Sub-damped Lyman α systems in the XQ-100 survey – I. Identification and contribution to the cosmological H I budget

Trystyn A. M. Berg [1,2]★ Sara L. Ellison,[2] Rubén Sánchez-Ramírez,[3,4]
Sebastián López,[5] Valentina D'Odorico,[6,7] George D. Becker [8] Lise Christensen [9]
Guido Cupani,[6] Kelly D. Denney[10] and Gábor Worseck[11,12]

[1]*European Southern Observatory, Alonso de Cordova 3107, Casilla 19001, Santiago, Chile*
[2]*Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia V8P 1A1, Canada*
[3]*INAF, Istituto di Astrofisica e Planetologia Spaziali, Via Fosso del Cavaliere 100, I-00133 Roma, Italy*
[4]*Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía s/n, E-18008 Granada, Spain*
[5]*Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile*
[6]*INAF-Osservatorio Astronomico di Trieste, Via Tiepolo 11, I-34143 Trieste, Italy*
[7]*Scuola Normale Superiore Piazza dei Cavalieri, 7, I-56126 Pisa, Italy*
[8]*Department of Physics & Astronomy, University of California, Riverside, CA 92521, USA*
[9]*DARK, Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK-2100 Copenhagen, Denmark*
[10]*Department of Astronomy, The Ohio State University, 140 West 18th Avenue, Columbus, OH 43210, USA*
[11]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[12]*Institut für Physik und Astronomie, Universität Potsdam, Karl-Liebknecht-Str 24/25, D-14476 Potsdam, Germany*

ABSTRACT

Sub-damped Lyman α systems (subDLAs; H I column densities of $19.0 \leq \log N(\text{H I}) < 20.3$) are rarely included in the cosmic H I census performed at redshifts $z_{\text{abs}} \gtrsim 1.5$, yet are expected to contribute significantly to the overall H I mass budget of the Universe. In this paper, we present a blindly selected sample of 155 subDLAs found along 100 quasar sightlines (with a redshift path-length $\Delta X = 475$) in the XQ-100 legacy survey to investigate the contribution of subDLAs to the H I mass density of the Universe. The impact of X-Shooter's spectral resolution on Ly α absorber identification is evaluated, and found to be sufficient for reliably finding absorbers down to a column density of $\log N(\text{H I}) \geq 18.9$. We compared the implications of searching for subDLAs solely using H I absorption versus the use of metal lines to confirm the identification, and found that metal-selection techniques would have missed 75 subDLAs. Using a bootstrap Monte Carlo simulation, we computed the column density distribution function ($f(N, X)$) and the cosmological H I mass density ($\Omega_{\text{H I}}$) of subDLAs and compared with our previous work based on the XQ-100 damped Lyman α systems. We do not find any significant redshift evolution in $f(N, X)$ or $\Omega_{\text{H I}}$ for subDLAs. However, subDLAs contribute 10–20 per cent of the total $\Omega_{\text{H I}}$ measured at redshifts $2 < z < 5$, and thus have a small but significant contribution to the H I budget of the Universe.

Key words: galaxies: high-redshift – galaxies: ISM – quasars: absorption lines.

1 INTRODUCTION

Studying the evolution of neutral gas reservoirs over cosmic time provides key constraints on aspects of galaxy evolution. The neutral atomic gas, primarily traced by H I, is thought to be an indicator of the gas reservoirs that eventually form stars in galaxies (Bird et al. 2015; Somerville & Davé 2015). Furthermore, the rate at which the H I reservoirs change places direct constraints on the processes that

can nourish or prevent future star formation (Davé et al. 2013; Lilly et al. 2013).

The comoving H I mass density ($\Omega_{\text{H I}}$; measured with respect to the critical mass density of the Universe) has typically been used to quantify the amount of neutral gas in galaxies. At $z \approx 0$, $\Omega_{\text{H I}}$ is quantified from 21-cm emission density maps (Zwaan et al. 2005; Martin et al. 2010; Braun 2012). However at redshifts $z \gtrsim 0.2$ where current radio facilities cannot easily detect the faint 21-cm emission, quasars (QSO) are used to probe sightlines through intervening H I clouds in the form of Ly α absorption (Lanzetta, Wolfe & Turnshek 1995; Rao & Turnshek 2000; Storrie-Lombardi & Wolfe

★ E-mail: trystyn.berg@eso.org

2000; Ellison, Ryan & Prochaska 2001; Péroux et al. 2003b; Prochaska, Herbert-Fort & Wolfe 2005; Rao, Turnshek & Nestor 2006; Prochaska & Wolfe 2009; Noterdaeme et al. 2012; Zafar et al. 2013; Crighton et al. 2015; Neeleman et al. 2016; Sánchez-Ramírez et al. 2016).

Of the many classes of Ly α absorption line systems seen towards QSOs, damped Lyman α systems (DLAs), defined as having H I column densities of logN(H I) $\geq$ 20.3 (Wolfe et al. 1986; Wolfe, Gawiser & Prochaska 2005), are typically used to measure $\Omega_{\mathrm{H I}}$. Although few in number compared to lower H I column density counterparts, DLAs dominate the H I column density distribution from $z_{\mathrm{abs}} \sim 5$ to the present epoch (particularly those with logN(H I) $>$ 21; Wolfe et al. 2005; Prochaska et al. 2005; Noterdaeme et al. 2012). However, subDLAs (19.0 $\leq$ logN(H I) $<$ 20.3) are thought to host a modest contribution of $\approx$10–20 per cent to the $\Omega_{\mathrm{H I}}$ budget of the Universe (O'Meara et al. 2007; Zafar et al. 2013). The difficulty in using subDLAs to probe the neutral gas reservoirs of the Universe is that these systems are often not fully self-shielded from the cosmic UV background, and thus do not completely trace neutral gas reservoirs (e.g. Zheng & Miralda-Escudé 2002; O'Meara et al. 2007). Nevertheless, there has been a substantial body of work calculating the contribution of H I from subDLAs to the column density distribution function ($f(N, X)$) and $\Omega_{\mathrm{H I}}$ (Péroux et al. 2003b,a, 2005; O'Meara et al. 2007; Guimarães et al. 2009; Zafar et al. 2013), as well as understanding whether subDLAs probe different types of gaseous systems than DLAs (Péroux et al. 2003a; Kulkarni et al. 2007; Dessauges-Zavadsky, Ellison & Murphy 2009; Quiret et al. 2016).

It is expected that the observed $f(N, X)$ can place strong constraints on feedback and ionization prescriptions in cosmological simulations of neutral gas reservoirs in the Universe (e.g. Rahmati et al. 2013a; Bird et al. 2014). Both observational (Péroux et al. 2003a,b; O'Meara et al. 2007; Guimarães et al. 2009; Zafar et al. 2013) and simulation-based studies (Rahmati et al. 2013b, 2015; Villaescusa-Navarro et al. 2018) of the nature of $f(N, X)$ for subDLA column densities have suggested that subDLAs are more common at redshifts $z_{\mathrm{abs}} > 3.5$ and contribute more than $\approx$10–20 per cent of $\Omega_{\mathrm{H I}}$ traced solely by DLAs (Péroux et al. 2005; Guimarães et al. 2009; Noterdaeme et al. 2009; Zafar et al. 2013). To date, samples of subDLAs have typically been selected using the presence of metal lines rather than using solely H I absorption as is done with DLAs. Such a selection effect can potentially lead to sample incompleteness from missing metal-poor systems, and introduce biases in the comparison with DLAs. Additionally, most subDLA samples are too modest in size to detect any redshift evolution at $>2\sigma$ confidence (O'Meara et al. 2007; Guimarães et al. 2009; Zafar et al. 2013, with redshift path-lengths of $\Delta X = 105$, 378, and 193, respectively).

This paper aims to investigate the contribution of subDLAs to the $\Omega_{\mathrm{H I}}$ budget at high redshifts (2.5 $\leq z_{\mathrm{abs}} \leq$ 4.5) using a completely blind sample of subDLAs identified in the XQ-100 survey (López et al. 2016) solely based on their Lyman series absorption. We searched for all H I absorbers in the XQ-100 spectra down to a column density threshold of logN(H I) $\geq$ 18.8 (i.e. the subDLA threshold minus the typical logN(H I) error of $\pm$0.2 dex) in order to evaluate the completeness of our absorber identification and properly account for misclassification of absorbers as subDLAs in our computation of $f(N, X)$ and $\Omega_{\mathrm{H I}}$. However, the standard subDLA H I column density limits are still adopted in all our computations. The reader is referred to Sánchez-Ramírez et al. (2016) for the H I statistics of the XQ-100 DLA sample. Throughout this paper we

assume a flat ΛCDM Universe with $H_0 = 70.0$ km s^{-1} Mpc^{-1} and $\Omega_{\mathrm{M, 0}} = 0.3$.

2 DATA

2.1 XQ-100 Survey design and data reduction

The XQ-100 Legacy Survey (PI S. López) observed 100 QSO sightlines at redshift $z_{\mathrm{em}} \sim$ 3.5–4.5 with the X-Shooter Spectrograph (Vernet et al. 2011) on the Very Large Telescope (VLT) in Chile. The QSOs were purposefully chosen to be blind to any intervening absorption line systems, thus providing a random sample of sub-DLAs and DLAs to study the cosmological implications of these intervening absorbers (Berg et al. 2016, 2017; Sánchez-Ramírez et al. 2016; Christensen et al. 2017). For more details about the XQ-100 science cases and survey design, see López et al. (2016).

For each QSO in the XQ-100 survey, the per arm exposures were either $\sim$0.5 or $\sim$1 h in length (depending whether the QSO was classified as 'bright' or 'faint', respectively), providing median signal-to-noise ratios of $\approx$30 pixel^{-1}. The spectra for each arm were reduced using an internal IDL package developed by G. Becker and G. Cupani. For more details on the data reduction, see López et al. (2016).

As the Ly α forest and intervening gas clouds along every sightline partially absorb all QSO photons below the Lyman limit, the QSO's flux is systematically absorbed for all wavelengths below the corresponding observed Lyman Limit at the redshift of the QSO ($\lambda_{\mathrm{LL}} = (1 + z_{\mathrm{em}}) \times 911.7633$ Å), making both accurate continuum placement and subDLA identification difficult below this wavelength. For this reason, the wavelength regime below the Lyman limit is consequently not used for identifying the subDLAs in this work. The resulting redshift path-length searched for subDLAs is $\Delta X = 475$.

2.2 Identification of intervening absorbers

Starting with the manually placed QSO continuum fits outlined in López et al. (2016), we visually identified candidate subDLAs by systematically scrolling through the XQ-100 spectra starting from the Ly α emission of the QSO to λ_{LL}. Candidate absorbers are first identified by any saturated absorption broad enough to fit a simulated Ly α Voigt profile of column density logN(H I) $\geq$ 18.8 (assuming a typical Doppler parameter of $b = 10$ km s^{-1}; e.g. Dessauges-Zavadsky et al. 2003). Note that our adopted logN(H I) $\geq$ 18.8 search threshold is below the minimum subDLA threshold (logN(H I) = 19.0) by the typical H I column density error (i.e. $\pm$0.2) in order to enable a complete bootstrap error analysis (Section 3.1). Simultaneously to the Ly α fitting, a simulated higher order Lyman series line profile (usually Ly β) was checked at the corresponding wavelength. For each of these saturated regions identified, one or more Voigt profiles were fitted by adjusting simultaneously the original continuum fits, the logN(H I), and z_{abs} until the modelled Ly α and Ly β Voigt profiles matched the continuum-normalized spectra. Continuum adjustments were common for large column density absorption systems on top or near QSO emission lines. In many of these adjustments, the best-fitting column logN(H I) was constrained by simultaneously fitting both the Ly α and Ly β profiles. Errors on the logN(H I) column density were estimated as the minimum and maximum possible logN(H I) that would provide a visually reasonable fit to the data. For cases where the continuum was uncertain, the assigned logN(H I) error was increased to incorporate this uncertainty in the continuum placement.

