KODIAQ-Z: Metals and Baryons in the Cool Intergalactic and Circumgalactic Gas at $2.2 \lesssim z \lesssim 3.6$

Nicolas Lehner[1] , Claire Kopenhafer[2] , John M. O'Meara[3] , J. Christopher Howk[4,5] , Michele Fumagalli[4,5] ,
J. Xavier Prochaska[6,7] , Ayan Acharyya[8,9,10] , Brian W. O'Shea[11] , Molly S. Peeples[8,9] , Jason Tumlinson[8,9] , and
Cameron B. Hummels[12]

[1] Department of Physics and Astronomy, University of Notre Dame, Notre Dame, IN 46556, USA; nlehner@nd.edu
[2] Department of Physics and Astronomy, Department of Computational Mathematics, Science, and Engineering, Michigan State University, East Lansing, MI 48824, USA
[3] W. M. Keck Observatory, 65-1120 Mamalahoa Highway, Kamuela, HI 96743, USA
[4] Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy
[5] INAF—Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy
[6] UCO/Lick Observatory, Department of Astronomy & Astrophysics, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[7] Kavli Institute for the Physics and Mathematics of the Universe (WIP), 5-1-5 Kashiwanoha, Kashiwa, 277-8583, Japan
[8] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[9] Department of Physics & Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA
[10] Research School of Astronomy and Astrophysics, Australian National University, Weston Creek, ACT 2611, Australia
[11] Department of Computational Mathematics, Science, and Engineering, Department of Physics and Astronomy, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, USA
[12] Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

Abstract

We present the KODIAQ-Z survey aimed to characterize the cool, photoionized gas at $2.2 \lesssim z \lesssim 3.6$ in 202 H I-selected absorbers with $14.6 \leqslant \log N_{\rm H\,I} < 20$ that probe the interface between galaxies and the intergalactic medium (IGM). We find that gas with $14.6 \leqslant \log N_{\rm H\,I} < 20$ at $2.2 \lesssim z \lesssim 3.6$ can be metal-rich ($-1.6 \lesssim [{\rm X/H}] \lesssim -0.2$) as seen in damped Ly$\alpha$ absorbers (DLAs); it can also be very metal-poor ($[{\rm X/H}] < -2.4$) or even pristine ($[{\rm X/H}] < -3.8$), which is not observed in DLAs but is common in the IGM. For $16 < \log N_{\rm H\,I} < 20$ absorbers, the frequency of pristine absorbers is about 1%–10%, while for $14.6 \leqslant \log N_{\rm H\,I} \leqslant 16$ absorbers it is 10%–20%, similar to the diffuse IGM. Supersolar gas is extremely rare ($<1\%$) at these redshifts. The factor of several thousand spread from the lowest to highest metallicities and large metallicity variations (a factor of a few to >100) between absorbers separated by less than $\Delta v < 500$ km s^{-1} imply that the metals are poorly mixed in $14.6 \leqslant \log N_{\rm H\,I} < 20$ gas. We show that these photoionized absorbers contribute to about 14% of the cosmic baryons and 45% of the cosmic metals at $2.2 \lesssim z \lesssim 3.6$. We find that the mean metallicity increases with $N_{\rm H\,I}$, consistent with what is found in $z < 1$ gas. The metallicity of gas in this column density regime has increased by a factor ~ 8 from $2.2 \lesssim z \lesssim 3.6$ to $z < 1$, but the contribution of the $14.6 \leqslant \log N_{\rm H\,I} < 19$ absorbers to the total metal budget of the universe at $z < 1$ is a quarter of that at $2.2 \lesssim z \lesssim 3.6$. We show that FOGGIE cosmological zoom-in simulations have a similar evolution of $[{\rm X/H}]$ with $N_{\rm H\,I}$, which is not observed in lower-resolution simulations. In these simulations, very metal-poor absorbers with $[{\rm X/H}] < -2.4$ at $z \sim 2$–3 are tracers of inflows, while higher-metallicity absorbers are a mixture of inflows and outflows.

Supporting material: figure sets, machine-readable tables

1. Introduction

The intergalactic medium (IGM) and the circumgalactic medium (CGM) are massive reservoirs of baryons (e.g., Fukugita et al. 1998; McQuinn 2016; Macquart et al. 2020) and major fuel sources for star formation in galaxies (e.g., Tumlinson et al. 2017). They play a vital role in the formation and evolution of galactic and large-scale structures in the universe. The empirical and theoretical characterizations of the CGM and IGM before, during, and after the peak of cosmic star formation ($z \sim 2$) are therefore of prime importance to

understanding their role in the evolution of galaxies and large-scale structures. Gathering data that can probe both the low- and high-redshift universe is critical for robustly constraining the evolution of these cosmic structures. This is also the best route to fully test always-improving cosmological simulations. Modern numerical simulations are now reaching new milestones in fidelity. Notably, a new generation of simulations boasts unprecedentedly high spatial resolution, even in the more diffuse regions of the simulation boxes (e.g., Hummels et al. 2019; Mandelker et al. 2019; Peeples et al. 2019; Suresh et al. 2019; van de Voort et al. 2019). This allows for far more realistic modeling of the CGM and IGM and their interrelationships with galaxies.

Over the past few years, our group has engaged in several surveys at both low and high redshifts to target absorbers in the

H I column density range $15 \lesssim \log N_{\rm H\,I} \lesssim 19$. This column density interval spans a range of physical environments from the diffuse IGM (Lyα forest, hereafter LYAF) to the denser portions of the CGM and the edges of galaxy disks. Following Lehner et al. (2018), we describe absorbers as strong LYAF systems (SLFSs; $\log N_{\rm H\,I} = 14.5$–16.2), as partial Lyman systems (pLLSs; $\log N_{\rm H\,I} = 16.2$–17.2), or as Lyman limit systems (LLSs; $\log N_{\rm H\,I} = 17.2$–19). Below this range is the LYAF ($\log N_{\rm H\,I} \lesssim 14.5$), and above it are the super-LLSs (SLLSs; $\log N_{\rm H\,I} = 19.0$–20.3, aka the subdamped Lyα absorbers) and the damped Lyα absorbers (DLAs; $\log N_{\rm H\,I} \geqslant 20.3$). With overdensities between the diffuse IGM (probed by the LYAF) and galaxies (probed by DLAs), the SLFSs, pLLSs, and LLSs spanning $15 \lesssim \log N_{\rm H\,I} \lesssim 19$ should be at the heart of the exchange of matter between galaxies, their CGM, and the diffuse IGM.

One of the main goals of our H I absorption surveys is to provide a census of the chemical enrichment of the absorbers with $14.6 \lesssim \log N_{\rm H\,I} \lesssim 20$ over cosmic time (e.g., the COS CGM Compendium (CCC) at low redshift, Lehner et al. 2018, 2019; Wotta et al. 2019; see also Lehner et al. 2013; Wotta et al. 2016; the HD-LLS survey, Prochaska et al. 2015; Fumagalli et al. 2016b; the KODIAQ and KODIAQ-Z surveys, Lehner et al. 2014, 2016; this paper). The metallicity of the absorbers is a direct measure of their metal enrichment and a key diagnostic of their origins. The metallicity provides direct information on how efficient (or not) galaxies are at enriching their immediate surroundings and beyond, as well as on the level of metal mixing in the diffuse regions of the universe. The selection of our surveys provides a relatively unbiased way to obtain a census of the metallicity, since selecting on hydrogen is sensitive to both very low and very high metallicities.

While pLLSs and LLSs have been found often in the CGM of galaxies at both low and high redshift (e.g., Lehner et al. 2009, 2013; Ribaudo et al. 2011; Kacprzak et al. 2012; Prochaska et al. 2017b; Cooper et al. 2021; Berg et al. 2022), there is growing evidence that at least some of the pLLSs and LLSs with very low metallicities may probe the denser IGM (very low metallicity being $[{\rm X/H}] \lesssim -1.4$ at $z \lesssim 1$, Berg et al. 2022; and $[{\rm X/H}] \lesssim -3$ at $2 \lesssim z \lesssim 3.5$, Fumagalli et al. 2011a, 2016a; Crighton et al. 2016; Lofthouse et al. 2020).[13] Recent high-resolution zoom-in cosmological simulations also show that metal-poor ($[{\rm X/H}] < -3$) LLS-like absorbers can be found in the IGM far from any galaxies (Mandelker et al. 2019, 2021). Determining the metallicity distributions of the SLFSs, pLLSs, and LLSs at low and high z directly informs us on the metallicity enrichment (or lack thereof) relative to galaxies and their inner regions probed by DLAs and SLLSs. This technique can also assess the amount of pristine gas in these overdense regions of the universe and how it compares to the more diffuse IGM probed by LYAF absorbers (Fumagalli et al. 2011a; Mandelker et al. 2019).

In previous papers, our group presented a sample of 261 $z < 1$ absorbers with $15 \lesssim \log N_{\rm H\,I} \lesssim 19$ (Lehner et al. 2018, 2019; Wotta et al. 2019, hereafter CCCI, CCCII, CCCIII, respectively). This sample includes an unexpectedly large fraction of metal-poor absorbers, many of which have metallicities $[{\rm X/H}] < -2$, implying little chemical enrichment since $z \sim 2$–3. The finding that metal-poor absorbers with

$[{\rm X/H}] \leqslant -1.4$ represents about half of the population of $z \lesssim 1$ absorbers with $15 \lesssim \log N_{\rm H\,I} \lesssim 19$ was unanticipated. This is because neither the LYAF (due, in retrospect, to a lack of sensitivity) nor the SLLSs/DLAs hinted at the presence and importance of such low-metallicity gas at low redshift. Although not fully appreciated at the time, there was, however, already some evidence of very low metallicity gas at low redshift based on individual studies of LLSs (Cooksey et al. 2008; Ribaudo et al. 2011; see also Tripp et al. 2005 for a rare example of a primitive SLLS at $z \sim 0.06$). CCC also reveals that, at $z < 1$, there is no evidence of pristine gas, implying that $>99\%$ of the gas with $15 \lesssim \log N_{\rm H\,I} \lesssim 19$ has been polluted by $z \lesssim 1$, even if it is only at a very low level. Low-redshift photoionized gas with $15 \lesssim \log N_{\rm H\,I} \lesssim 19$ has a wide range of metallicities observed ($-3 \lesssim [{\rm X/H}] \lesssim +0.4$) and can show large metallicity variations (factor up to 25) over small redshift/velocity separation ($\Delta v < 300$ km s^{-1}), implying that these absorbers sample a highly inhomogeneous medium probed by these absorbers. The CCC results raise the following questions: Are these properties unique to $z \lesssim 1$? How do they evolve from earlier times, across the peak of cosmic star formation at $z \sim 2.5$?

To address these questions, we present here the Keck Database of Ionized Absorbers toward Quasars (KODIAQ-Z) survey, a survey designed to identify H I absorbers with $\log N_{\rm H\,I} \gtrsim 14.5$ at $z > 2$ in Keck HIRES spectra available in the KODIAQ database (O'Meara et al. 2015, 2017) and determine their metallicities. Due to the higher sensitivity of the Keck HIRES data, compared to the Hubble Space Telescope (HST) Cosmic Origins Spectrograph (COS) data, we can probe low-$N_{\rm H\,I}$ absorbers at high redshift and still be sensitive to low-metallicity gas. In fact, due to the new sensitivity gained with 8–10 m ground-based telescopes, early studies on the metal enrichment at high redshift largely focused mostly on IGM absorbers with $\log N_{\rm H\,I} \lesssim 14.5$ (e.g., Songaila 1998; Ellison et al. 2000; Schaye et al. 2003; Aguirre et al. 2004; Simcoe et al. 2004). Our sample of H I-selected absorbers consists of 155 SLFSs, 24 pLLSs, 16 LLSs, and 7 SLLSs at $2.2 \lesssim z \lesssim 3.6$, totaling 202 absorbers, 195 of them with $14.5 \lesssim \log N_{\rm H\,I} < 19$. On account of the nature of the Keck HIRES data, the sample of pLLSs and LLSs has only marginally increased by a factor 1.3 compared to our earlier pilot survey (Lehner et al. 2016). However, in this paper we undertake a systematic search and characterization of H I-selected SLFSs.[14] In our analysis, we include the H I-selected absorbers from the HD-LLS survey (Prochaska et al. 2015; Fumagalli et al. 2016b), which adds another 46 LLSs, increasing the total sample to 241 absorbers with $14.5 \lesssim \log N_{\rm H\,I} \lesssim 19$. The HD-LLS survey is directly complementary to KODIAQ-Z because most of the HD-LLS absorbers have $\log N_{\rm H\,I} \gtrsim 17.5$ (including 73 SLLSs), whereas KODIAQ-Z largely samples gas with lower $N_{\rm H\,I}$.

The main goals of the present paper are to determine the metallicity distributions of these absorbers, assess how their metallicity changes with $N_{\rm H\,I}$ and z, and provide a robust limit on the amount of pristine gas at high redshift. We also present and discuss the physical properties of these absorbers and use these results to estimate the metal and baryon budgets of the cool photoionized CGM at $z \simeq 2.2$–3.6. Our paper is organized as follows. In Section 2, we describe the survey design of

---

[13] For the metallicity, we use conventional square bracket notation $[{\rm X/H}] \equiv \log N_{\rm X}/N_{\rm H} - \log({\rm X/H})_{\odot}$, where X is an α-element, unless otherwise stated.

[14] Simcoe et al. (2004) surveyed the metallicity of the IGM at $2.2 \lesssim z \lesssim 2.8$. While they focused on absorbers with $\log N_{\rm H\,I} < 14.5$, they also included about 30–40 SLFSs as part of their survey.

KODIAQ-Z, the search for the H I-selected absorbers, and the sample of SLFSs, pLLSs, and LLSs. In Section 3, we present the methods to derive the column densities of the metals and H I for each absorber, while in Section 4 we present basic statistical properties of our sample (z, $N_{\rm H\,I}$, and the Doppler parameter b). In Section 5, we present ionization modeling of the KODIAQ-Z absorbers and assumptions for estimating their metallicity. Our main results on the metallicities of the H I-selected absorbers are presented in Section 6. In Section 7, we derive the physical properties of the absorbers, including their density ($n_{\rm H}$), total column density ($N_{\rm H}$), temperature, and linear scale ($l \equiv N_{\rm H}/n_{\rm H}$), and we model the neutral fraction as a function of $N_{\rm H\,I}$ in the range $14.6 \leqslant \log N_{\rm H\,I} \leqslant 20$. In Section 8, we estimate the cosmic metal and baryon budgets of these absorbers. In Section 9, we compare the KODIAQ-Z results to the properties of the gas in cosmological zoom-in simulations from the Figuring Out Gas & Galaxies in Enzo (FOGGIE) project (Peeples et al. 2019; Corlies et al. 2020) and use these simulations to gain insight into the connection of the $14.5 \lesssim \log N_{\rm H\,I} \lesssim 20$ gas depending on its metallicity with galaxies. In Section 10, we discuss the results, and in Section 11, we summarize our main conclusions.

2. Survey Design, Database, and Observations

2.1. Sample Selection Criteria

Our survey is based on a search for H I absorption in the KODIAQ DR2 database (O'Meara et al. 2017) of 300 QSO spectra observed with HIRES. The QSOs useful for our survey lie at $2 \lesssim z_{\rm em} \lesssim 4.5$, limiting the sample size to 235 QSOs. We also searched in the spectra of higher-redshift ($z_{\rm em} > 4.5$) QSOs, but the LYAF is too dense to allow us to reliably determine the H I and metal-line column densities at these redshifts. The lower redshift threshold $z \gtrsim 2$ is required to have wavelength coverage of at least Lyα and Lyβ. In each of the QSO spectra, we search for absorbers with H I column density $\log N_{\rm H\,I} \gtrsim 14.5$. This contrasts with our previous KODIAQ-Z survey (Lehner et al. 2016), where our search for H I absorbers was limited to absorbers with $\log N_{\rm H\,I} \gtrsim 16.2$ (to match our initial lower z survey, Lehner et al. 2013) and was by no means a systematic and complete examination of the KODIAQ DR1 database (O'Meara et al. 2015).

A major difference (besides the redshift probed) with the low-redshift survey CCC (CCCI) is that all the HIRES spectra are flux normalized for the co-addition of the individual exposures and echelle orders (O'Meara et al. 2015, 2017). This makes determination of large-scale flux decrements (e.g., at the Lyman limit break or damping wings of DLAs) unreliable. Therefore, to estimate the H I column density, we can rely only on the absorption observed in the Lyman series transitions. This limits especially the sample size of absorbers with $17.2 \lesssim \log N_{\rm H\,I} \lesssim 18$ because it requires having enough weak Lyman transitions that are not too contaminated to robustly derive $N_{\rm H\,I}$ (see Section 3.2). For absorbers with $\log N_{\rm H\,I} \gtrsim 18$, damping wings start to be observable in the Lyα absorption, and therefore this transition can be used to derive $N_{\rm H\,I}$ depending on the level of contamination of the absorption from the LYAF. On the other hand, while CCC is limited to absorbers with $\log N_{\rm H\,I} \gtrsim 15.3$, due to the need to be sensitive to metallicities $[{\rm X/H}] \lesssim -1$, the typically higher signal-to-noise ratios (S/Ns) and higher spectral resolution of the Keck HIRES spectra allow us to probe lower $N_{\rm H\,I}$, down to the LYAF threshold of about $\log N_{\rm H\,I} \simeq 14.5$, and still be sensitive to metallicity with $[{\rm X/H}] \lesssim -2$. Although with very high S/N Keck HIRES spectra a strong limit ($[{\rm X/H}] \lesssim -2$) on the metallicities can be placed down to $\log N_{\rm H\,I} \simeq 13.6$ (e.g., Simcoe et al. 2004), our lower limit on $N_{\rm H\,I}$ was selected to correspond to the division between the diffuse IGM and denser IGM/diffuse CGM. Following Schaye (2001b), a 14.5 dex limit on the H I column density corresponds to an overdensity of $\delta \gtrsim 5$ at $z \sim 2.8$ (about the average redshift of the absorbers in our sample; see Section 4.1).

2.2. Search for H I Absorbers in the KODIAQ Database

To identify the strong H I absorption, we developed an automated search tool to help identify the Lyman series transitions from 1215 Å (Lyα) down to 917 Å. In Figure 1, the colored profiles show single-component Voigt models for absorbers with a Doppler parameter $b = 25$ km s^{-1} and $\log N_{\rm H\,I} = 15, 16, 17, 18$. That b-value is selected based on previous profile fitting (PF) of strong H I absorbers at similar redshifts (e.g., Lehner et al. 2016); as we will see below, that choice is close a posteriori to the mean b-value in our sample. We use these H I models to set the following criteria for our automatic search: (1) Lyα must have flux $F_\lambda \leqslant 0$ over $\gtrsim 0.4$ Å in contiguous pixels; (2) at the average velocity/redshift (v_c) of the identified Lyα, the flux in the Lyβ absorption must also reach $F_\lambda \leqslant 0$; and (3) at v_c, we also require that there is substantial absorption in at least Lyγ (and Lyδ if there is wavelength coverage of that transition). Additional conditions were added to ensure that when a DLA or an SLLS is present the search program would not continue to look for additional absorbers near the redshift of the DLA or SLLS. For this search, we therefore require for each QSO wavelength coverage of at least Lyα, Lyβ, and Lyγ. However, since the damping wings of Lyα can be used to determine $N_{\rm H\,I}$ if $\log N_{\rm H\,I} \gtrsim 18.5$, we also undertook a separate search that relies solely on Lyα where the flux $f_\lambda \simeq 0$ over a wavelength width of $\gtrsim 2$ Å. A posteriori, only a few of these very strong H I absorbers ended up in our sample, due to the fact that the KODIAQ spectra are normalized prior to co-addition, and the normalization was often not reliable enough to model the damping wings for these absorbers, especially the strong SLLSs and DLAs (DLAs not being an issue, since they are not part of our survey).

For each candidate absorber, our search tool created a stack of H I transitions as shown on Figure 1. Over 3000 candidates were created, many of these being false positives owing to the large contamination from the LYAF and other metal lines. That number is still small enough that a quick look at the stack of H I profiles could easily weed out most of the false positives, using the Voigt models as a guide. This vetting led the sample to be reduced by more than a third, to 949 candidate absorbers, the redshift distributions of which we show in Figure 2. In this figure, we also show the redshift distribution of the QSOs. About 83% of the candidate absorbers are at $2.2 \leqslant z_{\rm abs} \leqslant 3.6$; candidate absorbers at $z_{\rm abs} \leqslant 2$ are detected solely based on Lyα.

To assemble our final sample for metallicity determinations, we considered both the ability to estimate $N_{\rm H\,I}$ and the coverage of metal ions (in order to be able to estimate the metallicity). For each candidate absorber, we therefore created another stack of velocity profiles with Lyα, Lyβ, Lyδ, and some key metal transitions used in the estimation of the metallicity (including Si II, Si III, Si IV, C II, C III, C IV, Al II, Al III, Fe II, O I, and
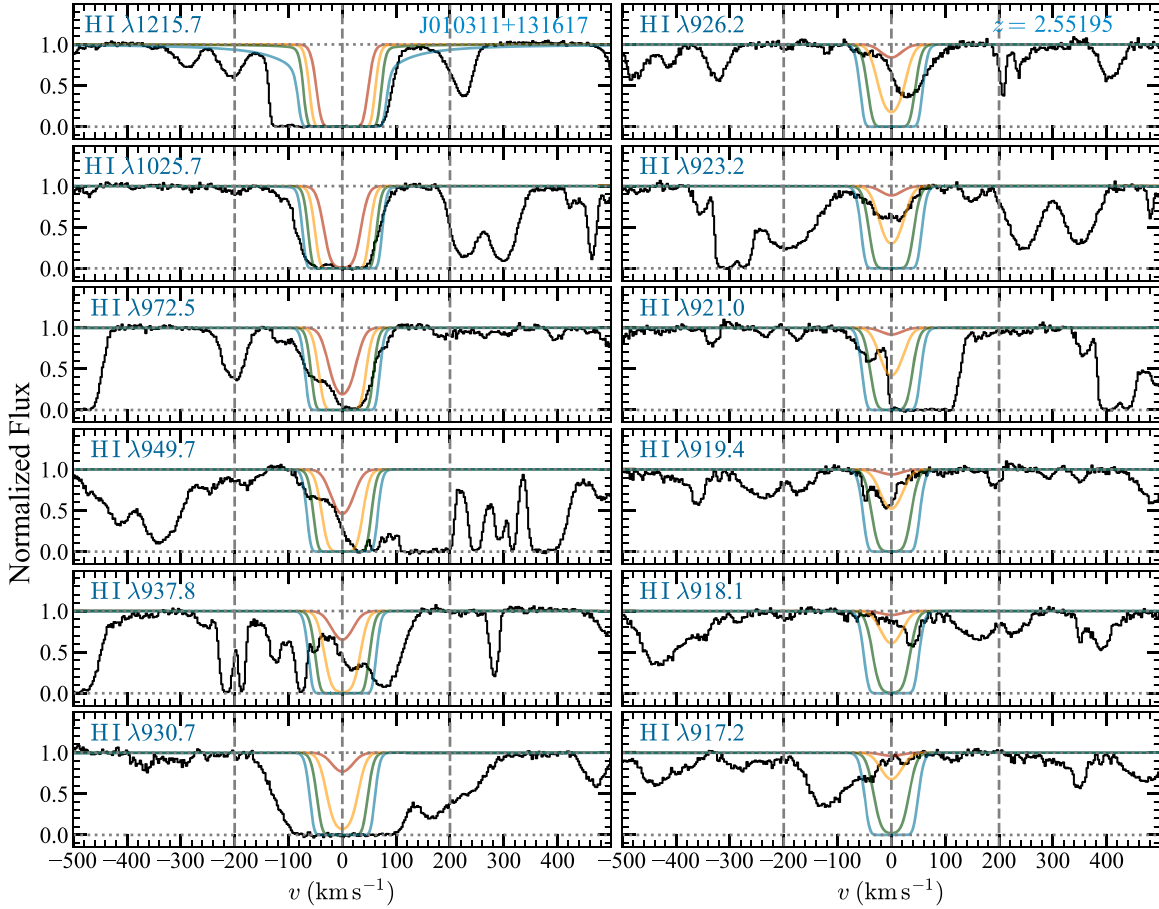
Figure 1. Example of an absorber identified by our automatic searching tool. For each identified candidate absorber, a similar figure is created that shows the normalized spectra (in black) of the H I transitions from the 1215 to 917 Å transitions as a function of the velocity in the redshift rest frame (indicated in the top right panel). The colored profiles are Voigt profiles for an absorber with $b = 25$ km s^{-1} and $\log N_{\rm H\,I} = 15, 16, 17, 18$. These are used to help assess the level of contamination in each H I transition. Ultimately, for this candidate absorber, two absorbers were identified at $z = 2.55125$ and 2.55216 (corresponding to $v \simeq -50$ and $+17$ km s^{-1} in this figure).
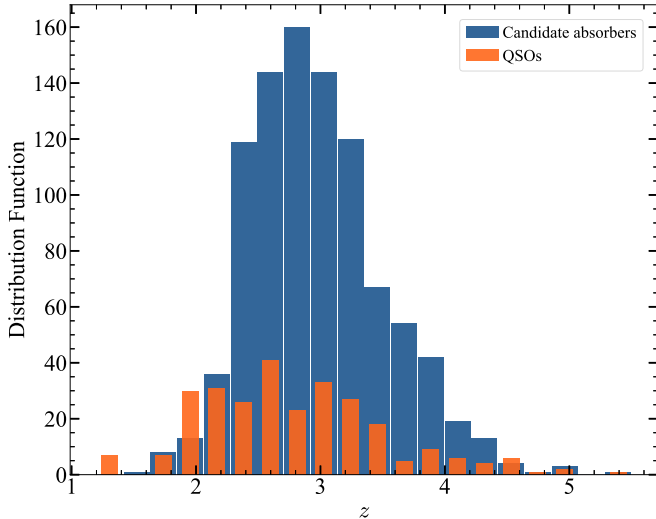

Figure 2. Distribution of the redshifts for the QSOs and the *candidate* absorbers from the automatic search in KODIAQ DR2.

O VI) to assess whether there is enough metal-ion coverage to determine the metallicity. After all the measurements are done (detailed in Section 3), the final sample size consists of 202 absorbers, i.e., ≈7% of the candidates from the initial automatic search were absorbers where we could place a

robust estimate on $N_{\rm H\,I}$ and metal-line column densities (including strong upper limits). In Table 1, we summarize the sample of QSOs with information on their redshifts, the wavelength coverage, S/N, and spectral resolution of the Keck HIRES observations. The final KODIAQ-Z sample includes 155 SLFSs, 24 pLLSs, 16 LLSs, and 7 SLLSs. We note that the redshift of each absorber was determined from the peak optical depth of the strongest H I component in our automatic search process. In most cases (89%), the velocities in the rest frame of the absorber redshift are $|v| < 20$ km s^{-1}, but in some cases they can be somewhat larger.

2.3. Comparison Samples

One of the goals of our survey is to determine how the metallicity changes with $N_{\rm H\,I}$ at similar redshifts and between low and high z. For the low redshift, we use results from our CCC survey, including the compilation of H I absorbers at $\log N_{\rm H\,I} \geqslant 19$ (see CCCI; CCCII; CCCIII and references therein). At $z \gtrsim 2$, we use two surveys: (1) the HD-LLS survey (Prochaska et al. 2015; Fumagalli et al. 2016b), a survey of absorbers with $17.3 \leqslant \log N_{\rm H\,I} < 20.3$ at $z = 2$–4.39, ∼95% of the absorbers having H I column densities in the range $18 \leqslant \log N_{\rm H\,I} < 20.3$ (which is directly complementary to KODIAQ-Z; HD-LLS also follows the same H I selection); and (2) the DLA survey by Rafelski et al. (2012), hereafter the

Table 1
Sample Summary

Target	z_{em}	λ_{min} (Å)	λ_{max} (Å)	S/N$_{935}$	S/N$_{1170}$	S/N$_{1450}$	Decker	Obs. Flag
J002127−020333	2.596	3365	6198	10	12	10	C1	0
J003501−091817	2.418	3123	5979	7	18	19	C5	0
J004530−261709	3.440	3972	8542	10	17	15	C1	0
J005700+143737	2.643	3366	6198	5	9	8	C1	0
J010311+131617	2.721	3122	8566	54	76	84	C5	1
J010806+163550	2.652	3047	8158	34	101	29	C5	1
J010925−210257	3.226	3304	6179	19	19	14	C5	0
J011150+140141	2.470	3132	6030	4	7	6	C5	0
J012156+144823	2.870	3243	7512	30	87	79	C5	1
J014516−094517	2.729	3022	7522	77	112	100	C5	1

Note. The median S/N is calculated by S/N $= 1.0/\sigma$ per pixel within ± 5 Å (where σ is the error on the normalized flux) of three rest wavelengths. A -1 in one of the S/N columns indicates that there is no wavelength coverage where the S/N is estimated. The pixels are a constant 2.1 km s^{-1} for pre-2004 observations or 2.6 km s^{-1} for post-2004 observations and for co-additions of observations mixing pre- and post-2004 data (see O'Meara et al. 2017). The observation flag "1" in the last column indicates if there was more than one run or a single run ("0") (see Table 2 in O'Meara et al. 2017). The decker setup corresponds to the following resolution: C1: $R = 45,000$; C5: $R = 34,000$.

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

R12-DLA survey. Our use of these surveys was mostly dictated by the ease of access to the data while working in the early stage of KODIAQ-Z, including the metal and H I column densities for the HD-LLS. We note that the Magellan DLA survey by Jorgenson et al. (2013) has essentially the same metallicity distribution as R12-DLA in our redshift of interest $2.2 \lesssim z \lesssim 3.6$ (see below). The recent XQ-100 survey of SLLSs by Berg et al. (2021) also shows the same metallicity distributions as in HD-LLS.

For the R12-DLA, we adopt the metallicities from the α-elements (or, if not available, Zn). For the HD-LLS, we reestimate all the metallicities using the same extreme ultraviolet background (EUVB) adopted in CCC, in order to avoid mismatches in the comparisons of the metallicities—although, as we will show in Section 5.3, the systematic error arising from the EUVB in the photoionization modeling of the absorbers is much smaller (in fact, it is negligible) at $z > 2$ than at lower redshift, especially for the strong LLSs and SLLSs that are part of the HD-LLS survey.

3. Estimation of the Column Densities

Our analysis follows closely that undertaken in CCC to estimate the column densities of H I and metal ions. Throughout we assume the atomic parameters (central wavelengths and f-values) listed in Morton (2003). For each absorber, we first attempt to estimate the H I column density since that step is necessary to estimate the metallicity of the absorbers. For H I, we use a combination of different methods—apparent optical depth (AOD) method (Savage & Sembach 1991), curve-of-growth (COG) method, and Voigt PF method—while for the metal ions, we use solely the AOD method. We employ several methods to estimate the H I column densities, on account of the contamination, blending, and saturation effects being more severe than for the metal lines (see Figure 1).

3.1. Metal Column Densities

Although we only estimate the column densities of the metal species once the H I transitions have been modeled (and hence the candidate absorber has become a selected absorber), we first describe the metal lines, as we use only the AOD method to extract the properties of the absorption. In this method, the absorption profiles are converted into apparent optical depth per unit velocity, $\tau_a(v) = \ln[F_c(v)/F_{obs}(v)]$, where $F_c(v)$ and $F_{obs}(v)$ are the modeled continuum and observed fluxes as a function of velocity, respectively. The AOD, $\tau_a(v)$, is related to the apparent column density per unit velocity, $N_a(v)$, through the relation $N_a(v) = 3.768 \times 10^{14} \tau_a(v)/(f\lambda(\text{Å}))$ cm^{-2}(km s^{-1})$^{-1}$, where f is the oscillator strength of the transition and λ is the wavelength in Å. The total column density is obtained by integrating the profile over the predefined velocity interval, $N = \int_{v_{min}}^{v_{max}} N_a(v)dv$, where $[v_{min}, v_{max}]$ are the boundaries of the absorption. We computed the average line centroids through the first moment of the AOD $v_a = \int v\tau_a(v)dv/\int \tau_a(v)dv$ km s^{-1}. The velocity range over which the line was integrated was determined from weak H I transitions and metal line.

