Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **514,** 3559–3578 (2022)
Advance Access publication 2022 June 10

https://doi.org/10.1093/mnras/stac1550

Discovery of three new near-pristine absorption clouds at $z = 2.6$–4.4

P. Frédéric Robert,[1] Michael T. Murphy ,[1]★ John M. O'Meara,[2,3] Neil H. M. Crighton[1] and Michele Fumagalli [4,5,6]

[1]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia*
[2]*Department of Chemistry & Physics, Saint Michael's College, One Winooski Park, Colchester, VT 05439, USA*
[3]*W. M. Keck Observatory, 65-1120 Mamalahoa Highway, Kamuela, HI 96743, USA*
[4]*Institute for Computational Cosmology and Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK*
[5]*Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*
[6]*INAF – Osservatorio Astronomico di Trieste, via G. Tiepolo 11, I-34143 Trieste, Italy*

Accepted 2022 May 31. Received 2022 April 24; in original form 2021 June 4

ABSTRACT

We report the discovery of three new 'near-pristine' Lyman limit systems (LLSs), with metallicities ≈1/1000 solar, at redshifts 2.6, 3.8, and 4.0, with a targeted survey at the Keck Observatory. High-resolution echelle spectra of eight candidates yielded precise column densities of hydrogen and weak, but clearly detected, metal lines in seven LLSs; we previously reported the one remaining, apparently metal-free LLS, to have metallicity $<1/10\,000$ solar. Robust photoionization modelling provides metallicities [Si/H] $= -3.05$ to -2.94, with 0.26 dex uncertainties (95 per cent confidence) for three LLSs, and [Si/H] $\gtrsim -2.5$ for the remaining four. Previous simulations suggest that near-pristine LLSs could be the remnants of PopIII supernovae, so comparing their detailed metal abundances with nucleosynthetic yields from supernovae models is an important goal. Unfortunately, at most two different metals were detected in each new system, despite their neutral hydrogen column densities ($10^{19.2}$–$10^{19.4}$ cm^{-2}) being two orders of magnitude larger than the two previous, serendipitously discovered near-pristine LLSs. Nevertheless, the success of this first targeted survey for near-pristine systems demonstrates the prospect that a much larger, future survey could identify clear observational signatures of PopIII stars. With a well-understood selection function, such a survey would also yield the number density of near-pristine absorbers that, via comparison to future simulations, could reveal the origin(s) of these rare systems.

Key words: line: profiles – galaxies: haloes – intergalactic medium – quasars: absorption lines.

1 INTRODUCTION

In the current standard model of cosmology, the first light elements were created through big bang nucleosynthesis: hydrogen (H), helium (He), deuterium (D), helium-3 (^{3}He), helium-4 (^{4}He), and a small amount of lithium-7 (^{7}Li). The first stars, referred to as PopIII stars, were made of this pristine material and, through stellar nucleosynthesis, created the elements heavier than helium: metals. These metals were then dispersed in the surrounding environment of the PopIII stars after they exploded as supernovae. The gas in the newly metal-polluted environment was used in the formation of the second generation of stars, PopII. This mechanism is commonly accepted, and within cosmological simulations it is possible to follow the initial collapse of pristine material leading up to the creation of PopIII stars, their death, and the following formation of PopII stars (e.g. Bromm & Larson 2004; Bromm et al. 2009). Yet, since we lack observational constraints, the chemical and physical properties of PopIII stars are not currently known. One obvious method to constrain their properties would be direct observations, but it is thought that most of the PopIII stars died within the first ∼0.5 billion years of the Universe. They are therefore too distant, too faint, and too short-lived to be directly observed with current telescopes.