To ensure higher redshift Lyman series lines were not confused with lower redshift Ly α absorption, a full model spectrum was generated to keep track of all fitted absorption systems along each sightline. Upon identifying a new absorber, the associated Voigt profiles for several key Lyman series lines (from Ly α to Ly δ) were added to the full model spectrum. The identification of systems was terminated when the observed-frame wavelength of additional Ly α absorbers was shorter than λ_{LL}.

Once the absorber identification process was complete, the initial list of absorption line systems was cleaned by adjusting or removing any absorbing systems whose modelled Lyman series lines (Ly α–Ly δ, and the highest order Lyman series lines between 911 and 920 Å) did not fit the continuum normalized spectra. The final cleaned sample of identified candidate absorbers consists of 87 $18.8 \leq \mathrm{logN(H\,I)} < 19.0$ Lyman limit systems, 155 sub-DLAs, and 34 DLAs. Fig. 1 shows the best-fitting Ly α and Ly β profiles for a representative sample of 14 absorbing systems, while the remaining profiles are displayed in Appendix A (Fig. A1; available online). Table 1 contains the full list of 276 candidate absorbers, including their z_{abs} and logN(H I).

Given the moderate resolution of the XQ-100 spectra, it is sometimes difficult to model regions of extended saturated absorption where the damping wings of the Voigt profile are not visible into individual absorption features. The troughs could be fit reasonably by one or more subDLAs, but also by several lower column density Lyman Limit systems ($17 < \mathrm{logN(H\,I)} < 19$) that are clustered together. Unless a higher order Lyman series line could disentangle this degeneracy (such as the absorbers seen towards J0034+1639 at $z_{abs} \approx 4.2$; see Fig. A1), no candidate absorbers were fitted to the data in these degenerate regions. In many cases, we note that the edges of the Lyman series absorption appear to be better fit by Lyman Limit systems with high Doppler parameters rather than subDLAs. We therefore caution that our identification of subDLAs is potentially incomplete.

2.2.1 Comparison to Sánchez-Ramírez et al. (2016)

In Sánchez-Ramírez et al. (2016), we identified 41 DLAs in the XQ-100 sample using a similar absorber identification and fitting method. However, the identification of DLAs was completed over a slightly longer redshift path in Sánchez-Ramírez et al. (2016, $\Delta X = 536$) in that it was not truncated at λ_{LL} as was done here. As a result, the initial search from Sánchez-Ramírez et al. (2016) includes eight more DLAs, which have been removed from the sample in this paper. We also identify two XQ-100 DLAs missed in the Sánchez-Ramírez et al. (2016) catalogue; one system is towards J0214−0517 ($z_{abs} = 3.7211$, logN(H I) $= 20.6 \pm 0.2$), and the other is along the J1111−0804 sightline ($z_{abs} = 3.60690$, logN(H I) $= 20.4 \pm 0.15$).

In order to facilitate a fair comparison between the H I contributions of subDLAs and DLAs in XQ-100, it is important that we use the same continuum fits for both the DLA and subDLA identification; thus we remeasured the N(H I) values for all the previously identified XQ-100 DLAs. All of the Sánchez-Ramírez et al. (2016) DLAs within the common redshift range of the two studies were recovered, and their column densities are consistent (within the H I column density errors) with those that we derived in our subDLA search. The largest discrepancies occurred in regions where the continuum is difficult to constrain, such as systems that are nearby the background QSO's Lyα emission line. For this paper,

we use the newly measured logN(H I) values of the XQ-100 DLAs for consistency.

2.3 Completeness analysis

2.3.1 Recovery test

To assess the completeness of the visual identification process and ensure that the modest XQ-100 spectral resolution (FWHM resolution $R \sim 5400$–8900) did not affect the determination of column densities, an absorption line recovery test was necessary. To ensure the Ly α forest statistics and X-Shooter instrumental effects are reproduced while maintaining the same continuum fits used in the original identification process, artificial absorption systems were randomly inserted into the XQ-100 spectra.

Prior to insertion, a list of artificial systems was first generated to ensure that an approximately uniform distribution of column densities was used. In bins of column density from logN(H I) $= 18.0$ to 20.3[1] (of bin width 0.1 dex), 15 ± 3 systems were generated with column densities uniformly selected within the bin's limits. Therefore the total number of systems (between 276 and 414 systems) was randomly generated, preventing any subconscious counting during the recovery process. For each of these fake systems, a model Lyman series (down to Ly δ) spectrum was generated, and randomly inserted into one of the XQ-100 QSO spectra at a random redshift. To simulate noise in the model spectrum, the flux of the model spectra was then varied randomly using the XQ-100 error spectrum flux as the standard deviation of a Gaussian distribution. Between 0 and 7 fake absorption systems were assigned to each XQ-100 sightline.

The identification process described in Section 2.2 was repeated on the spectra with fake absorbers inserted, starting with the locations of the previously identified real absorbers flagged in the model spectra. The top panel of Fig. 2 shows the column density distribution of fake systems inserted into the XQ-100 spectra (with observed-frame Ly α absorption at wavelengths larger than λ_{LL}; black line). The solid red bars indicate that we missed seven, logN(H I) ≤ 19.2 systems in our recovery analysis. However, the three missed systems with the higher column densities (logN(H I) $\geq$ 18.9) happened to be randomly inserted on top of already identified, higher column density Ly α absorbers. Since the H I contribution from real overlapping systems would have already be accounted in the original fit to the larger H I systems, we do not include these three missed systems in the recovery rate derivation. Therefore only $\approx$27 per cent of systems with logN(H I) $= 18.8$ were missed in the recovery analysis.

For all the recovered systems, the bottom panel of Fig. 2 shows the difference in logN(H I) between the measured column density of the fake absorber (N_{Meas}) and the randomly generated value (N_{Fake}). All but 10 of the recovered systems ($\approx$3 per cent of the total number of systems) have consistent column densities with N_{Fake}. For these 10 systems, the column densities are only inconsistent by $\sim$0.1 dex, which is statistically consistent with the random flux errors inserted into the fake model spectra.

---

[1]The recovery rate for DLAs in the XQ-100 data set was previously measured at 100 per cent in Sánchez-Ramírez et al. (2016). Given that we recovered all the previously identified DLAs in an identical recovery test in Sánchez-Ramírez et al. (2016), we did not repeat the experiment for DLAs.
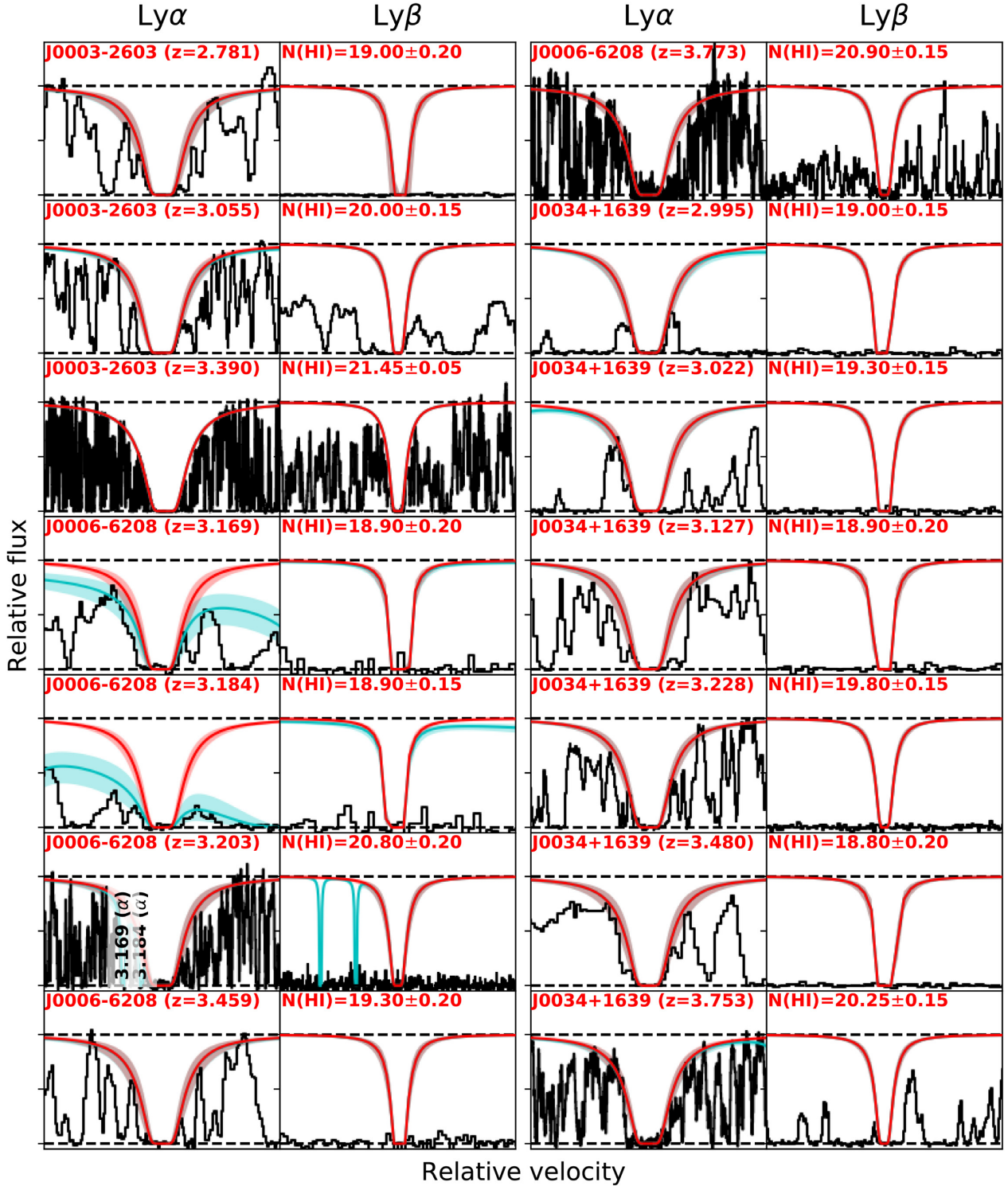
Figure 1. Each pair of panels displays the fitted Ly α (first and third columns) and Ly β (second and fourth columns) profiles (dark red lines) and associated errors (dark red shaded region) for 14 of the identified candidate system. The light blue line and shaded region denote the full model spectrum and associated error combined from all absorbers within the entire sightline. The red text denotes the properties of the absorber [QSO sightline name and redshift in the corresponding Ly α panel, and logN(H I) in Ly β panel], while the vertical black text labels the redshift and Lyman series lines associated with additional absorption included in the model. The x-axis of each panel is scaled such that integrated region of the Ly series line's Voigt profile displayed encompasses 98 per cent of the absorbed flux. The remaining 262 systems are displayed in Fig. A1 (available online).

Table 1. H I absorbers in XQ-100 (see online version for full table).