We emphasize that, as much as possible, the integration range $[v_{min}, v_{max}]$ corresponds to a single absorbing H I complex, as shown, e.g., in Figure 3, but this does not necessarily mean that the metal lines consist solely of a single component (often they are, but not always). For the absorbers considered in this survey, the main H I components can typically be separated when the central velocities between two H I components are at least separated by $|\Delta v| \gtrsim 40$ km s^{-1} and $\log N_{H I} \lesssim 18$. As we will see below, the mean Doppler parameter of the H I components is around 27 km s^{-1}, corresponding to a temperature of $T < 4 \times 10^4$ K, typical for photoionized gas. This means that the thermal broadening is important and hence the H I lines are broader than the metal lines, i.e., there are H I transitions that can appear as a single component (at the observed spectral resolution) despite the metal ions showing multiple components over a velocity width of ~ 50 km s^{-1}. In that case, we integrate the metal ions over a similar velocity width to the H I when that information is available from weak H I transitions; if the latter is not available (e.g., SLLSs where weak H I transitions are not covered, which is the case for $\lesssim 5\%$ of the sample), we guide our analysis using the low ions or O I (if it is detected), which are better tracers of denser regions probed by the strongest H I absorbers. In the latter case, we integrate over the entire profiles of the low ions that may have a single or several components; this velocity
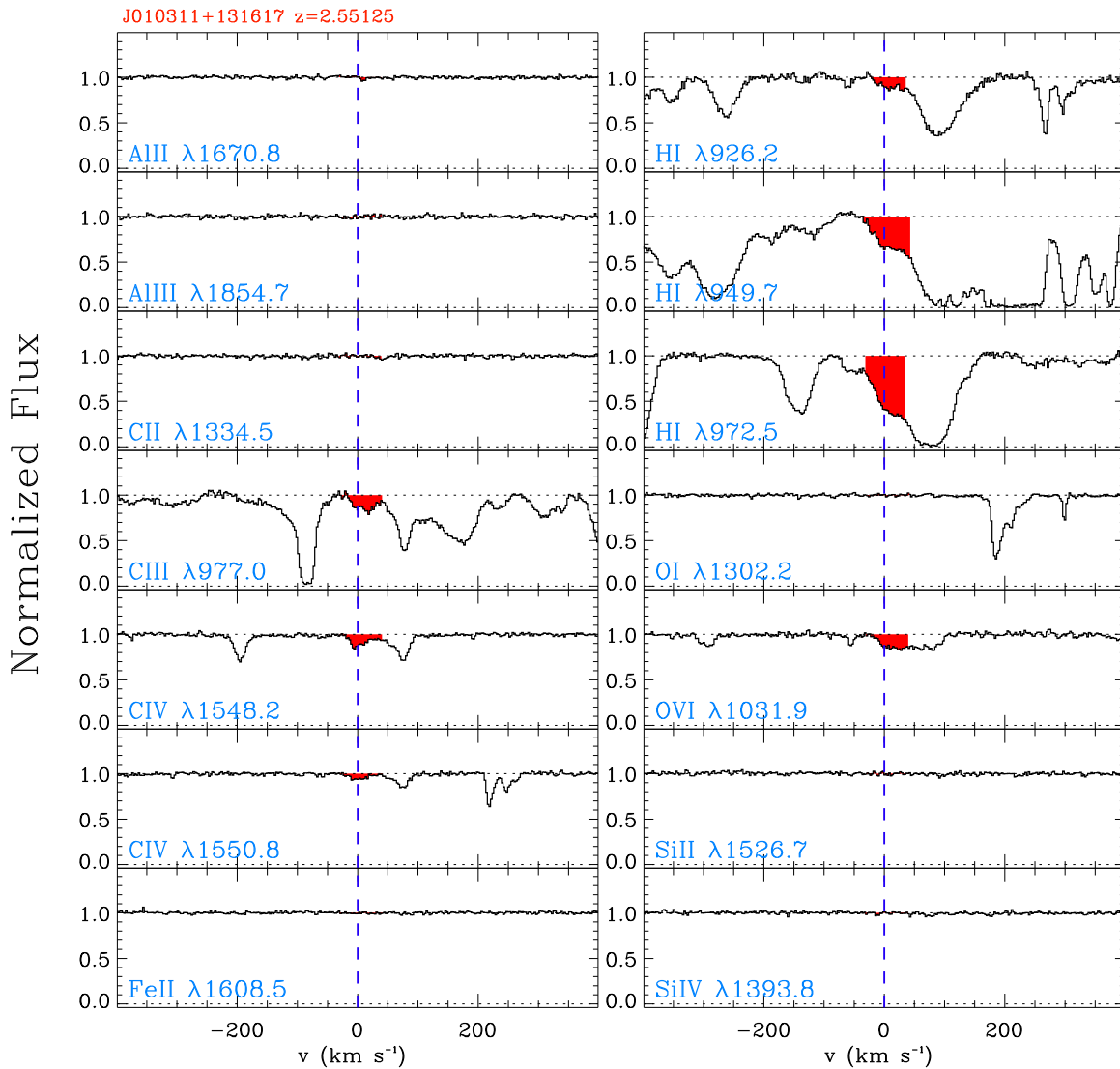
Figure 3. Example of normalized spectra against the rest-frame velocity of the absorber at $z = 2.55125$ toward J010311+131617 that show detection of H I, C III, C IV, and O VI (the integration ranges of the profiles are shown in red). Another absorber separated by 65 km s^{-1} is found at $z = 2.55216$; in this case, the column densities can be estimated reliably in each component/absorber (see also Figure 4).

integration range dictates the integration range of the higher ions, even though they may extend well beyond the absorption seen in the low ions (see, e.g., Lehner et al. 2014).

As illustrated in Figures 1 and 3, even deep in the LYAF, the original continuum placement (see O'Meara et al. 2015, 2017) is satisfactory and can be used in many cases to estimate the physical parameters of the absorption lines accurately. We therefore took full advantage of the Keck HIRES spectra already being normalized to use the AOD method in a fully automated way. For each identified absorber, we typically made two iterations to estimate the physical parameters of the lines (v_a, N_a, and equivalent width W_λ). In the first iteration, the velocity range is estimated based on the stack of velocity profiles used to preview the data. Our program created figures that show for each transition the normalized and apparent column density profiles (such as Figure 3 in CCCI), allowing us to refine the integration velocity range and determine whether the continuum needed any local refinement. In the second iteration, the refined velocity range is applied to the original or renormalized spectra. If we determine that the continuum model needs to be reestimated, we either use the

automated Legendre polynomial fitting described in CCCI or manually select the continuum region to be fitted by low-order Legendre polynomials.

We integrate the equivalent widths using the same [v_{min}, v_{max}] integration range adopted for the $N_a(v)$ profiles. The main use of the equivalent width for the metal lines in our survey is to determine whether the absorption is detected at the $\geqslant 2\sigma$ level. If it is not, we quote a 2σ upper limit on the column density, which is simply defined as twice the 1σ error derived for the column density assuming that the absorption line lies on the linear part of the curve of growth. The 1σ error is determined by integrating the spectrum over a velocity interval similar to that of a detected ion or very weak H I transition, or otherwise ± 20 km s^{-1}, based on the typical smallest velocity intervals in other absorbers with weak detection of metals.

When an absorption feature is detected at $\geqslant 2\sigma$ significance, the next step is to assess whether there are any contamination or saturation issues. The main ions detected in this work are C II, C III, C IV, Si II, Si III, Si IV, and O VI (see Figure 3 and the Appendix). We also consider other ions, such as Fe II, Fe III, Al II, and Al III, or atoms such as O I, but in most cases these

are not detected (Figure 3 and the Appendix). For C IV λλ1548, 1550, Si IV λλ1393, 1402, O VI λλ1031, 1037, and S II λλ1193, 1260, 1304, and 1526, these are either doublets or ions with multiple transitions, and both contamination and saturation can be readily checked by comparing their $N_a(v)$ profiles. Ions with $\lambda \lesssim 1215$ Å are more often likely to be contaminated than ions with transitions at longer wavelengths, due to the dense LYAF at these redshifts. For example, for the absorber shown in Figure 3, the C IV transitions are not contaminated, but the weak transition of O VI is. In contrast, the weakest H I absorber at about -55 km s^{-1} seen in Figure 3 (which is not analyzed as part of this work because its H I column density is $\log N_{H I} < 14.5$) shows absorption in both transitions of each doublet of C IV and O VI. The C II ion has two transitions at 1334 and 1036 Å, and we always check the latter transition if $\lambda1334$ is detected; more often than not, it is, however, contaminated by the LYAF.

On the other hand, ions such as C III λ977 and Si III λ1206 are in the LYAF and have only a single transition. In this case, our strategy for assessing whether a given transition is contaminated by another unrelated absorber is based on a combination of visual inspection and comparison of the relevant quantities (e.g., mean velocity and apparent optical depth) with other transitions. Strong contamination can often be diagnosed visually, by comparing both the central velocities of the peak optical depth and the shape of the velocity structure of the absorption profiles across transitions from the same or similar ions.[15] The mean velocities of the absorption profiles estimated from the integration of the absorption profiles can often point to strong contamination if the mean velocity from one transition is significantly different from others that are aligned. In the case of obvious contamination, the absorption feature is discarded. The only exceptions are as follows. If the contamination is very mild, only occurring in the wing of a relatively strong absorption, we slightly change the velocity integration range to avoid the contaminated portion for that absorption feature. In cases where we are not sure that the absorption arises solely from the understudied metal ion (e.g., detection of C III but no coverage or detection of C IV or other ions), but the feature could very likely be produced by that ion based on the same velocity as that of H I and narrow absorption (as expected from a metal, low ion), we treat the column density from that ion as an upper limit in our ionization modeling.[16]

For most of the metal ions in our survey, saturation is not a major issue, due to the lower column H I densities than those of DLAs and even SLLSs (there are only seven SLLSs in our sample and no DLA), the lower metallicities on average than at $z < 1$, and the high spectral resolution of the Keck data. To determine whether there is saturation, we proceed similarly to how we did for the contamination assessment. Where there is no obvious sign of contamination and the absorption from the metal ion reaches zero flux, the absorption is automatically marked as a lower limit. The comparison of the AOD profiles for doublet ions or ions with multiple transitions can readily help determine whether there is saturation if the profiles do not match at the peak optical depth and if the apparent column

density of the stronger transitions is systematically smaller than that of the weaker transitions.

For atomic or ionic transitions that have $\Delta \log(\lambda f) \simeq 0.3$ and the difference in apparent column densities of the weak and strong transitions is $\Delta(\log N_a) \equiv \log N_a^{weak} - \log N_a^{strong} \leqslant 0.13$ dex (see Wotta et al. 2016), we are able to correct for the mild saturation using the method described in Savage & Sembach (1991) and Wotta et al. (2016). In this case, we report the apparent column densities for each transition and then the adopted column density, corrected for saturation. In cases where we were not able to correct for saturation, we report the column density as a lower limit determined by the AOD method.[17]

For multiple transitions where there are no saturation or contamination issues, we give velocities and column densities as the weighted average. All the column densities, velocity integration ranges, average velocities, and redshifts for the metal ions and H I (see Section 3.2) are summarized in Table 2. We also provide all the results in a machine-readable format (see the Appendix).

3.2. Determination of the H I Properties

To derive the metallicity, we can rely in all the cases on having column densities for several metal ions (detections and/or nondetections). However, to have information on the amount of hydrogen atoms, we rely solely on H I. The advantage with H I is that we can access, in most cases, the entire Lyman series from 1215 Å all the way down to 913 Å (depending on wavelength coverage, redshift of the absorber, presence of an SLLS/DLA along the sight line, and strength of the absorption), spanning a difference in $f\lambda$ of over 2700. However, the entire Lyman series is in the LYAF (and Lyβ forest), and therefore many transitions can be and often are heavily contaminated, necessitating a close look at all the H I absorption features to assess the level of contamination. Therefore, our first task is to systematically look simultaneously at several H I profiles (from 1215 to 915 Å) to assess whether there is enough information (i.e., portion of the profiles of H I transitions that are uncontaminated) to derive $N_{H I}$. If it is judged that the contamination can be dealt with, we proceed by estimating the apparent column densities from the AOD profiles (and equivalent widths) and creating a mask of the H I profiles that are contaminated (i.e., flagging the pixels that are clearly contaminated), a mask that is then used in the PF (see below).

Due to the LYAF contamination, each transition is handled manually, and we often also reassess the continuum that can be more uncertain than in the redder part by modeling it with low-order Legendre polynomials. As mentioned above, we treated as much as possible an absorber as a single H I component. For the example shown in Figure 3, there are two components separated by 65 km s^{-1}, which correspond to two absorbers at $z_{abs} = 2.55125$ and 2.55216. We therefore integrate the AOD profiles over each component (for the absorber at $z_{abs} = 2.55125$, this corresponds to the velocity interval of about $[-30, 40]$ km s^{-1}). For absorbers with $\log N_{H I} \gtrsim 18$, this is not always possible, especially if there is no coverage of the weaker Lyman series transitions (for these absorbers, the damping wings of Lyα and Lyβ can be used to determine $N_{H I}$;

---

[15] Of course, if the investigated ion is not in the same gas phase as the ion that is uncontaminated (e.g., O VI in some cases), these criteria cannot be used.

[16] This upper limit can be differentiated from a nondetection upper limit, as the error bars of the former are not the standard nondetection upper limit errors +0.18, −0.30 dex.

[17] Contamination and saturation can be a priori both present at the same time and deceitful. However, this is unlikely to occur frequently over the same pixels, and therefore each can be disentangled reliably in most cases.

Table 2
Summary of the Column Densities of the KODIAQ-Z Absorbers

Target	z_{abs}	Ion	v_1	v_2	v	σ_v	$\log N$	$\sigma_{\log N}^1$	$\sigma_{\log N}^2$	Detection Flag	Reliability Flag	Unique ID
			(km s^{-1})	(km s^{-1})	(km s^{-1})		[cm^{-2}]					
J002127−020333	2.54918	Al II	−50.0	50.0	999.0	999.0	11.625	0.176	0.301	−1	3	1
J002127−020333	2.54918	C II	−50.0	50.0	999.0	999.0	13.379	0.176	0.301	−1	3	1
J002127−020333	2.54918	C IV	−50.0	50.0	0.4	3.9	13.175	0.058	0.066	0	2	1
J002127−020333	2.54918	Fe II	−50.0	50.0	999.0	999.0	13.228	0.176	0.301	−1	3	1
J002127−020333	2.54918	H I	−50.0	50.0	−2.4	2.7	15.680	0.019	0.022	0	2	1
J002127−020333	2.54918	N I	−50.0	50.0	999.0	999.0	13.329	0.176	0.301	−1	3	1
J002127−020333	2.54918	N V	−50.0	50.0	999.0	999.0	12.658	0.176	0.301	−1	3	1
J002127−020333	2.54918	O I	−50.0	50.0	999.0	999.0	13.860	0.176	0.301	−1	3	1
J002127−020333	2.54918	O VI	−50.0	50.0	−10.0	5.2	13.418	0.068	0.081	0	1	1
J002127−020333	2.54918	S II	−50.0	50.0	999.0	999.0	13.674	0.176	0.301	−1	3	1

Note. A 999.0 value corresponds to a nondetection, and the upper limit on the column density is quoted at the 2σ level. The detection flag has the following definition: 0 = detection; − 1 = upper limit (nondetection); − 2 = detection but a lower limit due to saturation of the line. Depending on whether a single transition or several were detected, a reliability flag is assigned as follows: 1 = most reliable, i.e., results based on several transitions or a nondetection (a nondetection is always reliable because it is estimated in an uncontaminated region of the spectrum); 2 = reliable, i.e., results based on at least two transitions, but where only one transition is detected at the 2σ level and the upper limits agree with that detection; 3 = less reliable, i.e., results based only on a single transition or several saturated transitions.

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

see below). However, these absorbers are rare in our sample, $\lesssim 5\%$ (see below). For each transition, the AOD results are checked for contamination and/or saturation. For the transitions where there is no sign of contamination and/or saturation, we use a weighted average of the integrated apparent column densities, which provides the $N_{H\,I}$ estimates from the AOD method.

The equivalent widths, estimated along with the apparent column densities and average velocities, are used for the COG analysis, which employs a χ^2 minimization and χ^2 error derivation approach outlined by Sembach & Savage (1992). A single-component COG is assumed. Our procedure solves for $\log N$ and b independently in estimating the errors. We start with all the H I transitions for which we have estimated the equivalent widths and that do not appear a priori to be contaminated based on the AOD analysis. We check the results from the COG model and then remove only the transitions clearly departing from the model. We then run our COG program again with the new set of transitions, to obtain the final estimate of $N_{H\,I}$ with the COG method.

Using the normalized and masked H I spectra, we also perform a Voigt profile fit of the H I transitions using an in-house PF program described in Lehner et al. (2011, 2014, 2018), which was updated and refined from Fitzpatrick & Spitzer (1997). A major difference from the previous methods is that the model profiles are convolved with the HIRES instrumental line-spread function, assumed to be Gaussian. The three parameters for each component i (typically $i = 1$, but in some cases $i = 2$ or 3, or in rare cases $i > 3$)—column density (N_i), Doppler parameter (b_i), and central velocity (v_i)—are input as initial guesses and then subsequently varied to minimize the χ^2 of the fit to all the fitted H I transitions. As for the COG, we start with all the H I transitions that do not appear a priori contaminated and iterated, removing any transitions that a posteriori appear contaminated (or sometimes revisiting the masked pixels). In Figure 4, we show an example of PF that consists of a two-component fit corresponding to the absorbers at $z_{abs} = 2.55125$ and 2.55216. As can be seen in each panel of Figure 4, there is some evidence of contamination especially in the component at 65 km s^{-1} ($z_{abs} = 2.55216$), but

even in that component there is enough information to derive robustly $N_{H\,I}$ when using the different H I transitions simultaneously.

In general, there is a good agreement between these different methods to estimate $N_{H\,I}$ for absorbers with $14.5 \lesssim \log N_{H\,I} \lesssim 17.2$. The three methods explore different parameters and different transitions, allowing us to robustly estimate $N_{H\,I}$. When multiple approaches are valid, we simply average the values from the different methods and propagate the errors accordingly. For systems with $\log N_{H\,I} \gtrsim 17.2$, the PF results are most of the times solely adopted as a consequence of the need to fit either the Lyα damping wings for absorbers with $\log N_{H\,I} \gtrsim 19$ or a combination of weaker and stronger H I transitions in order to correctly model the absorption that depends more strongly on both N and and b in this regime. The eight systems with $17.5 \leqslant \log N_{H\,I} \leqslant 18.65$ require a more hands-on approach owing to having mostly saturated absorption in all the H I transitions and only weak damping features in Lyα. These were all analyzed as part of Lehner et al. (2014) and Lehner et al. (2016), and we refer the reader to these papers for a full description, noting that very conservative errors were adopted with a minimum error on the $N_{H\,I}$ for any LLS of $\sigma_{\log N} = 0.15$ using this methodology. However, the metal lines associated with these absorbers were all reanalyzed as part of this work.

The adopted $N_{H\,I}$ results are summarized in Table 2, and we also report in Table 3 the PF results, i.e., the velocities, Doppler parameters, and column densities, as well as errors associated with these quantities. In that table, the listed redshift corresponds to the adopted redshift (see above), and the third column indicates the component number for each absorber.

4. Empirical Properties of the KODIAQ-Z Survey

Before describing the ionization modeling of the absorbers, it is helpful to summarize the basic properties of the KODIAQ-Z survey, in particular in terms of $N_{H\,I}$ and redshift distributions, as we will compare our sample of absorbers to other surveys (such as surveys of stronger H I absorbers at similar redshift or absorbers with similar $N_{H\,I}$ but at lower redshift). As we show below, the Doppler parameters also

Table 3
Summary of the H I Fit Results

Target	z_{abs}	Comp.	v (km s^{-1})	σ_v (km s^{-1})	b (km s^{-1})	σ_b (km s^{-1})	log N [cm^{-2}]	$\sigma_{\log N}$
J002127−020333	2.54918	1	−5.6	0.2	32.5	0.2	15.672	0.016
J003501−091817	2.35342	1	9.1	0.6	29.1	0.7	15.560	0.030
J004530−261709	2.81907	1	−25.8	1.1	74.0	0.5	19.664	0.010
J004530−261709	3.24823	1	−2.1	0.2	25.1	0.2	15.081	0.016
J004530−261709	3.36517	1	−18.6	0.4	53.2	0.6	15.538	0.016
J004530−261709	3.36517	2	5.3	1.4	16.0	1.3	15.368	0.055
J005700+143737	2.54721	1	−30.9	1.4	55.8	1.9	15.062	0.018
J010311+131617	2.54303	1	−0.5	0.1	27.9	0.1	15.629	0.010
J010311+131617	2.55125	1	11.5	0.6	31.9	0.5	14.900	0.010
J010311+131617	2.55125	2	76.9	0.2	28.9	0.3	15.502	0.010

Note. The redshift between parentheses in the second column indicates the actual redshift of the absorber associated with that velocity component.

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

motive the use of photoionization as the main ionization mechanism in the ionization modeling to determine the metallicity of the absorbers.

4.1. H I Column Density and Redshift Distributions

In Figure 5, we show the distributions of the absorber redshifts and H I column densities, as well as the scatter plot between these two quantities. There are 202 absorbers in KODIAQ-Z, and their full $N_{H I}$ and z intervals are $14.43 \leqslant \log N_{H I} \leqslant 19.87$ and $2.23 \leqslant z_{abs} \leqslant 4.35$. However, 99% of the absorbers have $14.57 \leqslant \log N_{H I} \leqslant 19.87$ and $2.23 \leqslant z_{abs} \leqslant 3.57$ with means and standard deviations $\langle \log N_{H I} \rangle = 15.8 \pm 1.1$ (the median being $\log N_{H I} = 15.5$) and $\langle z_{abs} \rangle = 2.79 \pm 0.31$ (the median being $z_{abs} = 2.75$), respectively. There is an evident peak in the redshift distribution at $2.4 \lesssim z_{abs} \lesssim 2.6$ where there are about 65 absorbers; otherwise, between $z_{abs} \simeq 2.6$ and 3.35 the z-distribution is relatively flat, with about 30–35 absorbers in each 0.2 redshift bin. This peak in the redshift distribution corresponds to redshifts where the optimal conditions are assembled to derive $N_{H I}$ (access to the entire Lyman series transitions and smaller contamination from the LYAF): as z decreases below 2.4, the number of H I transitions diminishes rapidly owing to the available wavelength coverage; as z increases above 3.6, the level of contamination from the LYAF increases (and indeed at $3.5 \lesssim z_{abs} \lesssim 4.4$, only for five absorbers were we able to robustly derive $N_{H I}$). As alluded to before, we use two other main surveys at $z > 2$ for comparison with KODIAQ-Z: the HD-LLS survey (Prochaska et al. 2015; Fumagalli et al. 2016b) and the R12-DLA survey (Rafelski et al. 2012). Based on the redshift distribution of KODIAQ-Z, we limit these surveys to the redshift range $2.2 \leqslant z \lesssim 3.6$ when comparing the absorbers as a function of $N_{H I}$ over the same redshift interval.

Due to the selection of the absorbers, the $N_{H I}$ distribution should follow to some extent the column density distribution of H I (Prochaska et al. 2014), which is why the $N_{H I}$ distribution is not flat. From approximately 10^{15} to $10^{16.5}$ cm^{-2}, we see in Figure 5 the expected drop in the number of absorbers as $N_{H I}$ increases. At $\log N_{H I} < 15$, our sample is not complete, as there is a sharp drop in the number of absorbers. This is because as $N_{H I}$ decreases, we rely more and more on high-S/N spectra in order to be sensitive to low metallicities, significantly reducing the sample. For absorbers with $16.5 \lesssim \log N_{H I} \lesssim 17.5$, we have to depend more on weak Lyman series transitions to estimate
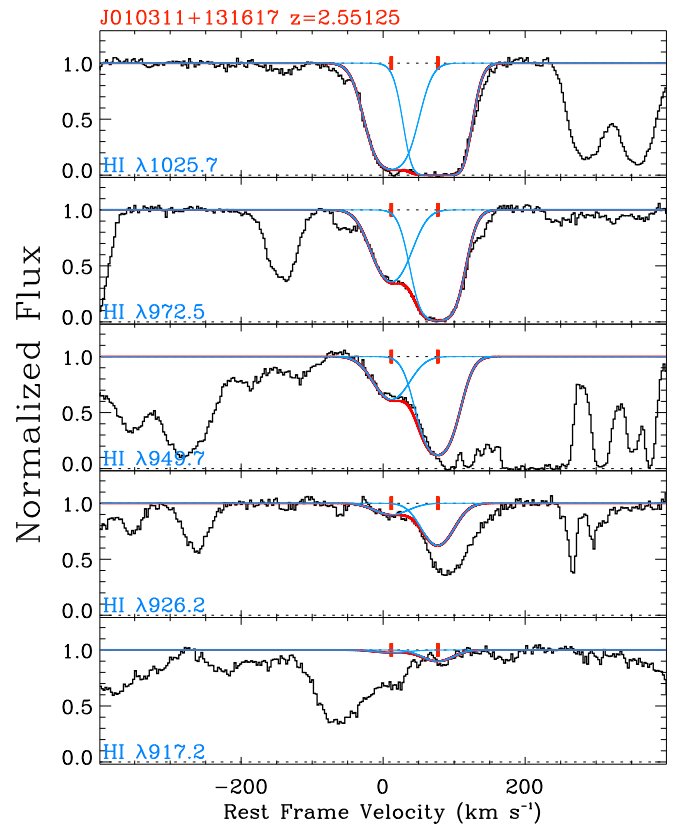

Figure 4. Example of the same absorber shown in Figure 3 where we show the Voigt profile fit (in red the composite profile fit, and in blue each individual component) to the individual transitions of H I (black spectra). Despite some contamination in each transition, there are enough uncontaminated regions in each profile to estimate robustly the H I column density in each component.

$N_{H I}$, limiting the number of absorbers especially as z increases beyond $z \gtrsim 3.2$. As described in Section 3.2, absorbers with $17.5 < \log N_{H I} \lesssim 18.65$ are rare because they are essentially on the flat part of the COG, dramatically restricting the number of absorbers where we can estimate $N_{H I}$. And indeed although absorbers with $18.7 \lesssim \log N_{H I} \lesssim 20$ should be less frequent than the previous category, there are more of these strong LLSs/ SLLSs in our sample owing to the fact that we can use the damping of Lyα and Lyβ to derive $N_{H I}$ (see also Section 3.2).
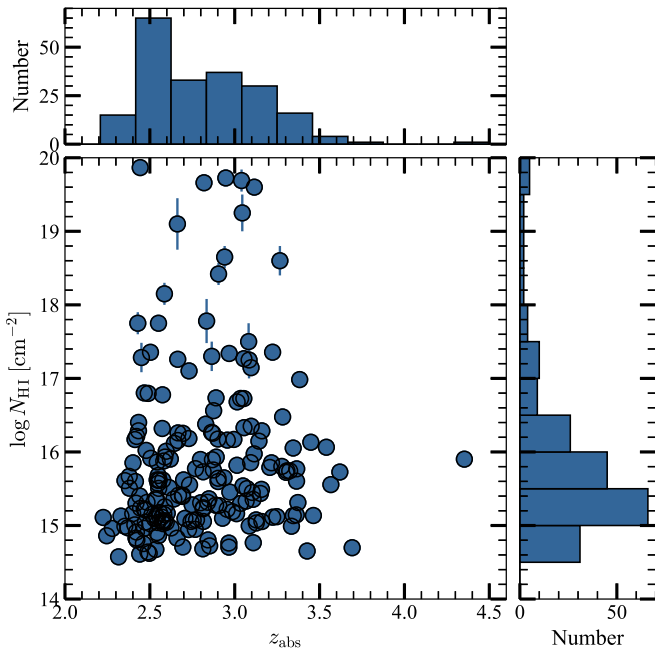
Figure 5. Distribution of the redshifts and H I column densities for the absorbers in the KODIAQ-Z sample. The top and right panels show the redshift and $\log N_{\rm H\,I}$ distributions, respectively.

4.2. H I Column Density and Doppler Parameter Distributions

In Figure 6, we compare the Doppler parameter, b, and $N_{\rm H\,I}$ derived from the PF in the individual components of H I. There is no strong trend between b and $N_{\rm H\,I}$, although broad components with $b > 40$ km s^{-1} are more frequent at $\log N_{\rm H\,I} < 16$ ($b = 40$ km s^{-1} implies $T < 10^5$ K, which is the threshold b-value between broad and narrow H I absorbers as defined by Lehner et al. 2007).[18] We, however, note that the majority of these broad components are found in more complex velocity profiles (two or more components), lower-S/N spectra, and/or strong H I absorbers where the b-value is not well constrained. It is quite possible that in most of these cases narrow components are blended, mimicking a broad component.

Considering the entire sample, the mean and standard deviation of b are $\langle b \rangle = 30 \pm 11$ km s^{-1}, and the median is 28 km s^{-1}. About 90% of the components have $13.3 \leqslant b \leqslant 40$ km s^{-1}, where $\langle b \rangle = 27 \pm 6$ km s^{-1} and the median is 27 km s^{-1}. The latter value corresponds to a temperature of the gas of $T < 4 \times 10^4$ K, consistent with the gas being primarily photoionized. At $z < 1$ for the CCC absorbers, the results are remarkably similar since 91% of the profile-fitted absorbers have also $b < 40$ km s^{-1} and a mean of 28 ± 8 km s^{-1} (CCCI and Figure 7 therein). This contrasts with the weaker LYAF absorbers with $\log N_{\rm H\,I} \lesssim 14$, where the median and mean b-values are systematically larger by 15%–30% at $z \lesssim 0.5$ than at $1.5 \lesssim z \lesssim 3.6$, implying that a larger fraction of the low-z IGM is hotter and/or more kinematically disturbed than the high-z IGM (Lehner et al. 2007).

---

[18] The Doppler parameter for H I is related to the gas temperature via $b_{\rm H\,I} = (2kT/m_{\rm H} + b_{\rm nt}^2)^{0.5}$, where k is the Boltzmann constant, T is the temperature of the gas, $m_{\rm H}$ is the mass of hydrogen, and $b_{\rm nt}$ is the (unknown) nonthermal component to the broadening.

5. Determining the Metallicity of the Absorbers

5.1. Motivations for Photoionization Modeling

Several empirical studies at both low and high redshifts have shown that SLFSs, pLLSs, and LLSs have properties that are consistent with the gas being predominantly photoionized (Prochaska 1999; Lehner et al. 2013; 2016; Prochaska et al. 2017b; CCCI; CCCII; Crighton et al. 2013; Fumagalli et al. 2016b; Cooper et al. 2021; Zahedy et al. 2021). In Section 4.2, we show that the KODIAQ-Z absorbers have temperatures on average $T \lesssim 4 \times 10^4$ K, consistent with the gas being photoionized (the same result was found for the low-redshift absorbers with similar $N_{\rm H\,I}$; CCCI). We also observe that when detected metal ions, such as the suite of carbon ions (C II, C III, C IV) or silicon ions (Si II, Si III, Si IV), align well with H I (see velocity profiles provided in the Appendix), this is consistent with these metal ions and H I probing a single ionization gas-phase (we, however, emphasize again that sometimes metal ions, due to their smaller b, can have more than one component within the breadth of the H I profiles even though the H I absorption shows only a single component — that also implies that the thermal broadening is a large contributor to the broadening).

In contrast to the low redshift universe, for the ionizing radiation field and typical lower metallicities at $z \sim 2$–3 (about 0.1% solar or [X/H] $= -2$; see below; see also Fumagalli et al. 2016b; Lehner et al. 2016), even strong transitions like C II $\lambda 1334$ and Si II $\lambda 1260$ are often not detected, so we have to rely more on intermediate (Si III, C III, S IV) and high (C IV) ions to determine the metallicity. Because the current survey probes lower-$N_{\rm H\,I}$ absorbers than our previous surveys at similar redshifts (Lehner et al. 2014, 2016), we find that the O VI absorption can often be narrow and align with the absorption seen in H I (and other lower metal ions if present). This contrasts with the very broad and strong O VI absorption frequently found in strong LLSs, SLLSs, and DLAs, which cannot be explained by photoionization (Lehner et al. 2014); it is quite plausible that this highly, collisionally ionized gas has also a different metallicity than the cooler, photoionized gas as discussed in Lehner et al. (2014). When O VI is narrow, it is used to constrain the photoionization models, and we find a posteriori that O VI is consistent with being produced by photoionization in many of these cases. The higher intensity of the EUVB at high z than at low z explains that O VI absorption associated with H I absorbers with $\log N_{\rm H\,I} \lesssim 17.2$ can be more often easily photoionized in low-density gas with densities that are reasonable (i.e., densities that do not imply extremely large path lengths for these absorbers, and hence a Hubble flow broadening that would be inconsistent with the observed broadening of the absorption lines).

5.2. Photoionization Modeling

To determine the metallicity for each absorber, we follow the same methodology and make similar assumptions to those described and discussed in CCCII and CCCIII. As laid out in the previous section, due to the higher-redshift absorbers and the change in the EUVB, we include high ions in the photoionization modeling. We model the photoionization using Cloudy (version C13.02; see Ferland et al. 2013), assuming a uniform slab geometry in thermal and ionization equilibrium. In all cases the slab is illuminated with a Haardt–Madau EUVB radiation field from quasars and galaxies (HM05 and HM12;
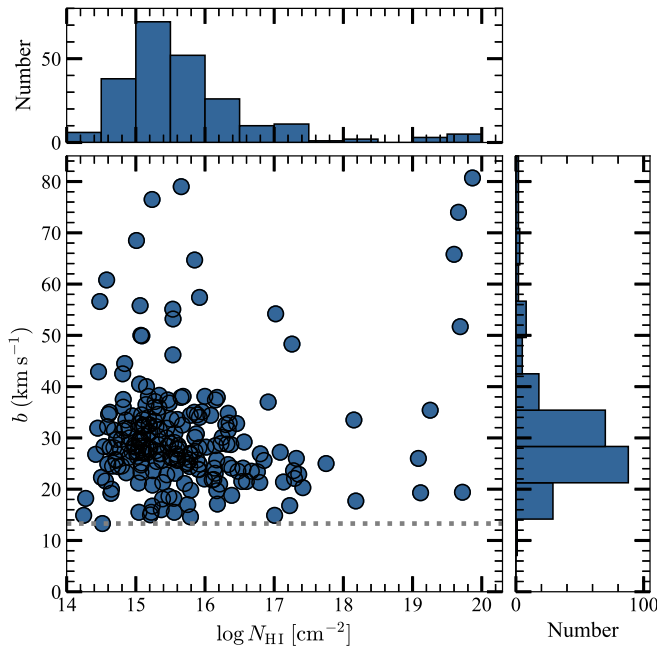
Figure 6. Distribution of the Doppler parameters (b) and H I column densities in the KODIAQ-Z sample obtained from the PF of individual components. The right and top panels show the b and $\log N_{\mathrm{H\,I}}$ distributions, respectively. The dotted line shows the lowest fitted b-value (corresponding $T \lesssim 10^4$ K); 90% of the components have $13.3 \leqslant b \leqslant 40$ km s^{-1}. Note that for the Keck HIRES spectra the instrumental b-value is about 4–5 km s^{-1}, well below the minimum observed b-value, i.e., narrower H I absorbers than $b < 12$ km s^{-1} could have been detected if present.

see Haardt & Madau 1996, 2001, 2012). We adopt HM05 (Haardt & Madau 2001, as implemented in Cloudy as the fiducial radiation field for KODIAQ-Z to allow for a direct comparison with the CCC results at lower redshift. However, we also use the HM12 EUVB to explore systematics associated with uncertainties in the radiation field. In Figure 7, we show the HM05 and HM12 EUVB at the average redshift of our survey ($z = 2.8$) along with the ionizing energy ranges for key ions used in KODIAQ-Z to constrain the metallicity of the absorbers. We also show in this figure the recent Khaire & Srianand (2019) EUVB models for their three favored slopes of the EUVB. While there are some differences between these EUVBs, these are not as large as at low z (Gibson et al. 2022). And indeed, we show a posteriori that the systematic errors arising from using a different EUVB are not as large as observed at $z \lesssim 1$. We note that for stronger absorbers ($\log N_{\mathrm{H\,I}} \gtrsim 18$) at $z \sim 3$, Fumagalli et al. (2016b) show that local ionizing sources did not affect much the metallicity estimates, and therefore we do not expect that local sources of (galactic) ionizing radiation would drastically change the metallicities in KODIAQ-Z. We adopt the same grid of Cloudy models as summarized in Table 3 of CCCII.