One way to infer the characteristics of the PopIII stars is to study element abundances in very old metal-poor stars found in the Galactic halo. Stars with extremely low iron abundances, [Fe/H] ~ -7 to -3, are believed to have formed from the remnants of PopIII stars (Christlieb et al. 2002; Frebel et al. 2005, 2007; Keller et al. 2014). Their abundance patterns should then reflect the environment in which they formed and could be compared with theoretical supernova yields of PopIII stars. However, that comparison is complicated by the interpretation of the observed stars' elemental abundance patterns. Indeed, metal-poor stars can have very different abundance patterns to each other (Yong et al. 2013). Moreover, these patterns could also be different from the time when the star formed, as its stellar interior may have been polluted by its environment and its nuclear burning. Therefore, linking the metal-poor stars of the Milky Way to the yields of PopIII stars is a difficult process. A complementary approach would be to study the remnants of PopIII stars: gas clouds that have been polluted by metals ejected during the death of PopIII stars as supernovae. Depending on the mass ranges and explosion energies of the PopIII progenitors, these remnants could remain in the intergalactic medium (IGM) or the circumgalactic environment (CGM) of a galaxy (e.g. Whalen et al. 2008; Wise et al. 2012; Liu & Bromm 2020).

★ E-mail: mmurphy@swin.edu.au

Some Lyman limit systems (LLSs) may be excellent candidates for such PopIII remnants. The physical properties of LLSs – usually defined as quasar absorbers with neutral hydrogen column densities $17.2 \leq \log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) < 20.3$ – have been investigated in several surveys by different groups at $z \leq 1$ (e.g. Lehner et al. 2013, 2016; Wotta et al. 2016) and $z \geq 2$ (e.g. Steidel 1990; Fumagalli et al. 2013; Cooper et al. 2015; Prochaska et al. 2015; Fumagalli, O'Meara & Prochaska 2016a; Lehner et al. 2016). These studies have shown that LLSs are relatively metal poor, with the metallicity distribution spanning $-2.0 \leq \log_{10}(Z/{\rm Z}_{\odot}) \leq 0.4$ at low redshift ($z_{\rm abs} \leq 2$) and $-4.0 \leq \log_{10}(Z/{\rm Z}_{\odot}) \leq 1$ at high redshift ($z_{\rm abs} \geq 2$). The high-redshift regime is the most relevant for the LLSs studied in this paper. Along with these observational surveys, simulations (e.g. Kereš et al. 2005; Dekel & Birnboim 2006; Dekel et al. 2009; Faucher-Giguère & Kereš 2011; Fumagalli et al. 2011b; van de Voort et al. 2012) have shown that the prime candidates for the metal-poor streams fuelling the formation of galaxies are LLSs, therefore making them potential candidates for circumgalactic PopIII remnants. Finally, three clouds with no apparent associated metal lines and metallicity upper limits of $\log_{10}(Z/{\rm Z}_{\odot}) \leq -4$ have been discovered: two at $z \sim 3$, serendipitously, by Fumagalli, O'Meara & Prochaska (2011a) (LLS1134 and LLS0958B), and the third, recently, at $z \sim 4.4$ (LLS1723) from a dedicated search by Robert et al. (2019). The very low metallicity of these systems, combined with the lack of detectable metal lines, implies that they could be intergalactic PopIII remnants. Another possibility, emphasized in Fumagalli et al. (2016a), is that they are entirely pristine, having experienced no pollution from nearby galaxies for ∼2 billion years.

Even LLSs with the lowest column densities [i.e. $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \approx 17.3$], and in the lowest part of the metallicity distribution at high redshift [$\log_{10}(Z/{\rm Z}_{\odot}) \leq -3$], still display detectable metal lines. Indeed, Crighton, O'Meara & Murphy (2016) discovered the lowest metallicity LLS with metal-line detections, LLS1249, with $\log_{10}(Z/{\rm Z}_{\odot}) = -3.41 \pm 0.26$. This very low metallicity falls in the range expected for PopIII remnants (Wise et al. 2012). With detections of C and Si, this LLS presents an abundance ratio of [C/Si] $= -0.26 \pm 0.17$, consistent with both predictions of nucleosynthesis models of PopIII stars (e.g. Heger & Woosley 2010; Cooke & Madau 2014) and PopII stars (Cen & Riquelme 2008). Its size, density, and temperature suggest that LLS1249 is more likely to be found in the IGM rather than in the CGM, although these parameters have large uncertainties and so the evidence for this is correspondingly weak. Following this first discovery of a 'near-pristine' LLS, Lehner et al. (2016) also reported a similar system (LLS0958A) and complementary integral field spectroscopy by Fumagalli et al. (2016b) mapped its galactic environment, along with another absorber in the same sightline, LLS0958B, which is one of the two apparently metal-free LLSs discovered by Fumagalli et al. (2011a). Interestingly, the environment of LLS0958A showed no nearby galaxies, potentially indicating an intergalactic environment. However, the environment of LLS0958B showed five Lyα-emitting galaxies at a similar redshift, with three appearing to be aligned in projection, suggesting a filamentary structure akin to a cold stream in the vicinity of galactic haloes. In the context of a simple expectation that the lowest metallicity absorbers arise furthest from galaxies, it is perhaps surprising that the opposite was found in this study, though it remains possible that LLS0958A arises in gas closely associated with galaxies below the detection limits of Fumagalli et al. (2016b).