QSO	z_{em}	z_{abs}	logN(H I)	Metal absorption detected?
J0003−2603	4.010	2.78140	19.00 ± 0.20	N
J0003−2603	4.010	3.05490	20.00 ± 0.15	Y
J0003−2603	4.010	3.16080	19.10 ± 0.15	Y
J0003−2603	4.010	3.39020	21.45 ± 0.05	Y
J0003−2603	4.010	3.43470	18.90 ± 0.10	N
J0006−6208	4.460	3.16890	18.90 ± 0.20	N
J0006−6208	4.460	3.18430	18.90 ± 0.15	N
J0006−6208	4.460	3.20250	20.80 ± 0.20	Y
J0006−6208	4.460	3.45890	19.30 ± 0.20	N
J0006−6208	4.460	3.77300	20.90 ± 0.15	Y
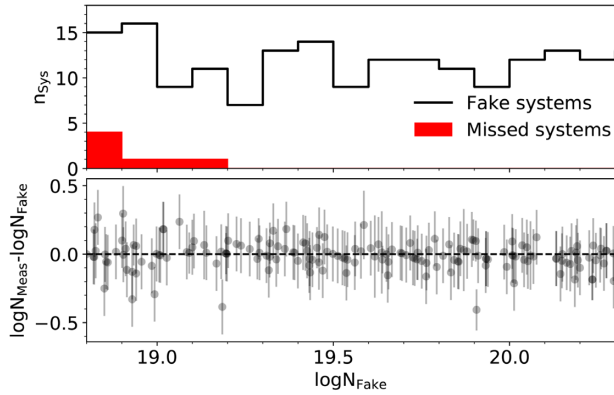
Figure 2. *Top panel:* Distribution of the number of fake systems of a given H I column density (logN$_{Fake}$) used for the recovery test (black line) and those systems that were missed in the recovery process (solid red histogram). Note that only systems whose Ly α is beyond $\geq \lambda_{LL}$ are included in this plot. *Bottom panel:* The difference in measured and fake logN(H I) column densities (i.e. logN$_{Meas}$−logN$_{Fake}$) as a function of logN$_{Fake}$.

In summary, the H I recovery test suggests that our identification of XQ-100 absorbers is 100 per cent complete for logN(H I) $\geq$ 18.9, and 73 per cent complete for logN(H I) = 18.8.

2.3.2 Effects from continuum placement

The location of the continuum becomes more uncertain towards bluer wavelengths as there is more absorption from additional Lyman series lines (i.e. Ly β and Ly γ) removing regions of clean continuum. To test whether or not our continuum placements introduce significant errors in our column density determinations, we artificially inserted a fake absorption line in the *un-normalized* spectrum and fitted the line using our continuum placements (without continuum adjustment). This experiment was done in three regions within the rest frame of the QSO, near the wavelengths: λ = 1120.7, 999.1, and 961.1 Å. These three regions correspond to locations in the QSO spectrum where there is only Ly α, Ly α + Ly β, and Ly α + Ly β + Ly γ absorption, respectively. In order to remove the effects of individual sightlines on the experiment, five QSO spectra were selected at random for each region. However, we required that each of the selected QSOs had no previously identified Ly α absorption within the region where the line was inserted to avoid any effects from previous continuum placement modifications during the Ly α fit. The column density of the fake absorption line started at logN(H I) = 18.8, and was increased in steps of 0.2 dex up

to a column density of logN(H I) = 20.0 (inclusive) to test the effects as a function of column density. For each fake absorption line, the wavelength was held fixed. For all but one QSO, we were able to recover the column density to within 0.2 dex (the typical quoted error on the H I columns). These errors were not systematically offset in one particular direction (especially in the Ly α + Ly β and Ly α + Ly β + Ly γ region), suggesting there is no significant effect from the increased absorption (and thus more uncertain continuum) on our column density determinations. We note that for the one QSO sightline which provided inconsistent fits, the identification of the line was difficult due to significant blending with Ly α forest for the inserted Ly α profiles with column densities logN(H I) = 19.0 and 19.2, and the offsets were measured to be +0.3 dex each (i.e. the column densities were overestimated), and is likely not strongly influenced from the continuum placement.

2.3.3 N(H I) comparison to higher resolution data

Although the resolution of the XQ-100 data (FWHM resolution $R \sim$ 5400–8900) is sufficient to resolve the absorbing systems, it may be inadequate to resolve close blends of multiple lines. In order to test the absorber recovery and fit accuracy, we compared the fits derived from our X-Shooter data with spectra obtained with high resolution ($R \approx 80\,000$) observations with the UVES instrument. Using all available archival UVES spectra[2] of the XQ-100 sightlines (19 sightlines total; with signal to noise of $\approx$5–15 pixel^{-1}), we repeated the Ly α profile fitting at the already identified redshifts of 62 H I absorbers with UVES coverage. Comparing the obtained logN(H I) from both the UVES and XQ-100 fits, all column densities are consistent within the errors (at most 0.2 dex difference), with two exceptions. For these particular exceptions, the XQ-100 absorbers (J0134+0400, z_{abs} = 3.999; J0137−4224, z_{abs} = 3.665) each appear to be better fit by two, smaller component systems using the UVES data. However, the total H I column measured contained within the UVES absorbers remains consistent with the single H I profile fit measured in the XQ-100 data.

Despite the overall consistency between the UVES and XQ-100 data, we note a systematic difference in logN(H I) exists between the two data sets. Fig. 3 shows the mean N(H I) offset between the UVES and XQ-100 (ΔlogN(H I); measured relative to the XQ-100 value) as a function of logN(H I), and indicates that the UVES data tend to provide lower logN(H I) fits than the XQ-100 data for absorbers with logN(H I) < 19.8 by up to 0.1 dex on average. All the offsets are tabulated in Table 2. Table 2 also includes the fraction of system which have a non-zero ΔlogN(H I) (n_{off}) relative to the total number of absorbers (n_{tot}) within each column density bin. As both the VIS ($R \approx 8900$) and UVB ($R \approx 5400$) arms of X-Shooter have different spectral resolutions, we checked for a dependence of the measured ΔlogN(H I) for absorbers identified in the respective arms. Absorbers detected in the VIS arm have a slightly smaller ΔlogN(H I) than those found in the lower resolution UVB arm by at most $\approx$0.05 dex, so the effects of the differing arm resolutions are minimal.

2.3.4 False-positives and metal lines

In addition to ensuring that all the subDLAs and DLAs were identified in the XQ-100 data set, it is possible that two closely separated Lyman limit systems or other Lyman series absorption can
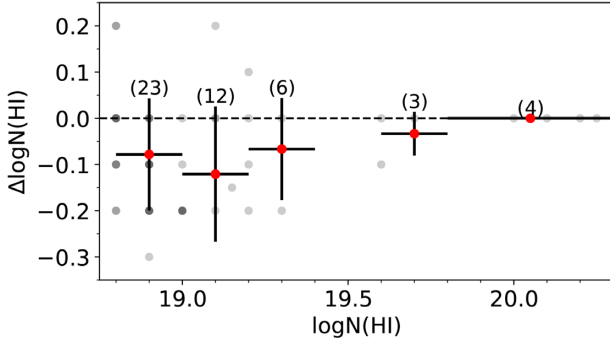
---

[2]Data accessed from http://archive.eso.org/wdb/wdb/eso/uves/form.

Figure 3. The red points show the mean offset of logN(H I) measured between UVES and XQ-100 (ΔlogN(H I); positive values imply a higher UVES-derived column density) in bins of column density. Error bars show the standard deviation in each bin, while the number in parenthesis above each data point indicate the number of absorbers within the bin. Offsets measured for individual absorbers are displayed as the grey circles, where a darker grey represent a higher density of systems with the same logN(H I) and ΔlogN(H I).

Table 2. Mean logN(H I) offsets between UVES and XQ-100 data.

XQ-100 logN(H I) range	ΔlogN(H I)	n_{off}/n_{tot}
$18.8 \leq$ logN(H I) < 19.0	-0.08	16/23
$19.0 \leq$ logN(H I) < 19.2	-0.10	8/13
$19.2 \leq$ logN(H I) < 19.4	-0.06	4/6
$19.4 \leq$ logN(H I) < 19.6	$-$	0/0
$19.6 \leq$ logN(H I) < 19.8	-0.03	1/3
$19.8 \leq$ logN(H I) < 20.3	0.0	0/4
logN(H I) ≥ 20.3	0.0	0/5

mimic a low column density subDLA, leading to falsely identified subDLA systems. For low-to-modest resolution spectra or low redshift absorption systems, the Lyman series lines are difficult to detect within the blanket Lyman limit absorption, or separate from the Ly β forest. One method to determine the false-positive rate is to use higher resolution data where it is easier to resolve distinct absorption lines. In Section 2.3, we found only two cases of the 63 systems where two Ly α lines were a better fit to the higher resolution UVES data, suggesting these false-positives are not a common occurrence in the XQ-100 data. However this is not a complete assessment of the false-positive rate as it is still possible that lines can be blended at even closer separations. We caution that our comparison with the UVES data does not uniformly probe the range in wavelength (with respect to the rest frame of the quasar) to properly assess the frequency of misidentification of systems due to blending.

Another common method to remove these 'false-positive' absorbers from a sample is to check for the presence of higher order Lyman series lines or metal lines associated with the absorber, particularly using high resolution data (Péroux et al. 2003a; O'Meara et al. 2007; Zafar et al. 2013). The typical metal lines used are: C II, C IV, Si II, Si IV, Mg II, or Fe II. However, using metal lines for confirmation of absorbers may either: (i) artificially remove systems
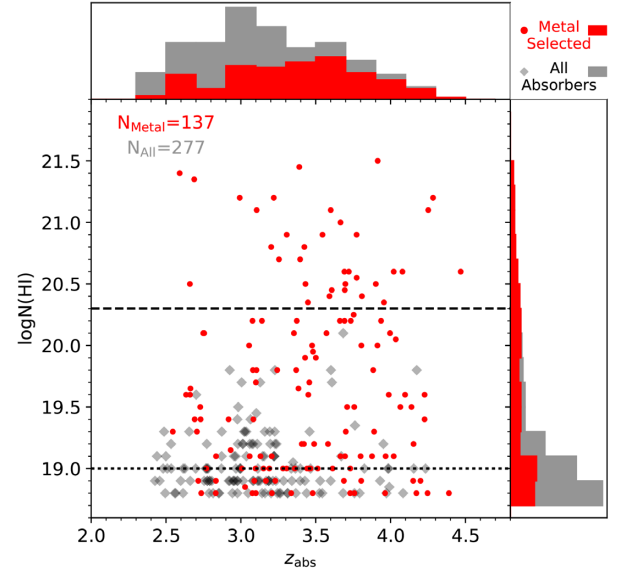

Figure 4. The H I column density (logN(H I)) of all candidate XQ-100 absorbers (light grey diamonds) as a function of their redshift (z_{abs}). The red circles indicate the XQ-100 metal-selected absorbers. The total number of absorbers (N) for each sample is given in the top left corner of the main panel. The distributions of z_{abs} and logN(H I) are given in the top and rightmost panels, using the same colour scheme for the two subsamples. For reference, the dotted and dashed horizontal lines denote the column density thresholds for subDLAs and DLAs, respectively.

with weak or no metal lines that are not easily identified (i.e. metal poor or systems with large dust or ionization corrections to neutral species), or (ii) select blends of lower N(H I) absorbers which still have strong metal lines. Either of these two cases can introduce different biases into the sample.