The main variables for the photoionization models are the ionization parameter—$U \equiv n_\gamma/n_{\mathrm{H}}$, the ionizing photon density/total hydrogen number density (neutral + ionized)—and the metallicity, [X/H]. We assume solar *relative|* heavy-element abundances from Asplund et al. (2009) but allow for possible variation between between C and α-elements (where α can be O or Si) since this ratio can be sensitive to nucleosynthesis effects (see, e.g., Cescutti et al. 2009; Mattsson 2010, for more details). [C/α] is therefore not

necessarily solar in the gas probed by absorbers studied here (see Aguirre et al. 2004; see also Lehner et al. 2013; CCCII, at $z < 1$).

To derive the metallicities, we compare the column densities for each absorber with a grid of photoionization models in a Bayesian context. We make use of the publicly available codes described in Fumagalli et al. (2016b; see also Prochaska et al. 2017b) and CCCII.[19] We perform the Markov Chain Monte Carlo (MCMC) analysis on an absorber-by-absorber basis. For each absorber we apply Gaussian priors on $N_{\mathrm{H\,I}}$ and redshift.[20] For all the absorbers, we first assume a flat prior on U. After examination of the model outputs (in particular, convergence of walkers, comparison of predicted and observed column densities, and "corner plots" showing the posterior probability density function (pdf's)), we can assess which absorbers can be reliably modeled with this flat prior on U and which may require a Gaussian prior on $\log U$ to help the models converge (see below). Similarly, during this first iteration, we also determine which absorbers can be modeled with a flat prior on [C/α] (i.e., for which we could estimate [C/α]), which absorbers require [C/α] = 0 (e.g., those with no detections for all the ions or no carbon ions), and which absorbers may require a Gaussian prior on [C/α] (absorbers with poor constraints on carbon ions and/or α-elements).

When the constraints from the metal ions are not optimal, we adopt a Gaussian prior on $\log U$ (CCCII). These concern only SLFSs and pLLSs; the 16 LLSs and 7 SLLSs have enough metal-ion constraints for the ionization models to converge. A total of 58% of the SLFSs and pLLSs do not require a Gaussian prior on $\log U$. In Figure 8, we show the $\log U$ median values as a function of $N_{\mathrm{H\,I}}$ for the absorbers that can be modeled assuming a flat prior on $\log U$. As observed at low z (CCCIII; CCCII), a strong anticorrelation is observed between the ionization parameter and $N_{\mathrm{H\,I}}$, with a slope steeper by a factor 2.3 compared to the low-redshift absorbers. To characterize the distribution of $\log U$ appropriate for use as a Bayesian prior, we split the $N_{\mathrm{H\,I}}$ range into two bins, $14.4 \leqslant \log N_{\mathrm{H\,I}} < 16.2$ and $16.2 \leqslant \log N_{\mathrm{H\,I}} < 17.2$. In the former, 52% (75/143) can be modeled with a flat $\log U$ prior, while for the latter that number is 81% (30/37). In both cases, we find that each pdf is well-fit with a normal distribution with $\langle \log U \rangle = -1.43 \pm 0.59$ and -2.06 ± 0.53. We adopt these results for our Gaussian prior on $\log U$ in the $14.4 \leqslant \log N_{\mathrm{H\,I}} < 16.2$ and $16.2 \leqslant \log N_{\mathrm{H\,I}} < 17.2$ intervals, respectively.

For the Gaussian prior on [C/α], we similarly use absorbers with relatively good constraints from carbon and α-element ions to determine a mean and standard deviation of $\langle [C/\alpha] \rangle = -0.31 \pm 0.42$. We apply these for the Gaussian prior on the [C/α] ratio for absorbers with poorly constrained [C/α]. We note that for a few (13) absorbers, [C/α] is clustered around $-1 \lesssim [C/\alpha] \lesssim -0.9$. Not including those in the mean would of course increase the mean [C/α] value, but it would not change the posterior pdf of [C/α] for these absorbers. We discuss in Sections 5.3.4 and 6.5 in more detail the [C/α] ratio.

---

[19] The code presented in Fumagalli et al. (2016b) and CCCII has been incorporated into the PyIGM Python package and is available at https://github.com/pyigm/pyigm (Prochaska et al. 2017a).
[20] For the metal ions, the lower (upper) limits, the likelihoods are based on a rescaled Q-function (cumulative distribution function).
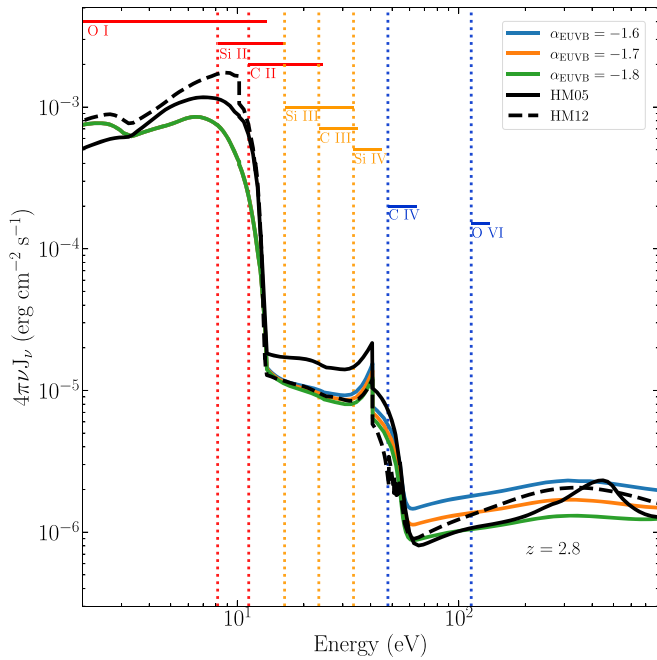
Figure 7. Comparison of several spectral energy distributions for the ionizing background radiation field. The HM05 is adopted in this work, although we systematically compare it with the results from HM12. We also show the recent EUVB models from Khaire & Srianand (2019) for comparison. The key ions used to constrain the metallicities are shown, where the dotted vertical bars show their minimum ionization energy, while their range of ionization energy is shown with the horizontal bar.
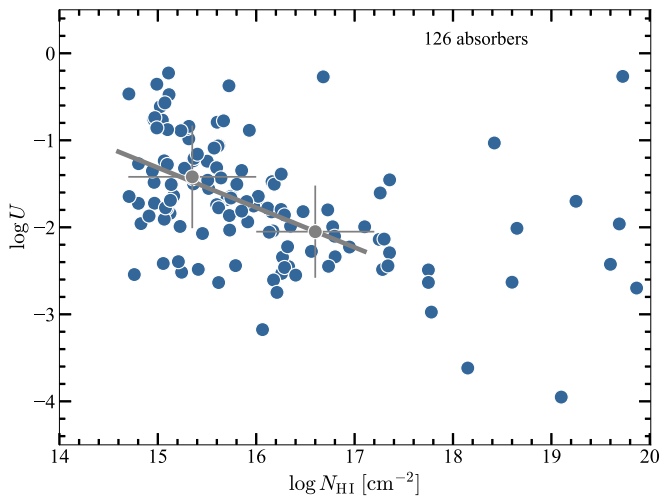

Figure 8. The $\log U$ median values as a function of $N_{\rm H\,I}$ for 126 absorbers in the KODIAQ-Z sample. The solid line shows a linear fit to the data with $14.6 \leqslant \log N_{\rm H\,I} \leqslant 17.2$ (with a slope -0.46 ± 0.10 and intercept $+5.56$). The plus signs show the mean values of the median values in each $N_{\rm H\,I}$ interval indicated by the horizontal bar.

We summarize the results of our photoionization modeling in Table 4 for all the absorbers in the KODIAQ-Z sample. For each absorber, we list the sight-line name, absorber redshift ($z_{\rm abs}$), $\log N_{\rm H\,I}$, metallicity, ionization parameter ($\log U$), and $[C/\alpha]$ ratio (when estimated). Each quantity is reported with the 68% CI and median values, except in the cases where we derive an upper or lower limit, in which case we report the 80% CI and median values. We emphasize that these values are attempts to summarize the posterior pdf's provided by the MCMC

sampling of the CLOUDY models. While those pdf's can be well behaved and nearly Gaussian, some are not as well behaved (see corner plots in the Appendix and supplemental materials).

For the HD-LLS absorbers used to compare with KODIAQ-Z, we apply the same overall methodology and the same HM05 EUVB in the photoionization models in order to reduce any systematic errors between the two surveys. However, we note that the absorption in the absorber in the HD-LLS and KODIAQ-Z is not treated in the same way. In the HD-LLS survey, a velocity interval of ± 500 km s^{-1} defines an absorber, which is driven by the methods to derive $N_{\rm H\,I}$ that use either the break at the Lyman limit or the damped wings from Lyα. For the metal lines, they sum the absorption components identified within that interval even though it is rarely larger than 200 km s^{-1}. As detailed in Section 3, we define an absorber using the Lyman series, and as much as possible we consider the absorption component by component, so in a window of ± 200–500 km s^{-1} KODIAQ-Z can have several absorbers. However, as $N_{\rm H\,I}$ increases and overlaps with the HD-LLS survey, i.e., for $\log N_{\rm H\,I} \gtrsim 18$, we lose our ability to derive the column densities in individual components, especially when we have information only from Lyα and Lyβ. In these cases, we rely on the low ions to define the integration range in the intermediate and low ions. In our approach, we make the physically motivated assumption that the optical depths of the low ions follow that of the stronger H I absorption while the higher ions follow better weaker H I absorbers that may not contribute much to the total H I column—but the extra absorption in the higher metal ions could dominate their column density (Section 3). However, a posteriori we will show that this different treatment does not lead to large different results.

5.3. Uncertainties on the Metallicities

All the errors for the output quantities (e.g., metallicity) reported in this work are statistical errors from the Bayesian MCMC ionization modeling. There are additional systematic and other uncertainties that we discuss here.

5.3.1 Ionizing Background

The first one is caused by the uncertainty in the ionizing EUVB as alluded to previously. At low redshift, Wotta et al. (2019) show that the metallicity is systematically shifted on average by about $+0.4$ dex for the SLFSs/pLLSs and $+0.2$ dex for the LLSs when the ionizing EUVB changes from HM05 to HM12 (see also Gibson et al. 2022 for a study with other EUVBs). We have done a similar experiment with the KODIAQ-Z and HD-LLS absorbers, showing that the impact of the EUVB is overall smaller at high z than observed at low z. In Figure 9, we show the comparison between the metallicities using HM12 and HM05 for the KODIAQ-Z absorbers where we did not apply a Gaussian prior on $\log U$ (which essentially removes all the limits, although including those would not change the results). The immediate conclusion from this figure is that the systematic differences between the metallicities derived using HM05 and HM12 are quite small, and much smaller than observed at low z. On average, considering all the absorbers, we find $\langle [X/H]_{\rm HM12} - [X/H]_{\rm HM05} \rangle = +0.10 \pm 0.17$. The differences are also not as systematic as low z since there are several cases where $[X/H]_{\rm HM12} \leqslant [X/H]_{\rm HM05}$.

Table 4
Metallicities of the Absorbers in KODIAQ-Z Sample

QSO Name	z_{abs}	log $N_{H\,I}$ [cm^{-2}]	[X/H]	log U	[C/α]
J010806+163550	2.31582	14.574 ± 0.004	<−4.738, <−3.675, <−2.475	−2.117, −1.521:, −0.945	···
J045213−164012	2.44044	14.611 ± 0.004	<−4.739, <−3.688,		
lt−2.568	−2.141, −1.540:, −0.910	···			
J170100+641209	2.49647	14.622 ± 0.007	−2.425, −2.069, −1.776	−1.743, −1.470, −1.222	−0.460, −0.099:, +0.264
J023359+004938	2.48836	14.640 ± 0.019	<−4.723, <−3.594, <−2.517	−2.098, −1.503:, −0.942	···
J073149+285448	3.42673	14.653 ± 0.008	<−4.743, <−3.677, <−2.418	−2.165, −1.557:, −0.954	···
J045213−164012	2.53542	14.674 ± 0.003	<−4.790, <−3.922, <−3.079	−2.109, −1.492:, −0.914	···
J012156+144823	2.81222	14.683 ± 0.002	<−4.777, <−3.941, <−3.014	−2.126, −1.567:, −0.976	···
J105756+455553	3.69194	14.698 ± 0.003	<−4.754, <−3.776, <−2.567	−2.198, −1.603:, −1.011	···
J010925−210257	2.96546	14.706 ± 0.008	−2.840, −2.632, −2.263	−0.797, −0.561, −0.362	−0.754, −0.428:, −0.071
J014516−094517	2.69430	14.708 ± 0.002	−2.809, −2.269, −1.832	−2.031, −1.685, −1.335	−0.291, +0.020:, +0.355

Note. The lower to upper bounds for each quantity represent the 68% CI for detections and the 80% CI for the upper or lower limits; the middle number is the median metallicity. When a colon is present after the median value of log U, this implies that a log U Gaussian prior was used to determine the properties of the absorber.

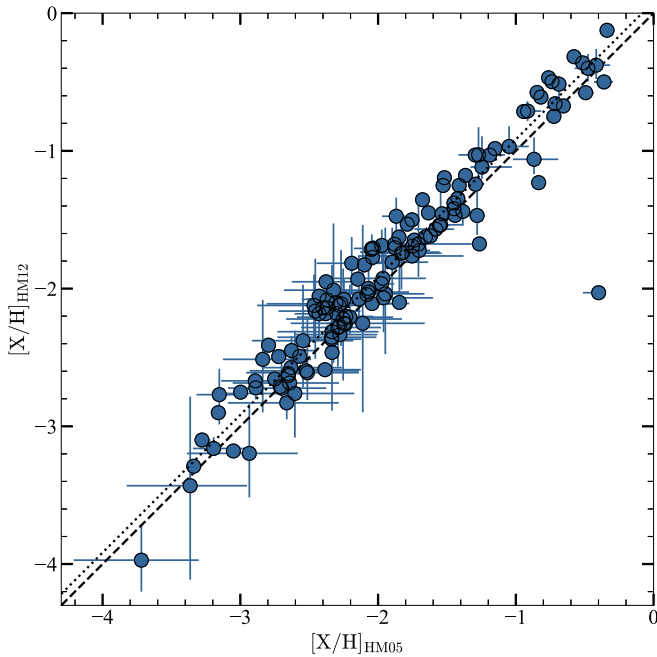(This table is available in its entirety in machine-readable form.)

Figure 9. Comparison of the median metallicities (with 68% CI) of KODIAQ-Z derived using the HM12 and HM05 EUVBs. The dashed line is the 1:1 relationship. The dotted line shows the mean of the differences between the metallicities derived using HM12 and HM05.

As illustrated in Figure 10, where the differences of the median metallicities between HM12 and HM05 versus $N_{H\,I}$ are shown for both the KODIAQ-Z and HD-LLS absorbers, that difference decreases with increasing $N_{H\,I}$ as observed for the low-redshift absorbers. In fact, for log $N_{H\,I} \gtrsim 17.2$, $\langle$[X/H]$_{HM12}$ − [X/H]$_{HM05}\rangle \simeq 0.0 \pm 0.1$, i.e., the HM12 and HM05 EUVBs essentially give the same metallicities and any small differences are essentially random. For the SLFSs and pLLSs, $\langle$[X/H]$_{HM12}$ − [X/H]$_{HM05}\rangle \simeq +0.12 \pm 0.15$. These results are not entirely surprising when considering the different EUVBs in Figure 7, with their difference being smaller than at low z over similar ionizing energies (Gibson et al. 2022). Based on these results and Figure 7, the metallicities would not change much if we had used the more recent EUVBs

from Khaire & Srianand (2019). For comparison with CCC, we adopt hereafter the results from the HM05 EUVB.

5.3.2 Multiple Ionization Phases

Throughout we also assume a single ionization phase. In cases where a higher ion (O VI or sometimes C IV in our survey) is clearly underproduced by the photoionizing model, this given high ion is not included in the final model. In the latter case, this would be evidence for an absorber probing multiple gas phases. As shown by Lehner et al. (2014), this can especially happen in the strong LLSs, where O VI can be very strong and can rarely be modeled with a single gas-phase ionization model with the lower ions. In these cases, the velocity profiles of the O VI profiles are also very different from those of the lower ions (including C IV), strongly hinting at multiple gas phases.

As we explore lower-$N_{H\,I}$ absorbers in this present survey, we find in many cases that a single ionization gas phase can in fact match well the observed metal-ion column densities, including those of O VI. In these cases and in contrast to stronger H I absorbers, the absorption is also often quite simple (typically a single component), and the velocity profiles of the different metal ions align well with each other (including O VI when present) and with the H I velocity profiles. This, of course, does not necessarily mean that a more complex ionization structure might not be present, but a single gas-phase model is often sufficient to reproduce the observed column densities, especially in the SLFS and pLLS regimes.

We adopt here therefore the simplest approach and assumption, a methodology also employed in CCC. Nevertheless, we note that, at least at low redshift, the effects on the metallicities between single versus multiple gas-phase modelings typically lead to small changes in metallicities of the order of < 0.1–0.2 dex (Howk et al. 2009; Haislmaier et al. 2021). Therefore, we do not expect that considering a more complex ionization structure for the absorbers in our sample would drastically change the metallicities.

5.3.3 Ions Included in the Models

Depending on wavelength coverage, contamination, and H I column density, the ions used to constrain the photoionization
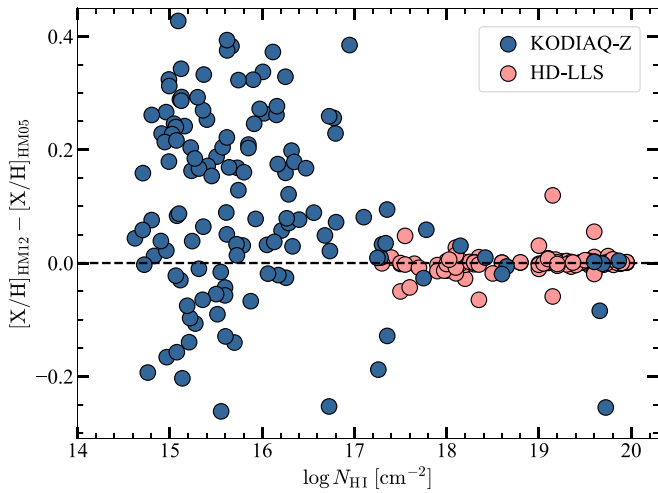
Figure 10. Difference between the median metallicities for the KODIAQ-Z and HD-LLS absorbers derived using the HM12 and HM05 EUVBs against the H I column densities of the absorbers.

models can be different. This may introduce some additional systematic uncertainties when comparing the metallicities across $N_{H I}$ and z. For example, the suite of carbon ions available in the redshift probed by our sample consists of C II, C III, and C IV. Sometimes all these ions are available, but it is not always the case. In particular, the inclusion of C IV can a priori change the density (ionization parameter) to lower (higher) values owing to the 48 eV ionization energy required to ionize C III to C IV. This change in log U could possibly affect also the metallicities.

At $z < 1$, Lehner et al. (2019) investigated the effect on the metallicity and other quantities (e.g., log U) if a high ion (in that case O IV) is included or not in the photoionization models. They estimated that this effect would only change the metallicities on average by 0.06 dex, with the largest metallicity difference being around 0.25 dex. However, and importantly, for the physical conditions of the absorbers such as U, n_H, and N_H, that effect can modify the values for these quantities on average by 0.5 dex.

Using our sample, we specifically undertook a similar experiment but using C IV instead of O IV. To select this sample, we require that the absorbers have coverage of C IV and C II and/or C III (and often there is also coverage of other ions, such as Si II, Si III, and Si IV). For 38 absorbers with $14.8 \lesssim \log N_{H I} \lesssim 20$, we systematically removed C IV in the photoionization models and ran the same photoionization models. In Figure 11, we show the comparison of [X/H] and log U as a function of $N_{H I}$ for these 38 absorbers with and without C IV included in the photoionization models. There is no trend in the difference between these quantities and $N_{H I}$. The errors are smaller when C IV is included in the models, which is expected. The overall impression from this experiment is that neither the metallicities nor the log U values (or quantities not shown here, such as n_H, N_H) drastically change with the inclusion or not of C IV in the models. The mean difference in metallicities between absorbers with and without C IV is -0.06 ± 0.21 dex. Only 8% of the 38 absorbers have metallicity differences larger than 0.4 dex (0.4–0.7 dex). For log U, the mean difference is, however, much smaller than at low redshift with -0.04 ± 0.29 dex. In that case, 10% of the absorbers have log U differences larger than 0.4 dex (0.4–1.0 dex). Therefore, the change in the ionic constraints between the different absorbers at $2.2 \lesssim z \lesssim 3.6$ is unlikely to
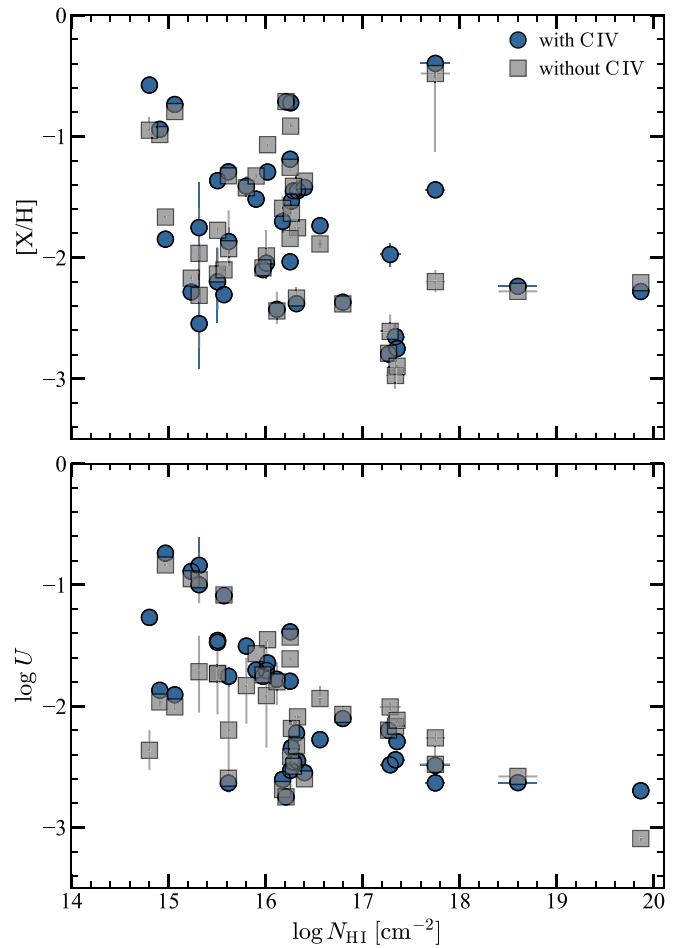

Figure 11. Comparison of the metallicities (top) and ionization parameters (bottom) derived without and with using C IV in the photoionization modeling as a function of the H I column density. The median values of posterior pdf's of the metallicity and log U are adopted as the central values with 68% CI.

produce large systematic uncertainties on the metallicities and other physical parameters such as the density or ionization parameter.

5.3.4 Nonsolar Relative Abundances

Other uncertainties may arise from dust or nonsolar relative abundances. Since carbon, silicon, and oxygen ions are typically used to constrain the ionization modeling at high redshift, we explicitly take account of possible nonsolar [C/α] by letting this ratio vary in the range $-1 \leqslant [C/\alpha] \leqslant +1$. In Figure 12, we show [C/α] as a function of $N_{H I}$ (left panel) and the neutral fraction $f_{H I} \equiv N_{H I}/(N_{H I} + N_{H II})$ (right panel) for the KODIAQ-Z and HD-LLS absorbers. We also show in this figure the results from CCC at $z \lesssim 1$ (CCCIII). For [C/α]$-N_{H I}$, a Spearman rank-order test shows no strong evidence for a correlation between [C/α] and $N_{H I}$ with a correlation coefficient $r_S = 0.14$ and a p-value $= 0.04$. On the other hand, the same statistical test shows a positive monotonic correlation between [C/α] and $f_{H I}$ with a correlation coefficient $r_S = 0.44$ and a p-value $\ll 0.05\%$. However, removing the 13 KODIAQ-Z absorbers with low [C/α] $\lesssim -0.9$ eliminates any evidence of correlation between [C/α] and $f_{H I}$ with a correlation coefficient $r_S = -0.05$ and a p-value $= 0.88$.
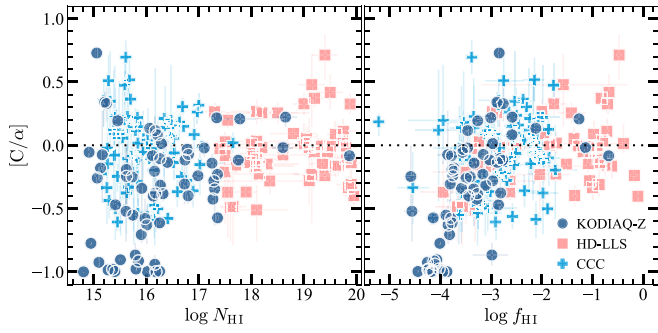
Figure 12. The ratio of [C/α] against the H I column density (left) and neutral fraction (right) from the KODIAQ-Z (2.2 $\lesssim z \lesssim$ 3.6), HD-LLS (2.2 $\lesssim z \lesssim$ 3.6), and CCC ($z \lesssim 1$) surveys. The median values of the [C/α] posterior pdf's are adopted as the central values with 68% CI.
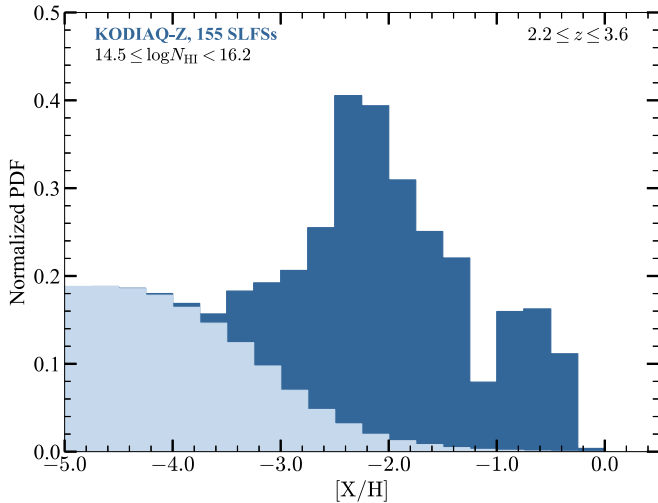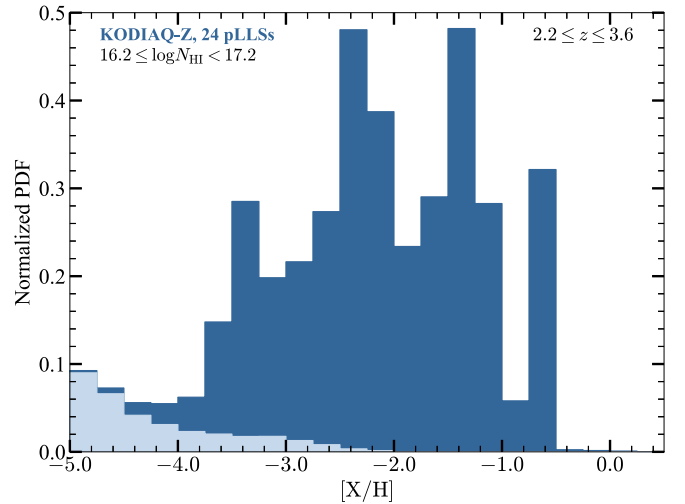

Figure 13. Posterior metallicity pdf of the SLFSs in the KODIAQ-Z sample. The light-colored region indicates the contribution from the upper limits.

So far CCC, HD-LLS, and KODIAQ-Z have been treated as a single sample. Considering the samples individually would not change the conclusions for the lack of correlation between [C/α] and $N_{\rm H\,I}$. However, while neither CCC nor HD-LLS shows any correlation between [C/α] and $f_{\rm H\,I}$, for KODIAQ-Z the Spearman rank-order test shows a strong positive monotonic correlation between [C/α] and $f_{\rm H\,I}$ with a correlation coefficient $r_{\rm S} = 0.72$ and a p-value $\ll 0.05\%$. In that case, removing the 13 absorbers with low [C/α] $\lesssim -0.9$ still yields $r_{\rm S} = 0.55$ and a p-value $\ll 0.05\%$. Therefore, in the KODIAQ-Z sample, [C/α] is more affected by ionization correction than in the other samples, most likely because larger ionization corrections are applied (for CCC, C II and C III are used to constrain [C/α] instead of C IV, C III, and C II for KODIAQ-Z, while for HD-LLS the ionization corrections are typically smaller). In the CCC and HD-LLS samples, we always find [C/α] $\geqslant -0.6$. Hence, [C/α] values with [C/α] < -0.6 are most likely predominantly caused by ionization rather than nucleosynthesis effects. Excluding the 13 absorbers with low [C/α] $\lesssim -0.9$, we find that $\langle$[C/α]$\rangle = -0.20 \pm 0.31$ in the KODIAQ-Z sample compared to $\langle$[C/α]$\rangle = -0.05 \pm 0.30$ in the CCC survey and $\langle$[C/α]$\rangle = -0.05 \pm 0.24$ in the HD-LLS survey. The similarity between CCC and HD-LLS is quite remarkable since different z, $N_{\rm H\,I}$, and [X/H] are probed. The 1σ dispersion is about the same in the three surveys, but on

Figure 14. Posterior metallicity pdf's of the pLLSs in the KODIAQ-Z sample. The light-colored region indicates the contribution from the upper limits.

average the KODIAQ-Z sample has a systematic effect of -0.15 dex on [C/α], which is most likely caused by ionization correction. We can test this further using the sample of 38 absorbers in Section 5.3.3, where C IV was removed in the ionization models. From this sample, there are 26 absorbers for which there is a good constraint on [C/α]. We find that the mean of the [C/α] median values increases from -0.38 ± 0.35 dex when C IV is included in the ionization models to -0.20 ± 0.35 when C IV has been removed. There is therefore a systematic ionization effect of about 0.18 dex on [C/α]. In conclusion, the variable C/α is important in that we can marginalize over that as a "nuisance" variable in order to get a better sense of the uncertainties in the metallicity, but it may not provide deep insights on the nucleosynthesis effects that may affect C/α (see also Section 6.5).

Finally, we note that for stronger LLSs and SLLSs such as those in the HD-LLS survey (Fumagalli et al. 2016b), dust depletion is found to be a small effect for strong LLSs and even SLLSs at high z and hence not very likely to have any appreciable effect on the metallicity estimates, especially considering that the typical species—carbon, silicon, oxygen —used to constrain the ionization models are known not to be strongly depleted onto dust (e.g., Savage & Sembach 1996; Jenkins 2009; Jenkins & Wallerstein 2017).

