(n_{\text{H}}/\text{cm}^{-3}) = -4$ to 0, with steps of 0.05, and $\log(T/\text{K}) = 4$ to 6, with steps of 0.05, all with stop neutral hydrogen column density $\log(N_{\text{H}}/\text{cm}^{-2}) = 16$, assuming solar abundance and the Haardt & Madau (2012) UV background at $z = 2.75$. Note that the mean hydrogen density at this redshift is about 10^{-5} cm^{-3} and the hydrogen density in the set of models covers 10 to 10^5 times the cosmic mean. For each combination of the low-metallicity and high-metallicity models, the relative fraction is solved to best fit the measurements. Our purpose here is to look for combinations that can provide a reasonable explanation of the overall trends of the column densities of various species. If we were to use the measurement uncertainties, the results would be completely driven by the few species with small error bars (e.g., 0.02 dex for C III and Si IV in Table 2, for $2.6 < z_{\text{abs}} < 2.8$). At such an uncertainty level, the model assumptions may also matter (e.g., the adopted solar abundance pattern). To focus on the overall trends, and given the possible systematics in deriving the column densities (e.g., seen in the differences from multiple lines), we adopt a uniform error bar of 0.2 dex in this exercise.

The solid line in the top panel of Figure 11 is the best-fit model, consisting of a small contribution (2.3%) of low-metallicity systems with $\log(n_{\text{H}}/\text{cm}^{-3}) = -2.9$ and $\log(T/\text{K}) = 5.4$ (red dotted line) and a large contribution (97.7%) of high-metallicity systems with $\log(n_{\text{H}}/\text{cm}^{-3}) = -1.5$ and $\log(T/\text{K}) = 4$ (blue dotted line). To have a better idea of the possible combinations, we show the constraints on n_{H} and T in the bottom panel. The high-metallicity component is well constrained to be around $\log(n_{\text{H}}/\text{cm}^{-3}) \sim -1.5$ and $T \sim 10^4$ K. The low-metallicity component, on the other hand, is only

¹⁵ We find that the discrepancies between the column densities from the different lines (e.g., doublets) of an ion are related to the difference in their cross sections (oscillator strengths). The line with the larger cross section is more easily saturated. With low-resolution spectra, the Voigt fit has the tendency to push the model toward the linear regime, by fitting it with a large b value. This would lead to a smaller fitted column density. Therefore, the column density from the line with the minimum oscillator strength is a better estimate. We performed tests using inverse variance-weighted mean column densities when doing the fitting with the Cloudy model, and the results remained similar.

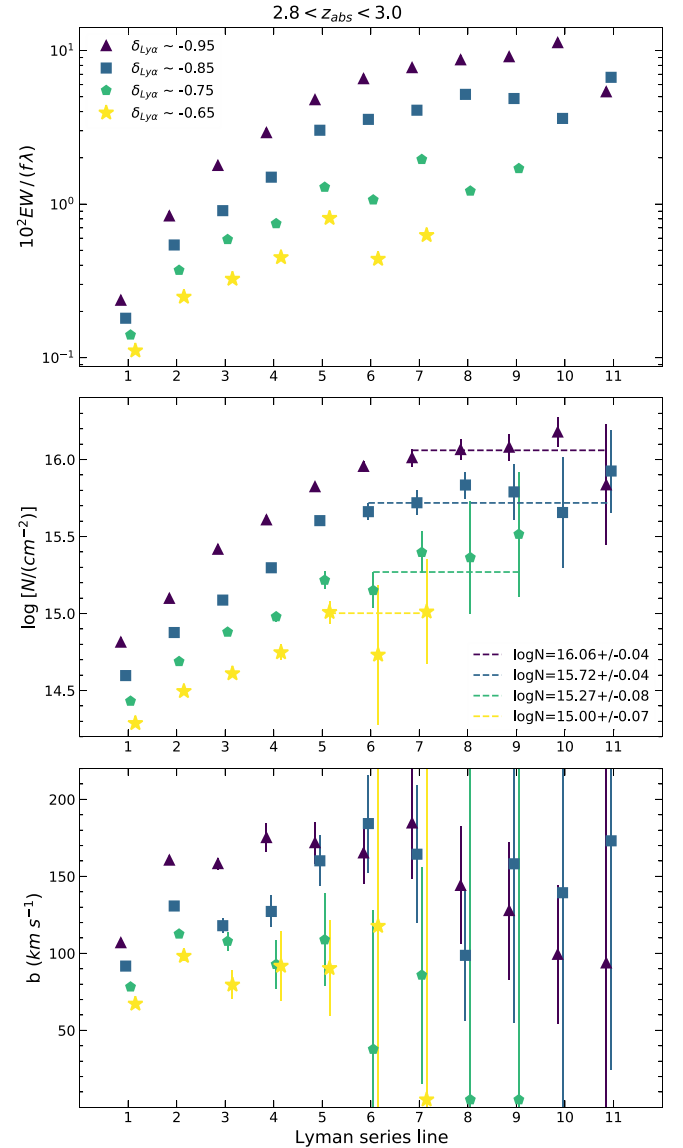


Figure 5. Measurements of the H I column density with Lyman-series lines at $z_{\text{abs}} \sim 2.9$. In each panel, the four sets of symbols correspond to absorbers with different values of $\delta_{\text{Ly}\alpha}$. The bottom two panels show the Voigt profile fitting constraints on column density N and Doppler parameter b . The top panel shows the derived equivalent width. In the middle panel, each dashed horizontal line represents the average column density in the plateau from higher-order Lyman series, taken as the inferred H I column density. See the details in the text.

loosely constrained. It can come from systems with either low density ($\sim 10^{-4} \text{ cm}^{-3}$) or high temperature ($\gtrsim 10^5 \text{ K}$), both driven by the high column densities of high-ionization species (e.g., O VI, N V, and C IV; see the parameter dependence in Figure 10). Such physical conditions appear to be in line with the prediction for highly ionized metals from hydrodynamic simulations (e.g., Rahmati et al. 2016). The high-density component likely comes from cool gas in the CGM, while the other one is contributed by the gas in the IGM. The two-component model is consistent with the findings based on high-resolution observations that metal absorbers show bimodal physical properties with low- and high-metallicity branches (Kim et al. 2016).

We note that we have performed a test with Cloudy by switching to the HM05 UV background model. We find that the

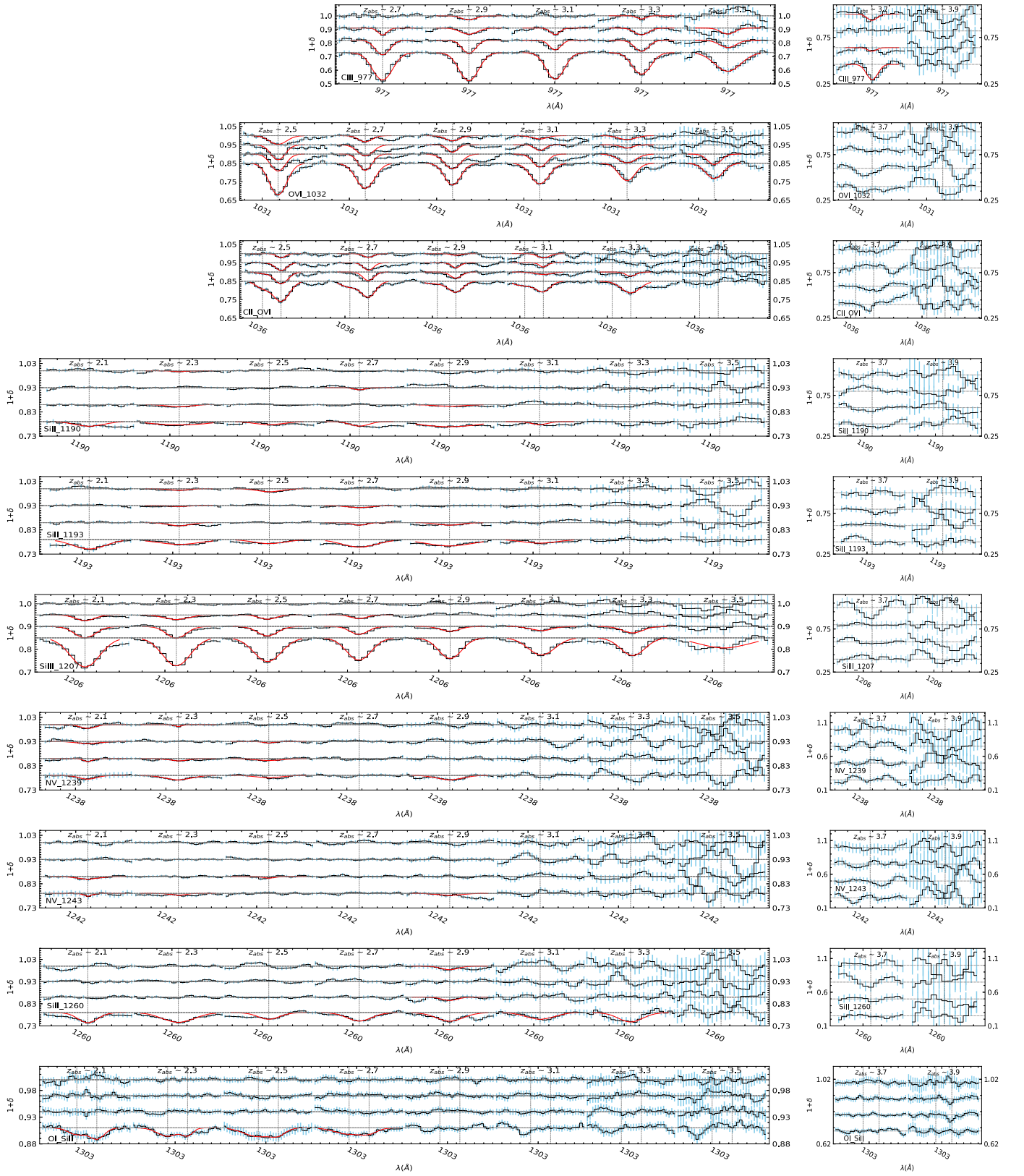


Figure 6. The dependence of the line profiles in the stacked spectra on $\delta_{\text{Ly}\alpha}$ and z_{abs} for 13 metal species. In each panel, the relevant species are labeled in the lower left corner, and the horizontal tick mark intervals are 1 Å. The $\delta_{\text{Ly}\alpha}$ and z_{abs} dependences are shown along the vertical and horizontal directions, respectively. Offsets are added for different values of $\delta_{\text{Ly}\alpha}$. The vertical lines mark the wavelengths of the relevant lines, and the red curves show the Voigt profile fits. In each row, the cases for the $z_{\text{abs}} < 3.6$ and $3.6 < z_{\text{abs}} < 4.0$ absorbers are plotted in the left and right panels, respectively, and the division is mainly driven by the potentially different dynamical ranges in the y-axes of the right panels, to accommodate the larger uncertainties in the profiles at higher redshifts.