From the full sample (FS) of candidate absorption systems with logN(H I) $\geq$ 18.8, we identified a subsample of metal-selected absorbers, where at least one of the following ionic species is detected: C IV ($\lambda\lambda$ 1548 and 1550 Å), Si IV ($\lambda\lambda$ 1393 and 1402 Å), Mg II ($\lambda\lambda$ 2796 and 2803 Å), or Fe II (λ 2344, 2382, and 2600 Å). The systems with observed metal lines are identified by a 'Y' in the 'Metal absorption detected?' column of Table 1. Fig. 4 shows the distribution of logN(H I) and z_{abs} for all the absorbers with logN(H I) $\geq$ 18.8 (grey diamonds) and metal-selected absorbers (red circles). We point out that all 34 DLAs have metal lines associated. It is clear that the metal selection process preferentially removes low redshift, low H I-column density (logN(H I) $\leq$ 19.4) candidate absorbers from the full list of absorbers; these removed systems tend to be found where higher order Lyman series blending is more likely to produce false-positive Ly α absorption.

To understand what is driving the large discrepancy between the number of absorbers between metal-selected sample (MS) and all identified absorbers at low z_{abs} and H I, Fig. 5 shows the cumulative number of subDLAs as a function of redshift for all the absorbers (solid black lines) and the MS (dashed black lines). The distributions are split up into three panels for three different column density ranges. For comparison, the shaded regions represent the number of absorbers whose commonly detected metal diagnostic lines (i.e. C IV 1548 Å and Si IV 1393 Å; light blue and magenta, respectively) are outside of the Ly α forest. Although only C IV and Si IV doublets have been displayed, other commonly observed low-ionization species (e.g. Si II 1260 Å, C II 1334 Å) are bluer than the Si IV doublet, or are often blended with telluric lines at
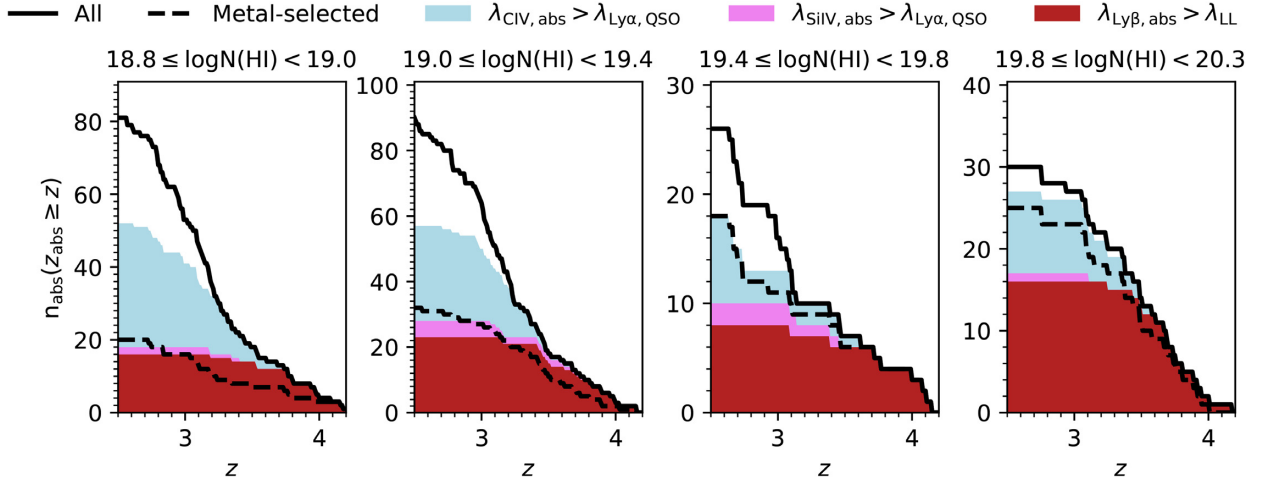
Figure 5. The cumulative distribution of XQ-100 absorbers (with $18.8 \le \log N(\text{H I}) < 20.3$) above a given redshift z ($n_{\text{abs}}(z_{\text{abs}} \ge z)$). The four panels split the absorbers into different $\log N(\text{H I})$ bins ([18.8,19.0), [19.0,19.4), [19.4,19.8), [19.8,20.3); from left to right). The solid and dashed black line denote the curves using all the identified absorbers and metal-selected systems, respectively. The shaded regions show the distribution of absorbers whose key diagnostic lines (Ly β, C IV, and Si IV; dark red, light blue, and magenta, respectively) are in uncontaminated regions of the spectrum (i.e. Ly β redward of the QSO's Lyman Limit; metal lines outside of the Ly α forest). Unshaded regions below the solid line denote the number of absorbers that cannot be confirmed as metal selected using either C IV or Si IV, and contains ≈ 40 per cent of all the absorbers identified (mostly below $z_{\text{abs}} \lesssim 3.2$ and $\log N(\text{H I}) \lesssim 19.4$). The shaded region above the dashed line estimates the inability to detect the given metal line (due to low signal to noise, blending, or being metal poor) combined with false-positives. This region is significantly populated by $\log N(\text{H I}) < 19.4$ absorbers.

these redshifts (e.g. Fe II 2586 Å, and the Mg II doublet at 2796 and 2803 Å); and thus these lines are used less frequently for metal-line identification in this work. 11 of the absorbers with confirmed metal lines ($19.3 \le \log N(\text{H I}) \le 20.2$) have been confirmed using only low-ionization metal lines; eight of these systems have blended C IV or Si IV, and one subDLA ($\log N(\text{H I}) = 19.8$) has no associated C IV and Si IV absorption.

The dark red shaded regions of Fig. 5 show the number of absorbers whose Ly β is redward of λ_{LL} to summarize the frequency in which can confirm the presence of H I absorption through Ly β. The redshift at which the metal-selected dashed line is greater than the red shaded region ($z_{\text{abs}} \lesssim 3.5$) denotes the highest redshift where it is no longer possible to remove false-positives in a metal-blind fashion using the Lyman series. This method for removing false-positives appears to be successful for the XQ-100 subDLAs, as the solid line for all $\log N(\text{H I}) \ge 19$ absorbers are very similar above $z_{\text{abs}} \gtrsim 3.5$. However, the large red-shaded region between the solid and dashed lines in the leftmost panels suggests that using Ly β absorption to remove false-positives becomes challenging for systems with low H I column densities ($\log N(\text{H I}) \lesssim 19.4$) up to $z_{\text{abs}} \approx 4$ in the XQ-100 data. Upon investigation, this difficulty in removing false-positive using Ly β is mostly due to blending of the weaker Ly β line of these absorbers with the Ly α forest. As a result it is more difficult to confirm the observed absorption is primarily from Ly β for $\log N(\text{H I}) < 19$ XQ-100 absorbers at all redshifts.

Of the 86 systems without detected metal lines, 77 of these systems have the corresponding C IV and Si IV absorption in the Ly α forest; this is represented by the unshaded regions below the solid black line. Thus for nearly 40 per cent of all identified subDLAs, *we cannot determine if the commonly detected metal lines are present and estimate the false-positive rate.* The subDLAs which populate the unshaded region are typically found at $z_{\text{abs}} \lesssim 3$ as a result of the XQ-100 QSO redshifts. We point out that this percentage is much less for subDLAs with $\log N(\text{H I}) \ge 19.8$ as neutral metal species typically outside the forest and not in telluric lines (e.g.

Fe II 1608 Å and Al II 1670 Å) are more often detected due to larger column densities of gas.

In Fig. 5, signatures of false-positives would manifest as shaded regions above the black dashed lines (i.e. no diagnostic absorption lines are detected). We caution that this shaded regime can also be driven by many different factors other than false-positives such as regions where the line is blended (particularly for Ly β; red shaded region), systems which are metal poor (blue and magenta regions only), or low signal-to-noise data. For the 14 absorbers which have undetected C IV or Si IV in uncontaminated regions of the spectrum, we measured detection thresholds of $\log N(\text{C IV}) \le 12.97$ and $\log N(\text{Si IV}) \le 12.76$. For C IV (which should be easily identified in absorbers within the shaded regime), the column density detection threshold is higher than the expected column densities observed in subDLAs from the literature ($\log N(\text{C IV}) \gtrsim 12.75$; Fox et al. 2007), and thus the XQ-100 sample can be limited by insufficient signal to noise to detect C IV for absorbers with $\log N(\text{H I}) < 19.4$ at lower redshift ($z < 3$). However, we note that generally higher metallicities are observed in low redshift subDLAs compared to DLAs (Pettini et al. 1999; Rafelski et al. 2014; Quiret et al. 2016).

To better understand whether the XQ-100 spectra have the capability of detecting low N(H I), metal-poor absorbers, we computed the signal-to-noise ratio (SNR) required to detect a neutral metal-line absorption line (with rest-frame wavelength λ and oscillator strength f), at 3σ significance, for an absorber with an observed metallicity [M/H] (i.e. no dust or ionization corrections) using

$$\text{SNR} = \frac{3 m_e c \text{ FWHM}}{\pi e_c^2 f \lambda^2} \left(N(H\,\text{I}) 10^{[M/H] + log\left(\frac{n_{\text{M}_\odot}}{n_{\text{H}_\odot}}\right)} \right)^{-1}; \quad (1)$$

where m_e and e_c are the electron mass and charge, and $\frac{n_{\text{M}_\odot}}{n_{\text{H}_\odot}}$ is the observed solar abundance of the species from Asplund et al. (2009). Equation (1) assumes the metal line is detected at the corresponding full width at half-maximum (FWHM; measured in the observed frame) of the instrumental resolution of the NIR arm ($R \approx 5600$).
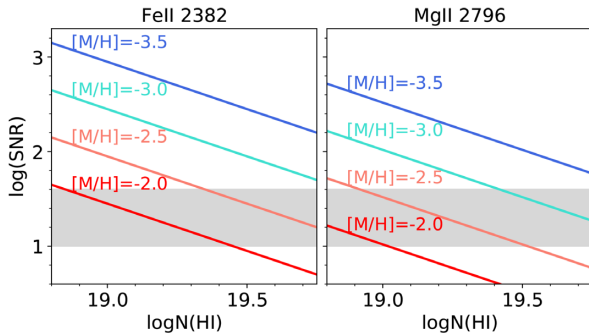
Figure 6. The predicted spectral SNR required to observe the metal lines
Fe II λ 2382 Å (left-hand panel) and Mg II λ 2796 Å (right-hand panel) as a
function of absorber logN(H I). A range in metallicities is represented by the
different coloured lines. The typical range in signal to noise of the XQ-100
spectra is denoted by the light grey region.

Fig. 6 shows the range of signal to noise required to detect the
strong neutral metal lines Fe II λ 2382 Å (left-hand panel) and
Mg II λ 2796 Å (right-hand panel), as a function of logN(H I) and
observed metallicity ([M/H]; coloured lines). The typical range of
SNR in the XQ-100 spectra is shown by the grey shaded regions.
Fig. 6 demonstrates that subDLAs should have detectable Fe II
2382 Å in our X-Shooter data for [Fe/H] > −2.0. Few subDLAs have
previously been detected with an observed [M/H] < −2 (Fumagalli,
O'Meara & Prochaska 2016; Quiret et al. 2016). Therefore, the
ability of our X-Shooter data to detect metal lines for typical
subDLA metallicities suggests that up to ≈46 per cent (128 systems
which do not have confirmed metal lines) of the candidate subDLAs
might be lower column density or blended systems. However the
lack of low-metallicity subDLAs in the literature could be a result
of metal-line selection, particularly as DLAs can probe down to
[M/H] ≈ −3.0 (Cooke, Pettini & Jorgenson 2015; Quiret et al.
2016). A detailed analysis of the metal abundances of the XQ-
100 subDLAs will be presented in a future paper (Berg et al., in
preparation).