6. Metallicity of the IGM/CGM Absorbers at 2.2 $\lesssim z \lesssim$ 3.6

6.1. Metallicity Distributions

In Figures 13–15, we show the posterior pdf's of the metallicities of the SLFSs, pLLSs, and LLSs for KODIAQ-Z. In Figure 15, we also overplot the metallicity pdf of the LLSs from the HD-LLS survey, showing similar pdf's between the two surveys, and hence implying that we can combine both samples to improve the statistics for the LLSs. These posterior pdf's are constructed by combining the normalized metallicity pdf's of all of the absorbers within a given H I column density regime. In Table 5, we summarize the medians, means, standard deviations, and interquartile ranges (IQRs) for the various absorbers, as well as a combination of these. We also include in this table the results for the HD-LLS, KODIAQ-Z +HD-LLS, and R12-DLA samples.
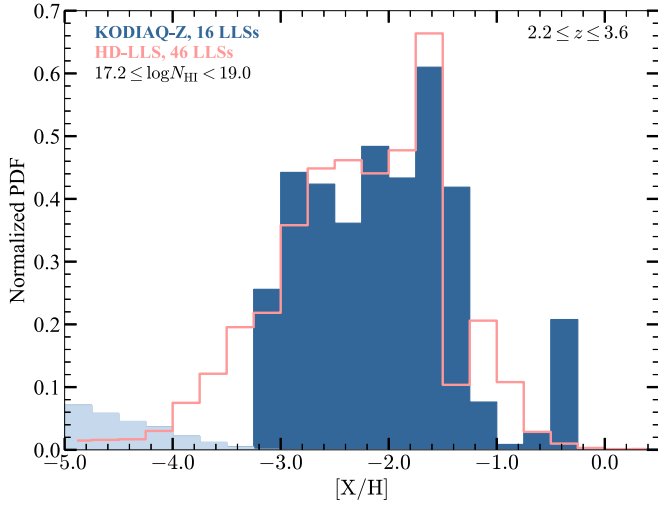
Figure 15. Comparison of the posterior metallicity pdf's of the LLSs in the KODIAQ-Z and HD-LLS sample. The light-colored region indicates the contribution from the upper limits.
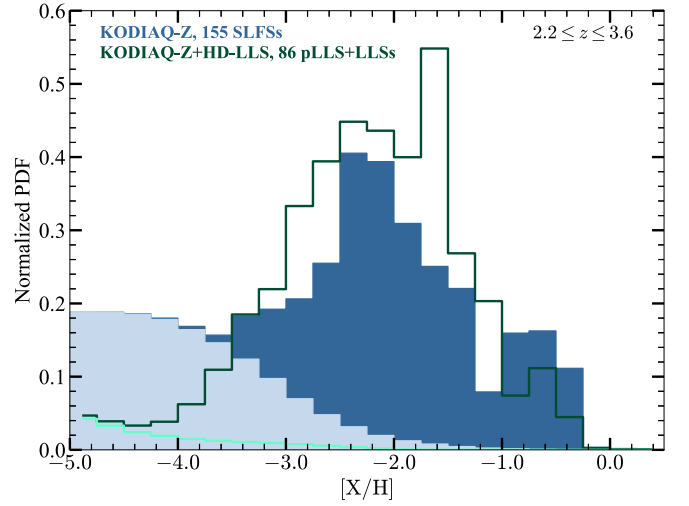

Figure 16. Comparison of the posterior metallicity pdf's of the SLFSs (KODIAQ-Z) and pLLSs+LLSs (KODIAD-Z+HD-LLS). The light-colored region indicates the contribution from the upper limits.
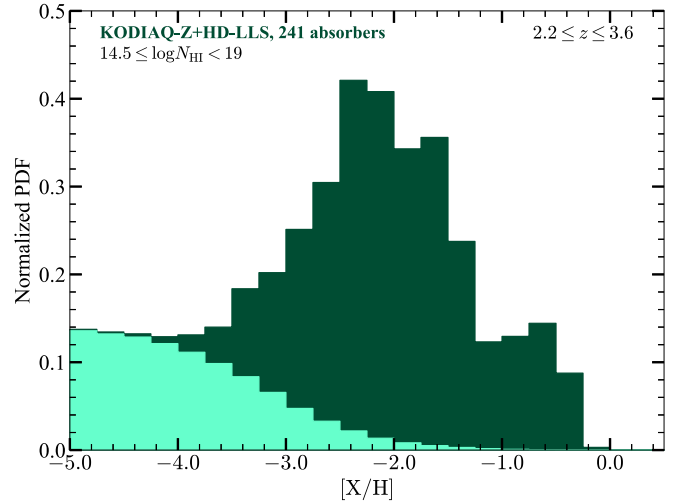
Table 5
Metallicity Statistics at $2.2 \leqslant z \leqslant 3.6$

Abs. Type	$N_{\rm abs}$	$\widetilde{[X/H]}$	$\overline{[X/H]}$	IQR
KODIAQ-Z Sample				
SLFSs	155	−2.42	−2.62 ± 1.23	[−3.61, −1.76]
pLLSs	24	−2.09	−2.27 ± 1.04	[−2.91, −1.45]
LLSs	16	−2.16	−2.17 ± 0.90	[−2.75, −1.58]
pLLSs+LLSs	40	−2.13	−2.23 ± 0.99	[−2.78, −1.51]
SLFSs+pLLSs+LLSs	195	−2.36	−2.54 ± 1.20	[−3.37, −1.69]
SLLSs	7	−2.09	−2.11 ± 1.29	[−2.70, −0.87]
HD-LLS Sample				
LLSs	46	−2.22	−2.28 ± 0.78	[−2.77, −1.71]
SLLS	73	−1.89	−1.91 ± 0.84	[−2.46, −1.33]
KODIAQ-Z+HD-LLS Sample				
LLSs	62	−2.20	−2.25 ± 0.81	[−2.75, −1.69]
pLLSs+LLSs	86	−2.19	−2.26 ± 0.88	[−2.77, −1.66]
SLFSs+pLLSs+LLSs	241	−2.34	−2.49 ± 1.13	[−3.23, −1.70]
SLLSs	80	−1.90	−1.93 ± 0.89	[−2.47, −1.33]
R12-DLA Sample				
DLAs	101	−1.41	−1.41 ± 0.50	[−1.70, −1.05]

Note. The tilde and bar above [X/H] represent the median and mean (with the standard deviation), respectively. To derive the mean, median, standard deviation, and IQR, we directly use the walkers for the KODIAQ-Z and HD-LLS absorbers. Note that these are means, standard deviations, and IQRs in logarithmic space.

Figure 17. Posterior metallicity pdf of the absorbers with $14.5 \lesssim \log N_{\rm H\,I} < 19$ in the combined sample of KODIAQ-Z and HD-LLS. The light-colored region indicates the contribution from the upper limits.

Considering first the SLFSs, the metallicity pdf is not unimodal. It is negatively skewed with a long and prominent tail extending well below $[X/H] < -3$. Its main peak is around the median value of −2.4 dex. The metallicity pdf of the SLFSs dips around −1.1 dex, with a second weaker peak around −0.6 dex. The IQR is nearly 2 dex from −3.6 to −1.8 dex. As $N_{\rm H\,I}$ increases, the overall median metallicity increases and the IQR decreases. For pLLSs+LLSs, the median metallicity is −2.2 and the IQR is only ∼1 dex from −2.8 to −1.7 dex. These features are distinctly illustrated in Figure 16, where we compare the pdf's of the SLFSs and pLLSs+LLSs: there are more frequently pLLSs+LLSs in the metallicity range $-3.2 \lesssim [X/H] \lesssim -1.2$ than SLFSs, but SLFSs become more frequent at extremely low metallicity $[X/H] \lesssim -3.5$. Compared to the SLFS and pLLS pdf's, the LLS pdf is not as negatively skewed, with about the same values for the mean and median.

For completeness, and if one ignores the H I column density dependence of the metallicity pdf's between the SLFSs, pLLSs, and LLSs, we show in Figure 17 the resulting metallicity pdf of the absorbers with $14.5 < \log N_{\rm H\,I} < 19$ at $2.2 \lesssim z \lesssim 3.6$. Since the SLFSs are about 2/3 of the sample, the metallicity pdf of the $14.5 < \log N_{\rm H\,I} < 19$ absorbers is not too different from that of the SLFSs, with a similarly skewed distribution toward low metallicities but also a second weaker peak around −0.6 dex.

6.2. Metallicity versus $N_{H\,I}$

As shown in the previous section, there is already evidence for an overall increase of the metallicities with $N_{\rm H\,I}$ in the range
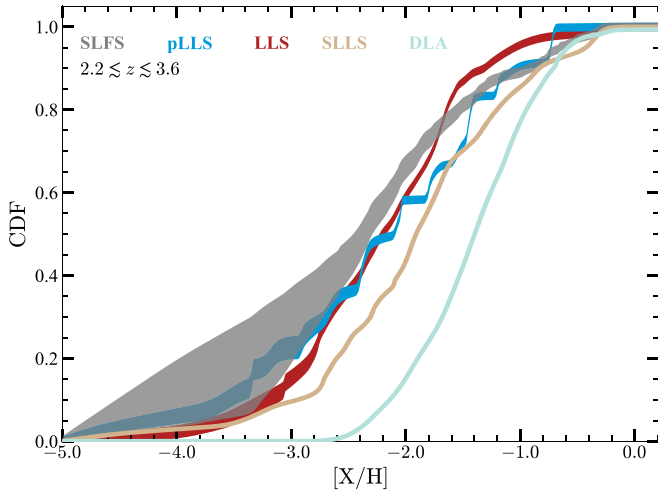
Figure 18. Cumulative probabilities of the SLFS, pLLS, LLS, SLLS, and DLA metallicity pdf's. The SLFSs and pLLSs are from KODIAQ-Z, the LLSs and SLLSs from KODIAQ-Z and HD-LLS, and the DLAs from R12-DLA.

Figure 19. Metallicities of the SLFSs ($14.5 < \log N_{\mathrm{H\,I}} < 16.2$), pLLSs ($16.2 \leqslant \log N_{\mathrm{H\,I}} < 17.2$), LLSs ($17.2 \leqslant \log N_{\mathrm{H\,I}} < 19$), SLLSs ($19 \leqslant \log N_{\mathrm{H\,I}} < 20.3$), and DLAs ($\log N_{\mathrm{H\,I}} \geqslant 20.3$) at $2.2 \lesssim z \lesssim 3.6$ as a function of $N_{\mathrm{H\,I}}$. The blue data are from the KODIAQ-Z survey (this paper). The orange data are from the HD-LLS survey (Fumagalli et al. 2016b), but we reestimated the metallicities of these absorbers following the same EUVB and methodology as in KODIAQ-Z. The purple data (DLAs) are from Rafelski et al. (2012).

Table 6
Fraction of Very and Extremely Metal-poor Absorbers at $2.2 \leqslant z \leqslant 3.6$

Abs. Type	N_{abs}	Fraction with $[\mathrm{X/H}] < -2.4$	Fraction with $[\mathrm{X/H}] < -3.0$	Fraction with $[\mathrm{X/H}] < -3.5$
SLFSs	155	>41.2%–54.3%	>14.7%–25.1%	>2.9%–8.9%
pLLSs	24	23.3%–54.2%	10.5%–37.1%	5.1%–27.5%
LLSs	62	26.3%–45.9%	8.6%–23.3%	2.0%–11.5%
SLLSs	80	20.1%–36.5%	3.1%–12.3%	0.3%–5.4%
DLAs	101	0.2%–4.3%	0.0%–2.6%	0.0%–2.6%

Note. The SLFSs and pLLSs are from KODIAQ-Z; the LLSs and SLLSs are from KODIAQ-Z+HD-LLS; the DLAs are from R12-DLA. The percentage ranges are estimated using the Wilson score with 90% confidence intervals. For upper limits, we use the 90th percentile to define the metallicity values. The SLFSs have an important fraction of upper limits, which is the reason that we mark them with a > sign in that table.

$14.5 \leqslant \log N_{\mathrm{H\,I}} \leqslant 19$. We now explore further how the metallicity varies with $N_{\mathrm{H\,I}}$ considering the higher-$N_{\mathrm{H\,I}}$ absorbers (the SLLSs and DLAs). From Table 5, the trend described above of increasing median/mean metallicity with increasing $N_{\mathrm{H\,I}}$ continues to apply in the SLLS and DLA regimes. The IQR and standard deviation for the SLLSs are about similar to those of pLLSs+LLSs, but those for the DLAs are substantially smaller (the IQR is 0.65 dex smaller than that of the SLFSs). The negative skewness is also less pronounced for the SLLSs and absent for the DLAs. In fact, the DLA metallicity pdf, contrary to the pdf's of the other weaker absorbers, is fully consistent with a Gaussian distribution according to the Kolmogorov−Smirnov test with a p-value $\ll 0.1\%$ (see also Rafelski et al. 2012).

These conclusions are further strengthened considering the cumulative distribution functions (CDFs) of the metallicity pdf's (see Wotta et al. 2019 for the description of the estimation of the CDFs including upper and lower limits). We show in Figure 18 the metallicity CDFs for the SLFSs, pLLSs, LLSs, SLLSs, and DLAs. With only 24 pLLSs, this regime remains the least sampled. Nevertheless, the CDF confirms the overall evolution of the metallicities as $N_{\mathrm{H\,I}}$ increases: very metal-poor (VMP) absorbers with $[\mathrm{X/H}] \leqslant -2.4$ are more likely to be found in the SLFS, pLLS, and LLS regimes than in the SLLS or DLA regime ($[\mathrm{X/H}]_{\mathrm{VMP}} = -2.4$ is the 2σ lower bound of the DLA metallicities). Notably the median metallicity of the SLFS is right at that threshold of $[\mathrm{X/H}]_{\mathrm{VMP}} = -2.4$, implying that 50% of these absorbers are VMP while only <5% of the DLAs may be VMP by definition.

Thus far, we have separately analyzed the absorbers in different $N_{\mathrm{H\,I}}$ categories. However, we can also simply plot the metallicities of the absorbers as a function of $N_{\mathrm{H\,I}}$, which is shown in Figure 19. This has the obvious advantage that one does not have to make an a priori differentiation between the different absorbers in classes of H I column densities. For the SLFSs, pLLSs, LLSs, and SLLSs with well-constrained metallicities (i.e., not including the lower and upper limits), the central values represent the median of the posterior pdf's, and the error bars represent the 68% CI. For the upper and lower limits, the downward- and upward-pointing triangles give the 90th and 10th quantiles, while the vertical bar gives
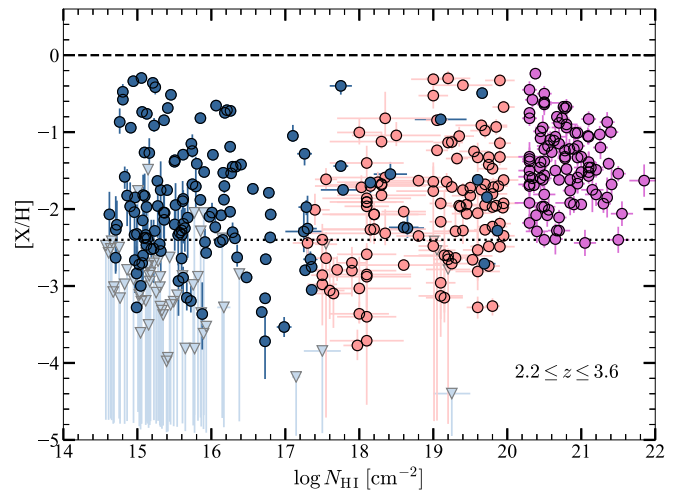
the 80% CI. For the DLAs, the best estimates with their 1σ error bars are shown. The horizontal dashed line at $[\mathrm{X/H}] = 0$ represents solar metallicity. The horizontal dotted line at $[\mathrm{X/H}] = -2.4$ represents the VMP limit as defined above. Note that the lack of data between $\log N_{\mathrm{H\,I}} = 20$ and 20.3 is purely artificial, resulting from our grid of photoionization models stopping at $\log N_{\mathrm{H\,I}} = 20$.

Figure 19 demonstrates even to a greater extent the striking metallicity changes as a function of $N_{\mathrm{H\,I}}$. Absorbers with $[\mathrm{X/H}] \gtrsim -2.4$ are observed at any $N_{\mathrm{H\,I}}$, but absorbers with $[\mathrm{X/H}] < -2.4$ are observed only at $\log N_{\mathrm{H\,I}} \lesssim 19.8$. While the VMP definition implies that this is VMP gas for DLAs, for low-$N_{\mathrm{H\,I}}$ absorbers the VMP definition corresponds to the median value for SLFSs and is about 0.3 dex below the median of pLLSs and LLSs. As already pointed out above, the frequency of VMP absorbers also increases as $N_{\mathrm{H\,I}}$ decreases, which is quantitatively summarized in Table 6. We therefore also define extremely metal-poor absorbers as absorbers with $[\mathrm{X/H}] < -3$, which are mostly observed at some level in
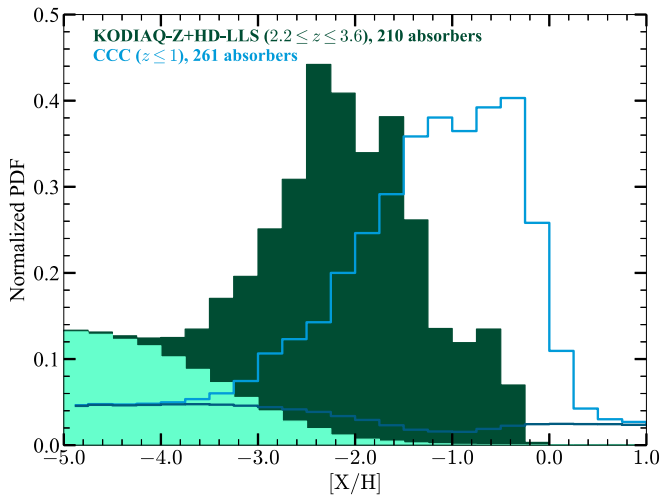
Figure 20. Comparison of the posterior metallicity pdf's of the absorbers with $15 \leqslant \log N_{\rm H\,I} < 19$ at high and low redshifts. The light-colored region indicates the contribution from the upper limits. For CCC, the light-colored histogram indicates the contributions from upper and lower limits.

Figure 21. The metallicity as a function of $N_{\rm H\,I}$ at low z (CCC) and high z (KODIAQ-Z, HD-LLS, R12-DLA). *For the $z < 1$ absorbers, we subtract the metallicity of each absorber by 0.9 dex* (see text for more details). For the black circles, lighter symbols are upper limits on the metallicity. For the blue plus signs, lighter symbols are upper limits if $[X/H] \leqslant -1$ and lower limits if $[X/H] > -1$. For clarity, we did not plot the error bars (which are available in Figure 19 and Figure 9 of CCCIII for the low-redshift sample). Median metallicities are plotted, but we treat upper and lower limits as in Figure 19, i.e., the lower and upper values represent the 10th and 90th percentiles, respectively.

absorbers with $\log N_{\rm H\,I} \lesssim 18$; see Table 6.[21] Extremely metal-poor absorbers have metallicities that are so low that only Population III stars may have contaminated them if they have some metals (see, e.g., Frebel et al. 2007; Crighton et al. 2015).

We finally note that the metallicity dip around -1.1 dex observed especially in the SLFS metallicity pdf is also evident in the nonbinned distribution shown in Figure 19. In that figure, there are a lack of data points at $-1.1 \lesssim [X/H] \lesssim -1$ for H I column density absorbers with $14.5 \lesssim \log N_{\rm H\,I} \lesssim 16.4$. This is reminiscent of the metallicity dip at a similar value of $[X/H] \simeq -1$ observed in the pLLS+LLS pdf at $0.45 \lesssim z \lesssim 1$ (Lehner et al. 2019; see also Section 6.3 for the cosmic evolution of the metallicity).

6.3. Cosmic Evolution of the Metallicity

In Lehner et al. (2016), for the first time, we studied the redshift evolution of the metallicity absorbers with $16.2 \lesssim \log N_{\rm H\,I} < 19$, but with a much smaller sample at both low and high z. The samples of absorbers at all surveyed z have increased at least by a factor 8 and provide far more robust results. Furthermore, we also now sample at low and high z the $N_{\rm H\,I}$ regimes below $\log N_{\rm H\,I} \lesssim 16.2$. In Figure 20, we show the metallicity pdf's at low ($z < 1$) and high ($2.2 \leqslant z \leqslant 3.6$) redshift for H I-selected absorbers with $15 \lesssim N_{\rm H\,I} < 19$. There is some overlap between the two distributions between $-3 \lesssim [X/H] \lesssim -1.4$, but evidently there are more absorbers with $[X/H] \lesssim -2$ at high z and many more with $[X/H] > -1.2$ at $z < 1$. In fact, the latter value corresponds to the median metallicity at $z < 1$ for $16.2 \lesssim \log N_{\rm H\,I} < 19$ absorbers, and only 10% of the high-redshift absorbers have metallicities $[X/H] \geqslant -1.2$. The metallicity median and mean at $z < 1$ are about $+1$ dex higher than at $2.2 \leqslant z \leqslant 3.6$. Both metallicity distributions are negatively skewed with similar IQRs $\simeq 1.5$ dex at low and high z.

In Figure 21, we compare the metallicities of the absorbers (following the same methodology used to produce Figure 19) at $z < 1$ and $2.2 \leqslant z \leqslant 3.6$ over the H I column density range $14.5 \lesssim N_{\rm H\,I} < 22$. For the low-redshift absorbers, we shifted the metallicity by -0.9 dex, which is about the average difference between the mean/median metallicity of the absorbers at low and high z (-0.8 dex for DLAs, -1 dex for lower-$N_{\rm H\,I}$ absorbers). The notable feature from this figure is that despite different metallicity behaviors at low and high H I column densities, the offset metallicities and the overall dispersions at low and high z show a striking overlap at any given $N_{\rm H\,I}$ (where there is a good sampling in both data sets). A closer look reveals a small excess of higher-metallicity absorbers in the SLFS, SLLS, and DLA regimes at high z (i.e., a slightly larger dispersion of the metallicity distribution). However, overall the major change between $2.2 \leqslant z \leqslant 3.6$ and $z < 1$ absorbers is an increase of their metallicity by a factor ~ 8.

In Figure 22, we now show the metallicities as a function of the redshift for the absorbers included in CCC, HD-LLS, and KODIAQ-Z (for clarity we do not show the R12-DLA sample in this figure). For the HD-LLS sample, we differentiate the LLSs and SLLSs. This figure strengthens the previous conclusions, showing the overall increase in metallicities as z decreases. At $z < 1$, there is a substantial fraction of $\log N_{\rm H\,I} < 19$ absorbers with $-0.15 \lesssim [X/H] \lesssim +0.15$, while these are mostly absent at higher redshift. On the other hand, at $z < 1$ there is no evidence of population of absorbers with $[X/H] \lesssim -3$, while there is at $2.2 \leqslant z \leqslant 3.6$. As already noted by Lehner et al. (2016) with a much smaller sample, while the VMP threshold increases by 1 dex over $2.2 \leqslant z \leqslant 3.6$ to $z < 1$ (the 5% quantile of DLA metallicities from -2.4 to -1.4 dex), a similar fraction of about 50% of $\log N_{\rm H\,I} \lesssim 18$ VMP absorbers at low and high z lie in the respective VMP regime.

In Figure 22, there is also some evidence that the metallicity distribution changes to a certain degree around $z \sim 2.85$. Considering the KODIAQ-Z and HD-LLS sample with

---

[21] We emphasize that VMP DLAs with $[X/H] < -2.4$ exist at $z \sim 2$–4 but are found only in targeted searches (e.g., Cooke et al. 2011, 2016)—they are uncommon and would require a much larger sample to be present in an H I selected sample.
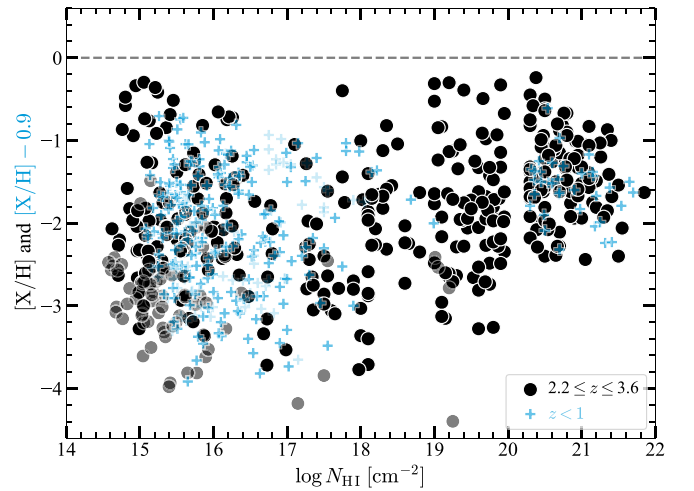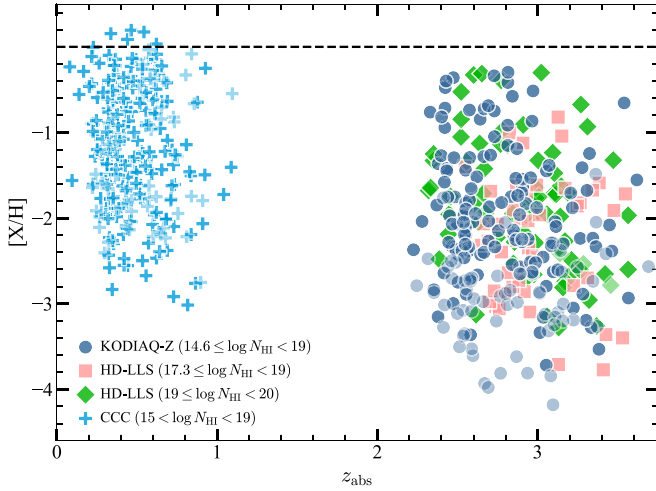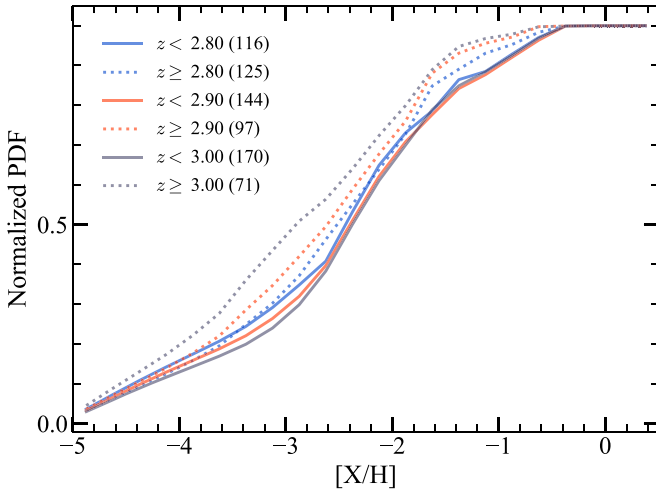
Figure 22. The metallicity of different absorbers as a function of the redshift. For the HD-LLS and KODIAQ-Z surveys, lighter symbols are upper limits on the metallicity. For CCC, lighter symbols are upper limits if $[X/H] \leqslant -1$ and lower limits if $[X/H] > -1$. Median metallicities are plotted for the CCC, KODIAQ-Z, and HD-LLS surveys, but we treat upper and lower limits as in Figure 19, i.e., the lower and upper values represent the 10th and 90th percentiles, respectively.

Figure 24. The metallicities of closely redshift-separated KODIAQ-Z absorbers (aka paired absorbers) as a function of their absolute velocity difference. For the upper limits (blue triangles), we use the upper bound of the 80% CI to be the most conservative. Each paired absorber is connected with a dotted line.

6.4. Metallicity Variation over Small Velocity Scales

Due to the treatment of the absorbers where we separate individual H I components as much as possible (see Section 3), our KODIAQ-Z survey also provides information on the metallicity variation over small redshift/velocity separation between absorbers. Following Lehner et al. (2019), we define here paired absorbers as absorbers that are closely separated in redshift space so that $\Delta v \equiv |(z_{abs}^2 - z_{abs}^1)/(1 + z_{abs}^1)|c| < 500$ km s^{-1}. We have a sample of 36 such paired absorbers.

In Figure 24, we show the median or upper limit on the metallicities of the 36 paired absorbers as a function of their absolute velocity separation, where the two absorbers of each pair are connected by a dotted vertical line. In Figure 25, we show in the left panel the absolute metallicity difference ($\Delta[X/H]$) between the higher and lower median metallicities of the paired absorbers against their absolute velocity separation and in the right panel the histogram distribution of $\Delta[X/H]$. From these two figures, at any velocity separation, there is a large scatter from about 0.2 to >2 dex. At $\Delta v \lesssim 150$ km s^{-1}, the paired absorbers are more clustered above $[X/H] \simeq -2.4$ and their metallicity difference is smaller on average than for larger Δv. Using the survival method to account for the limits (Feigelson & Nelson 1985), we estimate a mean $\langle \Delta[X/H] \rangle = 0.68 \pm 0.14$ (error on the mean), and the IQR in $\Delta[X/H]$ is [0.25, 0.70] for paired absorbers separated by $\Delta v \lesssim 150$ km s^{-1} (sample size 17). On the other hand, paired absorbers separated by $\Delta v > 150$ km s^{-1} have $\langle \Delta[X/H] \rangle = 1.67 \pm 0.21$ (error on the mean), and the IQR in $\Delta[X/H]$ is [0.64, 2.10] (sample size 19). Therefore, paired absorbers with larger velocity separation show larger metallicity variation on average. For the entire sample, we estimate a mean $\langle \Delta[X/H] \rangle = 1.20 \pm 0.16$, and the IQR in $\Delta[X/H]$ is [0.38, 1.81]. Ignoring the velocity separation, we also observe that there are two main clusters separated at $\Delta[X/H] \simeq 1$ with about 50% of the sample in each bin and with a mean $\langle \Delta[X/H] \rangle = 0.39 \pm 0.05$ and 2.15 ± 0.13 below and above this limit, respectively. Therefore, for about half of the paired absorbers, there is evidence for metallicity variations over $\Delta v \lesssim 500$ km s^{-1} of a factor 2–3 and for the other half of

Figure 23. Comparison of the metallicity CDFs of the absorbers in KODIAQ-Z and HD-LLS with $14.5 \leqslant \log N_{H I} < 19$ above and below a given redshift threshold, $z_{th} = 2.8$, 2.9, and 3.0. The numbers between parentheses indicate the number of absorbers in each redshift interval.

$\log N_{H I} < 19$ absorbers at $z > 2.85$, the fraction of absorbers with $[X/H] \geqslant -1$ is very small ($4/112 \simeq 3.6\%$), while at $z \leqslant 2.85$ it is a factor ~ 4 larger ($17/130 \simeq 13.1\%$). To explore that potential of the change of the metallicity with z, in Figure 23 we show the comparison between the metallicity CDFs of the absorbers with $14.5 \leqslant \log N_{H I} < 19$ from KODIAQ-Z and HD-LLS above and below several redshift thresholds, $z_{th} = 2.8$, 2.9, and 3.0. The large number of absorbers with $[X/H] \gtrsim -1$ at $z \lesssim 2.9$ compared to higher z is further demonstrated in this figure. However, this figure also shows that at $z \gtrsim 2.9$, and even more so at $z \geqslant 3.0$, there is a much larger fraction of metal-poor absorbers with $[X/H] \lesssim -2.8$ than at lower redshifts, demonstrating that within the redshift interval $2.2 \lesssim z \lesssim 3.6$ there is a change in the metallicity pdf's with an overall increase of the metallicity at $2.2 \lesssim z < 2.8$ compared to $2.8 \lesssim z \lesssim 3.6$.
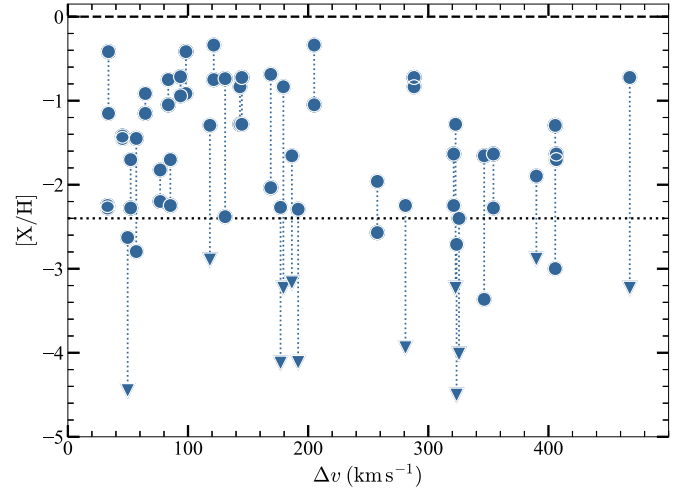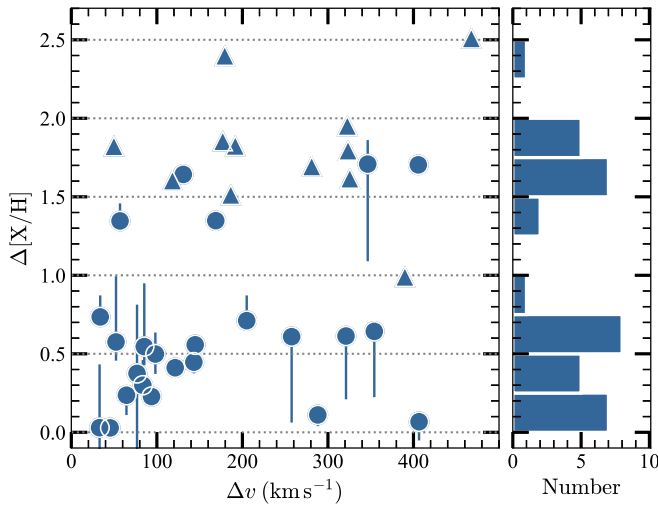
Figure 25. Left: scatter plot of the difference between the higher and lower median metallicities for the paired absorbers as a function of their absolute velocity difference. For the lower limits (blue triangles), we use the lower bound of the 80% CI to be the most conservative. Right: distribution of the metallicity differences between paired absorbers.
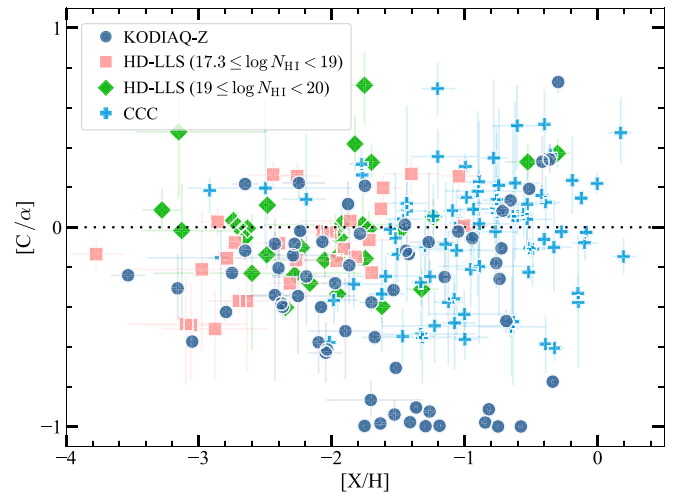

Figure 26. The [C/α] ratio against the metallicity of the absorbers from the KODIAQ-Z ($2.2 \lesssim z \lesssim 3.6$), HD-LLS ($2.2 \lesssim z \lesssim 3.6$), and CCC ($z \lesssim 1$) surveys. The median values of the [C/α] and [X/H] posterior pdf's are adopted with 68% CI. Note that [C/α] values with [C/α] < −0.6 are most likely predominantly caused by ionization corrections rather than nucleosynthesis effects (see text for more details).

a factor >140. For absorbers with $\Delta v < 150$ km s^{-1}, the metallicity variation is on average around a factor 5 but grows to ∼50 for absorbers with $\Delta v > 150$ km s^{-1}.