Hence, given this somewhat counter-intuitive situation from a very small sample, more 'near-pristine' LLSs need to be discovered to refine our understanding of their origins. While they appear, so far, to be rare, previous searches have not targeted them directly; instead, a dedicated search is required and now possible with existing LLS surveys. We reported the first such search in Robert et al. (2019), in which we established a list of very metal-poor candidates and focused on the discovery and properties of the apparently metal-free absorber, LLS1723. Here, we study the rest of the sample, i.e. seven systems. They all show detectable metal lines and could therefore be near-pristine LLSs and candidates for PopIII remnants. Importantly, the sample was deliberately biased towards LLSs with higher hydrogen column densities ($N_{\rm H\,I}$) than the first near-pristine absorber discovered by Crighton et al. (2016) (LLS1249) so that other metallic species, in addition to Si and C, could be detectable even at very low metallicities $\log_{10}(Z/{\rm Z}_{\odot}) \leq -3$. This should improve prospects for distinguishing between PopIII and PopII nucleosynthetic scenarios.

This paper is structured as follows. In Section 2, we summarize briefly how the sample of very metal-poor candidates was selected and describe our Keck observing campaign. In Section 3, we summarize our general approach to analysing all the candidates, and in Section 4, we describe the absorption-line features of each LLS and their physical properties obtained through photoionization modelling. This allows us to accurately reassess their metallicity, and check the efficiency of our dedicated search. Indeed, among the seven very metal-poor candidates, we ultimately find that they are near-pristine, i.e. have metallicities $\log_{10}(Z/{\rm Z}_{\odot}) \lesssim -3$. Their origins are discussed in Section 5, and we also present suggestions for a strategy to improve future dedicated searches for very metal-poor LLSs. Note that, unless otherwise stated, all column density measurements in this paper are quoted with 1σ uncertainties, all column density upper limits are 2σ, and all metallicity and abundance ratio estimates (including limits) from our photoionization analyses are quoted at 95 per cent confidence.

2 TARGET SELECTION, OBSERVATIONS, AND DATA REDUCTION

In Robert et al. (2019), we described a dedicated search for very metal-poor LLSs. Briefly, we targeted eight LLSs, for which no previous high-resolution ($R > 30\,000$) spectra existed. We observed these eight targets with Keck/HIRES during three runs in 2016–2017, obtaining signal-to-noise ratios (SNRs) $\gtrsim 20$ per $\approx$2.3 km s^{-1} pixel in the continuum near the expected wavelength of the strongest metal absorption lines (generally Si II λ 1260). The targets were selected to satisfy the following criteria: (i) their lower resolution quasar spectra, from Cooper et al. (2015) and Prochaska et al. (2015), showed either no metal lines at the LLS redshifts or only very weak (usually not very clear) detections; (ii) their metallicity estimates from these lower resolution spectra were consistent with $\log_{10}(Z/{\rm Z}_{\odot}) \leq -3$; and (iii) their H I column density estimates were higher than for LLS1249 discovered by Crighton et al. (2016), i.e. $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \geq 17.36$. The third criterion was introduced to attempt to identify near-pristine LLSs in which a larger variety of metallic species could be detected. Table 1 lists the metallicity estimates obtained by Cooper et al. (2015) from their Magellan/MagE spectra, and by Fumagalli et al. (2016a) for their Keck/ESI and Magellan/MIKE spectra of the others, along with the literature $N_{\rm H\,I}$ estimates. One of the systems listed, LLS1723, appeared to be free of metal lines in the HIRES spectra and indicated a significantly lower metallicity than the others; this LLS was studied in Robert et al. (2019).