In summary, the region between the solid (all identified absorbers
with logN(H I) ≥ 18.8) and dashed lines (metal-selected absorbers)
in Fig. 5 represents a combination of the false-positives in our FS
of absorbers, metal-poor absorbers ([M/H] ≲ −2), and absorbers
whose metal lines lie in regions suffering from blending or insuf-
ficient signal to noise to classify into the MS. As these two curves
encompass the two extremes in the number of subDLAs identified,
we use all the identified absorbers and MS as bounds to the true
number of subDLAs within the XQ-100 sample for our analysis. It is
worth emphasizing that the discrepancy in the number of absorbers
identified that do not have confirmed metal lines is strongest at
low column densities (logN(H I) ≤ 19.4) and for systems at low
redshifts where metal lines are shifted into the Ly α forest ($z_{abs} ≲$
3.2). Although the false-positive rate may be significant at low
column densities (19.0 ≤ logN(H I) ≤ 19.4), the difference in their
contribution to the total H I mass budget of subDLAs and DLAs is
only a couple of per cent (see Section 3.3).

3 RESULTS

In Sánchez-Ramírez et al. (2016), we computed both f(N, X) and
$Ω_{H I}$ for the XQ-100 DLA sample and analysed the evolution of
$Ω_{H I}$. In this section, we will present the added information obtained
from calculating f(N, X) and $Ω_{H I}$ from the XQ-100 subDLA sample.
For consistency with Sánchez-Ramírez et al. (2016), we restricted

the XQ-100 sample to exclude systems with an absorption redshift
within 5000 km s⁻¹ of the QSO emission redshift. These proximate
systems likely trace a different population of systems compared
to their intervening counterparts, (Ellison et al. 2002, 2010; Berg
et al. 2016; Perrotta et al. 2016) and are typically ignored when
computing $Ω_{H I}$. To properly measure f(N, X) for DLAs, a sample
size much larger than the 33 DLAs identified in the XQ-100 sample
must be used in order to statistically probe the high column density
end of f(N, X) (logN(H I) > 21.5). We thus combined the XQ-100
DLA sample with literature sample compiled in Sánchez-Ramírez
et al. (2016), which consists of 1019 DLAs measured in various
literature surveys across a redshift range of 1.6 < z_{abs} < 5.1. It is
important to note that this supplementary DLAs sample does not
include any subDLAs, and probes a very different redshift path than
the XQ-100 subDLAs. For this reason, we divide the entire XQ-100
absorber sample into DLAs and subDLAs in the rest of this analysis.

3.1 Bootstrap Monte Carlo resampling

To estimate the errors in deriving f(N, X) and $Ω_{H I}$ from the XQ-
100 data set, the bootstrap Monte Carlo (further abbreviated B-MC)
resampling technique presented in Sánchez-Ramírez et al. (2016)
was used. This technique provides a robust method of combining
the random errors quantified for measurements (such as logN(H I))
with the sampling errors inherent to a relatively small sample size.
In essence, the B-MC method takes the original absorber N(H I)
distribution and duplicates the sample a large number of times to
create N_{samp} simulated mock surveys. For each mock survey, the
logN(H I) of each absorbers is randomly varied within its errorbar
(assuming a Gaussian error distribution) to create a 'new' set
of logN(H I) measurements. From this new set of measurements,
absorbers are randomly drawn (with replacements) until a new
mock survey is generated with the same number of absorbers. As the
drawing of values is done with replacement, this mimics the effect
of sampling errors by randomly excluding or duplicating absorbers
in the sample. The choice of N_{samp} is arbitrary, but needs to be very
large (≳10 000) in order to provide a robust statistical sample.

We also experimented with statistically incorporating the effects
of using the modest resolution of X-Shooter to measure H I column
densities over high resolution data sets (e.g. UVES; see Section 2.3).
Using the frequency and average measured offset seen between the
X-Shooter and UVES column densities (as presented in Table 2),
we varied a random fraction of the column densities in the B-MC
simulation by the average measured N(H I) offset. After implement-
ing these offsets, we did not detect a significant difference in the
results of this paper. In addition to not having a sufficient sample
size of absorbers with UVES coverage to provide robust statistics
for different redshifts and column densities (see Section 2.3), we
did not implement this resolution experiment into the final B-MC
simulation.

As the XQ-100 absorber sample is split into two subsam-
ples based on column density (subDLAs and DLAs), the B-MC
technique also needs to take into account the effects of remov-
ing/including systems that could be misclassified into the wrong
H I absorber type (e.g. a subDLA with a measured logN(H I) =
20.1 ± 0.2 could be a DLA in reality, or conversely a logN(H I)
= 18.8 ± 0.2 absorber could be a subDLA). By design, this
misclassification effect is naturally taken into account by the random
variation in logN(H I) during each resampling of the identified
absorbers down to a column density threshold of logN(H I) = 18.8
(Section 2.2). However, as the supplementary literature DLA sample
from Sánchez-Ramírez et al. (2016) does not contain subDLAs

and covers a different redshift path; only absorbers along the associated redshift path of the XQ-100 sightlines are used in the subDLA computation, even if the supplementary literature DLA would be classified as a subDLA. We emphasize that even though all absorbers identified with $\log N(\text{H\,\textsc{i}}) \geq 18.8$ are used in the B-MC simulation, absorbers with a randomly varied $\log N(\text{H\,\textsc{i}})$ are still classified as subDLAs and DLAs based on the classical $\log N(\text{H\,\textsc{i}})$ cuts of $19.0 \leq \log N(\text{H\,\textsc{i}}) < 20.3$ and $\log N(\text{H\,\textsc{i}}) \geq 20.3$, respectively, within each mock generated survey. As already stated, the B-MC simulations were repeated using all identified absorbers with $\log N(\text{H\,\textsc{i}}) \geq 18.8$ and those that were metal selected to provide bounds on the quantities computed in this section. The subDLAs generated from these two sets of B-MC simulations will be further referred to as the FS and MS, respectively.

3.2 The frequency of subDLAs

3.2.1 f(N, X)

Using the B-MC resampling of the XQ-100 and supplementary DLA samples, $f(N, X)$ was computed using

$$f(N, X)\mathrm{d}N\mathrm{d}X = \frac{m_{\text{abs}}}{\Delta N \times \sum_i^{n_{\text{QSO}}} \Delta X_i} \mathrm{d}N\mathrm{d}X, \qquad (2)$$

where m_{abs} is the total number of absorbers with column densities between $N - \Delta N/2$ and $N + \Delta N/2$ along the observed absorption distance ΔX_i of the ith QSO sightline, and n_{QSO} is the number of QSO sightlines observed. The total redshift path covered by the QSO sightline (ΔX_i) is computed as

$$\Delta X(z) = \int_{z_{\text{min}}}^{z_{\text{max}}} (1 + z)^2 [\Omega_{\text{M}}(1 + z)^3 + \Omega_\Lambda]^{-0.5} \mathrm{d}z, \qquad (3)$$

where z_{min} and z_{max} are the minimum and maximum redshift observed along the QSO path, while Ω_{M} and Ω_Λ are the matter and dark energy densities of the Universe observed at the current epoch. z_{min} is set by the Lyman limit of the QSO, while z_{max} is set as $z_{\text{em}} - 5000\,\text{km s}^{-1}$ to remove proximate systems that can be associated with the QSO (Ellison et al. 2002, 2010; Berg et al. 2016; Perrotta et al. 2016).

Rather than using the standard bins of column density, which can hide or amplify features depending on the choice of bins, we elected to use a sliding $\log N(\text{H\,\textsc{i}})$ bin to quantify the median $f(N, X)$ in each column density bin. The sliding bin keeps a fixed bin width, but the centre of the bin shifts by a small amount in each step to capture the effects of individual data being included and excluded from the bin. Using a $\log N(\text{H\,\textsc{i}})$ bin width of 0.4 dex, we computed the median, 1σ and 2σ percentiles of the underlying $f(N, X)$ distribution of the 10 000 B-MC samples. For the subDLAs, the centre of the bin starts at $\log N(\text{H\,\textsc{i}}) = 19.2$ (such that bin edge includes the minimum subDLA threshold of $\log N(\text{H\,\textsc{i}}) = 19.0$), and is then shifted by 0.05 dex until the centre of the bin reaches $\log N(\text{H\,\textsc{i}}) = 20.1$ (such that bin edge includes $\log N(\text{H\,\textsc{i}}) = 20.3$). The same is done for the DLAs, but using $\log N(\text{H\,\textsc{i}})$ bin centres from 20.5 to 21.8. Note that the sliding bins of both the DLA and subDLA samples do not pass the $\log N(\text{H\,\textsc{i}}) = 20.3$ threshold due to the different redshift path coverage of the two samples. The B-MC computation of $f(N, X)$ was repeated using both the FS and MS absorbers from XQ-100.

Fig. 7 shows the resulting median column density distribution functions of the FS and MS subDLA samples (solid and dashed lines, respectively). The 1σ and 2σ errors are shown as the dark and light shaded regions, respectively, for the FS absorbers. Each sample was split into two redshift bins by the median redshift of the XQ-

100 absorbers ($z_{\text{abs}}=3.2$); $f(N, X)$ for the lower and higher redshift bins are denoted by the purple and orange colours, respectively. The numerical values of the FS and MS $f(N, X)$ curves are tabulated in the Appendix (Tables B1 and B2, respectively). For reference, the full $f(N, X)$ range for subDLAs and DLAs is shown in the inset panel of Fig. 7 with the same notation.

There is a clear difference in the shape of $f(N, X)$ between all candidate subDLAs (FS) and those absorbers which have metal lines (MS), resulting in up to a ≈ 0.4–0.6 dex difference in $f(N, X)$ at low column densities. It is interesting to note that the difference between the FS and MS curves changes with $\log N(\text{H\,\textsc{i}})$, and is strongest at the lowest column densities. This discrepancy could result from either incompleteness in the MS (i.e. lack of ability to detect metal lines at low column densities) or a high false-positive rate in the FS [i.e. from blending of low $\log N(\text{H\,\textsc{i}})$]. We emphasize again that using both the FS and MS likely bounds the true nature of $f(N, X)$ at these column densities. The derived $f(N, X)$ curves were compared with the values computed by Zafar et al. (2013, squares in Fig. 7) for subDLAs over approximately the same redshift range. The Zafar et al. (2013) sample consisted of a search for subDLAs and DLAs in archival high resolution spectra taken with UVES in addition to a literature compilation. The identification of absorbers presented in Zafar et al. (2013) started with an automated search for candidate subDLAs and DLAs characterized by saturated absorption over a large range of wavelength. The absorber was confirmed as a (sub)DLA if higher order Lyman series or metal lines were present; which is analogous to the methods used in determining the XQ-100 MS absorbers. It is therefore both encouraging and not surprising that $f(N, X)$ for the MS absorbers is roughly consistent with Zafar et al. (2013) over the similar high-redshift bin ($3.1 \leq z < 5.0$; orange squares in Fig. 7). The discrepancy seen between the low redshift data is likely a result of different redshift coverage, as nearly two thirds of the absorbers in the low redshift bin of Zafar et al. (2013, $1.5 \leq z < 3.1$; purple squares in Fig. 7) are at a redshift $z_{\text{abs}} < 2.5$. We note that requiring coverage Ly β redward of λ_{LL} to remove potential false-positives (see Section 2.3) does not significantly affect the resulting $f(N, X)$ curve generated for the $3.2 \leq z < 4.5$ redshift bin. Given the redshift range of the QSOs of the XQ-100 sample, there is insufficient spectral coverage of Ly β to perform the same experiment within the lower redshift bin ($2.3 \leq z < 3.2$).