Due to the large metallicity variation over $\Delta v \lesssim 500$ km s^{-1}, a posteriori it is justified to treat these paired absorbers as individual absorbers. The other direct consequence is that, when possible, components in absorbers should be treated individually to derive the true metallicity of the gas rather than a column-density-weighted average metallicity between very different types of gaseous components. This was a consequence first inferred by Prochter et al. (2010) at similar redshifts for one LLS and then in the CCC survey at $z < 1$ for 30 paired absorbers (Lehner et al. 2019; Wotta et al. 2019; see also Kacprzak et al. 2019; Zahedy et al. 2021). However, we also find that for the smaller $\Delta v \lesssim 150$ km s^{-1} separation the metallicity variation is substantially smaller than for the paired absorbers with larger Δv. We discuss the implications of these findings in Section 10.4 in more detail.

6.5. The [C/α] Ratio as a Function of the Metallicity

In the photoionization modeling used to determine the metallicities of the absorbers, the [C/α] ratio is allowed to vary in the range $-1 \lesssim [C/\alpha] \lesssim +1$ to take into account possible nucleosynthesis effects that likely arise from the time lag in the production between α-elements and carbon (e.g., Cescutti et al. 2009; Mattsson 2010). The general trend between [C/α] and [α/H] seen in stars (e.g., Akerman et al. 2004; Fabbian et al. 2010) and SLLSs/DLAs (e.g., Pettini et al. 2008; Penprase et al. 2010; Cooke et al. 2011) is (1) a decrease of [C/α] from about [C/α] $\simeq 0$ at [α/H] $\gtrsim 0$ to [C/α] $\simeq -0.6$ at [α/H] $\simeq -0.5$, (2) a plateau in [C/α] from [α/H] $\simeq -0.5$ to [α/H] $\simeq -2$, and (3) an upturn in [C/α] at [α/H] $\leqslant -2$, where [C/α] increases up to [C/α] $\gtrsim +0.2$ at [α/H] $\leqslant -2.8$.

As discussed in Lehner et al. (2019) at low redshift and Lehner et al. (2016) at high redshift, the SLFSs, pLLSs, and LLSs do not quite follow this overall trend. In fact, little trend is observed between [C/α] and [α/H] for these absorbers, with a large scatter in [C/α] at any metallicity. While no obvious trend is observed between [C/α] and [α/H] for these

absorbers, Lehner et al. (2019) note three characteristics: (1) a lower floor level for the distribution at about [C/α] $\simeq -0.6$, similar to that observed for the stars, H II regions, and SLLSs/DLAs; (2) an upper ceiling of the distribution at [C/α] $\simeq +0.5$, but with a concentration of data around $0 <$ [C/α] $\sim +0.3$, similar again to the highest values observed in stars, H II regions, and SLLSs/DLAs; and (3) a tentative upturn in [C/α] at [X/H] $\lesssim -2$, which would be similar to the one observed in stars and SLLSs/DLAs.

With KODIAQ-Z and our remodeling of the HD-LLS absorbers, we can now revisit these conclusions with a much larger sample at high redshift, especially in the low-metallicity regime. In Figure 26, we show [C/α] as a function of [X/H] for all the absorbers in KODIAQ-Z, HD-LLS, and CCC, where the relative abundance and metallicity are simultaneously constrained. In that figure, we use different symbols to differentiate LLSs from SLLSs in the HD-LLS survey (there is only one SLLS in KODIAQ-Z with information on [C/α]; see Figure 12; we did not highlight this in Figure 26). Excluding the 13 absorbers with [C/α] $\lesssim -0.9$ (because those are more largely affected by ionization; see Section 5.3 and below), the addition of the KODIAQ-Z and HD-LLS data confirms the first two characteristics described above and summarized in Lehner et al. (2019). However, at [X/H] $\lesssim -2$, there is no clear upturn in [C/α] for the absorbers with log $N_{\rm H\,I} < 19$, although the scatter in the [C/α] distribution becomes smaller. On the other hand, considering the SLLSs from HD-LLS (diamond symbols in Figure 26), the upturn in [C/α] is apparent when the metallicity becomes smaller than at [X/H] $\lesssim -2.2$.

Since the derived ratio of [C/α] is controlled by both ionization corrections and nucleosynthesis effects, it is plausible that the ionization corrections wash away some of the trends that would appear if only nucleosynthesis effects were at play (such as for the DLAs). In Section 5.3.4, we show that [C/α] in the KODIAQ-Z sample is more affected by ionization than in the two other samples in view of the correlation between [C/α] and $f_{\rm H\,I}$. However, no such correlation is observed in CCC and HD-LLS, and the overall

Table 7
Physical Properties of the KODIAQ-Z Sample

QSO Name	z_{abs}	$\log N_{H\,I}$ [cm^{-2}]	$\log N_H$ [cm^{-2}]	$\log n_H$ [cm^{-3}]	$\log T$ [K]	l (kpc)
J010806+163550	2.31582	14.574 ± 0.004	17.935, 18.635, 19.305	−3.737, −3.161:, −2.565	4.387, 4.521, 4.644	0.1, 2.0, 35.7
J045213−164012	2.44044	14.611 ± 0.004	17.944, 18.649, 19.378	−3.785, −3.155:, −2.554	4.381, 4.516, 4.645	0.1, 2.1, 47.2
J170100+641209	2.49647	14.622 ± 0.007	18.409, 18.725, 19.013	−3.479, −3.231, −2.958	4.456, 4.513, 4.566	0.8, 2.9, 10.1
J023359+004938	2.48836	14.640 ± 0.019	18.025, 18.721, 19.371	−3.758, −3.197:, −2.602	4.391, 4.524, 4.638	0.1, 2.7, 43.8
J073149+285448	3.42673	14.653 ± 0.008	17.955, 18.665, 19.333	−3.880, −3.277:, −2.670	4.363, 4.502, 4.584	0.1, 2.8, 52.9
J045213−164012	2.53542	14.674 ± 0.003	18.047, 18.769, 19.434	−3.792, −3.214:, −2.597	4.389, 4.527, 4.640	0.1, 3.1, 54.5
J012156+144823	2.81222	14.683 ± 0.002	18.034, 18.687, 19.368	−3.765, −3.174:, −2.616	4.382, 4.508, 4.621	0.1, 2.4, 44.0
J105756+455553	3.69194	14.698 ± 0.003	17.959, 18.656, 19.307	−3.865, −3.273:, −2.678	4.352, 4.485, 4.564	0.1, 2.8, 48.1
J010925−210257	2.96546	14.706 ± 0.008	19.558, 19.789, 19.956	−4.401, −4.202, −3.967	4.580, 4.612, 4.621	108.6, 318.0, 737.9
J014516−094517	2.69430	14.708 ± 0.002	18.159, 18.564, 18.980	−3.390, −3.040, −2.695	4.392, 4.472, 4.554	0.2, 1.3, 7.6

Note. The lower and upper bounds for each quantity represent the 68% CI, while the middle value is the median value. When a colon is present after the median value of $\log n_H$, this implies that a $\log U$ Gaussian prior was used to determine the properties of the absorber (see Section 5.2).

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

dispersion is similar in these three surveys. The mean values for [C/α] in CCC and HD-LLS are also the same (see Section 5.3.4), despite widely different ionization corrections. Therefore, the [C/α] ratio with [C/α] > −0.6 determined from the ionization modeling does not seem to have a strong dependence on the level of the ionization correction. Nevertheless, the ionization correction essentially adds noise in the [C/α] distribution at the level of ±0.2–0.3 dex at any $N_{H\,I}$ and metallicity, which is large enough to hide subtle changes of [C/α] with [X/H], which are expected to be at the level of 0.2–0.5 dex (see above). Hence, the large scatter at any metallicities observed in Figure 26 must be largely caused by ionization, while nucleosynthesis must still control the range $-0.6 \lesssim [C/\alpha] \lesssim +0.5$, where most of the absorbers lie based on the similar ranges in [C/α] observed in DLAs, stars, and H II regions.

7. Physical Properties of the SLFSs, pLLSs, LLSs, and SLLSs at $2.2 \lesssim z \lesssim 3.6$

The photoionization models discussed in Section 5 also predict the physical properties of the absorbers. In particular, these include the neutral fraction of the gas (f_{HI}), the density of the gas (n_H), the total H column density ($N_H \equiv N_{H\,I} + N_{H\,II}$), the length scale ($l \equiv n_H/N_H$), and the temperature of the gas (T). These can in turn be used to constrain the cosmological baryon and metal budgets for the absorbers (Fumagalli et al. 2016b; Lehner et al. 2019). An important caveat to keep in mind and discussed at length in Fumagalli et al. (2016b) is that there is a degeneracy between the density and intensity of the radiation field. This dependency produces much less robust inference on the density and sizes of absorbers than for the metallicity, which depends only on the shape of the radiation field. We also remind the reader that only photoionization by the EUVB has been used in this work, and it is quite plausible that more than one ionization process is at play, which would also affect more the physical quantities discussed in this section.

7.1. Densities, H Column Densities, Temperatures, and Length Scales

We summarize the physical quantities (n_H, N_H, l, and T) derived from our modeling for the KODIAQ-Z sample in Table 7. Values of $\log n_H$ with a trailing colon represent

absorbers for which we adopted a Gaussian prior on $\log U$ and therefore on $\log n_H$. In Table 8, we summarize the statistics of the physical properties derived for the entire sample (median, mean, and dispersion of $N_{H\,I}$, n_H, N_H, l, and T) and a restricted sample where only a flat prior on $\log U$ was used. Comparing the results from Table 8 between the entire and restricted samples, it is apparent that the results are not statistically different. We emphasize that most of the pLLSs, LLSs, and SLLSs did not require a Gaussian prior on $\log U$. However, for about 45% of the pLLSs we had to use a Gaussian prior on $\log U$, and yet the results are essentially the same in both samples. Hereafter, we therefore consider the results from the entire sample. In Figure 27, we show the posterior pdf's of N_H, n_H, and l for the SLLSs and LLSs (using results from KODIAQ-Z and HD-LLS) in the top two rows and the pLLSs and SLFSs (from KODIAQ-Z) in the bottom two rows. These pdf's contrast remarkably from the metallicity pdf's with much narrower distributions.

Temperatures of ionized gas predicted by CLOUDY are based on the balance of atomic heating (photoionization) and cooling (recombination and collisional effects). As such, temperatures are dependent on the radiation field shape and intensity, density, and metallicity of the models. The distribution of predicted temperatures has overall a small dispersion, and there is a clear trend of decreasing temperatures of about 0.1–0.2 dex between each absorber category from the SLFSs, pLLSs, LLSs, to SLLSs (see Table 8), i.e., the temperature of the gas decreases as the H I column density increases. For the SLFSs, the average temperature is around 3×10^4 K, while it is a factor 2 less on average for the SLLSs. In Section 4.2, we show that for about 90% of the H I components $\langle b \rangle = 27 \pm 6$ km s^{-1}. This places an upper limit on the temperature from the observations of $T < 4 \times 10^4$ K, consistent with the ionization modeling. It also implies that for the SLFSs, on average, the thermal broadening component is somewhat larger than the nonthermal component, while for the other absorbers nonthermal motions take over the thermal ones. Compared to Lehner et al. (2019), the predicted temperatures of the $z < 1$ SLFSs, pLLSs, and LLSs are about 1.4 times smaller than at $2.2 \lesssim z \lesssim 3.6$.

Considering the SLFSs, pLLSs, and LLSs, the total H column densities decrease with decreasing $N_{H\,I}$, but their distributions overlap within 1σ dispersion (which are large for

Table 8
Statistical Physical Properties of the Absorbers

Absorber Type	$\widetilde{\log N_{\rm H\,I}}$ [cm^{-2}]	$\overline{\log N_{\rm H\,I}}$ [cm^{-2}]	$\widetilde{\log n_{\rm H}}$ [cm^{-3}]	$\overline{\log n_{\rm H}}$ [cm^{-3}]	$\widetilde{\log N_{\rm H}}$ [cm^{-2}]	$\overline{\log N_{\rm H}}$ [cm^{-2}]	$\widetilde{\log l}$ [kpc]	$\overline{\log l}$ [kpc]	$\widetilde{\log T}$ [K]	$\overline{\log T}$ [K]
				Entire Sample						
SLFSs	15.36	15.41 ± 0.42	-3.23	-3.25 ± 0.57	19.59	19.47 ± 0.68	1.31	1.23 ± 1.19	4.46	4.46 ± 0.14
pLLSs	16.40	16.52 ± 0.27	-2.60	-2.69 ± 0.54	19.97	19.95 ± 0.70	1.33	1.15 ± 1.21	4.36	4.37 ± 0.11
LLSs	18.10	17.97 ± 0.43	-2.56	-2.55 ± 0.50	20.78	20.80 ± 0.64	1.86	1.85 ± 1.12	4.29	4.29 ± 0.09
SLLSs	19.58	19.50 ± 0.31	-2.25	-2.26 ± 0.76	20.69	20.80 ± 0.68	1.41	1.57 ± 1.43	4.19	4.19 ± 0.15
				Restricted Sample without the log U Constraint						
SLFSs	15.30	15.35 ± 0.42	-3.20	-3.22 ± 0.44	19.41	19.38 ± 0.58	1.05	1.12 ± 0.94	4.50	4.48 ± 0.11
pLLSs	16.44	16.55 ± 0.28	-2.62	-2.70 ± 0.53	19.99	19.99 ± 0.60	1.14	1.20 ± 1.19	4.38	4.38 ± 0.11
LLSs	18.10	17.96 ± 0.42	-2.56	-2.55 ± 0.49	20.80	20.80 ± 0.63	1.86	1.86 ± 1.10	4.29	4.30 ± 0.08
SLLSs	19.54	19.49 ± 0.31	-2.26	-2.28 ± 0.75	20.71	20.81 ± 0.68	1.46	1.60 ± 1.42	4.20	4.19 ± 0.15

Note. The tilde and bar above each parameter represent the median and mean (with the standard deviation), respectively, of the logarithmic values of the parameter. The entire sample consists of 321 absorbers: 155 SLFSs, 24 pLLSs, 62 LLSs, and 80 SLLSs. The restricted sample without the log U constraint contains 243 absorbers (i.e., absorbers that were modeled with a flat prior on log U): 85 SLFSs, 21 pLLSs, 59 LLSs, and 78 SLLSs. Note that $N_{\rm H\,I}$ is estimated directly from the data, while all other quantities are estimated from our photoionization modeling. We include the results from the HD-LLS survey for the LLSs and SLLSs.

these quantities; see Table 8 and Figure 27). The differences of the mean $\log N_{\rm H\,I}$ for SLFSs compared with the pLLSs and of the pLLSs compared with the LLSs are a factor 4 and 28, respectively. In contrast, the differences in total H are significantly smaller with differences of a factor $\sim$4 and 6.5 between the SLFSs and pLLSs and between the pLLSs and LLSs, respectively. Thus, despite their lower $N_{\rm H\,I}$ values, the contributions of pLLSs and SLFSs to the total mass and baryon budgets are higher compared with the LLSs than their H I columns would suggest, especially since they are more numerous (e.g., Prochaska et al. 2014; see also Section 8). The $N_{\rm H}$ column densities for SLLSs are similar to those for LLSs; $N_{\rm H}$ is about the same even though $N_{\rm H\,I}$ is on average a factor 34 smaller for the LLSs than for SLLSs. Given the higher incidence of LLSs, they make a larger contribution to the cosmic baryon budget than the SLLSs. Compared to Lehner et al. (2019), $N_{\rm H}$ at the $z < 1$ SLFSs, pLLSs, and LLSs are a factor $\sim$10–15 smaller than at $2.2 \lesssim z \lesssim 3.6$.

The estimated hydrogen densities increase as $N_{\rm H\,I}$ increases. The SLLSs have the largest dispersion on $n_{\rm H}$, which is caused by a long tail at low and high $n_{\rm H}$ values (see Figure 27). While there is an increase in $n_{\rm H}$ with increasing $N_{\rm H\,I}$, the bulk of the hydrogen density for the pLLSs, LLS, and LLSs is within about $\log n_{\rm H} \simeq -3.4$ and -1.8, while for SLFSs it is between $\log n_{\rm H} \simeq -4.2$ and -2.5. At $z \sim 2.8$ (the mean redshift of KODIAQ-Z), the mean cosmic density is about $\bar{n}_{\rm H} \simeq 10^{-5}$ cm^{-3} (Schaye 2001a). The pLLSs, LLSs, and SLLSs therefore probe overdensities of $\delta \equiv n_{\rm H}/\bar{n}_{\rm H} - 1 \sim 40$–1600, while for the lower column density SLFSs $\delta \sim 5$–320. A large part of the observed pLLSs, LLSs, and SLLSs largely probe virialized (mini)halos ($\delta > 100$; Schaye 2001a). For SLFSs, there is more a mixture of systems probing overdensities below and above 60, at the transition between the IGM and virialized halos (Schaye 2001a). Compared to the CCCIII, $n_{\rm H}$ at the $z < 1$ SLFSs, pLLSs, LLSs, SLLSs are a factor $\sim$1.5–2.5 smaller than at $2.2 \lesssim z \lesssim 3.6$. However, due to the redshift dependence of the mean cosmic density that is $\propto ((1+z)/4)^3$ (Schaye 2001a), the density contrast at $z < 1$ for the SLFSs and other stronger H I absorbers is always $\delta > 100$, even for the lower densities probed by SLFSs (although we note that at $z < 1$ only $\log N_{\rm H\,I} > 15$ SLFSs could be probed).

The length scales, which are directly related to $n_{\rm H}$ and $N_{\rm H}$, have the largest dispersion among all these physical quantities. For the SLFSs and pLLSs, the length-scale IQR is about $4 \lesssim l \lesssim 150$ kpc. For the LLSs, the IQR for l is $18 \lesssim l \lesssim 288$ kpc, and for the SLLSs it is $8 \lesssim l \lesssim 80$ kpc. In contrast, at $z < 1$, the length-scale IQR of these absorbers is $0.3 \lesssim l \lesssim 3.2$ kpc (Lehner et al. 2019). At $2.2 \lesssim z \lesssim 3.6$, the SLFSs, pLLSs, LLSs, and even the SLLSs may probe a variety of structures from gas embedded in virialized halos to large-scale structures. Recent MUSE galaxy surveys show strong hints that at least LLSs may probe a wide variety of environments at $z > 2$, including gaseous filaments, CGM of galaxies, or intragroup gas (e.g., Fumagalli et al. 2016a, 2017; Lofthouse et al. 2020).

7.2. Neutral Fractions

The behavior of the neutral fraction with $N_{\rm H\,I}$ can be used to constrain the cosmic baryon and metal budgets (see Section 8). In Figure 28, we show the neutral fraction of the absorbers estimated from the CLOUDY models as a function of $N_{\rm H\,I}$ from $\log N_{\rm H\,I} \simeq 14.6$ to 20 using KODIAQ-Z and HD-LLS (see Section 5.2; see also the recent work for the SLLSs from Berg et al. 2021). There is a clear correlation between $f_{\rm H\,I}$ and $N_{\rm H\,I}$. This is confirmed by the Spearman rank-order test that shows a strong positive monotonic correlation between $f_{\rm H\,I}$ and $N_{\rm H\,I}$ with a correlation coefficient $r_{\rm S} = 0.87$ and a p-value $\ll 0.05\%$. However, the slope of the correlation is shallower between the SLFS, pLLS, and LLS regimes compared to the LLS and SLLS regimes. The dashed red line in Figure 28 is a linear fit to the data with $14.6 \leqslant \log N_{\rm H\,I} < 19$, yielding $\log f_{\rm H\,I} = (0.51 \pm 0.04) \log N_{\rm H\,I} - 12.00$. The slope is quite similar to that found at $z < 1$ for absorbers with $15.1 \leqslant \log N_{\rm H\,I} < 19$ (Lehner et al. 2019). The dashed–dotted magenta line is a linear fit to the data with $17.2 < \log N_{\rm H\,I} \leqslant 20$, $\log f_{\rm H\,I} = (1.00 \pm 0.08) \log N_{\rm H\,I} - 20.76$, which is consistent with the linear fit derived in the HD-LLS survey for the LLSs +SLLSs (Fumagalli et al. 2016b). The Spearman rank-order test also produces a similar conclusion: for the pLLSs+LLSs, the correlation coefficient is $r_{\rm S} = 0.65$ and the p-value is $\ll 0.05\%$, while for the SLFSs, it drops to $r_{\rm S} = 0.27$ and a p-value $= 0.011$. There is a similar effect if we consider $N_{\rm H}$
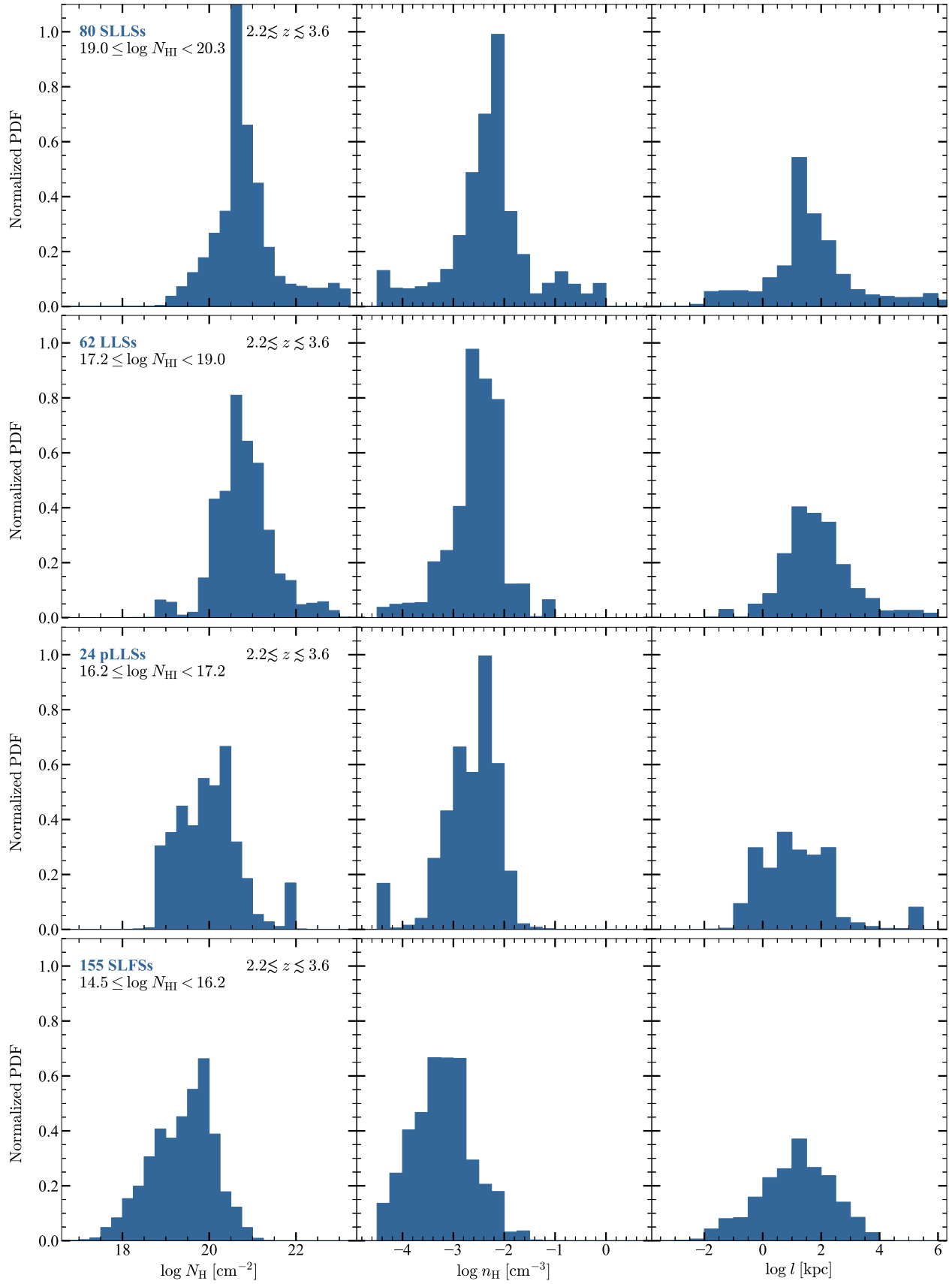
Figure 27. Posterior pdf's of the total H column density (left), H density (middle), and path length (right) for SLLSs, LLSs, pLLSs, and SLFSs for the entire sample from the KODIAQ-Z survey and HD-LLSs.
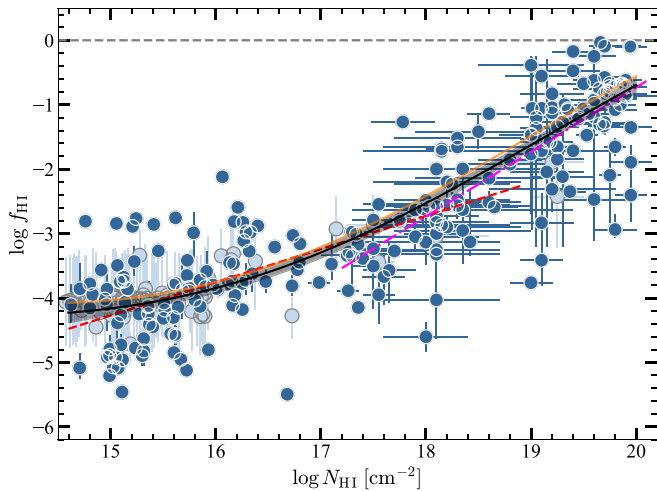
Figure 28. The neutral fraction of H I vs. $N_{\rm H\,I}$ using the KODIAQ-Z and HD-LLS surveys. The median values of the posterior pdf's are adopted with 68% CI. Light-blue data were estimated using a Gaussian prior on $\log U$, while the other data have a flat prior on $\log U$. The black curve is a GP-derived model to the entire $N_{\rm H\,I}$ range shown in the figure (the dark area corresponds to the dispersion derived from the GP model). The dashed red line is a linear fit to the data with $14.6 \leqslant \log N_{\rm H\,I} < 19$ (i.e., the SLFSs, pLLSs, and LLSs). The dashed–dotted magenta line is a linear fit to the data with $17.2 < \log N_{\rm H\,I} \leqslant 20$ (i.e., the LLSs and SLLSs). For all the models, we consider only data with a flat prior on $\log U$ (although the results are similar if the entire data set was considered instead). The orange curve is a GP-derived model to the entire $N_{\rm H\,I}$ range but considering the neutral fraction derived using the HM12 EUVB instead of HM05.

instead of $f_{\rm H\,I}$. This trend can be understood as follows: as the gas becomes more optically thick, $N_{\rm H\,I}$ increases more rapidly relative to the total H column density.

To model this overall trend between $f_{\rm H\,I}$ and $N_{\rm H\,I}$, we use a Gaussian process (GP) model, providing a generic supervised learning method designed to solve a regression. We refer the reader to Appendix F in Lehner et al. (2020) for more details, but in short, the advantage of this method is that the prediction interpolates the observations in a nonparametric way and is probabilistic so that empirical confidence intervals can be computed. We use the Python GAUSSIAN PROCESS REGRESSION package within SCIKIT-LEARN (Pedregosa et al. 2011; Buitinck et al. 2013) to model the data with a squared-exponential kernel (the use of a more complex kernel like a Matern kernel would yield similar results when using similar bounds). The overall scatter in the data is typically much larger than the errors on a single data point, and we use the 1σ dispersion derived for SLFSs, pLLSs, LLSs, and SLLSs to inform the GP model of the intrinsic scatter in the data (this can be understood as a prior factor to smooth out the scatter of the data). The black curve in Figure 28 is the GP model, with the dark area showing the standard deviation determined by the GP model. There is a large overlap with the two linear fits, but the GP fit has the advantage of providing a global model to the data without a priori selecting manually the inflection point in the curve. The black curve is well fitted by a third-order polynomial, $\log f_{\rm H\,I} = -0.0104x^3 + 0.6268x^2 - 11.62x + 64.23$ (where $x = \log N_{\rm H\,I}$). Finally, we show in Figure 28 with the orange line a GP model if instead we had used the results obtained with the HM12 EUVB. This model is about within the standard deviation of the GP model using the HM05 data, demonstrating that a change of the EUVB from HM05 to HM12 would not have a large impact on the relationship between $f_{\rm H\,I}$ and $N_{\rm H\,I}$ (and

therefore on the estimates of the cosmological baryon and metal budgets).

8. Cosmological Budgets

8.1. Cosmological Baryon Budget

Observations of the LYAF show that about 80%–90% of the baryons at $2 < z < 4$ are found in the cool photoionized phase of IGM (e.g., Rauch et al. 1997; Weinberg et al. 1997; Kim et al. 2001). In Lehner et al. (2014), we show that the photoionized and collisionally ionized gas found in $\tau_{\rm LL} > 1$ systems is very likely the second-largest contributor to the baryon budget at $z \sim 2$–3. However, this result could be affected by small number statistics, as the Lehner et al. (2014) sample included only seven LLSs and eight SLLSs. Our earlier study also did not consider the baryon contributions from the SLFSs and pLLSs. Here we review the impact of the cool, photoionized gas in these absorbers on the cosmic baryon budget at $2.2 \lesssim z \lesssim 3.6$.

The mass gas density relative to the critical density, Ω_g, can be estimated via (e.g., Tytler 1987; Lehner et al. 2007; O'Meara et al. 2007; Prochaska et al. 2010; CCCIII):

$$\Omega_g = \frac{\mu_{\rm H} m_{\rm H} H_0}{\rho_c\, c} \int \frac{N_{\rm H\,I}}{f_{\rm H\,I}} \mathcal{F}(N_{\rm H\,I}) dN_{\rm H\,I}, \qquad (1)$$

where $m_{\rm H} = 1.673 \times 10^{-24}$ g is the mass of hydrogen, $\mu_{\rm H} = 1.3$ corrects for the presence of helium, $c = 2.9979 \times 10^{10}$ cm s^{-1} is the speed of light, $\rho_c = 3H_0^2/(8\pi G) = 8.62 \times 10^{-30}$ g cm^{-3} is the $z = 0$ critical density, $H_0 = 2.20 \times 10^{-18}$ s^{-1} is the Hubble constant,[22] and $\mathcal{F}(N_{\rm H\,I})$ is the column density distribution function. For $\mathcal{F}(N_{\rm H\,I})$, we adopt the spline function model from Prochaska et al. (2014) at $z \approx 2.5$ (see their Table 2 and Figure 7). For the neutral fraction, $f_{\rm H\,I}$, we adopt the GP model from Section 7.2. Integrating the above equation over the different $N_{\rm H\,I}$ regimes of the SLFSs, pLLSs, LLSs, and SLLSs, we find that $\Omega_g \simeq (3.3, 1.1, 1.1, 1.2) \times 10^{-3}$, respectively. The contributions of Ω_g to the total baryon density are $\Omega_g/\Omega_b \simeq 6.8\%$, 2.3%, 2.3%, and 2.5% for the SLFSs, pLLSs, LLSs, and SLLSs, respectively ($\Omega_b = 0.0486$; Planck Collaboration et al. 2016). The values for the LLSs and SLLSs are consistent with the ranges in Lehner et al. (2014, their Table 6, column with Si IV).[23]

Importantly, these estimates do not include a contribution from the highly ionized, hotter gas that is not photoionized. Lehner et al. (2014) found that the O VI collisionally ionized gas in LLSs and SLLSs could contribute 0.4%–10% to the baryonic budget. As noted in Section 5.1, for the pLLSs and SLFSs we often find that the O VI absorption is narrow and aligns with the absorption seen in H I (and other lower metal ions if they are detected), i.e., the properties of the O VI are quite different as $N_{\rm H\,I}$ decreases to $\log N_{\rm H\,I} < 17$ compared to stronger H I absorbers. The observed column densities for these narrow O VI components are also well reproduced by single-phase photoionization models such as those described in Section 5 where the source of the ionizing photons is EUVB. Providing a full description of the O VI associated with the

---

[22] We adopt here the Planck 2015 results (Planck Collaboration et al. 2016).
[23] Although we remind the reader that here the SLLSs include only absorbers with $19 \leqslant \log N_{\rm H\,I} \leqslant 20$ owing to the grid of ionization models stopping at $\log N_{\rm H\,I} = 20$.

KODIAQ-Z absorbers is beyond the scope of this paper, but based on a visual inspection of the profiles and our ionization modeling, very broad, strong O VI absorbers associated with SLFSs and pLLSs are rare, whereas they are common in LLSs, SLLSs, and even DLAs. Therefore, the contribution of broad, strong O VI associated with SLFSs and pLLSs to the cosmic baryon budget is very likely small.