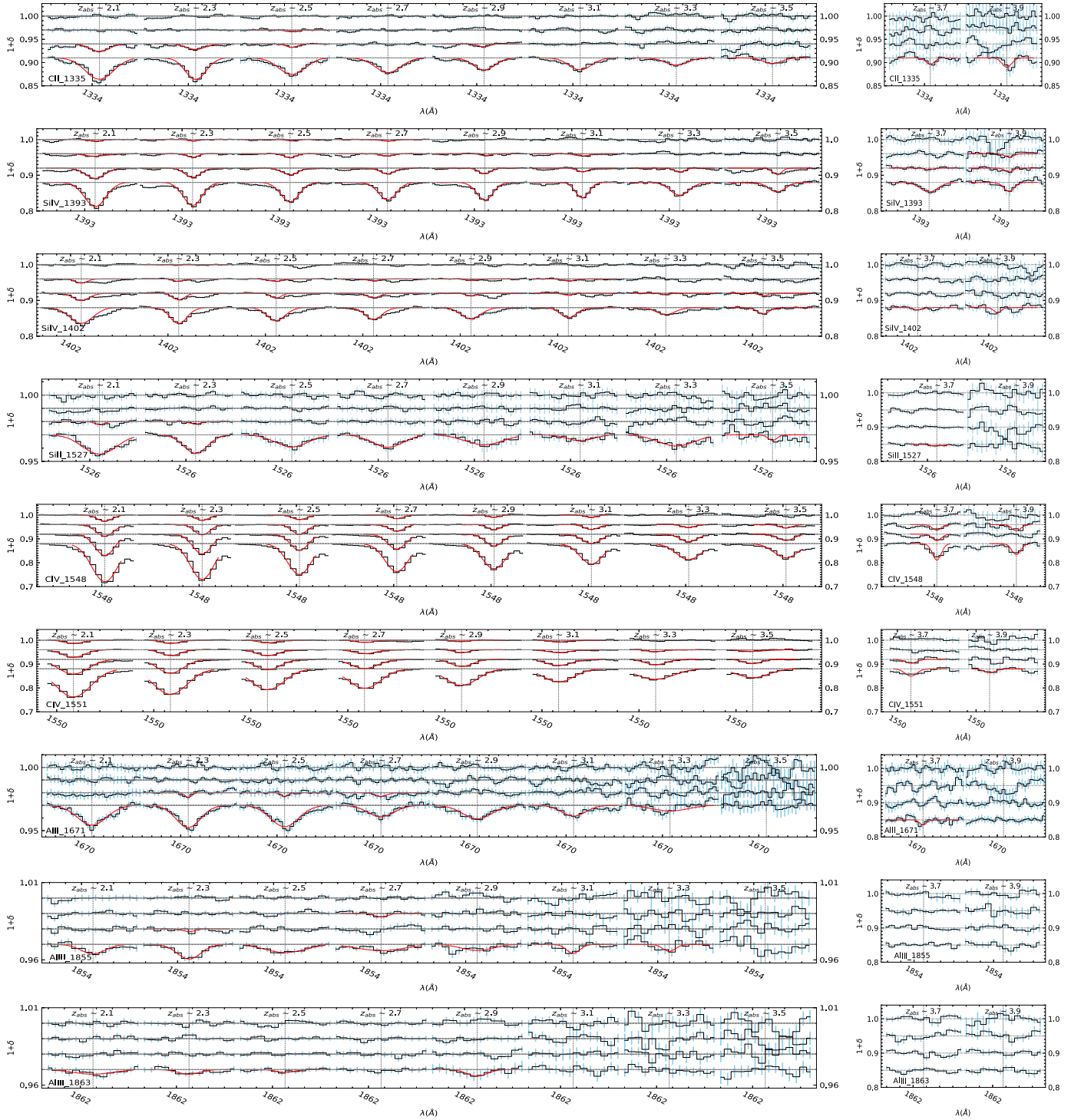


Figure 7. Continued: the dependence of the line profiles in the stacked spectra on $\delta_{\text{Ly}\alpha}$ and z_{abs} for 13 metal species.

overall trend, e.g., as seen in Figure 11, remains almost the same, with a slight shift in the number density direction. For example, the best-fit number density for the low-metallicity component shifts by about 0.2 dex, from $\log(n_{\text{H}}/\text{cm}^{-3}) \sim -1.5$ to -1.3 .

3.3.3. High- and Low-ionization Species and a Multicomponent Model

It should be emphasized that the exercise shown in Figure 11 indicates that the selected absorption systems cannot be

described by a simple single-component model. However, this does not necessarily mean that the right model is the two-component model, which is merely used here to illustrate the possible complexity of such systems. In fact, the fit shown in the top panel of Figure 11 misses a few data points, both in the high- and low-ionization species.

For example, the pattern in the high-ionization species is not fully reproduced, with the predicted C IV and Si IV lower than measured. While the column densities of most low-ionization

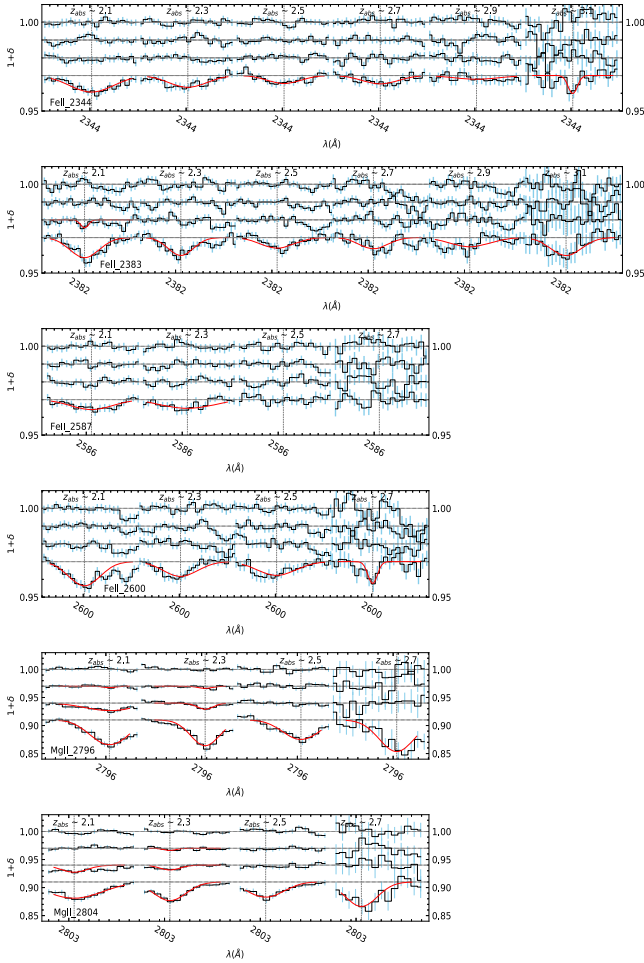


Figure 8. Continued: the dependence of the line profiles in the stacked spectra on $\delta_{\text{Ly}\alpha}$ and z_{abs} for 13 metal species.

species can be reasonably described by the model, the model underpredicts the column densities of Fe II and Mg II by about 0.5 dex and that of O I by about 1.5 dex.

An adjustment of the metallicity can result in the model better matching the column densities of low-ionization species. With $N_{\text{HI}} = 10^{16} \text{ cm}^{-2}$, we find that a model with supersolar metallicity ($[X/H] = 0.3$) and $\log(n_{\text{H}}/\text{cm}^{-3}) = -1$ can lead to a better match to the column densities of Fe II and Mg II, but the predicted O I column density still falls short of the measurement by ~ 0.75 dex. We note that in Pieri et al. (2014), a model with $[X/H] = 0.3$ was able to provide a match to their O I column density measurement (see their Figure 12). The difference may have two causes. First, we have a different selection of hydrogen absorption systems, while the measured O I column density, $\log(N_{\text{OI}}/\text{cm}^{-2}) = 13.77$, from our selection is higher than theirs (by about 0.37 dex). Second, the solar abundance pattern used in our work appears to be different from theirs. As it is unlikely to have a metallicity much higher than $[X/H] = 0.3$, in our adopted model we find it hard to explain the high O I column density measurement solely by the metallicity effect.

The mismatch between the measured and model O I column density could suggest that O is more abundant in real systems, i.e., the abundance is different from the solar abundance adopted in the calculation. This may also help to resolve the difference seen in the high-ionization species, which relates to

the CNO enrichment history. The high O/N ratio could be caused by a series of major star formation episodes in galaxies, with short intervals between them (e.g., Kobulnicky & Skillman 1998; Pettini et al. 2002; Pettini 2004), given that O is produced in core-collapse supernovae (>10 Myr after starburst) and N comes from intermediate-mass stars ($4\text{--}8M_{\odot}$; >100 Myr after starburst). Galactic winds may bring such an abundance pattern into the CGM/IGM, where the absorption systems likely reside.