To better quantify the significance of the difference in $f(N, X)$ between the MS and FS subDLAs, as well as search for any global evolution in $f(N, X)$ with redshift, we computed the best-fitting power law to $f(N, X)$ (i.e. $f(N, X)=K\text{N}^\alpha$; where K and α are the parameters of the fit). For each mock sample generated in the B-MC simulation, an unweighted least-squares fit to $f(N, X)$ was computed using the sliding bin data. The median and 1σ percentile errors of α and K of all 10 000 samples least-squares fits are presented in Table 3. We point out that a single power law does not appear to be a good fit for $f(N, X)$ for the $3.2 \leq z_{\text{abs}} < 4.5$ FS subDLAs. Comparing the power-law fit parameters α and K between the MS and FS data sets demonstrates that the power-law fits between the two subDLA ($19.0 \leq \log N(\text{H\,\textsc{i}}) < 20.3$) samples are inconsistent at the $\approx 2\sigma$ level, demonstrating that using metal selection could have a profound effect on the interpreted $f(N, X)$.

Comparing the power-law indices from each redshift bin, an insignificant evolution of α with redshift is detected at $\approx 1.3\sigma$ and $\approx 0.6\sigma$ significance for FS and MS subDLAs, respectively, and slight evolution for DLAs at $\approx 1.5\sigma$. The increased number of $\log N(\text{H\,\textsc{i}}) \gtrsim 19.7$ subDLAs at higher redshifts is consistent with previous observational studies (Péroux et al. 2003b; Zafar et al. 2013). However, we note that the simulations predict a consistently
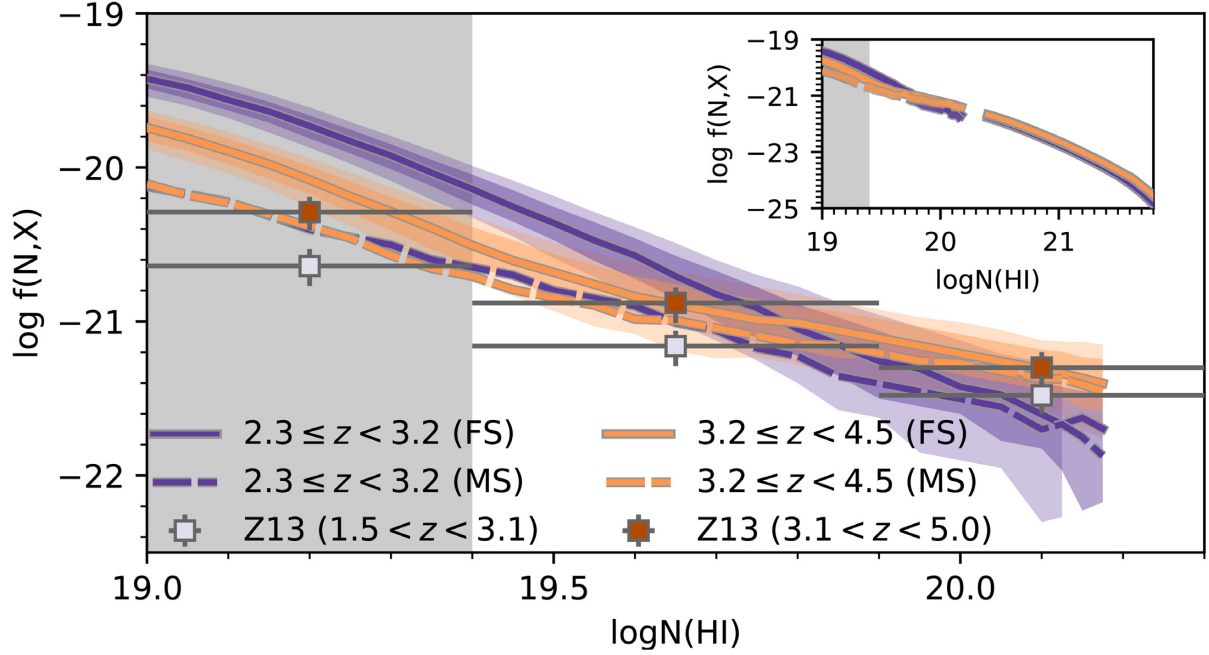
Figure 7. The column density distribution of subDLAs in the XQ-100 sample. The median $f(N, X)$ for the full sample of subDLAs (FS) is given by the solid line, while the dark and light shaded regions represent the 1σ and 2σ percentiles, respectively. The dashed line shows the median $f(N, X)$ for the metal-selected sample of subDLAs (MS). The confidence intervals for the MS selected sample are similar to those of the FS, but are not shown for clarity. Note that the dashed lines from both of the MS redshift bins overlap below $\log N(\mathrm{H\,I}) \approx 19.8$. Both subDLA samples were split by the median redshift of the XQ-100 absorbers ($z_{\mathrm{abs}} = 3.2$), such that those in the lower redshift bin ($2.3 \leq z_{\mathrm{abs}} < 3.2$) are in dark purple while those at higher redshifts ($3.2 \leq z_{\mathrm{abs}} < 4.5$) are in light orange. The greyed region below $\log N(\mathrm{H\,I}) = 19.4$ represents the regime where incompleteness and false-positive effects could significantly influence the determined $f(N, X)$ (Section 2.3). The square points show the data for Zafar et al. (2013, Z13) subDLAs in two similar redshifts ($1.5 \leq z_{\mathrm{abs}} < 3.1$ and $3.1 \leq z_{\mathrm{abs}} < 5$; light purple and dark orange squares, respectively). The inset panel shows the median $f(N, X)$ curves of both the subDLA and DLA samples for reference.

Table 3. Power-law fit parameters to $f(N, X) = K N^{\alpha}$.

$\log N(\mathrm{H\,I})$	z_{abs}	α		$\log(K)$	
		FS absorbers	MS absorbers	FS absorbers	MS absorbers
$\in [19.0, 20.3)$	$\in [2.3, 3.2)$	$-2.00^{+0.176}_{-0.176}$	$-1.46^{+0.276}_{-0.196}$	$18.64^{+3.381}_{-3.369}$	$7.74^{+3.847}_{-5.293}$
	$\in [3.2, 4.5)$	$-1.37^{+0.083}_{-0.119}$	$-1.12^{+0.186}_{-0.198}$	$6.18^{+2.294}_{-1.602}$	$1.15^{+3.752}_{-3.688}$
$\in [19.3, 20.3)$	$\in [2.3, 3.2)$	$-1.94^{+0.230}_{-0.399}$	$-1.41^{+0.337}_{-0.305}$	$17.46^{+7.831}_{-4.457}$	$6.72^{+6.050}_{-6.660}$
	$\in [3.2, 4.5)$	$-1.14^{+0.260}_{-0.160}$	$-0.90^{+0.373}_{-0.293}$	$1.68^{+3.077}_{-5.255}$	$-3.30^{+5.769}_{-7.461}$
$\in [20.3, 22.0)$	$\in [2.3, 3.2)$	$-2.10^{+0.114}_{-0.098}$	$-2.08^{+0.062}_{-0.084}$	$21.23^{+2.058}_{-2.392}$	$20.88^{+1.713}_{-1.319}$
	$\in [3.2, 4.5)$	$-1.96^{+0.131}_{-0.082}$	$-1.95^{+0.111}_{-0.116}$	$18.52^{+1.690}_{-2.729}$	$18.20^{+2.354}_{-2.308}$

higher $f(N, X)$ across all column densities at higher redshifts (Rahmati et al. 2013a; Villaescusa-Navarro et al. 2018), whereas both XQ-100 data sets suggest this is not the case at $\log N(\mathrm{H\,I}) \lesssim 19.7$. In the FS subDLAs, the increased number of low redshift, low column density subDLAs could partly be explained by a high false-positive rate due to blending of low redshift Ly α with the Ly β forest. However the MS also shows no redshift evolution in $f(N, X)$ for $\log N(\mathrm{H\,I}) \lesssim 19.7$, which is approximately consistent with the lack of $f(N, X)$ redshift evolution seen in simulations when local ionizing sources are included (Rahmati et al. 2013b).

Previous analyses comparing the power-law indices of DLAs and subDLAs have indicated a flattening in $f(N, X)$ in the subDLA regime (Péroux et al. 2005; O'Meara et al. 2007; Guimarães et al. 2009; Zafar et al. 2013), which has been attributed to the loss of Ly α self-shielding in low N(H I) subDLA systems (Zheng & Miralda-Escudé 2002). O'Meara et al. (2007) noted that this flattening in $f(N, X)$ at low column densities is dominated by $\log N(\mathrm{H\,I}) \leq 19.3$

systems; an effect which was tentatively seen in their power-law fits to $f(N, X)$ when only $\log N(\mathrm{H\,I}) \geq 19.3$ absorbers were included. We attempted to search for this flattening as well by restricting the subDLA sample only to absorbers with $19.3 \leq \log N(\mathrm{H\,I}) < 20.3$ and recomputing the power-law fit to limited column density range (see Table 3 for fit parameters). The power-law indices between the two $\log N(\mathrm{H\,I})$ regimes are inconsistent at 1.3σ and 0.6σ significance for the FS and MS, respectively, and thus we do not find significant evidence of this flattening.

3.2.2 The line density of subDLAs

In addition to $f(N, X)$, the sightline density of absorbers ($\ell(X)$) provides additional constraints on the nature of the absorbers. $\ell(X)$ is defined as

$$\ell(X) = \int_{N_{\mathrm{min}}}^{N_{\mathrm{max}}} f(N, X) \mathrm{d}N = \frac{m_{\mathrm{abs}}}{\sum_i^{n_{\mathrm{QSO}}} \Delta X_i}. \quad (4)$$
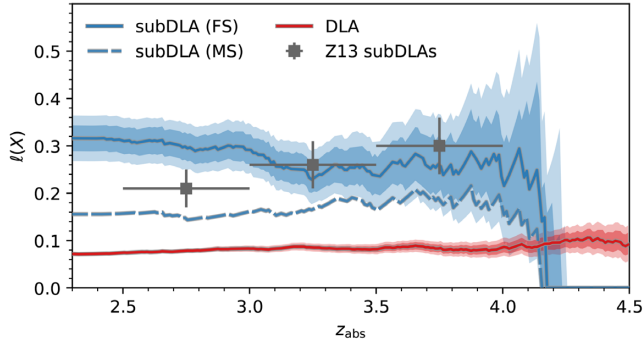
Figure 8. The redshift evolution of $\ell(X)$ as probed by DLAs (red) and subDLAs (blue) in the XQ-100 survey. The solid line shows the median $\ell(X)$ evolution for the full sample of subDLAs (FS) while the dashed line represents when only the metal-selected subDLAs (MS) is used. The shaded regions show the 1σ and 2σ confidence intervals for the FS; the confidence intervals for the MS are approximately the same size, but are not shown for clarity. The values of $\ell(X)$ computed for the Zafar et al. (2013, Z13) subDLAs are shown as the grey squares.

Using the same sliding bin technique described above, we computed $\ell(X)$ in bins of $\Delta X = 2.5$ with redshift steps of 0.01. Fig. 8 shows the resulting redshift evolution curves of $\ell(X)$ for the subDLAs (both FS and MS, blue solid and dashed lines) and DLAs (red line; from Sánchez-Ramírez et al. 2016). The curves presented in Fig. 8 are available in Table B3 (available online). Previous studies have disagreed on whether there is indeed any redshift evolution in $\ell(X)$ for subDLAs (O'Meara et al. 2007; Zafar et al. 2013). However, any redshift evolution present in $\ell(X)$ can physically manifests as a change in either (or both) the H I cross-sections of the absorbers or their number density (e.g. Prochaska et al. 2005).