The photoionized gas associated with pLLSs, LLSs, and SLLSs combined therefore contributes $\Omega_g/\Omega_b \simeq 7\%$ and the SLFSs contribute another 7% to the total baryon budget.[24] In contrast, the neutral gas that is mostly found in DLAs yields only $\Omega_g/\Omega_b \simeq 2\%$ at $z \sim 2$–5 (e.g., Prochaska et al. 2005; Wolfe et al. 2005; Lehner et al. 2014; Péroux & Howk 2020). At $z < 1$, Lehner et al. (2019) find fractional contributions to the baryonic budget for the SLFSs, pLLSs, and LLSs that are about 5 times smaller. At $z < 1$, the SLFSs, pLLSs, and LLSs are probing much larger overdensities, and, at these redshifts, only $\sim 5\%$ of the baryons are thought to be in the CGM of galaxies (Shull et al. 2012).

8.2. Cosmological Metal Budget

The metal density of the absorbers can be estimated with the baryon density via $\Omega_m = \Omega_g Z$, where $Z = 10^{[X/H]}\, Z_\odot$ is the metallicity of the gas in mass units and $Z_\odot = 0.0142$ is the bulk solar metallicity in mass units from Asplund et al. (2009). For the metallicity, we use the linear mean metallicities in an about 0.5 dex bin of $N_{H\,I}$ from $\log N_{H\,I} = 14.5$ to $\log N_{H\,I} = 20$. We note that the linear mean metallicities for SLFSs, pLLSs, and LLSs are 0.034, 0.032, and 0.022 $Z_\odot$, respectively. We find $\Omega_m = (15.9, 5.0, 3.6) \times 10^{-7}$ for the SLFSs, pLLSs, and LLSs, respectively. The latter value is consistent with that derived by Fumagalli et al. (2016b), $\Omega_m = 5.1 \times 10^{-7}$. Since we do not have SLLSs with $\log N_{H\,I} > 20$ in our sample, we use the result from the HD-LLS survey where Fumagalli et al. (2016b) derived $\Omega_m = 1.6 \times 10^{-6}$ for the SLLSs. For reasons (sample variance, using the arithmetic mean) discussed in detail in Fumagalli et al. (2016b), these values are uncertain by a factor ~ 3 (see also discussion in Berg et al. 2021). We note that the comoving metal-mass density can be estimated via $\rho_m = \Omega_m \rho_c$ (e.g., Péroux & Howk 2020). Therefore, for the combined photoionized SLFSs, pLLSs, and LLSs, we find $\rho_m = 3.1 \times 10^5\, M_\odot\, \mathrm{cMpc}^{-3}$, while combined with the SLLSs, this increases to $5.1 \times 10^5\, M_\odot\, \mathrm{cMpc}^{-3}$.

The total expected cosmic metal density from the output of stars at $z \sim 2.8$ is about $\Omega_m^{\mathrm{exp}} \simeq 9 \times 10^{-6}$ or, expressed in terms of cosmic metal mass density, $\rho_m^{\mathrm{exp}} \simeq 10^6\, M_\odot\, \mathrm{cMpc}^{-3}$ (Péroux & Howk 2020). Péroux & Howk (2020) also recently estimated for the DLAs, i.e., for the neutral gas, that $\Omega_m \simeq 3.5 \times 10^{-6}$ and for the dust $\Omega_m \simeq 1 \times 10^{-6}$ at $2.2 \lesssim z \lesssim 3.6$, i.e., about 40% of the metals are in the neutral gas and about 10% are in the form of dust at $2.2 \lesssim z \lesssim 3.6$. The SLLSs account for 17.7%, while the SLFSs, pLLSs, and LLSs account for 17.7%, 5.6%, and 4.0%, respectively, and all combined they account for about 45% of the metals at $2.2 \lesssim z \lesssim 3.6$. We note that the recent estimates from Péroux & Howk (2020) for the SLLSs and LLSs are about within 1σ errors from those derived here and Fumagalli et al. (2016b). As for the baryons, another contribution is from the hot O VI gas

associated with LLSs, SLLSs, and DLAs that can contribute another $\gtrsim 7\%$ (Lehner et al. 2014).

While most of the baryons at $2.2 \lesssim z \lesssim 3.6$ are in the LYAF, this is not the case for the metals since the DLAs, dust, SLLSs, LLSs, pLLSs, and SLFSs contribute to about $\gtrsim 95\%$ of the cosmic metal budget. Direct estimates of the metal budgets in the LYAF are uncertain owing to ionization correction and the fact that some of the metals are not solely associated with $\log N_{H\,I} \gtrsim 14.5$ absorbers. Using the recent results on $\Omega_{C\,IV}$ from D'Odorico et al. (2013, 2010; but see also, e.g., Schaye et al. 2003; Simcoe et al. 2004; Songaila 2005) and correcting their estimate from C IV to carbon and carbon to the total metal content yields $\Omega_m \gtrsim 3.4 \times 10^{-7}$ at $z \sim 2.8$ for the "IGM" or $\gtrsim 4\%$ of the total cosmic metal budget (although we note that this may include contributions from higher H I column density absorbers than LYAF absorbers). Overall the photoionized universe probed by absorbers $\log N_{H\,I} < 20$ contains about 50% of the metals at $2.2 \lesssim z \lesssim 3.6$. Combining these results with those from Péroux & Howk (2020), the majority of metals are in the neutral and ionized gas at $z \sim 2.8$.

At $z < 1$, $\Omega_m = (4.6,\ 6.5,\ 25.9) \times 10^{-7}$ for the SLFSs, pLLSs, and LLSs, respectively (Lehner et al. 2019).[25] In terms of comoving mass density, the combined SLFSs, pLLSs, and LLSs at $z < 1$ have $\rho_m = 4.7 \times 10^5\, M_\odot\, \mathrm{cMpc}^{-3}$. Therefore, the amount of metals in SLFSs, pLLSs, and LLSs has increased by a factor of 1.5 from $z \simeq 2.2$–3.6 to $z \lesssim 1$. However, the total expected cosmic metal density at $z \sim 0.5$ is about $\Omega_m^{\mathrm{exp}} \simeq 6 \times 10^{-5}$, corresponding to a cosmic metal mass density $\rho_m^{\mathrm{exp}} \simeq 8 \times 10^6\, M_\odot\, \mathrm{cMpc}^{-3}$ (Péroux & Howk 2020), a factor 7 larger than at $z \sim 2.8$. This implies that the (photoionized) metals associated with the SLFSs, pLLSs, and LLSs account for about 6% of the cosmic metal budget at low redshift, a factor of 4.4 lower than the contribution of the same absorbers at $2.2 \lesssim z \lesssim 3.6$. As shown by Péroux & Howk (2020), the majority of the metals at low redshift remain in galaxies (including stars), in their vicinity, and in the hot intracluster medium, demonstrating a major shift in the distribution of the metals from 2.3 to $\gtrsim 9$ Gyr after the big bang.

9. Comparison to the FOGGIE Simulations

9.1. Motivation and Methodology

In this section, we compare the KODIAQ-Z metallicity measurements as a function of H I column density to the simulated CGM from the FOGGIE project (Peeples et al. 2019, and https://foggie.science/). This project is composed of cosmological zoom-in simulations of six Milky Way–like halos. While KODIAQ-Z selects for absorbers without knowing their galaxy associations, absorbers with the highest column densities probe overdensities that are consistent with being virialized (mini)halos of galaxies (see Sections 7.1 and 9.3). The FOGGIE simulations let us address whether or not the metallicity–H I column density distribution seen in KODIAQ-Z could be explained by only the CGM and nearby IGM of progenitors to Milky Way–mass galaxies.

---

[24] We separate the SLFSs from the other absorbers because, e.g., Kim et al. (2001) included those in the LYAF contribution, where they estimate that 90% of all baryons reside in the LYAF at $1.5 < z < 4$.

[25] In CCCIII, we inadvertently used the mean logarithmic metallicity, instead of the linear mean, which does not change the value for the LLSs but does for the SLFSs and pLLSs, which are, respectively, 0.177 and 0.197 $Z_\odot$. This increases Ω_m by a factor 5 for these absorbers. As the LLSs dominate the metal budget, the effect on the total Ω_m for the SLFSs, pLLSs, and LLSs is only an increase by a factor 1.3, i.e., $\Omega_m \simeq 3.7 \times 10^{-6}$ for these absorbers at $z < 1$ compared to the value listed in CCCIII.

Many groups, including FOGGIE, have recently shown that better spatial resolution in the simulated CGM leads to a natural development of small clouds (Mandelker et al. 2019; Peeples et al. 2019; Rhodin et al. 2019; Suresh et al. 2019; van de Voort et al. 2019), largely because once the cooling length of the gas is resolved, the gas can cool and condense into smaller structures (Hummels et al. 2019; Zheng et al. 2020). Importantly for comparisons to observables such as metallicity, improved spatial resolution in the low-density gas also reduces the artificial mixing of metals, thereby allowing gas to remain at "high" and "low" metallicities even in the CGM, that is, while the median metallicities do not change much, the *scatter* about that median increases with improved resolution (Corlies et al. 2020). In the FOGGIE cosmological zoom-in simulations analyzed here the CGM resolution at $z = 2$ is 91 proper pc in the cooler gas, with a coarsest resolution of 365 proper pc. (The resolutions are correspondingly a factor of 0.85 and 0.75 smaller at $z = 2.5$ and $z = 3$, due to the spatial resolution being at a fixed comoving size.) This spatial resolution profile both enables the development of small, relatively dense and cool cloudlets and allows for the hotter, more rarefied, generally more metal-rich gas to be well resolved. Our analysis uses the $z = 2$, 2.5, and 3 outputs for each of the six FOGGIE halos. As presented in Simons et al. (2020), these halos are selected to be "Milky Way–like" in that they are roughly $\sim 10^{12} M_\odot$ at $z = 0$ and have no mergers of mass ratios more than 1:10 after $z = 2$ (see also Zheng et al. 2020 and Lochhaas et al. 2021 for the full histories and $z = 0$ properties of the Tempest, Squall, and Maelstrom halos). While the KODIAQ-Z absorbers are not selected to be associated with Milky Way–like halos, galaxy halos of this size account for most of the stellar mass in the universe and therefore most of the chemical enrichment.

We extracted H I absorbers from each sight line using the Synthetic Absorption Line Surveyor Application (SALSA; Boyd et al. 2020) and its Simple Procedure for Iterative Cloud Extraction (SPICE) method. Across all six FOGGIE galaxies, we found 8825 H I absorbers at $z = 3$, 8207 at $z = 2.5$, and 7694 at $z = 2$, for a total of 24,726 H I absorbers. SPICE identifies absorbers based on H I number density, without the need for synthetic spectra. First, a number density cutoff is determined such that 80% of the total column density lies above the cutoff.[26] An interval is defined encompassing the sight line's cells that lie above this cutoff. This interval is then masked, and a new number density cutoff is found. The intervals from these two generations are then combined if they overlap along the line of sight and their average line-of-sight velocities are within 10 km s^{-1} of each other. These steps are repeated until the total column density of the remaining, unmasked cells falls below $\log N_{\rm H\,I} = 12$. SALSA then calculates the total column density and column-density-weighted average metallicity of each identified absorber.

While KODIAQ-Z is an H I-selected survey, we also extracted C IV and Si IV absorbers from these same sight lines in order to assess how our sample of simulated absorbers would differ if we also required the presence of metal absorption. Carbon and silicon are not tracked by the FOGGIE simulations, so they are inferred from the metallicity field assuming solar abundances. Their ionization states are estimated using the

Trident code (Hummels et al. 2017), based on a table of CLOUDY equilibrium models (Ferland et al. 2013) using the Haardt & Madau (2012) UV background with self-shielding (Emerick et al. 2019). We pair our H I and metal absorbers based on their position within the simulations and their velocities. To be paired, the two absorbers must be within the same sight line, within 100 pc of each other, and have line-of-sight velocities within 10 km s^{-1} of each other. Pairing absorbers in this way is robust to minor changes in distance and velocity. Of our 24,736 H I absorbers at redshifts 2, 2.5, and 3, 1189 have an associated C IV absorber and 358 are associated with an Si IV absorber.

9.2. Comparison of the Metallicities

The column density and average metallicity of our extracted FOGGIE H I absorbers are shown in Figure 29. The plot bounds and dotted and dashed lines are the same as in Figure 19. The left panel shows all H I absorbers in cyan, with the observational samples overplotted. Grayscale contours have been added to illustrate the structure of the scatter plot where the density of points is highest. For the rightmost panel, no metallicity limit or metal-line requirement has been imposed on the H I absorbers. The middle panel only shows H I absorbers that are close to a C IV or Si IV absorber in position and velocity space, which—due to column density limits on our selection— is analogous to requiring a C IV or Si IV detection with column densities $> 10^{12}$ cm^{-2}. This limits the metallicity of the H I absorbers to $[{\rm X/H}] \gtrsim -2.4$, which is the limit defining the metal-poor regime for KODIAQ-Z. Imposing requirements on metal-line detection can therefore limit the observed metallicity range (reinforcing the idea that an H I selection is required to probe the entire metallicity distribution). It should be noted that the gradual increase in the absorber metallicity as $N_{\rm H\,I}$ decreases seen in the middle panel is caused by the column density requirement of $\log N_{\rm X} > 12$ placed on the simulated C IV and Si IV absorbers. The far right panel in this figure compares probability densities for the full, C IV-paired, and Si IV-paired H I absorber populations in FOGGIE.

There is a broad agreement between the FOGGIE H I absorbers and those seen by the KODIAQ-Z, HD-LLS, and R12-DLA surveys. Both the observed and simulated absorbers follow similar trends in metallicity with $N_{\rm H\,I}$. Table 9 gives the median, mean, standard deviation, and IQR range on the metallicity for different classes of H I absorber, which can be compared to the observed populations in Table 5 (although note that the observations probe a larger redshift range). Both the mean and median metallicities increase with $N_{\rm H\,I}$, with the spread tending to decrease as $N_{\rm H\,I}$ increases. This is remarkable because neither EAGLE nor FIRE zoom-in simulations could reproduce the similar trend seen at low redshift (Wotta et al. 2019; Lehner et al. 2019). In Figure 30, we compare the metallicity CDFs for the various absorbers at $2 \lesssim z \lesssim 3$ in the FOGGIE simulations and observations, further demonstrating the similar trends of increasing metallicity with increasing $N_{\rm H\,I}$.[27] However, this figure also shows that the metallicity of the FOGGIE absorbers is higher than that seen in the observations (even though their IQR or standard deviations are similar), especially for the LLSs and DLAs.

---

[26] Adjusting this cutoff has a slight effect on the total number of absorbers but does not significantly impact the resulting distribution of properties. Using a 70% threshold results in a 6% drop in the number of H I absorbers, and a 90% threshold results in an 11% increase.

[27] The FOGGIE H I absorbers can have metallicity as low as $\log(Z/Z_\odot) = -8$ (the FOGGIE metallicity floor). There are absorbers with $\log(Z/Z_\odot) < -5$ classified as SLFSs and pLLSs. These absorbers are not excluded from the CDFs in Figure 30, so these CDFs do not start at zero.

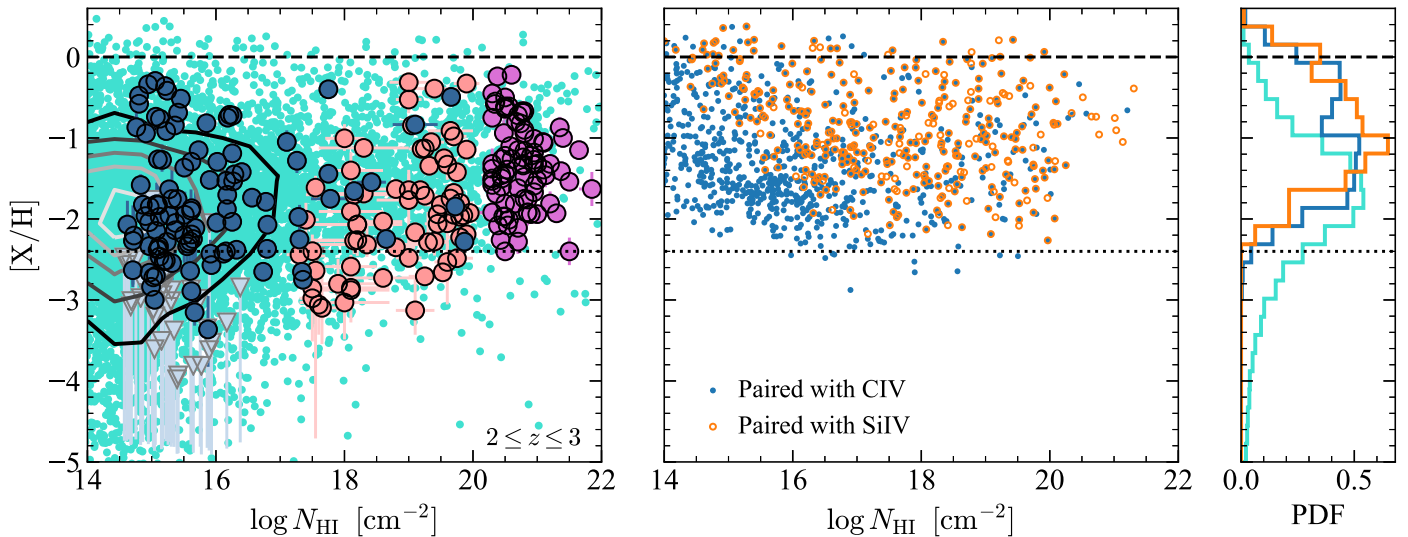**Figure 29.** Metallicities of the H I absorbers extracted from the FOGGIE cosmological zoom-in simulations (Peeples et al. 2019; Simons et al. 2020). On the left, all H I absorbers (cyan) are compared to the KODIAQ-Z, HD-LLS, and R12-DLA surveys (where we limit these surveys to $2 \leqslant z \leqslant 3$ to match the FOGGIE redshift range). Contours have been added where the density of cyan scatter plot points is highest. In the middle panel, we only show H I absorbers that have been paired with a C IV (blue filled circles) or Si IV (orange open circles) absorber, analogous to requiring detection of C IV or Si IV with $\log N_X > 12$. Probability densities of both the full FOGGIE H I absorber population (cyan) and the subset with paired metal absorbers (blue and orange) are shown on the right. The dashed and dotted lines are the solar and VMP metallicity levels, respectively.

Table 9
Metallicity Statistics of the FOGGIE Absorbers at $2 \leqslant z \leqslant 3$

Abs. Type	$\widetilde{[X/H]}$	$\overline{[X/H]}$	IQR
SLFSs	−2.43	−2.87 ± 1.68	[−3.22, −1.86]
pLLSs	−1.98	−2.13 ± 1.09	[−2.49, −1.56]
LLSs	−1.62	−1.73 ± 0.92	[−2.10, −1.18]
SLLSs	−1.41	−1.45 ± 0.76	[−1.87, −0.96]
DLA	−0.96	−1.19 ± 1.00	[−1.57, −0.60]

Note. The tilde and bar above [X/H] represent the median and mean (with the standard deviation), respectively.

The FOGGIE absorbers therefore mimic the general trend seen in the observations, where higher H I column density regimes are dominated by higher metallicities. In FOGGIE, however, the metallicity at which a particular H I regime is most prevalent is higher than seen in the observational data. FOGGIE's SLFSs have a higher metallicity than both the observed SLFSs and the observed LLSs. Moreover, the FOGGIE LLSs more closely match the CDF of the observed SLLSs. Similarly, the FOGGIE SLLS CDF is closer in metallicity to the observed DLAs. The FOGGIE DLAs have a substantially higher metallicity (the FOGGIE median metallicity is 0.5 dex higher, or a factor 3 higher, than in the R12-DLA sample).

The metallicities used for the FOGGIE absorbers are conceptually slightly different from the metallicities derived in KODIAQ-Z and other spectral surveys, in that the metallicities reported for FOGGIE are the column-density-weighted average of the gas cells identified as the absorber rather than the result of a fit to single-phase ionization models. However, Marra et al. (2021) show that these two approaches, when applied to the same physical system, do not lead to large discrepancies in the average metallicity of the gas. The precise metallicities of FOGGIE's absorbers likely depend strongly on its implementation of supernova feedback, and these would be expected to
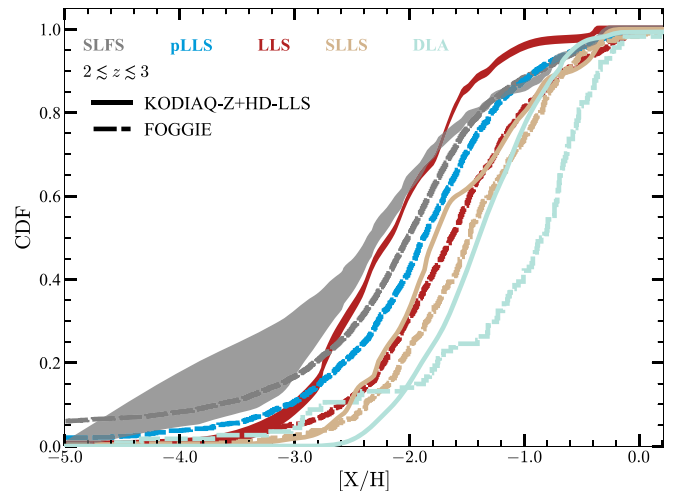

Figure 30. Comparison of the metallicity CDFs for the different categories of absorbers in the KODIAQ-Z+HD-LLS+R12 observations and FOGGIE simulations at $2 \leqslant z \leqslant 3$ (the sample of pLLSs in this redshift interval for the observation sample is too small to produce a reliable CDF and therefore is not shown here).

differ slightly (e.g., be generally lower or higher) for different implementations. Yet despite FOGGIE's overall higher metallicity, the general similarity between the observed and simulated absorbers in Figure 29 seems to be a good indication that the *trends* observed in the KODIAQ-Z, HD-LLS, and DLA surveys can in principle be explained by CGM and nearby IGM gas around typical galaxies at these redshifts.

9.3. Insights from FOGGIE Simulations

Given some similarities in the metallicity trends with $N_{H I}$ between our samples of observed and simulated absorbers, we can use the FOGGIE simulations to shed light on the origin(s) of these absorbers and their connection to galaxies. From the empirical results from KODIAQ-Z, the overdensities (at least
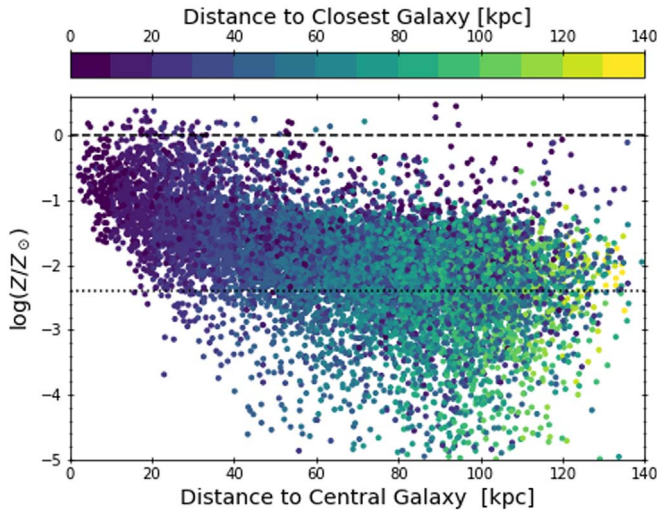
Figure 31. Metallicities of the FOGGIE H I absorbers vs. their distance from the central galaxy, colored by the distance to their closest galaxy. Possible "closest galaxies" include the central satellite and all satellites with more than 1% of the central's stellar mass. High-metallicity absorbers far from the central galaxy tend to be close to a satellite galaxy. Dashed and dotted lines are the same as in Figure 19.

for the pLLSs, LLSs, and SLLSs) imply that these absorbers should largely probe gas within virialized (mini)halos (see Section 7.1). The Keck Baryonic Structure Survey (KBSS) also shows that at $z \sim 2$–3 there is a strong incidence of absorbers with $\log N_{\rm H\,I} > 14.5$ with galaxies at transverse physical distance $\lesssim 300$ kpc and velocity separation between the absorber and galaxy redshifts $\lesssim 300$ km s^{-1} (Rudie et al. 2012). This connection is not observed for lower-$N_{\rm H\,I}$ absorbers, also implying some connection between $\log N_{\rm H\,I} > 14.5$ absorbers and galaxies. On the other hand, recent Very Large Telescope (VLT)/MUSE IFU observations of high-z LLSs also show that at least some of these absorbers may probe IGM filaments (Fumagalli et al. 2016a; Lofthouse et al. 2020). At $z < 1$, similar findings are found using Keck/KCWI and VLT/MUSE observations, where VMP pLLSs are found either in the CGM of galaxies or well outside (likely in the IGM), while metal-enriched are always found in the CGM of galaxies (Berg et al. 2022). With the FOGGIE simulations, we can explore the origins of these absorbers without observational constraints due to, e.g., the brightness limit of the galaxies. It is possible, however, that biases due to the selection criteria of the FOGGIE galaxies as Milky Way–like analogs at $z = 0$ could bias the results in some way. Importantly, we can also explore whether and how the origins may depend on the metal enrichment of the absorbers.

Each of the central FOGGIE galaxies is surrounded by a number of satellites at $z = 2$–3. These satellite galaxies likely contribute to the enrichment seen in our absorbers, and it is worth determining whether a given absorber is more likely to be associated with such a satellite than with the central galaxy. Including only those satellites with at least 1% of the central galaxy's stellar mass, we identify which galaxy each absorber is physically closest to. In Figure 31, we compare the absorber metallicities to their distance from the central galaxy and the distance to their closest galaxies, be that a satellite or the central galaxy. Generally, distance to an absorber's closest galaxy increases with distance from the central galaxy, which is the behavior we expect for absorbers whose closest galaxy is the
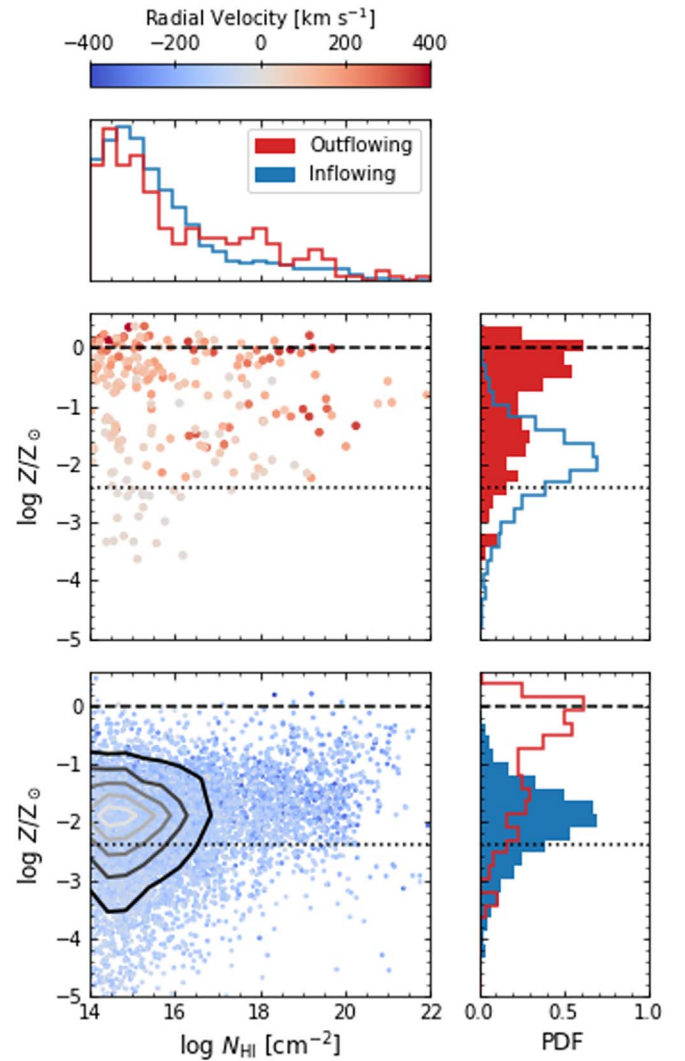

Figure 32. Metallicities of the FOGGIE H I absorbers in Figure 29, split into outflowing (red) and inflowing (blue) based on $|\hat{v} \cdot \hat{r}| > 0.5$ in the frame of the host central galaxy. Saturation shows the magnitude of the radial velocity. Probability densities are provided as histograms along the side and top. Both metallicity distributions are repeated in the side panels, with the filled histogram corresponding to the adjacent scatter plot. Contours have been added to show where the density of inflowing absorbers is highest.

central galaxy. Metallicity also tends to decrease with increasing distance from the central galaxy. Yet there are important exceptions to these two trends: some absorbers are far from the central galaxy yet close to a satellite, and these absorbers span the range of extracted metallicities. This includes some absorbers distant from the central galaxy that have metallicity close to and above solar. Therefore, not all high-metallicity absorbers should be assumed to be associated with the central galaxy.

In Figures 32 and 33, we select FOGGIE absorbers that are inflowing toward or outflowing from the central galaxy. This separation is made based on the angle between an absorber's 3D velocity vector and its position vector, both of which are calculated in the rest frame of the relevant central FOGGIE galaxy. Outflowing gas must have an angle $\theta > 60°$, while inflowing absorbers must have $\theta < 120°$. This is equivalent to $\hat{v} \cdot \hat{r} = \cos \theta = \pm 0.5$. This cut prevents gas that is neither strongly inflowing nor outflowing from being selected.

Figure 32 shows the metallicity and H I column densities of our inflowing (blue) and outflowing (red) FOGGIE absorbers.
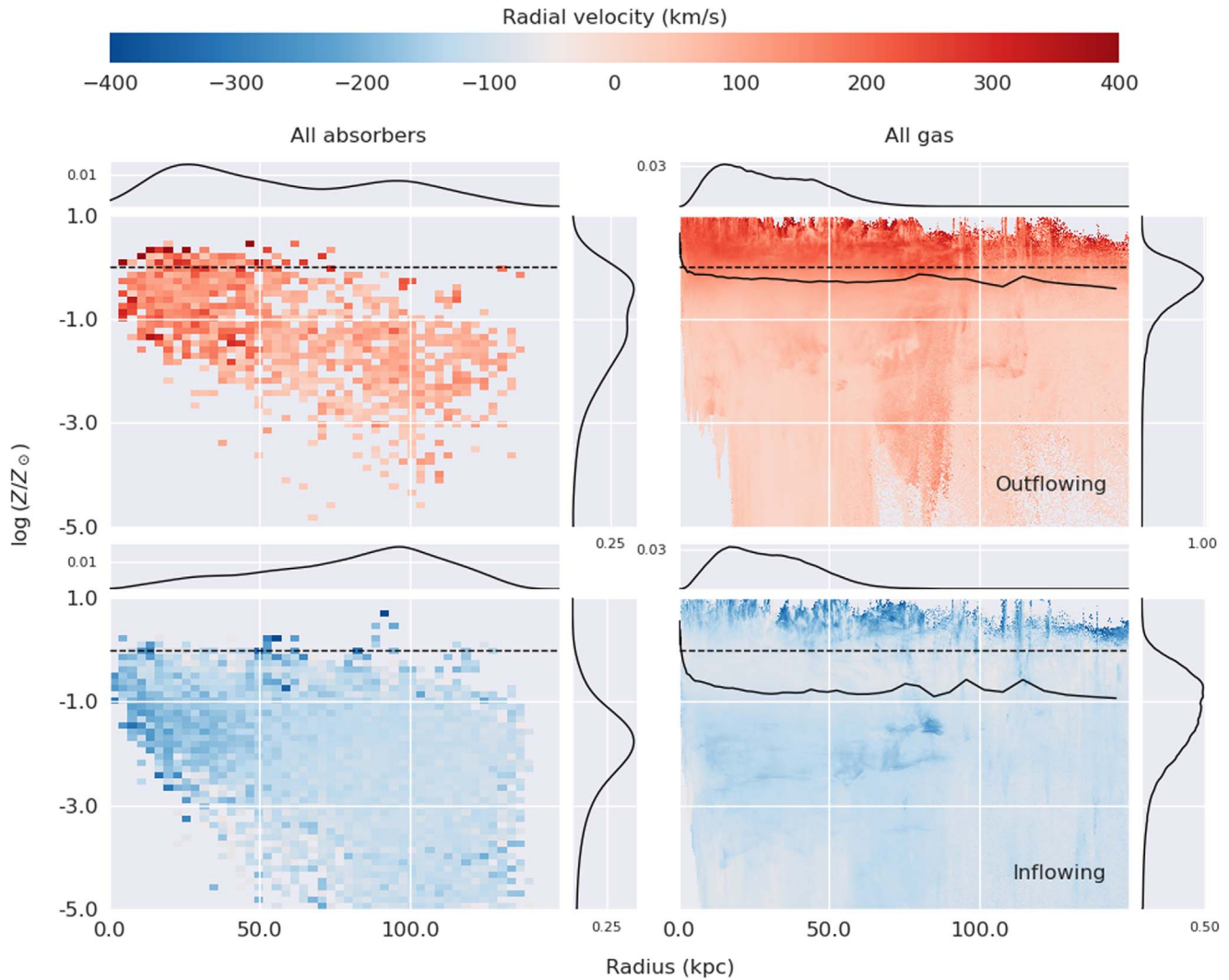
Figure 33. On the left, the metallicities of outflowing (red) and inflowing (blue) FOGGIE absorbers are binned vs. their distance from their host central galaxies, colored by the average radial velocity for that bin. On the right, the gas in all 18 FOGGIE snapshots is shown in the same configuration. The mean metallicity ($\log \langle Z/Z_\odot \rangle$) is marked with a black line. The dashed line indicates the solar metallicity level. Probability densities are shown at the top and sides of each panel.