There are certainly other possibilities. Given the resolution of the eBOSS spectra, the strong absorption systems that we select, in this case, $\delta_{\text{Ly}\alpha} \sim -0.95$, may have a small fraction of contamination from high-column density systems. In high-column density systems (e.g., Lyman-limit systems), the column density of O I, the one with the lowest ionization potential for the species considered here, could become enhanced, owing to the low photoionization rate inside the systems from the attenuated ionizing background (see the bottom two panels of Figure 14 in Appendix A). Such an enhancement in O I column density is seen in DLAs (see, e.g., Cooke et al. 2011a, 2011b). Figure 12 illustrates a three-component model (not from model fitting). Loosely motivated by Figure 11, the components with $([X/H], \log n_{\text{H}}, \log T) = (-2.0, -4.0, 5.0)$ and $(-1.0, -3.0, 4.0)$, both with $N_{\text{HI}} = 10^{16} \text{ cm}^{-2}$, are adapted to reasonably produce the trend seen in about half of the species with higher ionization potentials. The third component represents a contamination from $N_{\text{HI}} = 10^{19} \text{ cm}^{-2}$ systems ($[X/H] = -1.0$ and $\log n_{\text{H}} = -1.5$), which is able to produce the high column density of O I and improves the match to those of Fe II and Mg II. In combination, the three components lead to a trend that follows the measurements reasonably well. Increasing the metallicity, n_{H} , and N_{HI} all have the effect of boosting the column density of O I, the species with the lowest ionization potential, and the latter two have the additional effect of causing a steeper jump between the Fe II/Mg II and O I column densities. The contamination from the third component is assumed to be 5% in the illustration. While the metallicity $[X/H] = -1$ is motivated by the measurements in Lyman-limit systems (e.g., Fumagalli et al. 2016; Lehner et al. 2016), given the degeneracy among $[X/H]$, n_{H} , and N_{HI} , the values used here are only for the purpose of illustrating the possible contribution from high-column density systems. In addition, there is hardly any reason to limit to three components for the selected absorbers.

If we broadly take the contamination from the Lyman-limit systems to be at a level of a few percent, is it reasonable? To answer this, we must understand the systems that meet the selection function. Similar to the case of flux transmission $F < 0.25$ in Pieri et al. (2014), isolated lines without damping wings do not reach $-1 < \delta_{\text{Ly}\alpha} < -0.9$ in moderate-resolution spectra (such as those observed in eBOSS). The selection function requires lines that are both saturated in high resolution and clustered with one another on scales within the FWHM of the SDSS resolution element (see Figure 2 of Pieri et al. 2014). For this reason, we cannot simply integrate the column density distribution function in order to assess the proportion of Lyman-limit systems present in the selected absorbers. Put in logic terms, the presence of $N_{\text{HI}} \gtrsim 10^{14} \text{ cm}^{-2}$ is a necessary but not sufficient condition for selection, since clustering is also required; however, $N_{\text{HI}} > 10^{17.5} \text{ cm}^{-2}$ is neither necessary nor sufficient, since such systems are not necessarily sufficiently

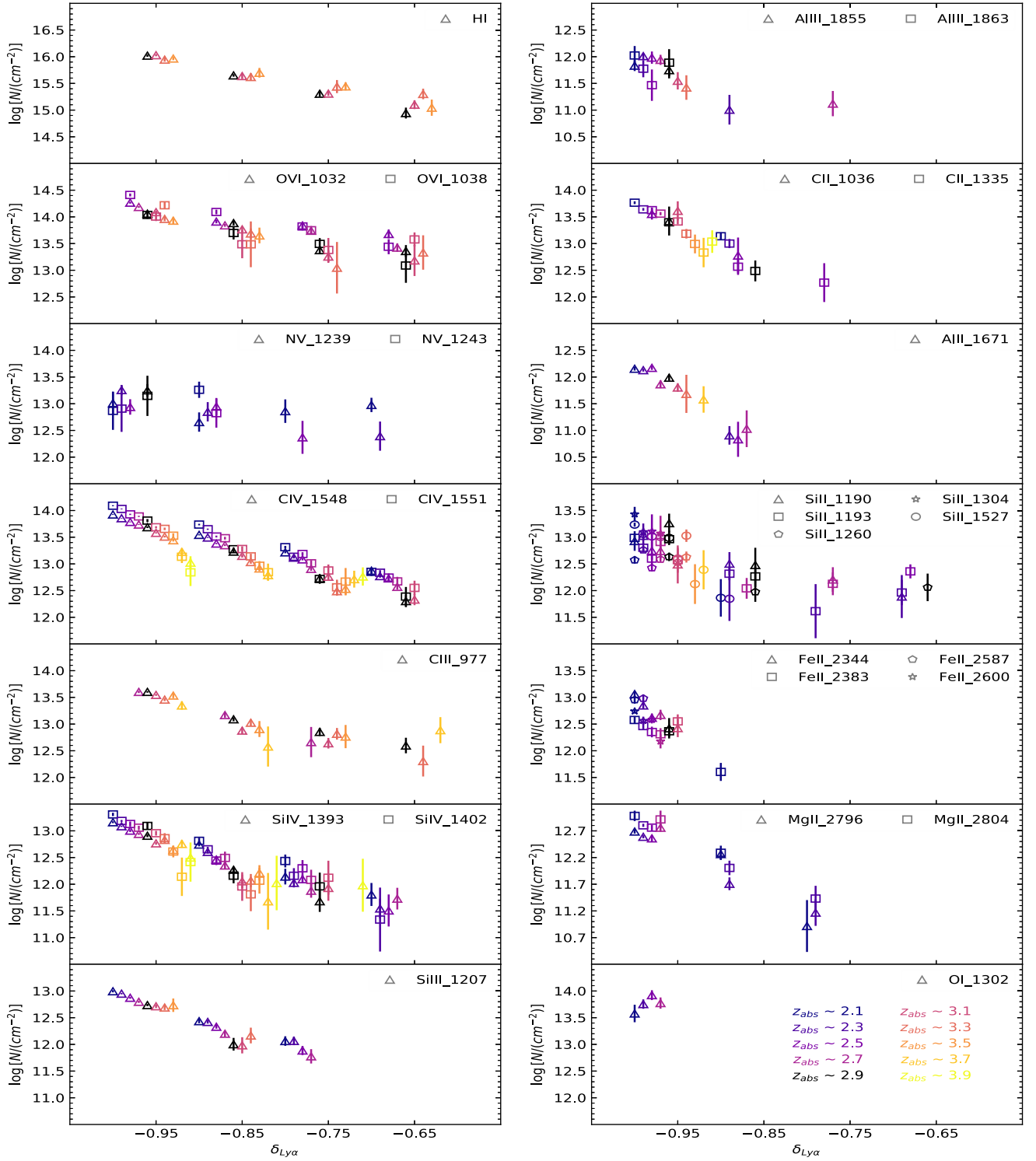


Figure 9. The dependence of the column density on $\delta_{\text{Ly}\alpha}$ and absorption redshift for 13 metal species, as well as neutral hydrogen. In each panel, there are up to four sets of data points for the four values of $\delta_{\text{Ly}\alpha}$ (-0.95 , -0.85 , -0.75 , and -0.65). Each set of data points at a given $\delta_{\text{Ly}\alpha}$ is color-coded by the absorber redshift. If there are multiple lines of a species, they are distinguished by different symbols.

clustered with other $N_{\text{HI}} > 10^{14} \text{ cm}^{-2}$ lines on $\sim 150 \text{ km s}^{-1}$ scales. Hence, the column density distribution function is largely orthogonal to our selection function.

In order to understand the populations of individual lines selected here, we must compare with the results from line fits of high-resolution data and verify what the selection function provides. This was done in Pieri et al. (2014), with the

conclusion that $N_{\text{HI}} > 10^{17.5} \text{ cm}^{-2}$ systems can contribute no more than 3.7% of the systems with a true flux transmission of $F < 0.25$ with SDSS resolution. While our selection function is for stronger features than those in Pieri et al. (2014), the selection function is otherwise the same. We have no basis for concluding that the somewhat stronger blended absorption here differs from the line model test in Pieri et al. (2014). We

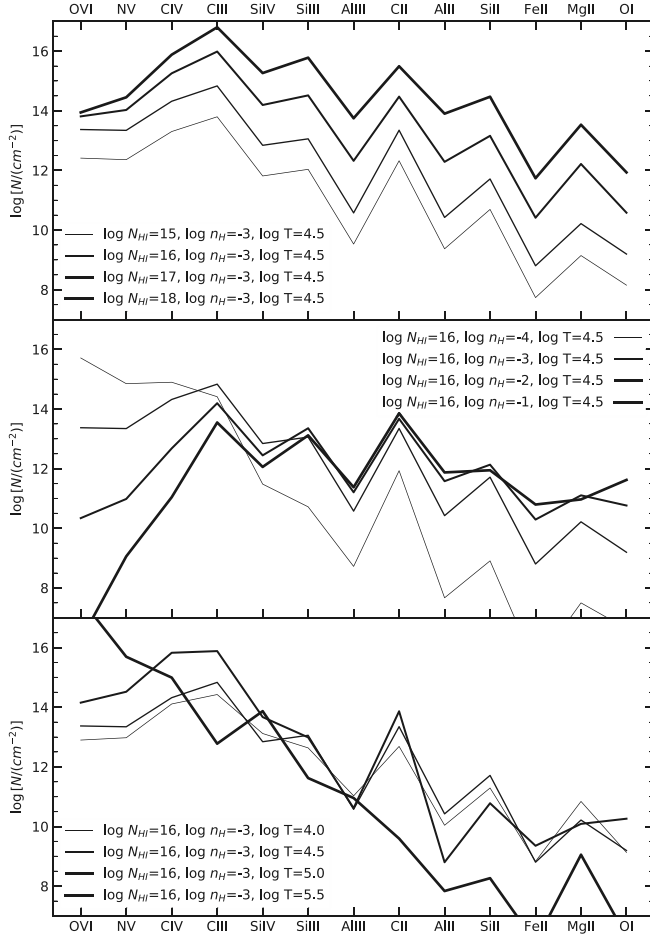


Figure 10. The dependence of the column densities of various species on the neutral hydrogen column density (top), total hydrogen number density (middle), and temperature (bottom), based on Cloudy models. The species are ordered from high-ionization to low-ionization potential, from left to right. The metallicity is fixed to $[X/H] = -1$ for all cases.

therefore also assume their conclusion, that systems with $N_{\text{HI}} > 10^{17.5} \text{ cm}^{-2}$ should be no more than 3.7% of our sample.