After applying a unweighted least-squares linear fit to the redshift evolution of $\ell(X)$ for each of the 10 000 iterations in the MC-B simulation, we find the median slope of the redshift evolution is almost negligible; $-0.02^{+0.011}_{-0.013}$ for the FS subDLAs, $0.01^{+0.006}_{-0.009}$ for the MS subDLAs, and $0.01^{+0.009}_{-0.009}$ for both DLA samples. We point out that the evolution of $\ell(X)$ in the MS is shallower than the evolution detected by Zafar et al. (2013, grey points in Fig. 8). The difference in the evolution of the $\ell(X)$ curves of the MS and FS (both in terms of the slope and relative offset) results from the increased number of low column density FS absorbers seen in Fig. 7.

The lack of evolution in $\ell(X)$ seen for DLAs has been suggested to be due to a balance between the decreasing density of the Universe and a decrease in the strength of ionizing UV background in order to maintain an approximately constant $\ell(X)$ (Prochaska et al. 2005; O'Meara et al. 2007). Despite subDLAs being about 2–4 times more common than DLAs along a QSO sightline, the redshift evolution of both the subDLAs frequency and cross-sections appear to follow the evolution of the respective DLA quantities over these redshifts.

3.3 The H I mass contribution of subDLAs

$\Omega_{\rm H I}$ was estimated using

$$\Omega_{\rm H I}(X) = \frac{H_0 m_{\rm H}}{c\rho_0} \int_{N_{\rm min}}^{N_{\rm max}} N f(N, X) {\rm d}N, \qquad (5)$$

where H_0 is the Hubble constant at the current epoch, c is the speed of light, $m_{\rm H}$ is the atomic mass of hydrogen, and ρ_0 is the critical density of the Universe.

Following Sánchez-Ramírez et al. (2016), $\Omega_{\rm H I}$ is computed in sliding bins of redshift. Rather than holding the bin width fixed

in redshift space, Sánchez-Ramírez et al. (2016) elected to use a fixed redshift path-length $\Delta X = 2.5$, with redshift bin steps of 0.01. For consistency, we elected to use the same sliding bin set-up to reproduce the observed $\Omega_{\rm H I}$ for the XQ-100 DLAs. For the derivation of $\Omega_{\rm H I}$ (equation 6), we start with the same $z_{\rm min}$ and $z_{\rm max}$ used in Sánchez-Ramírez et al. (2016) for all the XQ-100 and supplementary QSO sightlines. However, due to the incompleteness of subDLA absorbers below $\lambda_{\rm LL}$ (Section 2.2), we increased $z_{\rm min}$ when necessary to remove all XQ-100 absorbers whose Ly α fell below $\lambda_{\rm LL}$.

Similar to Fig. 7, the top panel of Fig. 9 shows the redshift evolution of $\Omega_{\rm H I}$ for both the DLAs (red) and subDLAs (blue), using both the FS (solid line) and MS absorbers (dashed line; subDLAs only). The $\Omega_{\rm H I}$ curves presented in Fig. 9 are available in the Appendix (Table B4). Although there is a slight but significant difference in the magnitude of $\Omega_{\rm H I}$ for the subDLAs between the FS and MS (detected at the $\approx 1\sigma$ level), no significant evolution in $\Omega_{\rm H I}$ with redshift is detected for subDLAs in the range 2.5 $\lesssim z_{\rm abs} \lesssim 4.2$. We remind the reader that beyond $z_{\rm abs} \gtrsim 4.2$ the redshift path coverage of the XQ-100 survey significantly drops (fig. 5 of Sánchez-Ramírez et al. 2016), so it is unclear if the drop in $\Omega_{\rm H I}$ beyond $z_{\rm abs} \gtrsim 4.2$ for subDLAs is real.

The relatively constant $\Omega_{\rm H I}$ as a function of redshifts as traced by subDLAs is remarkably similar to the gradual decrease in $\Omega_{\rm H I}$ measured in DLAs over the same redshift range (2.5 $\lesssim z_{\rm abs} \lesssim$ 4.2). The canonical picture that has emerged to explain the gradual evolution in $\Omega_{\rm H I}$ in DLAs is a constant replenishment of gas in star-forming galaxies to counteract the consumption of gas from star formation and ejection from galactic outflows (Noterdaeme et al. 2009, 2012; Zafar et al. 2013). The similar lack of significant redshift evolution in the H I statistics of both DLAs and subDLAs ($f(N, X)$, $\ell(X)$ and $\Omega_{\rm H I}$) is suggestive that the same astrophysical processes are responsible in the host galaxies of these absorbers.

Previous cosmological simulations have demonstrated that only higher mass galaxies contribute significantly to $\Omega_{\rm H I}$ below $z_{\rm abs} \lesssim$ 3 while lower mass galaxies dominate the contribution at $z_{\rm abs} \gtrsim$ 3, which leads to a relatively constant $\Omega_{\rm H I}$ with redshift (Davé et al. 2013; Lagos et al. 2014). We speculate that the lack of significant evolution in $\Omega_{\rm H I}$ for both subDLAs and DLAs for 2.5 $\lesssim z_{\rm abs} \lesssim$ 4.2 in Fig. 9 could potentially suggest that the average subDLA and DLA absorbers are probing the same range of host galaxy halo masses. As such, this potential equality in halo mass range for subDLA and DLA host galaxies would be in tension with previous works which use the observed higher metallicity of subDLAs (relative to DLAs) as a proxy for a higher host galaxy mass (e.g. Kulkarni et al. 2007, 2010; Quiret et al. 2016).

To estimate the relative mass contribution of subDLAs to the H I budget, we computed the relative ratio of $\Omega_{\rm H I}$ from subDLAs to the total $\Omega_{\rm H I}$ (i.e. DLAs and subDLAs). To reduce the effects of the vastly different redshift sampling of the DLA and subDLA samples, we elected to use the median $f(N, X)$ curves (Fig. 7) for the two redshift bins using the discrete limit of equation (5), namely

$$\Omega_{\rm H I}(X) = \frac{H_0 m_{\rm H}}{c\rho_0} \frac{\sum_j^{m_{\rm abs}} N_i}{\sum_j^{n_{\rm QSO}} \Delta X_i}. \qquad (6)$$

The bottom two panels of Fig. 9 shows the cumulative fraction contribution to $\Omega_{\rm H I}$ from absorbers of a given H I column density or lower ($f_{\Omega_{\rm HI}}[< {\rm N(H I)}]$). Since we are only focusing on a range of N(H I) column densities, we restrict $f_{\Omega_{\rm HI}}[< {\rm N(H I)}]$ to have the
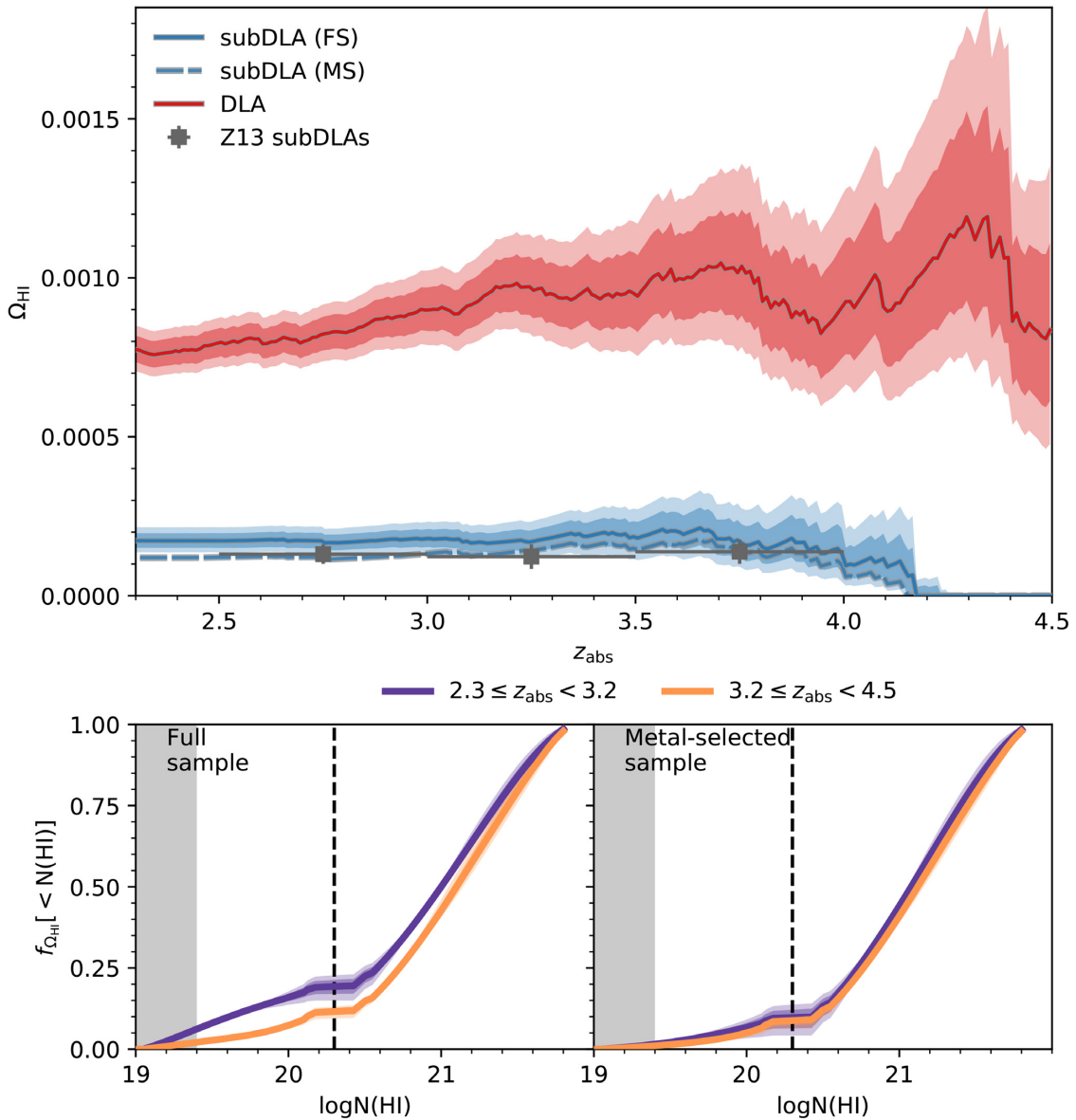
Figure 9. *Top panel:* The redshift evolution of $\Omega_{\rm H\,I}$ as probed by DLAs (red) and subDLAs (blue) in the XQ-100 survey. The notation for the DLA and subDLA samples are the same as Fig. 8. The values of $\Omega_{\rm H\,I}$ computed for subDLAs in Zafar et al. (2013, Z13) are shown as the grey squares. *Bottom panels:* The cumulative contribution of absorbers to the total $\Omega_{\rm H\,I}$ for a given maximum column density ($f_{\Omega_{\rm H\,I}}[< {\rm N(H\,I)}]$). $f_{\Omega_{\rm H\,I}}[< {\rm N(H\,I)}]$ is defined to be zero when only including absorbers with $\log {\rm N(H\,I)} < 19.0$, and has the value one when including all absorbers with $\log {\rm N(H\,I)} \leq 22.0$. The left-hand and right-hand panels in the bottom row show $f_{\Omega_{\rm H\,I}}[< {\rm N(H\,I)}]$ for the respective full (FS) and metal-selected (MS) subDLAs, while the colours show the two redshift ranges $2.3 \leq z_{\rm abs} < 3.2$ (purple) and $3.2 \leq z_{\rm abs} < 4.5$ (orange). The vertical dashed line indicates the $\log {\rm N(H\,I)} = 20.3$ threshold that separates DLAs and subDLAs, while the greyed region below $\log {\rm N(H\,I)} = 19.4$ represents the regime where incompleteness and false-positives could significantly affect the calculation (Section 2.3).

value 0 for all absorbers that have $\log {\rm N(H\,I)} < 19.0$, and is 1 for all absorbers with $\log {\rm N(H\,I)} < 22.0$.