The journal of our new Keck/HIRES observations of the seven LLSs studied here is provided in Table 2. HIRES was configured with the red cross-disperser, with a slit width of 1.148 arcsec (C5 decker, with length 7 arcsec) to provide a resolving power

Table 1. Literature information used to select the eight very metal-poor candidate LLSs targeted in our survey. LLS1153 and LLS1304 are from the sample of Cooper et al. (2015) that provided ionic column densities and metallicity estimates for 17 LLSs at $z \geq 2$ observed with the Magellan/MagE spectrograph. The other six LLSs are from the survey of Prochaska et al. (2015) that provided ionic column densities for 157 LLSs at $z \geq 2$ observed with Keck/ESI or Magellan/MIKE. Their metallicities were estimated in a subsequent study described in Fumagalli et al. (2016a).

Quasar name	LLS name	Spectrograph	QSO redshift	LLS redshift	$\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2})$	$\log_{10}({\rm Z/Z_\odot})$
SDSS J010619.20+004823.3	LLS0106	MIKEb	4.430	4.17157	19.05 ± 0.20	≤ -3.19
SDSS J034402.80−065300.0	LLS0344	MIKEb	3.957	3.84280	19.55 ± 0.15	≤ -3.01
SDSS J095256.41+332939.0	LLS0952	ESI	3.396	3.26180	19.90 ± 0.20	≤ -3.00
SDSS J115321.68+101112.9	LLS1153	MagE	4.127	4.03800	$17.00 \leq N_{\rm H\,I} \leq 19.00$	≤ -2.90
SDSS J115659.59+551308.1	LLS1156	ESI	3.110	2.61594	19.10 ± 0.30	≤ -3.10
SDSS J130452.57+023924.8	LLS1304	MagE	3.651	3.33690	$17.90 \leq N_{\rm H\,I} \leq 18.70$	-2.81 ± 0.15
SDSS J172323.20+224358.0	LLS1723	ESI	4.520	4.39100	18.30 ± 0.30	≤ -3.25
SDSS J224147.70+135203.0	LLS2241	MIKEb	4.470	3.65393	20.20 ± 0.20	≤ -3.45

Table 2. Journal of HIRES observations. 'EA' and 'XDA' are the echelle and cross-disperser angles; λ range is the wavelength coverage. The slit-width was 1.148 arcsec for all exposures, providing a nominal resolving power $R = 37\,500$.

Quasar	EA (°)	XDA (°)	λ range (Å)	Date (UT)	Exposure (s)	Seeing (arcsec)
SDSS J010619.20+004823.3	−0.330	0.830	4740–9360	2016-09-22	3600, 2733	∼0.8[a, b]
	−0.300	0.700	4680–9120	2017-08-19	3210	∼0.65–0.8
SDSS J034402.80−065300.0	0.200	0.400	4250–8700	2017-08-19	2 × 2710	∼0.9
	0.400	0.200	4010–8500	2017-08-20	3000	∼0.65[a]
SDSS J095256.41+332939.0	0.000	1.878	3750–6600	2015-11-14	3000	∼0.8
	0.000	1.752	3600–6480	2015-11-14	2700	∼0.8
SDSS J115321.68+101112.9	−0.100	0.400	4340–8690	2017-02-18	2 × 3600	∼0.7–0.9 (mostly ∼0.8)
	−0.100	0.400	4340–8690	2017-02-18	3600	∼0.7–0.8 (mostly ∼0.75)
SDSS J115659.59+551308.1	0.000	0.010	3640–8100	2017-02-17	3600	∼0.9–1.2 (mostly ∼1.0)
	0.000	0.010	3640–8100	2017-02-17	2 × 3600	∼1.0–1.3 (mostly ∼1.1)
SDSS J130452.57+023924.8	0.190	0.000	4000–8490	2017-02-17	2 × 3300	∼1.0–1.3 (mostly ∼1.1)
	0.190	0.000	4000–8490	2017-02-18	3300, 3300	∼0.75, ∼0.75–1.0 (mostly ∼0.9)
	0.190	0.000	4000–8490	2017-06-13	3600	∼0.7–1.1 (mostly ∼0.7)
	0.190	0.000	4000–8490	2017-06-13	3600	∼0.8–1.0 (mostly ∼0.8)
PSS J172323.10+224357.14	0.080	0.900	4786–9230	2017-06-13	3750, 3315[c, d]	∼0.7–0.8
	−0.150[e]	0.900	4786–9217	2017-06-13	2 × 3300	∼0.75–0.8
	−0.150	0.900	4786–9217	2017-06-13	3400, 3600	∼0.8–1.0, ∼0.9–1.5 (mostly ≤1.2)
SDSS J224147.70+135203.0	−0.250	0.730	4680–9130	2017-08-19	2 × 3210	∼0.6–0.7