The discussion here focuses on $\delta_{\text{Ly}\alpha} \sim -0.95$ systems. The OI line becomes too weak to be measured robustly for $\delta_{\text{Ly}\alpha} \gtrsim -0.85$ systems (bottom panel of Figure 6), and this is nearly so for Fe II and Mg II as well. Such a behavior lends support to the scenario of contamination from high-column density systems to the strongest absorption systems ($\delta_{\text{Ly}\alpha} \sim -0.95$) in this study with eBOSS spectra.

In Appendix C, we perform further tests through selecting presumably stronger absorbers within the $\delta_{\text{Ly}\alpha} \sim -0.95$ systems. The change in the metal column densities and the feature at the corresponding Lyman limit also provide further evidence of the contribution from high-column density systems.

To summarize, the column densities of various low-ionization species in $\delta_{\text{Ly}\alpha} \sim -0.95$ absorption systems at $2.7 < z_{\text{abs}} < 2.9$ can be reasonably explained by a model cloud with $n_{\text{H}} \sim 10^{-1.5} \text{ cm}^{-3}$ and $T \sim 10^4 \text{ K}$, photoionized by a UV background (Haardt & Madau 2012) to have a neutral hydrogen column density of $N_{\text{HI}} \sim 10^{16} \text{ cm}^{-2}$. The relatively high column densities of the high-ionization species suggest an additional component, with either low n_{H} or high T , or a deviation from solar abundance. The high column density of OI, the one with the lowest ionization potential of the species considered in this work, may signal

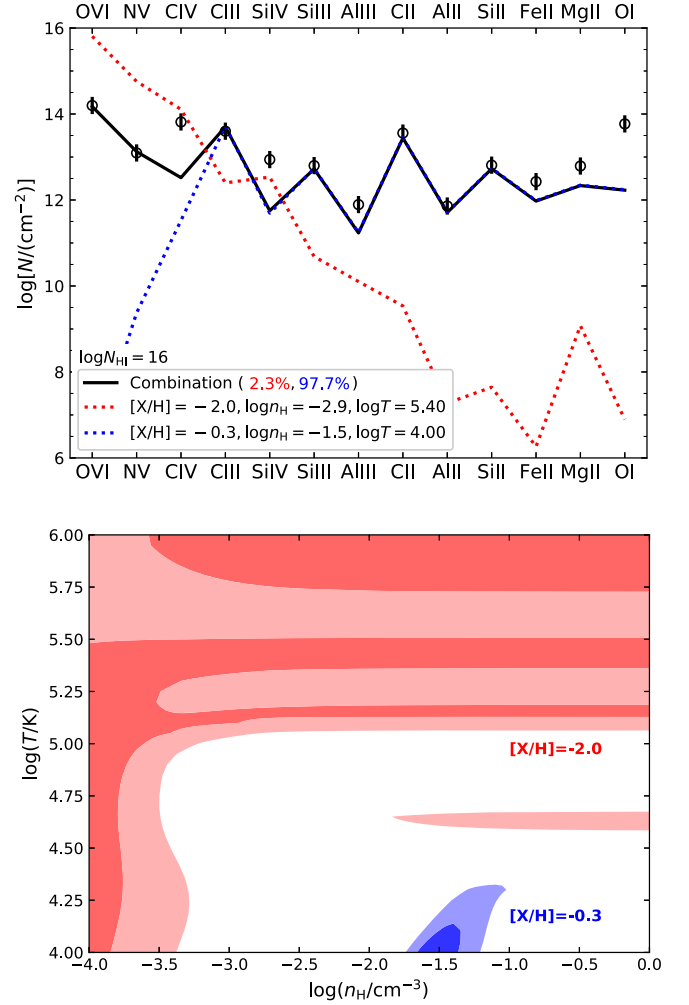


Figure 11. Illustration of the combination of two absorption systems for comparing to the observed column densities of various species with $\delta_{\text{Ly}\alpha} \sim -0.95$ and $2.6 < z_{\text{abs}} < 2.8$. Top: the best fit (black) from a combination of a low-metallicity system (red) and a high-metallicity system (blue). Bottom: the 1σ and 2σ constraints in n_{H} and T for the combination of a low-metallicity system and a high-metallicity system.

contamination from high-column density systems, in view of the spectral resolution of eBOSS.

Multiple components also appear to apply to other $\delta_{\text{Ly}\alpha}$ systems and to systems at different redshifts. In Figure 13, the column densities of various species at different $\delta_{\text{Ly}\alpha}$ values are shown in different panels, and in each panel the redshift dependence is color-coded. The Cloudy model (not a fit) with a high-metallicity and a low-metallicity component of different values of n_{H} and T is used to illustrate the possible multiple components. A histogram, instead of lines connecting the column density of each species, is used here for the model and its components, to avoid possible confusion with the redshift dependence. The models across the four values of $\delta_{\text{Ly}\alpha}$ are the same, except for the stop neutral hydrogen column density. The OI line is not robustly measured, except for the strongest absorption system with $\delta_{\text{Ly}\alpha} \sim -0.95$, consistent with it being contaminated by systems of high column density. Given the uncertainties discussed above, we make no attempt to further model the redshift dependence, which can be related to the metallicity and density evolution of the selected systems.

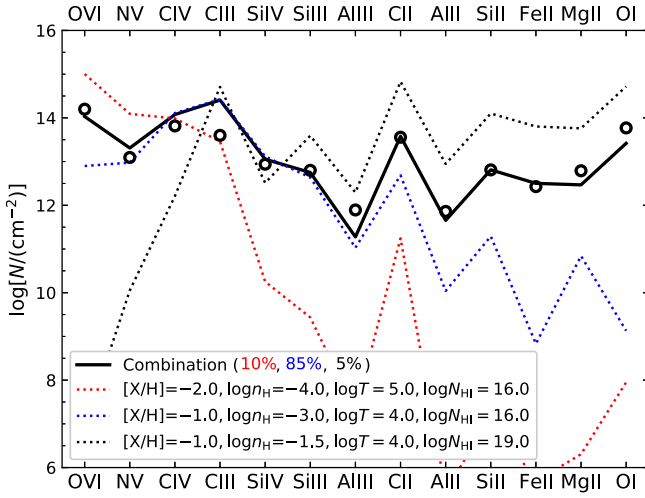


Figure 12. Illustration of a three-component model toward explaining the high measured column density of O I (or, more generally, species with low ionization potential). The open circles show the observed column densities of various species, with $\delta_{\text{Ly}\alpha} \sim -0.95$ and $2.6 < z_{\text{abs}} < 2.8$. A high-column density system (black dotted line) is added, and its combination with the other two systems (the red and blue dotted lines) is shown as the solid line, with the combination fractions labeled in the legend (not model fits). This is only for illustration purposes, as there is degeneracy among the model parameters (e.g., column density and metallicity), and the system is unlikely to be limited to three components. See the text for more details.

4. Summary and Discussion

We study metal absorption lines associated with relatively strong Ly α absorbers in the Ly α forest using quasar spectra from the eBOSS survey. Various metal species are revealed in the stacked spectra, and their column densities are derived as a function of Ly α absorber strength and redshift. We find that multiple components, including low- and high-metallicity gas of different densities and temperatures, are needed to interpret the inferred column density pattern of the metal species. The results can be used to study the chemical enrichment of the CGM and IGM. They also provide inputs for modeling the metal contamination in the Ly α forest BAO measurement.

We construct a catalog of Ly α absorbers in 10 redshift bins over $2 < z_{\text{abs}} < 4$, selected according to the Ly α flux fluctuation $\delta_{\text{Ly}\alpha}$. We focus on the relatively strong absorbers, with $\delta_{\text{Ly}\alpha} \sim -0.95, -0.85, -0.75$, and -0.65 . The Ly α absorbers in each redshift bin are cross-correlated with the flux fluctuation in quasar spectra over the full wavelength range (shifted in accordance with the Ly α absorber redshift) to obtain the stacked spectra. The neutral hydrogen column density of these absorbers, from analyzing higher-order Lyman-series lines, ranges from $\sim 10^{15} \text{ cm}^{-2}$ (for $\delta_{\text{Ly}\alpha} \sim -0.65$) to $\sim 10^{16} \text{ cm}^{-2}$ (for $\delta_{\text{Ly}\alpha} \sim -0.95$), which has only weak redshift evolution.

In the stacked spectra, up to 13 metal species can be identified, including high-ionization ones (e.g., O VI, N V, and C IV) and low-ionization ones (e.g., C II, Al II, Si II, Fe II, Mg II, and O I). The column densities of these species drop as $\delta_{\text{Ly}\alpha}$ increases toward higher redshift.