The bottom left-hand and right-hand panels of Fig. 9 show $f_{\Omega_{\rm H\,I}}[< {\rm N(H\,I)}]$ for both the FS and MS, respectively, with the different colours representing the two redshift bins: $2.3 \leq z_{\rm abs} < 3.2$ (purple) and $3.2 \leq z_{\rm abs} < 4.5$ (orange). We note that the discontinuity in $f_{\Omega_{\rm H\,I}}[< {\rm N(H\,I)}]$ at ${\rm N(H\,I)} \approx 20.3$ is a result from the switch between the subDLA and DLA $f(N, X)$ curves. We remind the reader that this discontinuity is a result of the different redshift paths covered by the DLA and subDLA samples, resulting in a truncated $f(N, X)$ at $\log {\rm N(H\,I)} = 20.3$. Overall, the relative contribution of mass from subDLAs appears to be between ≈ 12–22 per cent (for the MS and

FS, respectively), which is consistent with estimates in the literature (Péroux et al. 2005; Guimarães et al. 2009; Noterdaeme et al. 2009; Zafar et al. 2013). Despite the potentially high false-positive rate discussed in Section 2.3 for subDLAs with $\log {\rm N(H\,I)} \leq 19.4$ at $z_{\rm abs} \lesssim 3$, there is at most a 3 per cent contribution to the H I mass contribution (i.e. the difference between the FS and MS curves in the left-hand panel) from the potential false-positives identified. We also note the lack of a redshift evolution in $f_{\Omega_{\rm H\,I}}[< {\rm N(H\,I)}]$ for the MS, as previously seen by Zafar et al. (2013). Despite the increase in the number of $\log {\rm N(H\,I)} \gtrsim 19.7$ absorbers at higher redshifts (Fig. 7), there is little effect on the $\Omega_{\rm H\,I}$ contribution from these systems.

4 CONCLUSION

We have presented a metal-blind catalogue of 242 candidate absorbers with H I column densities of $18.8 \leq \log N(\text{H I}) < 20.3$, identified solely using their H I absorption, in the XQ-100 legacy survey. By comparing to archival UVES data, we demonstrated that X-Shooter has sufficient spectral resolution (FWHM resolution $R \sim 5400$–8900) to identify potential subDLAs and measure their H I column density accurately to within the typically quoted errors (± 0.2 dex), which are primarily due to continuum placement. Using associated metal absorption lines as a method of removing false-positive systems, we note that the 128 absorbers without confirmed metal lines were preferentially found to have $\log N(\text{H I}) \lesssim 19.4$ and are at lower redshifts ($z_{abs} \lesssim 3.0$). At these low column densities and redshifts, the corresponding prominent metal lines (C IV and Si IV) are shifted into the Ly α forest, Ly β shifts beyond the Lyman limit of the QSO, and blending of the Ly α and Ly β forests can mimic low column density subDLAs. Therefore there is a competing effect between incompleteness from using a metal-selected subsample and the true false-positive identification rate of $\lesssim 49$ per cent of the sample. The false-positive rate in XQ-100 is likely significant at low redshifts ($z_{abs} \lesssim 3.5$) and low H I column densities ($\log N(\text{H I}) \lesssim 19.4$).

Using a Monte Carlo Bootstrap simulation applied to the candidate absorber sample, we computed the column density distribution function ($f(N, X)$; Fig. 7), line density ($\ell(X)$, Fig. 8), and cosmic H I mass density ($\Omega_{\text{H I}}$; Fig. 9) for subDLAs selected using the standard column density range $19.0 \leq \log N(\text{H I}) < 20.3$. We compared the $f(N, X)$ computed for two different redshift bins ($2.3 \leq z_{abs} < 3.2$ and $3.2 \leq z_{abs} < 4.5$; split by the median absorber redshift of XQ-100) for both the metal-selected and full catalogue of subDLAs (dashed and solid lines in Fig. 7, respectively). Irrespective of which subDLA catalogue is used, we find that solely high column density subDLAs ($\log N(\text{H I}) \gtrsim 19.7$) are more frequently found at higher redshifts ($z_{abs} \geq 3.2$), whereas all column densities are expected to increase in frequency in simulations (Rahmati et al. 2013a; Villaescusa-Navarro et al. 2018). However, the power-law fits to the entire subDLA $f(N, X)$ curves suggest there is no significant redshift evolution in $f(N, X)$ ($\lesssim 1.3\sigma$ significance).

Using the FS and MS of subDLAs to bound the true value of $\Omega_{\text{H I}}$ and $\ell(X)$, we demonstrated that there is little evolution in $\ell(X)$ and $\Omega_{\text{H I}}$ as a function of redshift for DLAs; a result that is also seen for DLAs (Prochaska et al. 2005; Noterdaeme et al. 2012; Sánchez-Ramírez et al. 2016). In comparison with simulations (Davé et al. 2013; Lagos et al. 2014) where the contribution from varying galaxy mass influences the nature of the redshift evolution of $\Omega_{\text{H I}}$, the lack of significant redshift evolution in $\ell(X)$ and $\Omega_{\text{H I}}$ is suggestive that the subDLAs and DLAs could potentially trace host galaxies with roughly the same range of halo masses and that the evolution of their absorption cross-sections trace each other. Despite subDLAs being about 2–4 times more common than DLAs (Fig. 8), subDLAs contribute a modest amount (between ≈ 10–20 per cent) to the total $\Omega_{\text{H I}}$ budget over redshifts $2.3 \leq z_{abs} < 4.5$ (Fig. 9).

ACKNOWLEDGEMENTS

REFERENCES

Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Berg T. A. M. et al., 2016, MNRAS, 463, 3021
Berg T. A. M. et al., 2017, MNRAS, 464, L56
Bird S., Vogelsberger M., Haehnelt M., Sijacki D., Genel S., Torrey P., Springel V., Hernquist L., 2014, MNRAS, 445, 2313
Bird S., Haehnelt M., Neeleman M., Genel S., Vogelsberger M., Hernquist L., 2015, MNRAS, 447, 1834
Braun R., 2012, ApJ, 749, 87
Christensen L. et al., 2017, A&A, 608, A84
Cooke R. J., Pettini M., Jorgenson R. A., 2015, ApJ, 800, 12
Crighton N. H. M., Hennawi J. F., Simcoe R. A., Cooksey K. L., Murphy M. T., Fumagalli M., Prochaska J. X., Shanks T., 2015, MNRAS, 446, 18
Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, MNRAS, 434, 2645
Dessauges-Zavadsky M., Péroux C., Kim T.-S., D'Odorico S., McMahon R. G., 2003, MNRAS, 345, 447
Dessauges-Zavadsky M., Ellison S. L., Murphy M. T., 2009, MNRAS, 396, L61
Ellison S. L., Ryan S. G., Prochaska J. X., 2001, MNRAS, 326, 628
Ellison S. L., Yan L., Hook I. M., Pettini M., Wall J. V., Shaver P., 2002, A&A, 383, 91
Ellison S. L., Prochaska J. X., Hennawi J., Lopez S., Usher C., Wolfe A. M., Russell D. M., Benn C. R., 2010, MNRAS, 406, 1435
Fox A. J., Ledoux C., Petitjean P., Srianand R., 2007, A&A, 473, 791
Fumagalli M., O'Meara J. M., Prochaska J. X., 2016, MNRAS, 455, 4100
Guimarães R., Petitjean P., de Carvalho R. R., Djorgovski S. G., Noterdaeme P., Castro S., Poppe P. C. D. R., Aghaee A., 2009, A&A, 508, 133
Kulkarni V. P., Khare P., Péroux C., York D. G., Lauroesch J. T., Meiring J. D., 2007, ApJ, 661, 88
Kulkarni V. P., Khare P., Som D., Meiring J., York D. G., Péroux C., Lauroesch J. T., 2010, New A, 15, 735
Lagos C. d. P., Davis T. A., Lacey C. G., Zwaan M. A., Baugh C. M., Gonzalez- Perez V., Padilla N. D., 2014, MNRAS, 443, 1002
Lanzetta K. M., Wolfe A. M., Turnshek D. A., 1995, ApJ, 440, 435
Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
López S. et al., 2016, A&A, 594, A91
Martin A. M., Papastergis E., Giovanelli R., Haynes M. P., Springob C. M., Stierwalt S., 2010, ApJ, 723, 1359
Neeleman M., Prochaska J. X., Ribaudo J., Lehner N., Howk J. C., Rafelski M., Kanekar N., 2016, ApJ, 818, 113
Noterdaeme P., Petitjean P., Ledoux C., Srianand R., 2009, A&A, 505, 1087
Noterdaeme P. et al., 2012, A&A, 547, L1
O'Meara J. M., Prochaska J. X., Burles S., Prochter G., Bernstein R. A., Burgess K. M., 2007, ApJ, 656, 666
Péroux C., Dessauges-Zavadsky M., D'Odorico S., Kim T.-S., McMahon R. G., 2003a, MNRAS, 345, 480
Péroux C., McMahon R. G., Storrie-Lombardi L. J., Irwin M. J., 2003b, MNRAS, 346, 1103
Péroux C., Dessauges-Zavadsky M., D'Odorico S., Sun Kim T., McMahon R. G., 2005, MNRAS, 363, 479
Perrotta S. et al., 2016, MNRAS, 462, 3285
Pettini M., Ellison S. L., Steidel C. C., Bowen D. V., 1999, ApJ, 510, 576
Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543
Prochaska J. X., Herbert-Fort S., Wolfe A. M., 2005, ApJ, 635, 123
Quiret S. et al., 2016, MNRAS, 458, 4074
Rafelski M., Neeleman M., Fumagalli M., Wolfe A. M., Prochaska J. X., 2014, ApJ, 782, L29
Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013a, MNRAS, 430, 2427
Rahmati A., Schaye J., Pawlik A. H., Raičević M., 2013b, MNRAS, 431, 2261
Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, MNRAS, 452, 2034
Rao S. M., Turnshek D. A., 2000, ApJS, 130, 1

Rao S. M., Turnshek D. A., Nestor D. B., 2006, ApJ, 636, 610

Sánchez-Ramírez R. et al., 2016, MNRAS, 456, 4488

Somerville R. S., Davé R., 2015, ARA&A, 53, 51

Storrie-Lombardi L. J., Wolfe A. M., 2000, ApJ, 543, 552

Vernet J. et al., 2011, A&A, 536, A105

Villaescusa-Navarro F. et al., 2018, ApJ, 866, 135

Wolfe A. M., Turnshek D. A., Smith H. E., Cohen R. D., 1986, ApJS, 61, 249

Wolfe A. M., Gawiser E., Prochaska J. X., 2005, ARA&A, 43, 861

Zafar T., Péroux C., Popping A., Milliard B., Deharveng J.-M., Frank S., 2013, A&A, 556, A141

Zheng Z., Miralda-Escudé J., 2002, ApJ, 568, L71

Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L., 2005, MNRAS, 359, L30

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table 1. H I absorbers in XQ-100.
msv5_appendix.pdf

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.