[a] Very poor weather most of the night; snow during preceding day.
[b] Dome closed after 2733 s of second exposure.
[c] Exposure paused at 2425 s because target transited and could not be tracked.
[d] Stopped at 3315 s because an incorrect echelle angle was used which did not cover the important Si II λ 1260 line.
[e] Echelle angle was changed to focus on Si II λ 1260.

of $R = 37\,500$ for all quasars. The data reduction is described in detail in Robert et al. (2019). Briefly, we performed the initial reduction steps with the MAKEE pipeline: flat-fielding, order tracing, background subtraction and extraction of a 1D spectrum for each exposure. We wavelength-calibrated the data in the standard way by using spectra of a ThAr lamp but, unlike the case of J1723+2243, MAKEE automatically identified a sufficient number of ThAr lines during the calibration process. We then used UVES_POPLER (Murphy 2016) to merge together the extracted exposures, with the same process as Murphy et al. (2019); again see Robert et al. (2019) for more details of this procedure, and in particular how the continuum was manually estimated in the Lyman forest region of each spectrum.

For the seven LLSs studied here, our reassessment of their absorption features and metallicities with the newly acquired HIRES spectra showed that four have metallicity consistent with $\log_{10}(Z/Z_\odot) \lesssim -3$. As stated above, the analysis of LLS1723 is the focus of Robert et al. (2019) due to its apparent lack of metal lines. The three others – LLS0344, LLS1153, and LLS1156 – do show weak metal absorption lines and have $\log_{10}(Z/Z_\odot) \lesssim -3$: they are new examples of 'near-pristine' LLSs. Therefore, they are the focus of Section 4 where the details of the absorption line properties and photoionization analysis are provided. The absorption feature properties and metallicity estimates of the remaining four LLSs are reported briefly in Appendix A. None the less, these higher metallicity systems are important for considering an improved selection strategy for low-metallicity LLS candidates in future; we discuss this in Section 5.2.

3 GENERAL ANALYSIS APPROACH FOR ALL ABSORBERS

Here, we introduce the general analysis approach used for all the very metal-poor candidates. We start with the study of their absorption features to derive the metal and H I column densities, followed by the photoionization analysis that returns the physical properties of the absorbers – most importantly, the metallicity.

3.1 Spectral analysis

The starting point for all absorbers is to identify the metal lines detected at, or very near, the LLS redshift determined by previous authors in Table 1. They are generally the strongest metal transitions detectable for LLSs at $z \geq 2$ and correspond to the different ionization stages of carbon and silicon: C II, C III, C IV, Si II, Si III, and Si IV. At the resolution of Keck/HIRES, the detection of these transitions is unambiguous, especially if using the strongest available transition (highest oscillator strength) for an ion; e.g. Si II λ 1260. Therefore, the determination of the absorption redshift, a key component of the LLS's model, is robust.