We use a Cloudy model with gas cloud photoionized by an external UV background to help understand the results. In agreement with previous work (Pieri et al. 2010, 2014), we find that the column density distribution of the various metal species cannot be explained by a single gas component. In the model that we consider, the low-ionization species point to absorption

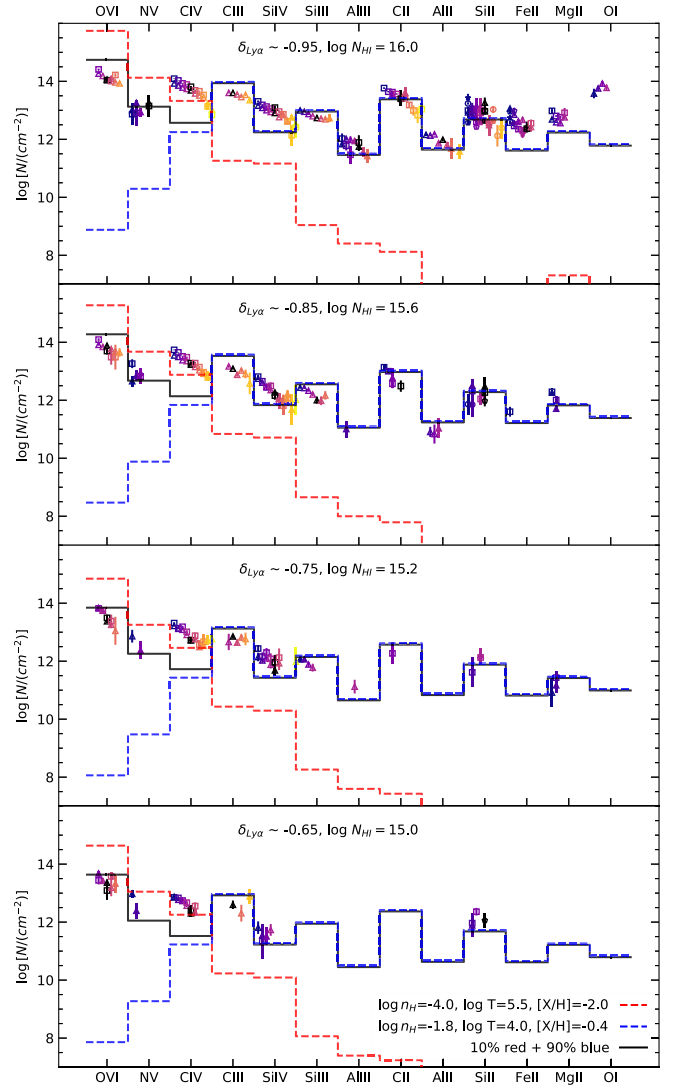


Figure 13. The column densities of 13 metal species as a function of $\delta_{\text{Ly}\alpha}$ and z_{abs} . The four panels are for different values of $\delta_{\text{Ly}\alpha}$. In each panel, the column density of each species is color-coded according to absorber redshift, low z_{abs} to high z_{abs} , from left to right within each species. The solid black line illustrates the model combination of a low-metallicity component (low n_{H} and high T ; red dashed line) and a high-metallicity component (high n_{H} and low T ; blue dashed line), and the difference in the models across the four panels is in the neutral hydrogen column density, as labeled.

systems of metallicity $[X/H] \sim -0.3$, with high density ($n_{\text{H}} \sim 10^{-1.5} \text{ cm}^{-3}$) and low temperature ($T \sim 10^4 \text{ K}$). The high-ionization species are likely low metallicity ($[X/H] \sim -2$), with either low density ($n_{\text{H}} \sim 10^{-4} \text{ cm}^{-3}$) or high temperature ($T > 10^{4.5} \text{ K}$). The two components likely correspond to cool gas in the CGM and gas in the IGM, respectively. For the selected absorbers to have $\delta_{\text{Ly}\alpha} \sim -0.95$, there is also a possibility of contamination from Ly α absorbers of higher neutral hydrogen column density.

The model explored in this work is for illustration purposes, highlighting the need for more than one component. In reality, for systems selected by $\delta_{\text{Ly}\alpha}$, there may be more than two or three components with a distribution of gas properties (n_{H} , T , abundance pattern, etc.). Large-volume hydrodynamic simulations of galaxy formation (e.g., Oppenheimer et al. 2012), with photoionization computation for metal-enriched gas (e.g., Oppenheimer & Schaye 2013), can help to guide the modeling

effort by identifying Ly α absorption systems in the synthetic spectra (matched with eBOSS resolution) and connecting them to the physical properties of the underlying gas (e.g., Turner et al. 2016). Meanwhile, high-resolution observations of the Ly α forest (e.g., Cowie et al. 1995; D’Odorico et al. 2016) can be used to study the components in the Ly α absorbers identified under eBOSS resolution and their associated metal properties. The complementary use of high-resolution spectra and stacked eBOSS spectra would provide an ideal approach for understanding the metal content in the Ly α forest.

This work is supported by the National Key R&D Program of China (grant No. 2018YFA0404503). L.Y. gratefully acknowledges the support of the China Scholarship Council (No. 201804910563) and the hospitality of the Department of Physics and Astronomy at the University of Utah during her visit. Z.Z. is supported by NSF grant AST-2007499. The support and resources from the Center for High Performance Computing at the University of Utah are gratefully acknowledged. G.R. acknowledges support from the National Research Foundation of Korea (NRF), through Grant No. 2020R1A2C1005655, funded by the Korean Ministry of Education, Science and Technology (MoEST).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.

SDSS-IV acknowledges support and resources from the Center for High Performance Computing at the University of Utah. The SDSS website is www.sdss.org.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration, including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian, the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

Appendix A Column Density Measurements for Metal Species and Model Trends

In Table 2, we list the column density measurements of all the metal species in the 10 absorber redshift bins over $2 < z_{\text{abs}} < 4$, and in the four $\delta_{\text{Ly}\alpha}$ bins at each redshift. The measurements result from Voigt fits with the stacked spectra. The vacuum wavelengths and oscillator strengths used in the

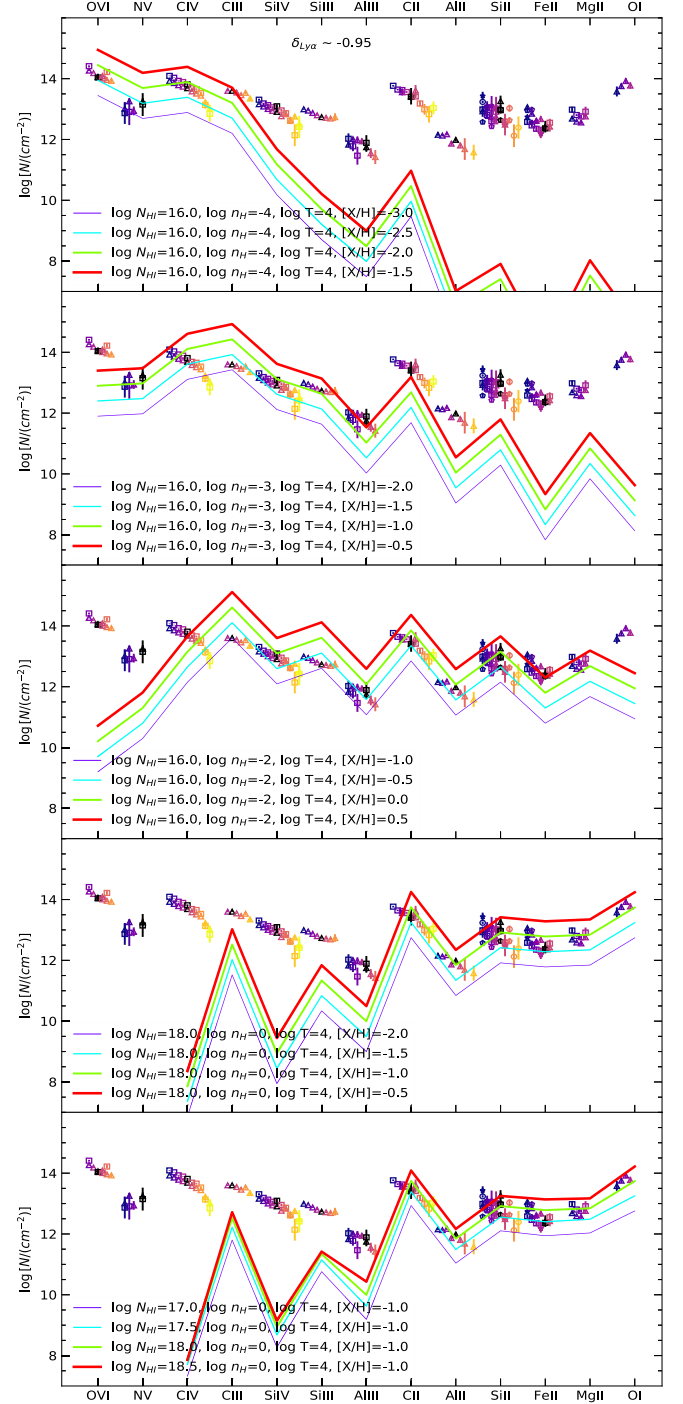


Figure 14. The dependence of the column densities of the metal species from the Cloudy model. The gas temperature is fixed at $T = 10^4$ K. Across the top four panels, the total hydrogen density n_H increases, and in each of these panels, different lines show the dependence on metallicity. Note that in the fourth panel, the neutral hydrogen column density N_{HI} changes from 10^{16} cm^{-2} to 10^{18} cm^{-2} . The bottom panel shows the dependence on N_{HI} , with other parameters fixed. For comparison, the measured column densities for $\delta_{\text{Ly}\alpha} \sim -0.95$ absorbers at different absorber redshifts (color-coded) are overlaid.

fits are taken from the Atomic Line List¹⁶ (Krogager 2018).

¹⁶ <https://github.com/jkrogager/VoigtFit/blob/master/VoigtFit/static/linelist.dat>

In Figure 14, we extend the exploration of the Cloudy models in the main text, overlaid with the measured column densities. The temperature is fixed at $T = 10^4$ K. From the top to bottom panels, the density n_{H} increases from 10^{-4} cm^{-3} to 1 cm^{-3} . Note that in the fourth panel, the neutral hydrogen column density N_{HI} changes from 10^{16} cm^{-2} to 10^{18} cm^{-2} . From the top four panels, we see that as n_{H} increases, the overall trend is for an increase in the column densities of low-ionization species and a decrease in those of high-ionization species, a result of the increase in the recombination rate. The dependence on metallicity changes the amplitude. The bottom panel shows the dependence on neutral hydrogen column density N_{HI} . At a higher N_{HI} , the species with the lowest ionization potential tend to have stronger increases in column density.