The pdf's are also included for these two quantities, with both metallicity pdf's repeated in the top and bottom panels for comparison. The inflowing absorbers dominate by number, but both categories have similar H I column distributions, with absorbers becoming more numerous at lower H I column densities. The outflowing absorbers skew toward the high-metallicity end of the distribution. This is not surprising, as in FOGGIE winds are launched by the same supernova feedback that also enriches the surrounding gas. There is a plateau in the outflowing metallicity distribution around $-2 \lesssim \log Z/Z_\odot \lesssim -1$, which may be tracing older outflows that have slowed and mixed with the ambient CGM. These absorbers number significantly less than the inflowing absorbers at the same metallicity, however, so it would be difficult for observations to disentangle these less enriched outflows. The inflowing absorbers follow a more symmetric metallicity distribution, with a peak about 0.5 dex above the metal-poor limit for KODIAQ-Z; however, as stated above, the FOGGIE absorbers generally tend to have higher metallicities than seen in the random sight lines observed by KODIAQ-Z. The inflowing absorbers also extend to the lowest metallicities. Even though Figure 31 reveals that some absorbers may be closer to a satellite than to the central galaxy, the trends in Figure 32 are unchanged

when the absorbers are limited to those whose closest galaxy is the central. The biggest difference is that there are no outflowing absorbers with $\log Z/Z_\odot < -2.4$.

Although the outflowing absorbers dominate the high end of the metallicity pdf, they are far outnumbered by the inflowing absorbers even at the highest metallicities. This is largely due to the selection on H I and the underlying physics included in the FOGGIE simulations. There is very little cool gas that would form H I within the outflows, as these are driven only by thermal pressure. Simulations with additional nonthermal pressure sources such as magnetic fields and cosmic rays (Butsky & Quinn 2018; Ji et al. 2020) demonstrate an increase in H I at large radii within the CGM. It is therefore quite likely that the number of H I absorbers classified as either inflowing or outflowing following our definition would change if these physical processes were included in the FOGGIE simulations, though it is likely that the numbers would still favor inflowing gas.

With Figure 33, we look at the metallicity and velocity of the FOGGIE absorbers as a function of their distance from the central simulated galaxy. We also compare the absorbers to the underlying gas distributions. This distribution compiles the gas in all 18 FOGGIE outputs we sampled. Probability densities are

shown for the metallicity and radial distributions, and the mean metallicity of the gas is shown with a black line. Colors indicate the mean radial velocity at that metallicity and radius; we note that, unlike with the absorbers, the gas radial velocity has no selection based on the angle toward or away from the galaxy.

First, we consider outflowing material: outflowing absorbers with the highest metallicities tend to be at or within the virial radius of these central halos, which at $z = 2$–3 is about 30–50 kpc, and have the highest velocities. These velocities are generally higher than what is seen in the gas at similar radii and metallicities, and they appear consistent with fast-moving enriched material launched by stellar feedback. On the other hand, high-metallicity gas can be found even beyond the virial radius, implying that galactic outflows become more diffuse as they travel farther away from the galaxy. The outflowing absorbers with the lowest metallicity are found well outside the virial radius, and their velocities are comparable to the gas in this regime. This supports the interpretation that these absorbers come from previously ejected gas.

Inflowing absorbers can be found at all radii, with their metallicity generally increasing the closer they are to their central galaxy, similar to the outflowing gas. The velocity of these inflowing absorbers is also higher near the galaxy. The FOGGIE zoom-in simulations contain several satellite and premerger galaxies that produce their own enriched outflows, which explains why high-metallicity absorbers can be found at large distances from the central galaxy. These absorbers do not appear to reflect the fastest inflowing gas, which has very high metallicity and can again be found at all radii.

Together, Figures 32 and 33 imply that, given a set of H I-selected absorbers that are associated with a galaxy system, most of these absorbers are likely to be inflowing toward a central galaxy. The precise number of inflowing and outflowing absorbers predicted by simulation will depend on the included physics, especially nonthermal pressure sources, though inflowing absorbers will likely still dominate. This dominance is true across column densities and metallicities, as satellites and nearby companions enrich gas flowing toward the central galaxy. Whether inflowing or outflowing, more enriched absorbers are more likely to be located physically close to the central galaxy.

10. Discussion

10.1. Bridging the LYAF and DLAs

The history of metallicity in the universe provides an important constraint on models of galaxy formation and evolution. The metals in the universe represent a fossil record of star formation. Characterizing how the metallicities change with H I column densities (and hence overdensities or densities) yields information on the transport of metals and efficiency of that transport from the densest to the most diffuse regions of the universe. At high redshift, much of the effort has focused on the metal enrichment of the most diffuse regions probed by the LYAF (e.g., Cowie et al. 1995; Songaila 1998; Ellison et al. 2000; Schaye et al. 2000, 2003; Aguirre et al. 2004; Simcoe et al. 2004; Simcoe 2011) and the densest regions of the universe probed by DLAS (e.g., Pettini et al. 1997, 1999; Prochaska 1999; Prochaska et al. 2003; Rafelski et al. 2012; Jorgenson et al. 2013). The HD-LLS survey (Prochaska et al. 2015; Fumagalli et al. 2016b; see also Berg et al. 2021 for the

XQ-100 survey of SLLSs) has surveyed H I column density absorbers mostly in the range $17.8 \lesssim \log N_{\rm H\,I} < 20.3$. With KODIAQ-Z (this work, and earlier KODIAQ surveys; Lehner et al. 2014, 2016), we bridge the gap between the DLA/SLLS regimes and the LYAF.[28]

In Section 6.2, we show that the scatter of absorber metallicities at $[{\rm X/H}] \geqslant -2.4$ is quite similar for the DLAs and the $14.6 \lesssim \log N_{\rm H\,I} \lesssim 20$ regime, i.e., metal-enriched gas not only is observed in the densest regions (in or near galaxies) but also has spread to the more diffuse regions down to overdensities $\delta \lesssim 10$ (see Figure 19). However, as $N_{\rm H\,I}$ decreases below $\log N_{\rm H\,I} \lesssim 14.4$ (diffuse IGM), most of the absorbers have metallicities $[{\rm X/H}] \lesssim -1.6$ (see Figure 5 in Simcoe et al. 2004), while gas with $-1.6 \lesssim [{\rm X/H}] \lesssim -0.2$ is commonly observed over the entire range $14.6 \lesssim \log N_{\rm H\,I} \lesssim 22$. Therefore, the diffuse IGM probed by LYAF absorbers has not been enriched at the same level as stronger H I absorbers.

In contrast, for metallicities $[{\rm X/H}] \lesssim -2.4$, blind surveys of DLAs do not reveal a significant population of VMP DLAs (e.g., Rafelski et al. 2012; Jorgenson et al. 2013). Yet VMP absorbers are regularly observed at $\log N_{\rm H\,I} \lesssim 20$, and their frequency increases with decreasing $N_{\rm H\,I}$ (see Table 6 and Figure 19). These VMP absorbers are, of course, also observed in the LYAF. The median metallicity of the LYAF at $z \sim 2.5$ is around $[{\rm X/H}] = -2.8$ (Simcoe et al. 2004; Simcoe 2011; see also, e.g., Ellison et al. 2000; Schaye et al. 2003; Aguirre et al. 2004), which is a factor 2.5 lower than the median metallicity of the SLFSs at $\langle z \rangle \simeq 2.8$ (Table 4). However, due to sensitivity issues, the lowest metallicities found for LYAF absorbers are around $[{\rm X/H}] \sim -3.5$, while our survey reveals a population of absorbers (most with $\log N_{\rm H\,I} < 18$) with $[{\rm X/H}] < -3.5$ that is nearly completely metal-free (see Section 10.2; see also Fumagalli et al. 2011a; Lehner et al. 2016). Simcoe et al. (2004) estimated that about 70%–80% of the LYAF has been enriched to $[{\rm X/H}] \gtrsim -3$, but 20%–30% might be chemically pristine gas at low densities, which is about the same as the fraction of SLFSs with $[{\rm X/H}] < -3$.

Therefore, gas in the intermediate overdensity regime between the DLAs and LYAF has metallicities that are found in both the most diffuse and densest regions of the universe. Gas probed by absorbers with $14.6 \lesssim \log N_{\rm H\,I} \lesssim 20$ has a much larger dispersion in metallicity than observed in either $\log N_{\rm H\,I} \lesssim 14.4$ or $\geqslant 20.3$ gas. The simultaneous presence of abundant pristine, metal-rich, intermediate-metallicity gas is unique to the H I column density range $14.6 \lesssim \log N_{\rm H\,I} \lesssim 20$. Below we discuss in more details these various levels of chemical enrichment and the mixing of metals in gas with H I column densities in the range $14.6 \lesssim \log N_{\rm H\,I} \lesssim 20$.

10.2. Pristine Gas

From the KODIAQ-Z ($2.2 \lesssim z \lesssim 3.6$) survey and CCC at low redshift ($z < 1$), we find that there is an abundance of VMP absorbers in relatively high overdensity regions (except for the DLAs by definition). While the metallicities have increased by about a factor 8–10 from $2.2 \lesssim z \lesssim 3.6$ to $z < 1$, the fractions of

---

[28] We show in Section 6.3 that similar trends of the metallicities with $N_{\rm H\,I}$ are observed at low and high redshifts in the range $15 \lesssim \log N_{\rm H\,I} \lesssim 22$. However, the current UV observations at low z do not have the sensitivity to probe much lower metallicities than $[{\rm X/H}] < -1$ in the LYAF at $z < 1$ (this would need to await a future $\sim$6 m UV space-based telescope), and we cannot yet assess the full range of metallicities for absorbers with $\log N_{\rm H\,I} < 15$ at $z < 1$. Therefore, we keep our discussion focused on the high-redshift universe.

VMP absorbers are strikingly similar at low and high redshifts in $\log N_{\rm H\,I} < 19$ absorbers. On the other hand, owing to the overall increase of the metallicity from $2.2 \lesssim z \lesssim 3.6$ to $z < 1$, it is not surprising that extremely metal-poor gas with $[{\rm X/H}] < -3.5$ or even < -3.0 is scant at $z < 1$ (Lehner et al. 2019). With KODIAQ-Z, we determine that the fractions of SLFSs and pLLSs with $[{\rm X/H}] < -3.5$ and < -3.0 are about 3%–10% and 15%–25% (90% confidence intervals; see Table 6). Combining KODIAQ-Z and HD-LLS, similar numbers are derived for the LLSs. For the SLLSs, the fractions of absorbers with these metallicities are smaller by about a factor 2.

The pristine LLSs reported by Fumagalli et al. (2011a) at $z \sim 3.5$ have $[{\rm X/H}] < -3.8$ and < -4.2, but at the time of this discovery it was impossible to say how common this population was. With KODIAQ-Z, we can revisit this question at $2.2 \lesssim z \lesssim 3.6$ (see also Lehner et al. 2016). The lowest metallicity where some metals are detected is around $[{\rm X/H}] \lesssim -3.8$ in KODIAQ-Z (see Figure 19). We use that value of $[{\rm X/H}] \simeq -3.8$ to separate pristine (no metal) from metal-enriched (even at very low levels) absorbers. Considering Figure 19, we can separate the population of H I absorbers into two broad categories owing to the impact that upper limits on $[{\rm X/H}]$ can have for interpreting the frequency of pristine absorbers:

(1) For absorbers with $16 \leqslant \log N_{\rm H\,I} \leqslant 20$ at $2.2 \lesssim z \lesssim 3.6$, there are only 11 upper limits on $[{\rm X/H}]$: 3 out 179 of these absorbers have $[{\rm X/H}] < -3.8$, implying a frequency of pristine absorbers in the range 0.7%–4.2% (90% CI) in this column density interval. This number cannot increase by much more than a factor ~ 2 including data with higher upper limits. Even if all the upper limits represent pristine gas for absorbers with $\log N_{\rm H\,I} \geqslant 16$, this would increase the frequency of pristine absorbers to only 3.7%–9.6% (90% CI). Hence, the fraction of pristine $\log N_{\rm H\,I} \geqslant 16$ absorbers (i.e., gas with overdensities $\delta \gtrsim 50$–100) is at least 1% and at most 10% at $2.2 \lesssim z \lesssim 3.6$.

(2) The second regime is for absorbers with $14.6 \lesssim \log N_{\rm H\,I} < 16$ at $2.2 \lesssim z \lesssim 3.6$, where the upper limits on the metallicities are much more numerous. Taking strictly $[{\rm X/H}] < -3.8$, there are 4 absorbers out of 142, implying a frequency of pristine absorbers of 1.3%–6.1% (90% CI), a similar percentage to that of the absorbers with higher $N_{\rm H\,I}$. However, if all the $\log N_{\rm H\,I} < 16$ absorbers with upper limits such as $[{\rm X/H}] < -3$ probe pristine absorbers, then the frequency of pristine $\log N_{\rm H\,I} < 16$ absorbers would increase to 13.3%–23.9% (90% CI). This would be similar to the possible population of metal-free LYAF absorbers at similar redshifts (see previous section and Simcoe et al. 2004). The fraction of pristine gas in overdensities $\delta \gtrsim 50$–100 at $2.2 \lesssim z \lesssim 3.6$ is therefore at the level of a few percent, while for smaller overdensities it is likely in the range 10%–20%.

The frequency of mock absorbers with $[{\rm X/H}] < -3.8$ in the FOGGIE simulations is very similar to the frequencies found in KODIAQ-Z over similar redshifts, despite that the FOGGIE absorbers have on average higher metallicities than observed. In contrast, cruder resolution simulations show a tendency to have overenriched pLLSs and LLSs (see discussion in Fumagalli et al. 2011a). In the EAGLE HiRes cosmological hydrodynamical simulations at $z \sim 2$–4, Rahmati & Oppenheimer (2018) find a fraction of pristine gas with $[{\rm X/H}] < -3.8$ of about 10% in the pLLS and LLS regimes. However, contrary to FOGGIE and the observations reported here or at low

redshift (Wotta et al. 2019; Lehner et al. 2019), the EAGLE simulations show little change of the metallicity with $N_{\rm H\,I}$ (see also Lehner et al. 2019; Wotta et al. 2019); the FIRE simulations also show similar results at $z < 1$ (Hafen et al. 2017). Consistent with the lack of $N_{\rm H\,I}$ dependence, these simulations find a significant population of $[{\rm X/H}] < -3$ DLAs that is not observed.

This highlights some issues in simulations that may be related to the implementation of feedback physics but is probably also related to insufficient resolution in the more diffuse gas. As discussed in Section 9.2, low-metallicity inflowing gas is able to naturally get much closer to the galaxy when the CGM structures are more resolved than in the standard-resolution simulations because the metals are not forced to overmix in the regions around galaxies (see also Hummels et al. 2019; Peeples et al. 2019; Suresh et al. 2019; van de Voort et al. 2019). As discussed in Section 9.3 and shown in Figures 31–33, the pristine and the VMP absorbers predominantly probe inflowing cool gas in the context of these simulations (see also Hafen et al. 2017; Rahmati & Oppenheimer 2018; Suresh et al. 2019), consistent with the idea that these absorbers are ideal candidates for the cold-mode accretion (e.g., Kereš et al. 2005; Dekel et al. 2009; Fumagalli et al. 2011b).

We finally note that achieving high spatial resolution not only in the CGM but also in the IGM is critical to fully capture the various structures probed by strong H I absorbers surveyed in KODIAQ-Z or CCC at low redshift. In recent high-resolution simulations of an intergalactic sheet between two massive halos at $z \sim 3$–5, Mandelker et al. (2019, 2021) show that pristine LLSs with $[{\rm X/H}] < -3$ such as those discussed in this work can be found in the IGM well outside the virial radius of any galaxy. Increasing the resolution in the IGM results in more H I in smaller, dense clouds, which are essentially absent in lower resolution. In their highest-resolution simulations (which have not converged yet), Mandelker et al. (2021) find that the covering fraction of LLSs in the IGM with $[{\rm X/H}] < -3$ is about 3%. Therefore, the IGM itself could also provide an alternative origin for the observed pristine strong H I absorbers. Observationally, galaxy counterpart searches for two pristine LLSs with $[{\rm X/H}] < -3.3$ at $z > 2$ show no associated galaxies for one and a group of galaxies for the other one (Fumagalli et al. 2016a). At $z < 1$, the recent BASIC galaxy survey of 19 pLLSs shows two populations for the VMP pLLSs, one associated with galaxy halos and another one well outside the virial radius of any galaxy, likely originating in the IGM (Berg et al. 2022). It is therefore likely based on these observations and high-resolution simulations that a combination of these dual origins (IGM and cold inflow) can most likely explain the pristine and VMP absorbers.

10.3. Extremely High Metallicities and Metal-rich Gas

The metal-enriched diffuse gas constitutes a record of the transport of metals from the densest to the most diffuse regions of the universe. In Section 10.1, we discuss that while the LYAF shows evidence for metal enrichment, it is not enriched to levels seen in higher H I column absorbers. Metallicities in the range $-1.6 \lesssim [{\rm X/H}] \lesssim -0.2$ are commonly observed over the entire range $14.6 \lesssim \log N_{\rm H\,I} \lesssim 22$, but not in the LYAF. Therefore, at $2.2 \lesssim z \lesssim 3.6$, galaxies have not yet polluted the diffuse IGM with metals at the same level as denser regions of the universe.

In Lehner et al. (2016), we reported the discovery of one supersolar pLLS at $z \simeq 2.5$ with $[X/H] \simeq +0.2$.[29] Besides its metallicity, this pLLS is unique on several other levels, including the detection of O I, its small physical size (0.35 pc), its relatively high density ($n_H \simeq 0.2$ cm^{-3}), and its multiphase nature, with C IV having a very different velocity profile compared to the low ions (see Figure 14 in Lehner et al. 2016). We argued in that paper that its high metallicity and multiphase nature strongly suggest that it directly probes an active outflow from a protogalaxy at $z \simeq 2.5$. At $z \sim 1.8$, Prochaska et al. (2006) uncovered two supersolar SLLSs, suggesting that these metal-rich absorbers may represent a significant metal reservoir in the young universe.

However, the HD-LLS survey does not report any supersolar SLLSs (see Figure 19), and the XQ-100 survey only reports one supersolar SLLS at $z \sim 2.5$ (see Figure 2 in Berg et al. 2021; we exclude poorly constrained upper limits on the metallicity that are above solar in that survey). Adding that supersolar pLLS to the KODIAQ-Z+HD-LLS survey would imply that only 1 out of 242 absorbers has a supersolar metallicity at $2.2 \lesssim z \lesssim 3.6$, i.e., $0.41\%^{+1.42\%}_{-0.32\%}$ (90% CI) of the absorbers have $[X/H] > 0$. None are found in the R12-DLA either, implying that the fraction of supersolar absorbers in the H I column density range $14.6 \lesssim \log N_{\mathrm{H\,I}} < 22$ at $2.2 \lesssim z \lesssim 3.6$ drops to $0.29\%^{+1.00\%}_{-0.23\%}$. Supersolar-metallicity absorbers are therefore very rare at $2.2 \lesssim z \lesssim 3.6$ for any $N_{\mathrm{H\,I}}$ absorber.

At low redshift, the situation is quite different, with at least 2%–6% of supersolar absorbers at $z < 1$ (see Figure 22 and Lehner et al. 2019). This number could increase by a factor ~ 3 if a harder EUVB was used instead of the HM05 EUVB to model the absorbers (at low redshift the effect of the EUVB is more important; see Wotta et al. 2019 and Section 5). This increase of supersolar absorbers is consistent with the overall increase of the metallicities by a factor 8–10 from $2.2 \lesssim z \lesssim 3.6$ to $z < 1$ (see Figure 21).

The larger fraction of metal-enriched and even super-metallicity absorbers at low redshift indicates that a much larger volume of the universe has been exposed to metal pollution than at $2.2 \lesssim z \lesssim 3.6$. Ferrara et al. (2000) developed a model that predicts that at $z < 1$ essentially all absorbers should have associated metal absorption, and the spread in metallicity should be less than 1 dex. This is not the case. Even at $z < 1$, the enrichment of the most diffuse gas in the universe is still very inhomogeneous in view of the nearly 3 dex spread in metallicities at $z < 1$ for absorbers with $15 \lesssim \log N_{\mathrm{H\,I}} \lesssim 19$ (see Figure 22 and next section).

10.4. Inhomogeneous Metal Mixing

There is ample evidence for inhomogeneous abundances in H I column density $14.6 \lesssim \log N_{\mathrm{H\,I}} \lesssim 20$ gas where the metallicities range from pristine ($[X/H] < -4$) to about $[X/H] \simeq -0.2$ dex, a factor of 6000 spread from the lowest to highest metallicities. Similar results are found at $z < 1$, but a factor somewhat smaller with 1000 variation between the highest and lowest metallicities (Lehner et al. 2019 and Figure 22). However, our results also show that there is inhomogeneous metal mixing in single halos as directly evidenced by the large variation in metallicity in closely redshift-separated absorbers (see Figure 25). Such large metallicity variation over small redshift separations was initially reported by

Prochter et al. (2010) between an SLFS, a pLLS, and an LLS at $z \sim 3.5$ separated in velocities along the same QSO from about 130 to 180 km s^{-1}, where the metallicity differences were a factor 3, 40, and 158. Now with a sample of 36 paired absorbers, we find similarly large metallicity variations between absorbers separated by less than $\Delta v < 500$ km s^{-1} along a given QSO sight line where about half of paired absorbers have metallicity variations of a factor 2–3 and for the other half of a factor > 140. However, on average the absorbers separated by $\Delta v < 150$ km s^{-1} have a metallicity variation that is on average around a factor 5, substantially smaller than in paired absorbers with larger Δv, where on the average the difference in metallicities is about 10 times larger. This would be consistent with the smaller variations being related to the variation within a halo and the larger variations being related to gas in interhalos.

Thus, within the overdensities of a few to several hundreds, the gas probed by absorbers with $14.6 \lesssim \log N_{\mathrm{H\,I}} \lesssim 20$ is chemically inhomogeneous where both metal-poor gas and metal-rich gas are observed. This again appears unique to the $14.6 \lesssim \log N_{\mathrm{H\,I}} \lesssim 20$ column density range. While for DLAs determining the metallicity in individual components is complicated since the H I velocity structure cannot be recovered over $\Delta v < 500$ km s^{-1}, the metallicity spread for the DLAs is much smaller than observed at $14.6 \lesssim \log N_{\mathrm{H\,I}} \lesssim 20$ (see Section 10.1). It is also much smaller for the LYAF (e.g., Simcoe et al. 2004), implying overall a more chemically homogeneous gas in the LYAF and DLA regimes. Therefore, in and quite near galaxies, in the gas traced by DLAs, the metal enrichment from supernovae is quite efficient, but beyond the immediate vicinity of the galaxies, the volume filling factor for metals must be low, which is consistent with models of metal ejection from supernovae that is not expected to be efficient and confined to small regions (e.g., Ferrara et al. 2000; Scannapieco 2005).

11. Summary and Concluding Remarks

Using the KODIAQ DR2 data from the KOA (O'Meara et al. 2015), we have built a sample of 202 H I-selected absorbers with $14.6 \leqslant \log N_{\mathrm{H\,I}} \leqslant 20$ at $2.2 \leqslant z \leqslant 3.6$ comparable in size to our companion survey, CCC, at $z < 1$. The H I selection and the H I column density range both ensure that no bias is introduced in the metallicity distribution of these absorbers, i.e., we are sensitive to any absorbers with $[X/H] \gtrsim -3.5$. The H I selection also provides a clean separation of the absorbers based on H I column density, which is largely used to separate the LYAF, absorbers with $\log N_{\mathrm{H\,I}} \lesssim 14.5$ (IGM), from the denser regions of the universe probed by stronger H I absorbers. By definition, the SLFSs all have overdensities $\delta > 3$ in our sample ($z = 3.6$; the maximum redshift in our statistical sample corresponds to this δ value using the analytical expression in Schaye 2001a). In contrast, a metal selection using C IV or O VI can include a wide range of H I column density absorbers from the LYAF to the pLLS regime.

In the KODIAQ-Z survey, we have 155 SLFSs, 24 pLLSs, 16 LLSs, and also 7 SLLSs, for a total of 202 absorbers. To increase the number of LLSs and SLLSs in our sample, we use the H I-selected absorbers from the HD-LLS survey that have an H I column density range $17.2 \leqslant \log N_{\mathrm{H\,I}} \leqslant 20$, with a majority having $17.7 \lesssim \log N_{\mathrm{H\,I}} \leqslant 20$ (Prochaska et al. 2015; Fumagalli et al. 2016b). We also use the results from the survey from Rafelski et al. (2012) to study the metallicity changes with $N_{\mathrm{H\,I}}$ in the DLA regime. For both of these surveys, we restrict them to absorbers at $2.2 \leqslant z \leqslant 3.6$, unless otherwise stated.

---

[29] The HIRES spectrum of this QSO was not released in KODIAQ DR2 and therefore is not part of the present KODIAQ-Z survey.

For all the absorbers with $\log N_{H\,I} \lesssim 20$ in the KODIAQ-Z and HD-LLS, we derive the posterior pdf's of the metallicities and other physical quantities using a Bayesian formalism that employs an MCMC sampling of a grid of CLOUDY photoionization models (where we assume that photons from the HM05 EUVB provide the source of photoionization). This follows directly from the methodology used in HD-LLS (Fumagalli et al. 2016b) and CCC (Wotta et al. 2019; Lehner et al. 2019). For absorbers with less-than-ideal constraints (about half of the SLFS sample and a few pLLSs), we adopt the "low-resolution" method developed at low redshift by Wotta et al. (2016) and refined in Wotta et al. (2019) to estimate the metallicities. Using SLFSs and pLLSs with reliable constraints from the metal ions, we find that $\log U$ can be reasonably well modeled by a Gaussian for the SLFSs and pLLSs, which can then be used reliably as a prior in the Bayesian MCMC modeling. We explore the effects of changing the EUVB from HM05 and HM12 to estimate the metallicities (Figures 9, 10), and we find that the changes in the metallicities between these two EUVBs are negligible for the absorbers with $\log N_{H\,I} \gtrsim 17.2$. For the SLFSs and pLLSs, the effect is only an increase of $+0.12 \pm 0.15$ dex. The effect from changing the EUVBs on the metallicities is therefore much smaller than at $z < 1$.

Our main findings on the properties of our statistical sample of H I-selected absorbers with $14.6 \leqslant \log N_{H\,I} \leqslant 20$ at $2.2 \leqslant z \leqslant 3.6$ can be summarized as follows.

1. From the comparison of the absorption profiles and the width of the profiles, we conclude that singly, doubly, and triply ionized species (e.g., Si II, Si III, Si IV, C II, C IV) and H I often trace the same gas in absorbers with $\log N_{H\,I} < 19$. For SLFSs and pLLSs, we also find that when O VI absorption is present, it has often a similar velocity structure to lower ions, i.e., the O VI absorption is commonly narrow in these absorbers. This contrasts from the O VI that is frequently strong and broad when detected in absorbers with $\log N_{H\,I} \gtrsim 17.8$.

2. From the PF of the H I transitions, we show that 90% of the components have $13.3 \leqslant b \leqslant 40$ km s^{-1} with a mean $\langle b \rangle = 27 \pm 6$ km s^{-1}. This implies a temperature of the gas of $T < 4 \times 10^4$ K, consistent with the gas being primarily photoionized.

3. We find that a single-phase photoionization model is appropriate to match the column densities of the low ions to high ions (including O VI) for the majority of the SLFSs and pLLSs. For the LLSs and SLLSs, when O VI is detected, a single-phase photoionization model cannot commonly reproduce the observed O VI column density, implying that as $N_{H\,I}$ increases, the multiple gas-phase nature becomes more important.

4. In our ionization models, $[C/\alpha]$ is allowed to vary in the range $[-1, +1]$ to accommodate for nonsolar relative abundances between carbon and α-elements caused by nucleosynthesis effects. This approach allows us to have a better sense of the uncertainties in the metallicity caused by the variability of C/α. Overall, we find that $[C/\alpha]$ is in the range $-0.6 \lesssim [C/\alpha] \lesssim +0.5$ as observed in other environments (DLAs, stars, H II regions), but robustly studying the variation of $[C/\alpha]$ with $[X/H]$ for these absorbers is hindered by ionization corrections that add enough noise in the $[C/\alpha]$ distribution to conceal any subtle changes between $[C/\alpha]$ with $[X/H]$ that are expected to be at the level of a factor 2–3.

5. The 155 H I-selected SLFSs ($14.6 \leqslant \log N_{H\,I} < 16.2$) probe a wide range of metallicities from $[X/H] < -4$ to $[X/H] \simeq -0.2$. The metallicity posterior pdf is negatively skewed, with a prominent tail extending well below $[X/H] < -3$. It has a main peak around the median value $[X/H] \simeq -2.4$ and another smaller peak around -0.6 dex solar (the dip around -1.1 is observed in both binned and unbinned data).

6. The 24 H I-selected pLLSs ($16.2 \leqslant \log N_{H\,I} < 17.2$) probe a range of metallicity from $[X/H] < -4.2$ to about $[X/H] \simeq -1$. The sample is still too small to robustly characterize the distribution, and the lack of pLLSs with $[X/H] > -1$ could be mainly due to small number statistics. However, both the median (-2.1 dex) and IQR metallicities imply an overall increase in metallicities compared to the SLFSs despite the lack of $[X/H] > -1$ pLLSs.

7. Our sample of H I-selected LLSs ($17.2 \leqslant \log N_{H\,I} < 19$) consists of 16 absorbers that we combine with the HD-LLS survey to reach a size sample of 62 LLSs. The full range and median of the LLS metallicity pdf are quite similar to those of the pLLSs. Combining the pLLSs +LLSs, the median metallicity is -2.2 and the IQR is only 1 dex, compared to -2.42 and 2 dex, respectively, for the SLFSs. The pLLSs+LLSs are more frequent in the metallicity range $-3.2 \lesssim [X/H] \lesssim -1.2$ but less frequent at very low metallicity $[X/H] \lesssim -3.5$ than the SLFSs, showing that there is a shift in the metallicity enrichment properties of these absorbers below and above $\log N_{H\,I} \simeq 16.2$ (at $2.2 \leqslant z \leqslant 3.6$, this corresponds to overdensities of $\delta \simeq 140$–50).

8. Combining the 7 SLLSs from KODIAQ-Z and 73 SLLSs from HD-LLS ($19 \leqslant \log N_{H\,I} \leqslant 20$) and using the Rafelski et al. (2012) sample of 101 DLAs ($\log N_{H\,I} \geqslant 20.3$), the overall metallicity trend observed with $N_{H\,I}$ continues in these regimes: the median/mean metallicity increases and IQR decreases with increasing $N_{H\,I}$. For the SLLSs, the median and IQR metallicities are -1.9 and 1.1 dex, while these are -1.4 and 0.65 dex for the DLAs. DLAs have therefore rarely VMP gas (i.e., $[X/H] < -2.4$ gas), while at $\log N_{H\,I} \lesssim 20$, VMP absorbers are not rare at $2.2 \leqslant z \leqslant 3.6$.

9. The fractions of extremely metal-poor systems with $[X/H] < -3.5$ for the SLFSs, pLLSs, and LLSs are about the same, around 3%–10% (90% confidence interval). For the SLLSs, it is somewhat smaller with a fraction in the range 0.3%–5.4%. Yet the most metal-poor absorber in our sample is an SLLS with $\log N_{H\,I} = 19.25 \pm 0.25$ and $[X/H] < -4.4$ with a cosmic overdensity of several thousands.

10. We find that (pristine) gas clouds with no metals down to the limit $[X/H] < -3.8$ constitute about 1%–10% of the $16 < \log N_{H\,I} < 20$ absorbers at $2.2 \leqslant z \leqslant 3.6$ and increases to 10%–20% for $14.6 \leqslant \log N_{H\,I} \leqslant 16$ absorbers, the latter being similar to the amount of pristine gas found in the diffuse IGM at similar redshifts. On the other hand, supersolar absorbers at any $N_{H\,I}$ at $2.2 \lesssim z \lesssim 3.6$ are very uncommon ($<1\%$).

11. Using paired absorbers with velocity separations of $\Delta v \lesssim 500$ km s^{-1} along the same QSO sight lines, we find that there is a large scatter in the metallicity from about 0.2 to >2 dex. For about half of the paired absorbers, there is evidence for metallicity variations over

$\Delta v \lesssim 500$ km s^{-1} of a factor 2–3, while for the other half of a factor > 140. The larger variations (on average a factor 50 difference in metallicities) are, however, observed more often in paired absorbers with $\Delta v > 150$ km s^{-1} than in smaller velocity separated absorbers (where the metallicity variation is on average a factor 5). It is therefore plausible that the smaller variations correspond to the change within a single galaxy halo, while the large variations correspond to differences between galaxy halos. Both the large metallicity range and metallicity variation between paired absorbers imply that the transport of metals from their formation sites (galaxies) into the CGM and the IGM is very inhomogeneous.

12. The photoionized gas associated with pLLSs, LLSs, and SLLSs contributes to the cosmic baryon budget $\Omega_g/\Omega_b \simeq 7\%$, and the SLFSs contribute another 7%. From the first KODIAQ survey (Lehner et al. 2014), the O VI-bearing hot collisionally ionized gas associated with LLSs and SLLSs is likely to contribute about the same level. These are the second-largest contributors to the cosmic baryon budget at high redshift behind the LYAF.

13. We show that about 18%, 6%, and 4% of the metals ever produced at $2.2 \lesssim z \lesssim 3.6$ are in photoionized gas associated with SLFSs, pLLSs, and LLSs, respectively. The SLLSs account for another 18%, and therefore about 45% of the metals at $2.2 \lesssim z \lesssim 3.6$ in photoionized absorbers with $14.6 \leqslant \log N_{\mathrm{H\,I}} < 20.3$. Combining the SLFSs, pLLSs, LLSs, and SLLSs, their comoving metal mass density of the photoionized gas probed by these absorbers is $\rho_{\mathrm{m}} = 5.2 \times 10^5$ $M_\odot$ cMpc^{-3}. Another possible $\gtrsim 5\%$ of the cosmic metal budget may also be in the form of highly ionized metals in collisionally ionized gas with $\log N_{\mathrm{H\,I}} \gtrsim 17.8$ (Lehner et al. 2014). Therefore over 50% of the metals produced at $2.2 \lesssim z \lesssim 3.6$ are in the diffuse ionized gas.