The LLSs are modelled using the VPFIT software (version 9.5) (Carswell & Webb 2014) which fits ideal Voigt profiles to absorption lines and minimizes χ^2 between the data and model to return their best-fit column densities N, Doppler parameters b (line widths), absorption redshifts $z_{\rm abs}$, and the associated uncertainties. All the LLSs analysed in this paper show detectable metal absorption lines and the fitting process for them is rather straightforward. For instance, the LLS redshift, $z_{\rm abs}$, is determined from the metal lines and then fixed for the fit to the H I lines. The main purpose of the Voigt profile modelling is to measure the total column density of each ion, integrated across all fitted velocity components. The same velocity components are fitted to all singly ionized species simultaneously. We then assess whether the same components can be fitted to the higher-ionization species, using changes in χ^2 per degree of freedom as a rough guide. This aids our assessment of whether the low and high ion absorption likely arises in the same gas phase or not, which informs our approach to the photoionization model of each LLS. However, it should be emphasized that fitting the high ions with the same or different velocity structure as the low ions does not appreciably change the total column density returned for any species.

The total H I line column density $N_{\rm H\,I}$ and representative b parameter is then determined for the LLS. But unlike metal absorption lines, it is challenging to fit ideal Voigt profiles to the different lines of the Lyman series: e.g. at $z \geq 2$, the Ly α forest is very thick and several unrelated systems can contribute to the absorption features seen; the Ly α line associated with an LLS can be on the flat part of the curve-of-growth and not have strong damping wings; and the Lyman series lines can be saturated. Due to these possible issues, we use a simple one-component Voigt profile fit for H I and follow one of the options below:

(i) When the presence of any weak damping wings in the profile of its Ly α absorption line cannot be confidently identified due to Lyman forest absorption, then $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \lesssim 19$. The low order lines such as Ly α or Ly β can only provide an upper limit on $N_{\rm H\,I}$. However, establishing a *lower* limit is much more important as this provides the highest possible (i.e. most conservative) metallicity upper limit. The common approach is to use the flux level bluewards of the Lyman limit to establish a lower limit on $N_{\rm H\,I}$ that provides a metallicity upper limit (e.g. Prochaska, O'Meara & Worseck 2010; O'Meara et al. 2013; Prochaska et al. 2015). This has the advantage that the measured $N_{\rm H\,I}$ is independent of b. We used this method for LLS0106 and LLS1304, and our analysis of LLS1723 in Robert et al. (2019) provides a detailed example of this approach.

(ii) When an LLS's Ly α profile has clear damping wings, then $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \geq 19$. $N_{\rm H\,I}$ is easily measured, and strongly constrained, by the observed transmission near the Ly α line and its damping wings. We used this method for LLS0952, LLS1156, and LLS2241. Again, the estimated b value does not have a significant impact on the $N_{\rm H\,I}$ measurement.

Finally, we note that two of our target LLSs have apparent D I Ly α detections at the expected -82 km s^{-1} offset from their associated H I lines. In both cases, the H I Ly α shows only weak damping wings [i.e. $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \leq 19$], so the D I lines provide an opportunity to refine the $N_{\rm H\,I}$ measurement: combined with a precise determination of the [D/H] ratio in low-metallicity LLSs and damped Ly α systems (DLAs), taken from Cooke et al. (2014), the estimate of $N_{\rm D\,I}$ provides a more precise, and likely more accurate, estimate of $N_{\rm H\,I}$ (with $\sim$0.1 dex uncertainty) that is consistent with the value determined form the H I Ly α line itself (with $\sim$0.3 dex uncertainty). However, the $N_{\rm D\,I}$ and b-parameter estimates in our two LLSs are not accurate enough for a competitive constraint on the primordial [D/H] value: other transitions of the D I Lyman series are not detected in our LLSs, and the D I Ly α lines are at least partially blended with other, unrelated H I lines, so they will not provide an accuracy in [D/H] near $\sim$0.02 dex which is typical for the best measurements (e.g. Cooke, Pettini & Steidel 2018).