Appendix B Shadow Lines

In the stacked spectrum shown in Figure 3, we label the major metal lines associated with the selected Ly α absorber. Many shadow lines can be seen in the stacked spectrum. These shadow lines are caused by metal lines in Ly α absorbers with wavelengths close to Ly α —the Ly α absorption used to select the types of absorbers of interest (e.g., with $-1.0 < \delta_{\text{Ly}\alpha} < -0.9$) can be contaminated by metal absorptions from neighboring Ly α absorbers. Then, in the stacked spectrum of the specific absorbers of interest, the metal lines in the neighboring absorbers show up with shifted wavelengths, emerging as shadow lines.

As the shadow lines are not directly related to what we study in this work, they are not labeled in Figure 3, to avoid any confusion. In Figure 15, we provide a version of the stacked spectrum with the major shadow lines identified and labeled, to aid any possible investigations of the shadow lines.

First, in the second row from top, we label the five metal lines in the Ly α absorber with wavelengths close to Ly α , including Si II 1190 and 1193 Å (blue), Si III 1207 Å (cyan), Si II 1260 Å (magenta), and C II 1335 Å (red). If one of these lines with rest-frame wavelength $\lambda_{\text{contamination}}$ contaminates the selected Ly α absorption, a metal line with rest-frame wavelength λ_{line} associated with the absorber where the contaminating line originates will show up as a shadow line at the following wavelength:

$$\lambda_{\text{shadow line}} = \frac{\lambda_{\text{Ly}\alpha}}{\lambda_{\text{contamination}}} \lambda_{\text{line}}, \quad (\text{B1})$$

where $\lambda_{\text{Ly}\alpha} = 1215.67 \text{ Å}$ is the Ly α wavelength. We then use dashed (dotted) vertical lines of the corresponding color to indicate the shadowed metal lines (Lyman-series lines).

As an example, consider the shadow lines caused by Si II 1190 Å and 1193 Å contaminating Ly α absorption. Here, $\lambda_{\text{contamination}} = 1190 \text{ Å}$ or 1193 Å . The corresponding Si IV 1394 Å and 1403 Å shadow lines show up at wavelengths indicated by the four blue dashed vertical lines around 1420–1430 Å. Redward of 1550 Å, we see the corresponding shadow lines of Si II 1527 Å and the C IV 1548/1551 Å doublet. Around 2850 Å are the shadow lines of the Mg II 2796/2804 Å doublet. The shadow Ly β lines are indicated by the two blue dotted vertical lines around 1045 Å.

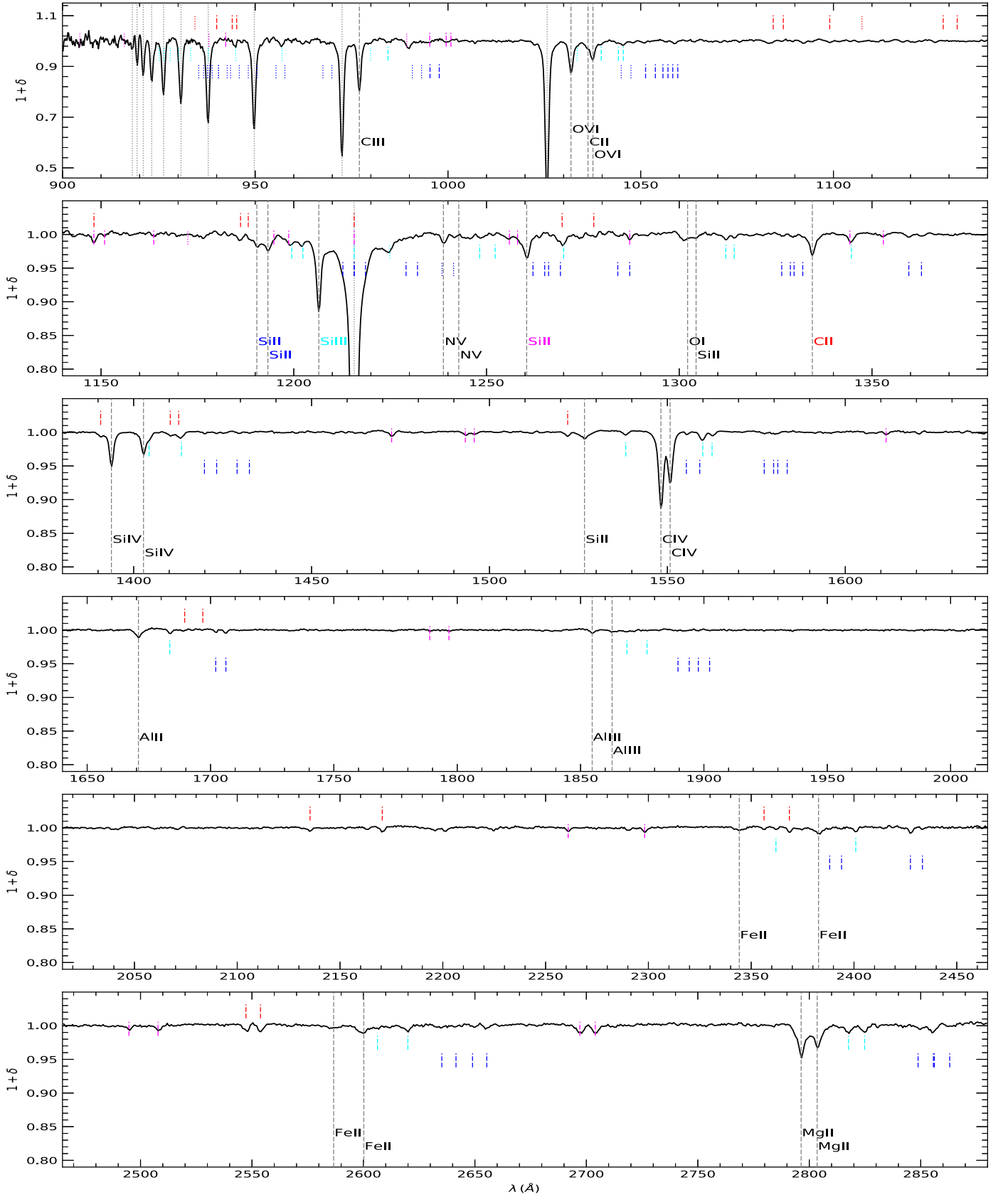


Figure 15. Similar to Figure 3, but with the positions of the major shadow lines indicated. Four species contaminating Ly α absorption selection are considered, including Si II 1190 $\text{\AA}$ and 1193 $\text{\AA}$ (blue), Si III 1207 $\text{\AA}$ (cyan), Si II 1304 $\text{\AA}$ (magenta), and C II 1335 $\text{\AA}$ (red). For each contaminating line with wavelength $\lambda_{\text{contamination}}$, the shadow of a given line λ_{line} (e.g., indicated by one of the long vertical lines) is shifted to a nearby position λ_{shadow} line at a fixed wavelength ratio, $\lambda_{\text{shadow}} = \lambda_{\text{Ly}\alpha} / \lambda_{\text{contamination}} \times \lambda_{\text{line}}$. These shadow lines are indicated by short dashed (dotted) vertical lines of the corresponding color for metals (the Lyman series). Vertical offsets are added to the vertical lines for different contaminating species.

Appendix C

Further Tests on the Contribution of High-Column Density Systems

We perform further tests to investigate the contribution of high-column density systems in the $-1 < \delta_{\text{Ly}\alpha} < -0.9$ absorbers.

By default, an absorber is selected to be in the sample by requiring the corresponding pixel with $-1 < \delta_{\text{Ly}\alpha} < -0.9$ to be the local minimum. With such a selection, the $\delta_{\text{Ly}\alpha}$ values of its neighboring pixels can either be below or above -0.9 . If the absorber has a higher column density, its neighboring pixels more likely have $\delta_{\text{Ly}\alpha}$ below -0.9 . We therefore impose an additional selection criterion, by requiring the $\delta_{\text{Ly}\alpha}$ values of the pixels to the left and the right of the above-selected pixel to be less than -0.9 . More exactly, we require $\delta_{\text{Ly}\alpha,i} < -0.9 + \sigma_\delta$, where σ_δ is the uncertainty in $\delta_{\text{Ly}\alpha}$, and i indicates the i th pixel next to the selected absorber, negative (positive) i for pixels to the left (right).

We create three subsamples—subsample-1, subsample-2, and subsample-3—by requiring 1, 2, and 3 left/right neighboring pixels to satisfy the additional selection above, respectively. These subsamples potentially have increasingly higher column densities. As illustrated in Figure 16, in the case of a single absorber under the SDSS resolution, the subsample-1 selection prefers to have systems of $\log(N_{\text{HI}}/\text{cm}^{-2}) > 18$ or those of $\log N_{\text{HI}} > 17$ with large b values, while the subsample-2 and subsample-3 selections would select absorbers of $\log N_{\text{HI}} > 18.5$. The absorbers in the three subsamples follow the same redshift distribution as the original sample, as shown in Figure 17.

Figure 18 compares the stacked spectra of the original $-1 < \delta_{\text{Ly}\alpha} < -0.9$ sample and the three subsamples. It is clear that the stacked Ly α line becomes increasingly broader, from the original sample to subsamples-1/2/3, and the sequence also displays an increasingly stronger wing-like feature. The Lyman-series lines appear to become broader and deeper. All of these are consistent with the expectation that the subsamples consist of absorbers with an increasing contribution from systems of high neutral hydrogen column density.