To study the cosmic evolution of these absorbers, we combine our results with the CCC survey that explores the properties of similar absorbers at $z < 1$ (Lehner et al. 2018, 2019; Wotta et al. 2019). The cosmic evolution of absorbers from $2.2 \lesssim z \lesssim 3.6$ to $z \lesssim 1$ with $\log N_{\mathrm{H\,I}} \gtrsim 15$ is summarized as follows.

14. We find that from $2.2 \leqslant z \leqslant 3.6$ to $z < 1$, there is an overall increase of the metallicity of the gas probed by SLFSs, pLLSs, LLSs, SLLSs, and DLAs by a factor ~ 8.

15. While the metallicity threshold for the VMP absorbers increases by about 1 dex from $2.2 \leqslant z \leqslant 3.6$ to $z < 1$, a similar fraction of about 50% of absorbers with $\log N_{\mathrm{H\,I}} \lesssim 18$ are VMP at low and high z.

16. Although there is plenty of *primitive* gas at $z \lesssim 1$, i.e., gas that has largely not been polluted, the fraction of pristine SLFSs, pLLSs, and LLSs with $[\mathrm{X/H}] < -3$ at $z \lesssim 1$ is $< 1\%$ at the 90% confidence level over the range $15 < \log N_{\mathrm{H\,I}} < 19$. In contrast, 10%–25% of similar absorbers at $2.2 \lesssim z \lesssim 3.6$ have $[\mathrm{X/H}] < -3$, implying that although the transport of metals outside galaxies is still very inhomogeneous at $z < 1$, all regions with $\log N_{\mathrm{H\,I}} \gtrsim 15$ have been polluted to some level by $z < 1$.

17. The hydrogen column density (N_{H}) is a factor 10–15 smaller at $z < 1$ than at $2.2 \lesssim z \lesssim 3.6$. The contributions to baryonic budget from the SLFSs, pLLSs, and LLSs are about 5 times smaller at $z < 1$ than at $2.2 \lesssim z \lesssim 3.6$.

18. The photoionized metals associated with the SLFSs, pLLSs, and LLSs take about 6% of the cosmic metal budget at low redshift. This is about a factor 4.4 decrease compared to the contribution of the same absorbers at $2.2 \lesssim z \lesssim 3.6$, i.e., at low redshift most of the metals are found in or near galaxies and in the intracluster medium (see Péroux & Howk 2020). This is a major change in the distribution of metals in the universe from $z \sim 2.8$ to $z \sim 0.5$.

We set our empirical results in the context of the cosmological zoom-in simulations using simulations from the FOGGIE project, and the main conclusions from that study are as follows.

19. Contrary to cruder resolution simulations (especially in the CGM), a striking feature of the FOGGIE cosmological zoom-in simulations is that the behavior of the metallicity as a function of $N_{\mathrm{H\,I}}$ is broadly similar to the observed empirical relationship: as $N_{\mathrm{H\,I}}$ increases, the overall metallicity increases and the dispersion of the metallicity decreases. However, as in other cosmological simulations, FOGGIE appears to have too many metals at any $N_{\mathrm{H\,I}}$ (e.g., supersolar-metallicity gas is not uncommon in these simulations, but observationally it is), implying that metals are more homogeneously distributed than observed.

20. In the FOGGIE simulations, outflowing absorbers with the highest metallicities tend to be at or within the virial radius of these central halos and probe active or recent galaxy outflows, while outflowing absorbers with the lowest metallicity are found well outside the virial radius of the central galaxies and therefore the remnants of ejected gas. On the other hand, inflowing absorbers can be found at all radii, with their metallicity generally increasing the closer they are to their central galaxy. VMP absorbers with $[\mathrm{X/H}] < -2.4$ are excellent probes of inflowing gas in these simulations. We finally note that recent high-resolution simulations of the IGM show that some of the observed pristine absorbers could also directly originate in the IGM, i.e., in regions well outside the virial radii of galaxies.

Facility: Keck(HIRES).

Software: Astropy (Astropy Collaboration et al. 2018), emcee (Foreman-Mackey et al. 2013), Matplotlib (Hunter 2007), PyIGM (Prochaska et al. 2017a), scikit-learn (Pedregosa et al. 2011; Buitinck et al. 2013), Enzo (Bryan et al. 2014; Brummel-Smith et al. 2019), yt (Turk et al. 2011), Trident (Hummels et al. 2017), SALSA (Boyd et al. 2020).

Appendix

In this appendix, we provide information regarding the supplemental files. For each KODIAQ-Z absorber, we produced a figure as shown in Figure A1 (see also Figure 3), where we plot the normalized profiles of metals and some H I transitions for which we estimated the column densities. The red portion in each profile shows the velocity range of the absorption over which the velocity profile was integrated to derive the column densities and average velocity (or upper limits on the column densities if no absorption is detected). The vertical dashed lines mark the zero velocity. We also provide for each KODIAQ-Z absorber (and sometimes paired or multiple closely redshift spaced absorbers) a set of figures (see Figure A2 and also Figure 4 in the main text) where the Voigt profile fits (in red the composite profile fit, and in blue each individual component fit) to the individual transitions of H I (black spectra) are shown. Note that multiple absorbers for which their velocity separation is $\Delta|v| < 500$ km s^{-1} may appear in a single figure (explaining why the number of figures in this set in smaller).

Next, we make available the visualization results from the MCMC Bayesian photoionization modeling. For each KODIAQ-Z absorber, we provide the comparison plots between the observations and models and corner plots as shown in Figures A3 and A4, respectively. In Figure A3, we compare the measured column densities for each ion (red) and the predicted column densities from the median MCMC model (blue). Upward-pointing triangles (when present) show lower limits (i.e., saturated transitions), while downward-pointing triangles show 2σ upper limits. Red data points with error bars (sometimes smaller than the circles) denote well-constrained column densities. From the corner or comparison plots, one can determine readily which modeling was used: (1) if there is no entry for $\log U$ ($\log U$ prior = False), then a flat prior on the ionization parameter was used; (2) if a value to $\log U$ is given, then a Gaussian prior on $\log U$ was used with the listed mean and dispersion values; (3) if $[C/\alpha]$ is present, the absence of value indicates that a flat prior was used; otherwise, a Gaussian prior was used on that ratio with the listed mean and dispersion values. The EUVB used in the modeling is also provided (in this case, HM05). Finally, the comparison plots show which ions were used in the ionization modeling.
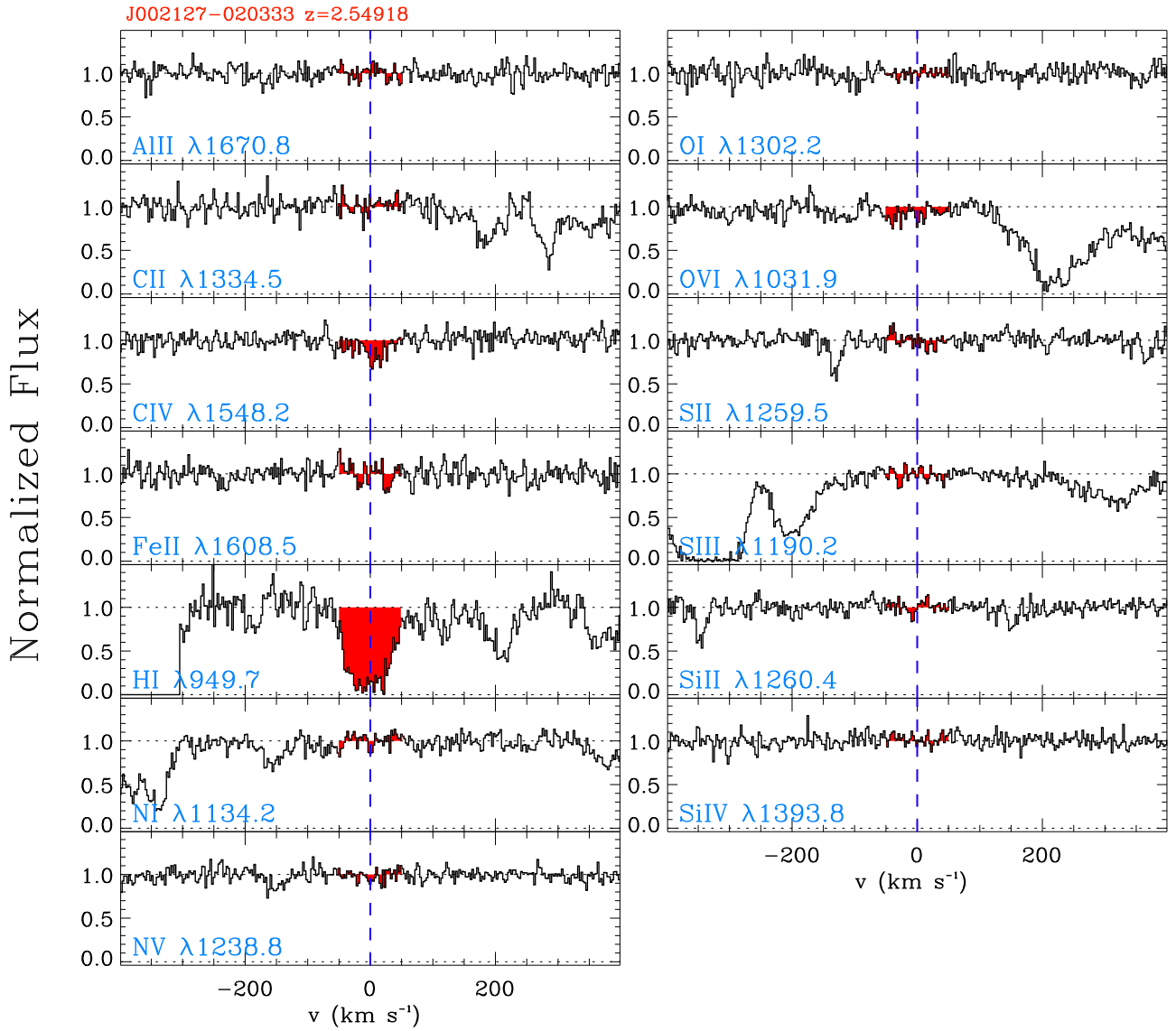
Figure A1. Normalized spectra against the rest-frame velocity of the absorber at $z = 2.54918$ toward J002127–020333. The integration range of the profiles is shown in red. The complete figure set (202 figures) is available in the online journal.

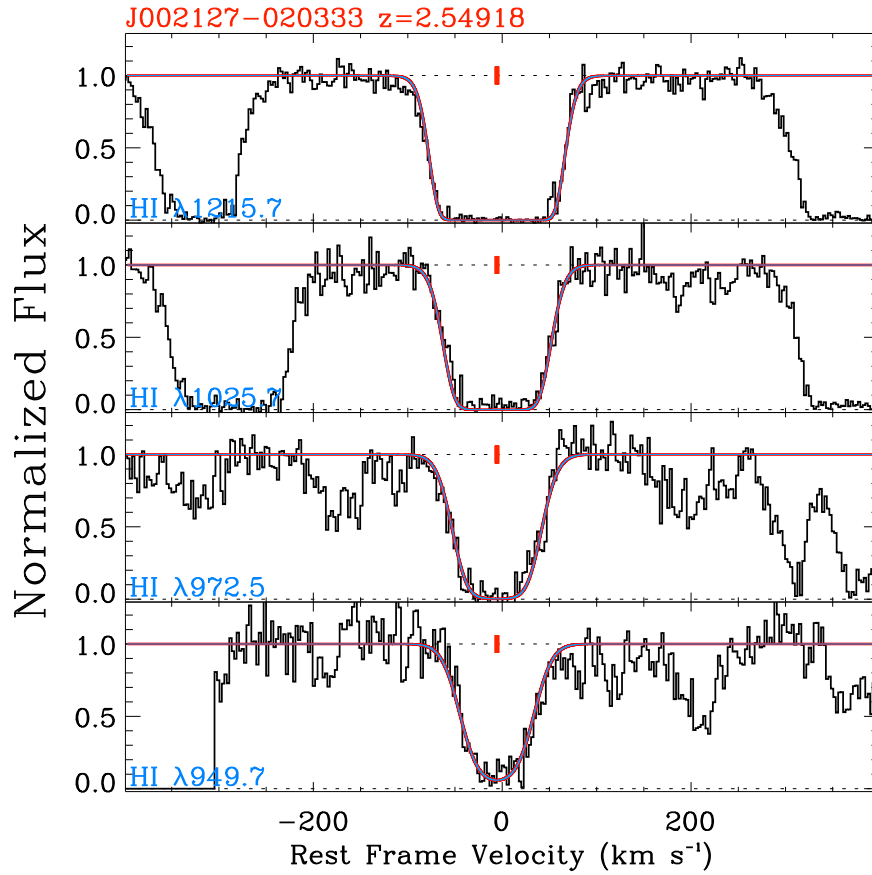(The complete figure set (202 images) is available.)

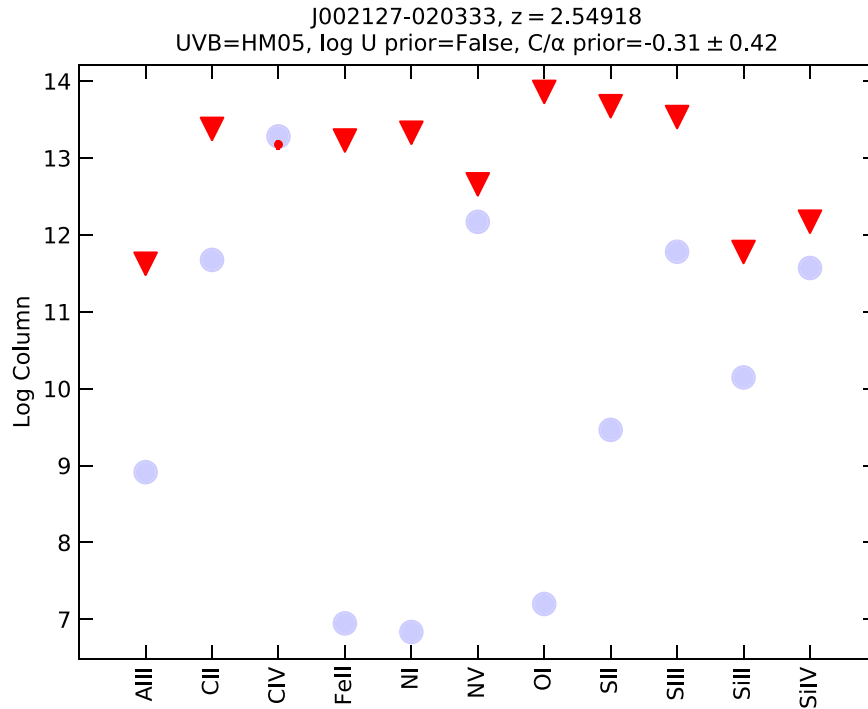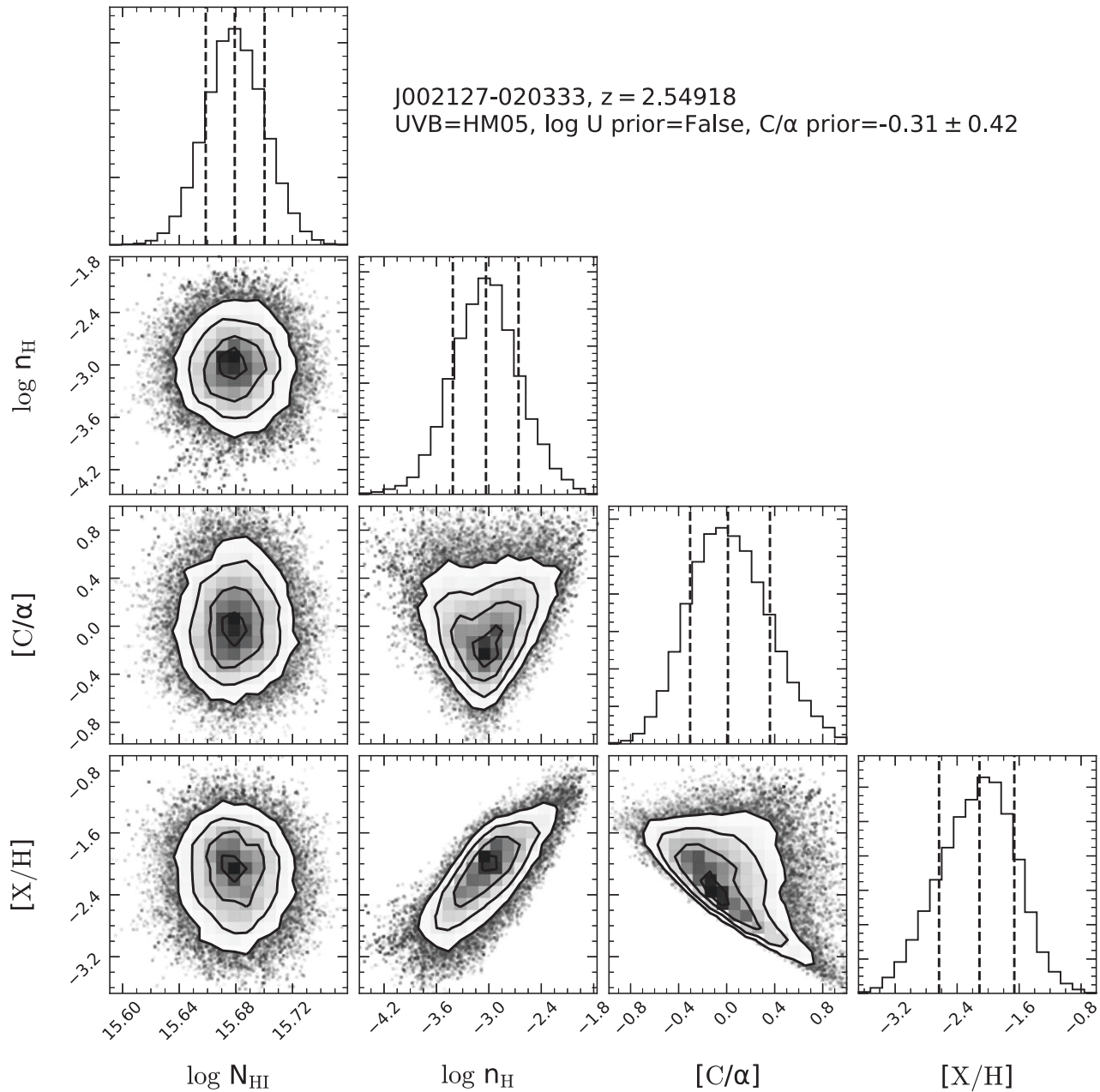**Figure A2.** Normalized spectra (black) and Voigt profile fit (in red the composite profile fit, and in blue each individual component) of the absorber at $z = 2.54918$ toward J002127–020333. The complete figure set (181 figures) is available in the online journal.

(The complete figure set (181 images) is available.)

Figure A3. MCMC comparison plot for the absorber at $z = 2.54918$ toward J002127–020333. It shows the measured column densities for each ion (red) and the predicted column densities from the median MCMC model (blue). If present, downward-pointing triangles show 2σ upper limits, while upward-pointing triangles are lower limits. The complete figure set (202 figures) is available in the online journal.

(The complete figure set (202 images) is available.)

Figure A4. MCMC corner plot for the absorber at $z = 2.54918$ toward J002127–020333. The histograms along the diagonal show the pdf's for $N_{H\,I}$, hydrogen number density (n_H), [C/α], and metallicity, respectively. The contour plots below the diagonal show the joint posterior pdf's of the given row and column. For this absorber, there is not enough constraint on the metal column densities to derive the metallicity and [C/α] without using priors (see Figure A3). The complete figure set (202 figures) is available in the online journal.

(The complete figure set (202 images) is available.)

ORCID iDs

Nicolas Lehner https://orcid.org/0000-0001-9158-0829
Claire Kopenhafer https://orcid.org/0000-0001-5158-1966
John M. O'Meara https://orcid.org/0000-0002-7893-1054
J. Christopher Howk https://orcid.org/0000-0002-2591-3792
Michele Fumagalli https://orcid.org/0000-0001-6676-3842
J. Xavier Prochaska https://orcid.org/0000-0002-7738-6875
Ayan Acharyya https://orcid.org/0000-0003-4804-7142
Brian W. O'Shea https://orcid.org/0000-0002-2786-0348
Molly S. Peeples https://orcid.org/0000-0003-1455-8788

Jason Tumlinson https://orcid.org/0000-0002-7982-412X
Cameron B. Hummels https://orcid.org/0000-0002-3817-8133

References

Aguirre, A., Schaye, J., Kim, T.-S., et al. 2004, ApJ, 602, 38
Akerman, C. J., Carigi, L., Nissen, P. E., Pettini, M., & Asplund, M. 2004, A&A, 414, 931
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Berg, M. A., Lehner, N., Howk, J. C., et al. 2022, arXiv:2204.13229

Berg, T. A. M., Fumagalli, M., D'Odorico, V., et al. 2021, MNRAS, 502, 4009
Boyd, B., Silvia, D., O'Shea, B., et al. 2020, JOSS, 5, 2581
Brummel-Smith, C., Bryan, G., Butsky, I., et al. 2019, JOSS, 4, 1636
Bryan, G. L., Norman, M. L., O'Shea, B. W., et al. 2014, ApJS, 211, 19
Buitinck, L., Louppe, G., Blondel, M., et al. 2013, arXiv:1309.0238
Butsky, I. S., & Quinn, T. R. 2018, ApJ, 868, 108
Cescutti, G., Matteucci, F., McWilliam, A., & Chiappini, C. 2009, A&A, 505, 605
Cooke, R., Pettini, M., Steidel, C. C., Rudie, G. C., & Nissen, P. E. 2011, MNRAS, 417, 1534
Cooke, R. J., Pettini, M., Nollett, K. M., & Jorgenson, R. 2016, ApJ, 830, 148
Cooksey, K. L., Prochaska, J. X., Chen, H.-W., Mulchaey, J. S., & Weiner, B. J. 2008, ApJ, 676, 262
Cooper, T. J., Rudie, G. C., Chen, H.-W., et al. 2021, MNRAS, 508, 4359
Corlies, L., Peeples, M. S., Tumlinson, J., et al. 2020, ApJ, 896, 125
Cowie, L. L., Songaila, A., Kim, T.-S., & Hu, E. M. 1995, AJ, 109, 1522
Crighton, N. H. M., Bechtold, J., Carswell, R. F., et al. 2013, MNRAS, 433, 178
Crighton, N. H. M., Hennawi, J. F., Simcoe, R. A., et al. 2015, MNRAS, 446, 18
Crighton, N. H. M., O'Meara, J. M., & Murphy, M. T. 2016, MNRAS, 457, L44
Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Natur, 457, 451
D'Odorico, V., Calura, F., Cristiani, S., & Viel, M. 2010, MNRAS, 401, 2715
D'Odorico, V., Cupani, G., Cristiani, S., et al. 2013, MNRAS, 435, 1198
Ellison, S. L., Songaila, A., Schaye, J., & Pettini, M. 2000, AJ, 120, 1175
Emerick, A., Bryan, G. L., & Mac Low, M.-M. 2019, MNRAS, 482, 1304
Fabbian, D., Khomenko, E., Moreno-Insertis, F., & Nordlund, Å. 2010, ApJ, 724, 1536
Feigelson, E. D., & Nelson, P. I. 1985, ApJ, 293, 192
Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, RMxAA, 49, 137
Ferrara, A., Pettini, M., & Shchekinov, Y. 2000, MNRAS, 319, 539
Fitzpatrick, E. L., & Spitzer, L., Jr. 1997, ApJ, 475, 623
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Frebel, A., Johnson, J. L., & Bromm, V. 2007, MNRAS, 380, L40
Fukugita, M., Hogan, C. J., & Peebles, P. J. E. 1998, ApJ, 503, 518
Fumagalli, M., Cantalupo, S., Dekel, A., et al. 2016a, MNRAS, 462, 1978
Fumagalli, M., Mackenzie, R., Trayford, J., et al. 2017, MNRAS, 471, 3686
Fumagalli, M., O'Meara, J. M., & Prochaska, J. X. 2011a, Sci, 334, 1245
Fumagalli, M., O'Meara, J. M., & Prochaska, J. X. 2016b, MNRAS, 455, 4100
Fumagalli, M., Prochaska, J. X., Kasen, D., et al. 2011b, MNRAS, 418, 1796
Gibson, J. L., Lehner, N., Oppenheimer, B. D., et al. 2022, AJ, 164, 9
Haardt, F., & Madau, P. 1996, ApJ, 461, 20
Haardt, F., & Madau, P. 2001, in Clusters of Galaxies and the High Redshift Universe Observed in X-rays, ed. D. M. Neumann & J. T. T. Van (Paris: CEA), 64
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Hafen, Z., Faucher-Giguère, C.-A., Anglés-Alcázar, D., et al. 2017, MNRAS, 469, 2292
Haislmaier, K. J., Tripp, T. M., Katz, N., et al. 2021, MNRAS, 502, 4993
Howk, J. C., Ribaudo, J. S., Lehner, N., Prochaska, J. X., & Chen, H.-W. 2009, MNRAS, 396, 1875
Hummels, C. B., Smith, B. D., Hopkins, P. F., et al. 2019, ApJ, 882, 156
Hummels, C. B., Smith, B. D., & Silvia, D. W. 2017, ApJ, 847, 59
Hunter, J. D. 2007, CSE, 9, 90
Jenkins, E. B. 2009, ApJ, 700, 1299
Jenkins, E. B., & Wallerstein, G. 2017, ApJ, 838, 85
Ji, S., Chan, T. K., Hummels, C. B., et al. 2020, MNRAS, 496, 4221
Jorgenson, R. A., Murphy, M. T., & Thompson, R. 2013, MNRAS, 435, 482
Kacprzak, G. G., Churchill, C. W., Steidel, C. C., Spitler, L. R., & Holtzman, J. A. 2012, MNRAS, 427, 3029
Kacprzak, G. G., Pointon, S. K., Nielsen, N. M., et al. 2019, ApJ, 886, 91
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Khaire, V., & Srianand, R. 2019, MNRAS, 484, 4174
Kim, T. S., Cristiani, S., & D'Odorico, S. 2001, A&A, 373, 757
Lehner, N., Berek, S. C., Howk, J. C., et al. 2020, ApJ, 900, 9
Lehner, N., Howk, J. C., Tripp, T. M., et al. 2013, ApJ, 770, 138
Lehner, N., O'Meara, J. M., Fox, A. J., et al. 2014, ApJ, 788, 119
Lehner, N., O'Meara, J. M., Howk, J. C., Prochaska, J. X., & Fumagalli, M. 2016, ApJ, 833, 283
Lehner, N., Prochaska, J. X., Kobulnicky, H. A., et al. 2009, ApJ, 694, 734
Lehner, N., Savage, B. D., Richter, P., et al. 2007, ApJ, 658, 680
Lehner, N., Wotta, C. B., Howk, J. C., et al. 2018, ApJ, 866, 33
Lehner, N., Wotta, C. B., Howk, J. C., et al. 2019, ApJ, 887, 5
Lehner, N., Zech, W. F., Howk, J. C., & Savage, B. D. 2011, ApJ, 727, 46
Lochhaas, C., Tumlinson, J., O'Shea, B. W., et al. 2021, ApJ, 922, 121
Lofthouse, E. K., Fumagalli, M., Fossati, M., et al. 2020, MNRAS, 491, 2057
Macquart, J. P., Prochaska, J. X., McQuinn, M., et al. 2020, Natur, 581, 391

Mandelker, N., van den Bosch, F. C., Springel, V., et al. 2021, ApJ, 923, 115
Mandelker, N., van den Bosch, F. C., Springel, V., & van de Voort, F. 2019, ApJL, 881, L20
Marra, R., Churchill, C. W., Doughty, C., et al. 2021, MNRAS, 508, 4938
Mattsson, L. 2010, A&A, 515, A68
McQuinn, M. 2016, ARA&A, 54, 313
Morton, D. C. 2003, ApJS, 149, 205
O'Meara, J. M., Lehner, N., Howk, J. C., et al. 2015, AJ, 150, 111
O'Meara, J. M., Lehner, N., Howk, J. C., et al. 2017, AJ, 154, 114
O'Meara, J. M., Prochaska, J. X., Burles, S., et al. 2007, ApJ, 656, 666
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JMLR, 12, 2825
Peeples, M. S., Corlies, L., Tumlinson, J., et al. 2019, ApJ, 873, 129
Penprase, B. E., Prochaska, J. X., Sargent, W. L. W., Toro-Martinez, I., & Beeler, D. J. 2010, ApJ, 721, 1
Péroux, C., & Howk, J. C. 2020, ARA&A, 58, 363
Pettini, M., Ellison, S. L., Steidel, C. C., & Bowen, D. V. 1999, ApJ, 510, 576
Pettini, M., Smith, L. J., King, D. L., & Hunstead, R. W. 1997, ApJ, 486, 665
Pettini, M., Zych, B. J., Steidel, C. C., & Chaffee, F. H. 2008, MNRAS, 385, 2011
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Pontzen, A., & Tremmel, M. 2018, ApJS, 237, 23
Prochaska, J. X. 1999, ApJL, 511, L71
Prochaska, J. X., Gawiser, E., Wolfe, A. M., Castro, S., & Djorgovski, S. G. 2003, ApJL, 595, L9
Prochaska, J. X., Herbert-Fort, S., & Wolfe, A. M. 2005, ApJ, 635, 123
Prochaska, J. X., Madau, P., O'Meara, J. M., & Fumagalli, M. 2014, MNRAS, 438, 476
Prochaska, J. X., O'Meara, J. M., Fumagalli, M., Bernstein, R. A., & Burles, S. M. 2015, ApJS, 221, 2
Prochaska, J. X., O'Meara, J. M., Herbert-Fort, S., et al. 2006, ApJL, 648, L97
Prochaska, J. X., O'Meara, J. M., & Worseck, G. 2010, ApJ, 718, 392
Prochaska, J. X., Tejos, N., Wotta, C. B., et al. 2017a, pyigm/pyigm: Initial Release for Publications, v1.0, Zenodo, doi: 10.5281/zenodo.437188
Prochaska, J. X., Werk, J. K., Worseck, G., et al. 2017b, ApJ, 837, 169
Prochter, G. E., Prochaska, J. X., O'Meara, J. M., Burles, S., & Bernstein, R. A. 2010, ApJ, 708, 1221
Rafelski, M., Wolfe, A. M., Prochaska, J. X., Neeleman, M., & Mendez, A. J. 2012, ApJ, 755, 89
Rahmati, A., & Oppenheimer, B. D. 2018, MNRAS, 476, 4865
Rauch, M., Miralda-Escudé, J., Sargent, W. L. W., et al. 1997, ApJ, 489, 7
Rhodin, N. H. P., Agertz, O., Christensen, L., Renaud, F., & Fynbo, J. P. U. 2019, MNRAS, 488, 3634
Ribaudo, J., Lehner, N., Howk, J. C., et al. 2011, ApJ, 743, 207
Rudie, G. C., Steidel, C. C., & Pettini, M. 2012, ApJL, 757, L30
Savage, B. D., & Sembach, K. R. 1991, ApJ, 379, 245
Savage, B. D., & Sembach, K. R. 1996, ARA&A, 34, 279
Scannapieco, E. 2005, ApJL, 624, L1
Schaye, J. 2001a, ApJ, 559, 507
Schaye, J. 2001b, ApJL, 562, L95
Schaye, J., Aguirre, A., Kim, T.-S., et al. 2003, ApJ, 596, 768
Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., & Sargent, W. L. W. 2000, MNRAS, 318, 817
Sembach, K. R., & Savage, B. D. 1992, ApJS, 83, 147
Shull, J. M., Smith, B. D., & Danforth, C. W. 2012, ApJ, 759, 23
Simcoe, R. A. 2011, ApJ, 738, 159
Simcoe, R. A., Sargent, W. L. W., & Rauch, M. 2004, ApJ, 606, 92
Simons, R. C., Peeples, M. S., Tumlinson, J., et al. 2020, ApJ, 905, 167
Songaila, A. 1998, AJ, 115, 2184
Songaila, A. 2005, AJ, 130, 1996
Suresh, J., Nelson, D., Genel, S., Rubin, K. H. R., & Hernquist, L. 2019, MNRAS, 483, 4040
Tripp, T. M., Jenkins, E. B., Bowen, D. V., et al. 2005, ApJ, 619, 714
Tumlinson, J., Peeples, M. S., & Werk, J. K. 2017, ARA&A, 55, 389
Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, ApJS, 192, 9
Tytler, D. 1987, ApJ, 321, 49
van de Voort, F., Springel, V., Mandelker, N., van den Bosch, F. C., & Pakmor, R. 2019, MNRAS, 482, L85
Weinberg, D. H., Miralda-Escudé, J., Hernquist, L., & Katz, N. 1997, ApJ, 490, 564
Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, ARA&A, 43, 861
Wotta, C. B., Lehner, N., Howk, J. C., et al. 2019, ApJ, 872, 81
Wotta, C. B., Lehner, N., Howk, J. C., O'Meara, J. M., & Prochaska, J. X. 2016, ApJ, 831, 95
Zahedy, F. S., Chen, H.-W., Cooper, T. M., et al. 2021, MNRAS, 506, 877
Zheng, Y., Peeples, M. S., O'Shea, B. W., et al. 2020, ApJ, 896, 143