3.2 Estimating metallicities with photoionization analysis

In this paper, we measure metallicity using silicon as the probe, and so we equate $\log_{10}(Z/Z_\odot)$ with [Si/H], i.e. the silicon abundance in the absorber measured relative to the solar values. None the less, one cannot assume that the observed ions are fully representative of the hydrogen and metal content of the LLSs. To convert the measured ionic column densities into atomic ones, ionization corrections are derived with the photoionization simulation software CLOUDY (version 13.03)[1] (Ferland et al. 2013) that performs radiative transfer calculations for an ideal gas cloud representative of an LLS.

We follow closely the method described in Crighton et al. (2015) so provide only a brief summary here. The CLOUDY models are generated assuming that an LLS is a single-phase slab with a constant density and illuminated on one side by an ionizing spectrum, the HM12 ultraviolet (UV) background of Haardt & Madau (2012) set at the LLS's redshift.[2] We generate a large grid of models as a function of $N_{\rm H\,I}$, the hydrogen volume density $n_{\rm H}$, [Si/H] and a free parameter $\alpha_{\rm UV}$ that tunes the hardness of the ionising spectrum, effectively representing the uncertainty in its shape (see Crighton et al. 2015). The column densities of the different ions are then calculated by interpolation on the grid.

To more efficiently use our computing resources, for each LLS we first generated a coarse CLOUDY grid covering $-4.2 < \log_{10}(n_{\rm H}/{\rm cm}^{-3}) < -1$, $15.5 < \log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) < 20.5$,

[1] See https://www.nublado.org/.

[2] The earlier 'HM05' spectrum (Haardt & Madau 1996) included background radiation from quasars and galaxies, while the HM12 spectrum also includes a contribution from ionizing photons escaping galaxies close to the gas slab under consideration. There is considerable discussion in the literature about which background spectrum is most appropriate for modelling LLSs (as summarized in e.g. Crighton et al. 2015; Fumagalli et al. 2016a; Lehner et al. 2019). Wotta et al. (2019) found that metallicities derived through CLOUDY modelling with the HM12 background were ≈0.2–0.4 dex higher than those derived using the HM05. However, in their work, the difference reduced strongly as a function of $N_{\rm H\,I}$, and for $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) > 19$ there is no evidence for a systematic difference. Similar results were also inferred by Pointon et al. (2019). We use the HM12 background in this work for consistency with Crighton et al. (2016) and to ensure our metallicity estimates remain conservative for the three near-pristine systems we identify with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) > 19$; i.e. we do not spuriously identify higher metallicity systems as having $\log_{10}(Z/Z_\odot) \lesssim -3$.

Table 3. Column densities and metallicity measurements for the three near-pristine LLSs. The neutral hydrogen and metallic ion column density measurements are provided for all detected lines with 1σ uncertainties, or 2σ upper limits for non-detections. The metal column density values correspond to the sum of all velocity components across the absorption profile. The upper limits were derived using the apparent optical depth method (Savage & Sembach 1991). The last row provides the fiducial metallicities, [Si/H], inferred from the photoionization analyses, with 95 per cent confidence intervals.

Ion	$\log_{10}(N/\mathrm{cm}^{-2})$		
	LLS0344	LLS1153	LLS1156
Si II	12.48 ± 0.02	12.22 ± 0.03	12.35 ± 0.15
Si IV	12.88 ± 0.04	12.46 ± 0.04	–
C II	13.37 ± 0.03	–	–
C IV	13.70 ± 0.02	13.37 ± 0.04	14.83 ± 0.09
Fe II	≤ 12.85	–	–
Al II	≤ 11.45	≤ 11.06	11.45 ± 0.09
Al III	–	–	≤ 11.55
D I	14.63 ± 0.10	14.80 ± 0.10	–
H I	19.23 ± 0.10	19.40 ± 0.10	19.30 ± 0.10
[Si/H]	-3.00 ± 0.26	-3.05 ± 0.26	-2.94 ± 0.26

$-6 <$ [Si/H] < -2 (all with steps of 0.5 dex) and set α_{UV} to 0, i.e. the un-tilted HM12 UV background. This allows us to quickly reassess the LLS's metallicity, filtering out those inconsistent with $\log_{10}(Z/Z_{\odot}) \leq -3$. For those with metallicities near, or clearly lower than this threshold, grids with finer steps (requiring more computing time) were then generated. The details of each grid are provided in Section 4 for each very metal-poor LLS. For all those grids, α_{UV} was varied from -2.50 to 1.50 in steps of 0.4 dex.