To further test the column density, in the gray shaded region in Figure 18 we normalize the stacked spectra to the mean continuum level within $960 \pm 1 \text{ \AA}$ in the absorber's rest frame, so that the flux decrements in the Lyman continuum ($< 912 \text{ \AA}$) can be revealed. These parts of the stacked spectra come from absorbers with $z_{\text{abs}} \gtrsim 2.95$. Similar to what is seen in the spectra of Lyman-limit systems (e.g., Figure 8 in Fumagalli et al. 2020), we see marked flux decrements blueward of $912 \text{ \AA}$. If the decrement were caused by a single absorber, the implied $\log N_{\text{HI}}$ would be about 16.9, 17.2, 17.5, and 17.5 for the original sample and subsamples-1/2/3, respectively, estimated using the decrement around $912 \text{ \AA}$. We note that a similar exercise in Pieri et al. (2014) leads to an estimate of $\log N_{\text{HI}} \sim 16.7$ (16.4) for absorbers with flux transmission $-0.05 < F < 0.05$ ($0.05 < F < 0.15$), a similar level to that inferred for our original sample. This is argued in Pieri et al. (2014) to be an upper limit, because of effects such as the large-scale excess in forest absorption associated with the selected absorbers. Anyway, the increasing decrement over these samples lends support to the scenario that high-column density systems contribute to the $-1 < \delta_{\text{Ly}\alpha} < -0.9$ absorbers selected in this work. A more realistic model, which is beyond the scope of this work, should consist of absorbers

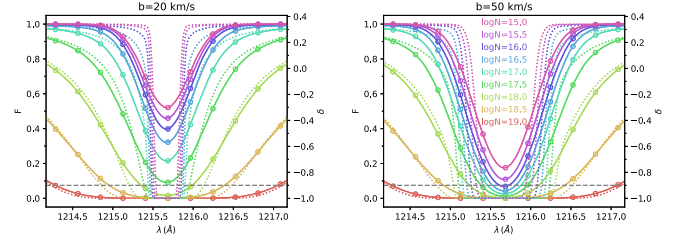


Figure 16. The dependence of the Ly α absorption profile on the neutral hydrogen column density $\log(N_{\text{HI}}/\text{cm}^{-2})$ and the Doppler parameter b for the case of a single Ly α absorber ($b = 20$ and 50 km s^{-1} in the left and right panels, respectively). In each panel, the dotted curves are the intrinsic profiles. The solid curves are those with eBOSS resolution ($R \sim 2000$), and the open circles indicate the eBOSS pixels ($\Delta \log \lambda \sim 10^{-4}$). The left y-axis is the normalized flux, and the right y-axis is the corresponding flux fluctuation $\delta = F/F_m - 1$, where the mean transmitted flux $F_m(z_{\text{abs}} = 2.7) \approx 0.745$ (Becker et al. 2013). The dashed horizontal line represents the $\delta = -0.9$ level for our selection of $z_{\text{abs}} \sim 2.7$ absorbers with $-1.0 < \delta_{\text{Ly}\alpha} < -0.9$.

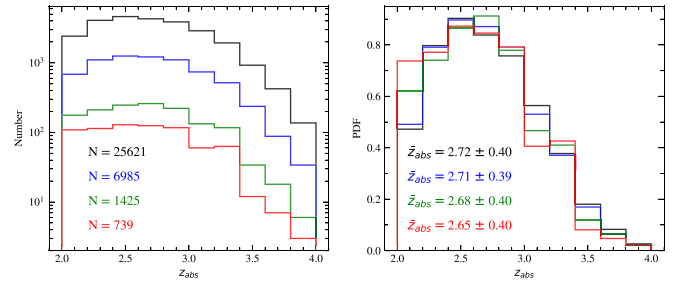


Figure 17. Redshift distributions for the Ly α absorbers in the original sample (black) with $-1.0 < \delta_{\text{Ly}\alpha} < -0.9$ and in the subsamples (blue/green/red), with 2/4/6 neighboring pixels around the selected absorbers in the original sample satisfying $\delta_{\text{Ly}\alpha} < -0.9 + \sigma_\delta$, with σ_δ being the uncertainty in $\delta_{\text{Ly}\alpha}$. The number distributions are shown in the left panel, and the normalized distributions (i.e., the probability distribution functions, or PDFs) are shown in the right panel.

with a distribution of column densities, and the decrement should be related to the average absorption $\langle \exp(-\tau) \rangle$ over such a distribution of absorbers. Without such a model, we cannot assess the fraction of Lyman-limit systems and the relative importance of the high-column density systems and the strongly clustered low-column density systems.

Next, we compare the metal species. Following the sequence of the original sample and subsamples-1/2/3, we see that the metal absorptions become increasingly stronger. In addition, a few species not clearly detected in the original sample start to emerge in the subsamples, which are labeled in Figure 18 in the gray text. For example, around $989 \text{ \AA}$ there are blended O I ($988.58/988.65/988.77 \text{ \AA}$), N III ($989.80 \text{ \AA}$), and Si II ($989.87 \text{ \AA}$), and around $1200 \text{ \AA}$ we see N I. These metal species are similar to those commonly seen in DLAs (e.g., Mas-Ribas et al. 2017).

The measured column densities of the metal species for the subsamples are shown in Figure 19, along with those for the original sample. From the original sample to subsamples-1/2/3, the column densities of low-ionization species (like O I and Fe II) generally increase by more than one order of magnitude, a much stronger change than the high-ionization species (e.g., O VI, C IV, and Si IV; with N V being an exception caused by blending with the Ly α shadow lines from the contamination of Si II $1190/1193 \text{ \AA}$ in Ly α , as shown in Figure 15). Such a relative change between the column densities of low- and high-ionization species is a typical feature in high-column density

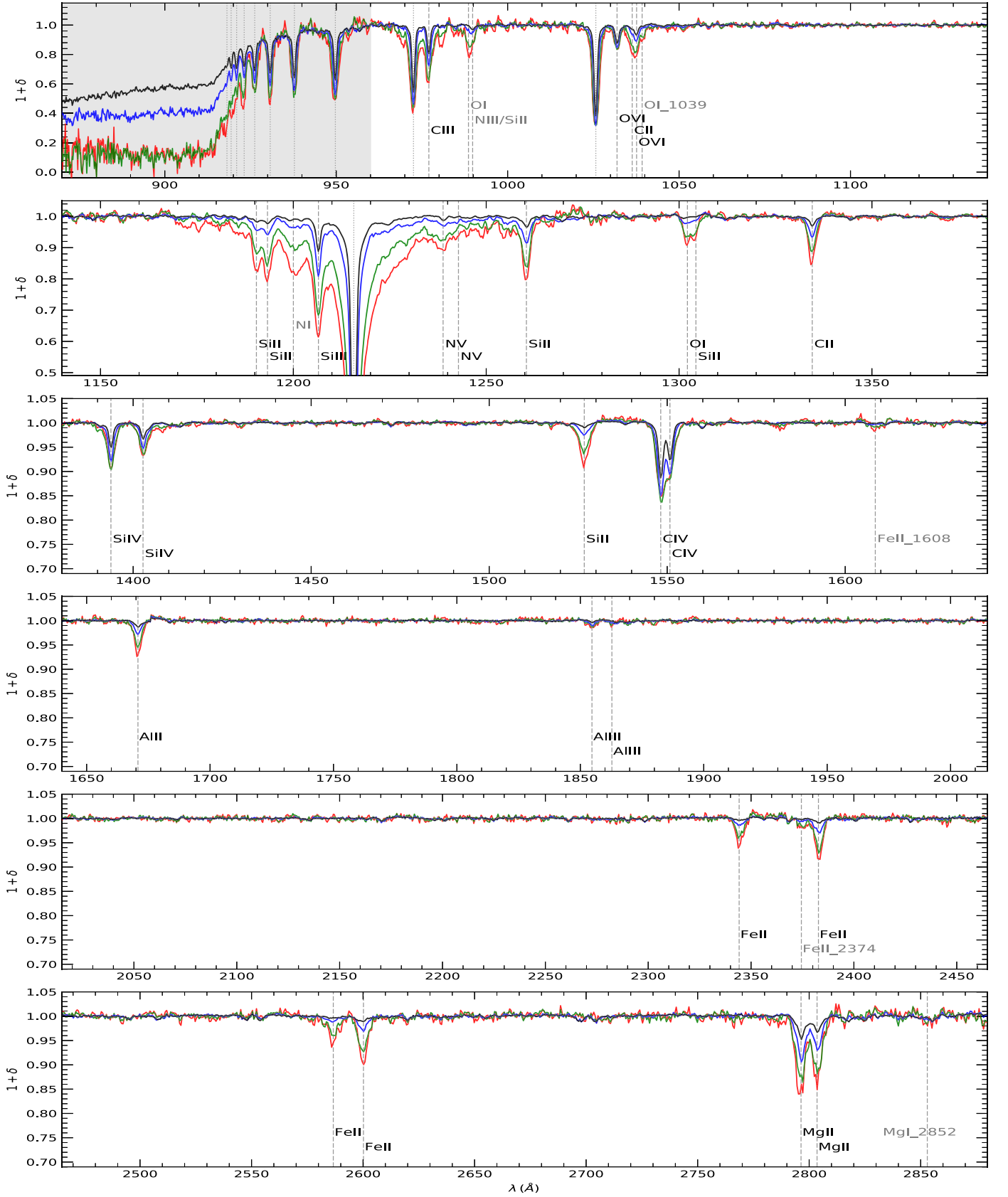


Figure 18. Similar to Figure 3, but with the stacked spectra of the original sample (black) and the three subsamples (blue/green/red). The original sample consists of absorbers with $-1.0 < \delta_{\text{Ly}\alpha} < -0.9$ in the redshift range of $2 < z_{\text{abs}} < 4$. The subsamples (blue/green/red) result from selecting stronger absorbers, by requiring the $\delta_{\text{Ly}\alpha}$ values of the neighboring 2/4/6 pixels around the absorbers in the original sample to lie below $-0.9 + \sigma_\delta$, where σ_δ is the uncertainty in $\delta_{\text{Ly}\alpha}$. In the shaded region, the flux is normalized to the mean continuum level within 960 ± 1 Å, to show the absorption near the Lyman break (912 Å).

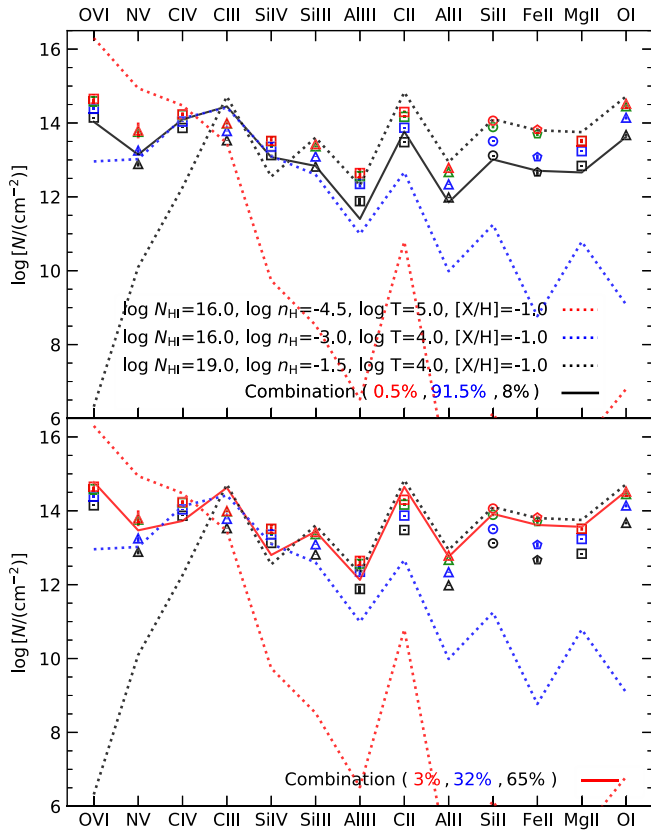


Figure 19. The column densities of the metal species in the original sample (black symbols) and in the subsamples (blue/green/red symbols). The colors of the symbols follow the same notation as in Figure 18. In each panel, three Cloudy model curves are shown—a component with a low hydrogen number density and a high temperature (red dotted line), a component with a medium hydrogen number density (blue dotted line), and a component with a high hydrogen number density and a neutral hydrogen column density (black dotted line). Different combinations (not model fits) of the three components are shown as the solid black line (top panel) and the solid red line (bottom panel), to illustrate the difference between the absorbers in the original sample and those in the subsample of the strongest absorption. See the text for details.

systems (see the bottom two panels of Figure 14), and is seen in the observational study with high-resolution quasar spectra (e.g., Lehner et al. 2021). Similar to Figure 12, in the top panel of Figure 19 we show a three-component Cloudy model (not a fit), to illustrate the role of adding a high-column density component in reproducing the trend in the low-ionization species. In the bottom panel, we boost the contribution of the high-column density system, which provides a reasonable match to the column densities of the low-ionization species in subsample-3, the one with presumably the highest contribution from high-column density systems. While the fractions in the illustration are not to be taken seriously, the numbers are self-consistent: the $\sim 65\%$ of high-column density systems in subsample-3 would make $\sim 2\%$ of the original sample, based on the ratio of the number of absorbers in the two samples (Figure 17), and this is treated as a lower limit of the contribution to the original sample.

To summarize, the relative changes in the column densities of low- and high-ionization metal species clearly favor the scenario that, from the original sample to subsamples-1/2/3, the average neutral hydrogen column density of the absorbers increases (with possible modulations by metallicity), instead of there being a pure effect of absorber clustering. The change in

the Lyman-series absorption profiles and the decrement at the Lyman limit also support such a scenario. Future investigations based on high-resolution observations and hydrodynamic simulations will help to verify and quantify the finding here.

ORCID iDs

Li Yang <https://orcid.org/0000-0001-5353-2957>

Zheng Zheng <https://orcid.org/0000-0003-1887-6732>

Hélien du Mas des Bourboux <https://orcid.org/0000-0001-8955-3573>

Kyle Dawson <https://orcid.org/0000-0002-0553-3805>

Donald P. Schneider <https://orcid.org/0000-0001-7240-7449>

References

- Ahumada, R., Allende Prieto, C., Almeida, A., et al. 2020, *ApJS*, **249**, 3
- Bautista, J. E., Bailey, S., Font-Ribera, A., et al. 2015, *JCAP*, **2015**, 060
- Bautista, J. E., Busca, N. G., Guy, J., et al. 2017, *A&A*, **603**, A12
- Becker, G. D., Hewett, P. C., Worseck, G., & Prochaska, J. X. 2013, *MNRAS*, **430**, 2067
- Bi, H., & Davidsen, A. F. 1997, *ApJ*, **479**, 523
- Blanton, M. R., Bershad, M. A., Abolfathi, B., et al. 2017, *AJ*, **154**, 28
- Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, *ApJL*, **437**, L9
- Chabanier, S., Etourneau, T., Le Goff, J.-M., et al. 2022, *ApJS*, **258**, 18
- Cooke, R., Pettini, M., Steidel, C. C., Rudie, G. C., & Jorgenson, R. A. 2011a, *MNRAS*, **412**, 1047
- Cooke, R., Pettini, M., Steidel, C. C., Rudie, G. C., & Nissen, P. E. 2011b, *MNRAS*, **417**, 1534
- Cowie, L. L., Songaila, A., Kim, T.-S., & Hu, E. M. 1995, *AJ*, **109**, 1522
- D’Odorico, V., Cristiani, S., Pomante, E., et al. 2016, *MNRAS*, **463**, 2690
- Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, *AJ*, **151**, 44
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, **145**, 10
- du Mas des Bourboux, H., Dawson, K. S., Busca, N. G., et al. 2019, *ApJ*, **878**, 47
- du Mas des Bourboux, H., Le Goff, J.-M., Blomqvist, M., et al. 2017, *A&A*, **608**, A130
- du Mas des Bourboux, H., Rich, J., Font-Ribera, A., et al. 2020, *ApJ*, **901**, 153
- du Mas des Bourboux, H., Rich, J., Font-Ribera, A., et al. 2021, Picca: Package for Igm Cosmological-Correlations Analyses, Astrophysics Source Code Library, record ascl:2106.018
- Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *AJ*, **142**, 72
- Ellison, S. L., Songaila, A., Schaye, J., & Pettini, M. 2000, *AJ*, **120**, 1175
- Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, *RMxAA*, **53**, 385
- Font-Ribera, A., Kirkby, D., Busca, N., et al. 2014, *JCAP*, **2014**, 027
- Fumagalli, M., Fotopoulou, S., & Thomson, L. 2020, *MNRAS*, **498**, 1951
- Fumagalli, M., O’Meara, J. M., & Prochaska, J. X. 2016, *MNRAS*, **455**, 4100
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
- Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, *AJ*, **131**, 2332
- Haardt, F., & Madau, P. 2012, *ApJ*, **746**, 125
- Iršič, V., Slosar, A., Bailey, S., et al. 2013, *JCAP*, **2013**, 016
- Kim, T. S., Carswell, R. F., & Ranquist, D. 2016, *MNRAS*, **456**, 3509
- Kobulnicky, H. A., & Skillman, E. D. 1998, *ApJ*, **497**, 601
- Krogager, J.-K. 2018, VoigtFit: Absorption Line Fitting for Voigt Profiles, Astrophysics Source Code Library, record ascl: 1811.016
- Lehner, N., Koppenhafer, C., O’Meara, J., et al. 2021, arXiv:2112.03304
- Lehner, N., O’Meara, J. M., Howk, J. C., Prochaska, J. X., & Fumagalli, M. 2016, *ApJ*, **833**, 283
- Lehner, N., Wotta, C. B., Howk, J. C., et al. 2019, *ApJ*, **887**, 5
- Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, *ApJS*, **250**, 8
- Mas-Ribas, L., Miralda-Escudé, J., Pérez-Ràfols, I., et al. 2017, *ApJ*, **846**, 4
- McDonald, P., Seljak, U., Burles, S., et al. 2006, *ApJS*, **163**, 80
- Myers, A. D., Palanque-Delabrouille, N., Prakash, A., et al. 2015, *ApJS*, **221**, 27
- Oppenheimer, B. D., Davé, R., Katz, N., Kollmeier, J. A., & Weinberg, D. H. 2012, *MNRAS*, **420**, 829
- Oppenheimer, B. D., & Schaye, J. 2013, *MNRAS*, **434**, 1043
- Parks, D., Prochaska, J. X., Dong, S., et al. 2018, *MNRAS*, **476**, 1151
- Pettini, M. 2004, in *Cosmochemistry, The Melting Pot of the Elements*, ed. Esteban C. et al. (Cambridge: Cambridge Univ. Press), 257
- Pettini, M., Rix, S. A., Steidel, C. C., et al. 2002, *ApJ*, **569**, 742

- Pieri, M. M., Frank, S., Weinberg, D. H., et al. 2010, [ApJL](#), **724**, L69
- Pieri, M. M., Mortonson, M. J., Frank, S., et al. 2014, [MNRAS](#), **441**, 1718
- Rahmati, A., Schaye, J., Crain, R. A., et al. 2016, [MNRAS](#), **459**, 310
- Rauch, M. 1998, [ARA&A](#), **36**, 267
- Schaye, J., Aguirre, A., Kim, T.-S., et al. 2003, [ApJ](#), **596**, 768
- Schneider, D. P., Richards, G. T., Hall, P. B., et al. 2010, [AJ](#), **139**, 2360
- Simcoe, R. A., Sargent, W. L. W., & Rauch, M. 2004, [ApJ](#), **606**, 92
- Slosar, A., Iršič, V., Kirkby, D., et al. 2013, [JCAP](#), **2013**, 026
- Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, [AJ](#), **146**, 32
- Turner, M. L., Schaye, J., Crain, R. A., Theuns, T., & Wendt, M. 2016, [MNRAS](#), **462**, 2440
- York, D. G., Adelman, J., Anderson, J. E., et al. 2000, [AJ](#), **120**, 1579