Still following the approach of Crighton et al. (2015), a Markov chain Monte Carlo (MCMC) technique is used to find the metallicity probability distribution of each LLS. Briefly, the predictions of CLOUDY for the ionic column densities are compared to the measurements, and we construct a likelihood function $\ln \mathcal{L}$ with the CLOUDY grid parameters: $N_{\mathrm{H\,I}}$, n_{H}, [Si/H], and α_{UV}. We maximize $\ln \mathcal{L}$ to determine the posterior distributions of its parameters using an MCMC sampler provided by the EMCEE code (Foreman-Mackey et al. 2013). During the sampling of the coarse and fine grids, we impose a flat prior on the density n_{H} of $-3.5 < \log_{10}(n_{\mathrm{H}}/\mathrm{cm}^{-3}) < -1.0$. Briefly, this range of densities represents the likely extremes that characterize LLSs arising in intergalactic and circumgalactic environments; see Fumagalli et al. (2016a) (fig. 3) and Robert et al. (2019) for a detailed discussion. While solar abundances are assumed for the metals when constructing the CLOUDY grids, non-solar patterns can be allowed a posteriori by allowing deviations over a range set via a flat prior. This is generally done in the sampling of the fine grids when a metal ion is detected, but it is also possible to infer an upper limit on a ratio from the non-detection of a metal species. As the metallicity is determined using the α element Si, a deviation from solar abundance pattern for an element X is expressed as [X/Si] and this ratio is allowed to vary, such as $-2 \leq$ [X/Si] ≤ 2.

From the analysis with the coarse grids, three LLSs appeared to be near-pristine candidates, i.e. with detected metal lines and metallicities consistent with [Si/H] ≤ -3: LLS0344, LL1153, LL1156. The analysis of these LLSs is described below in Section 4, while the properties of the remaining four LLSs are provided in Appendix A. Table 3 summarizes the metal and H I column density measurements and upper limits for each near-pristine LLS, together with the final metallicity estimate from the photoionization analysis.

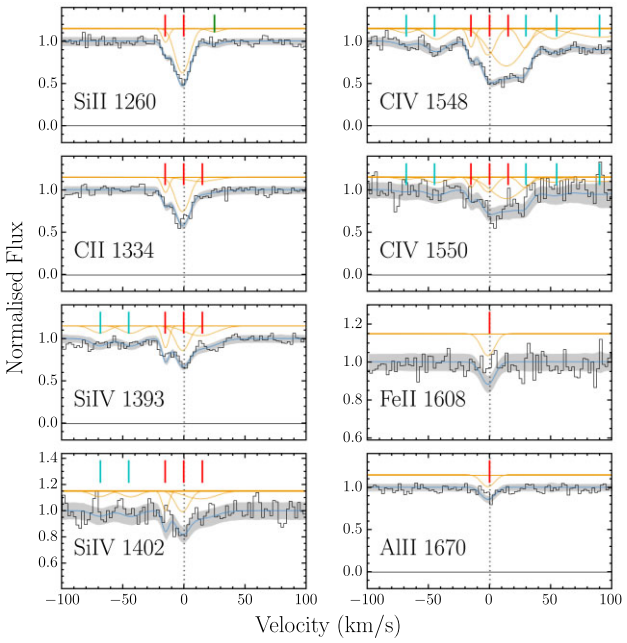

Figure 1. Regions of the J0344−0653 HIRES spectrum (black histogram) where the strongest metal absorption lines of LLS0344 are expected, centred at $z_{\mathrm{abs}} = 3.842\,8323$. The solid blue line in each panel represents the combined Voigt profile fit, with a total column density given in Table 3. The solid orange lines represent the Voigt profiles of the different velocity components contributing to the total absorption profile. These are indicated by vertical red and blue ticks. The red ticks are the main contributions to the total column density in each ion, while those represented by the cyan ticks contribute only weakly; in this absorber, the latter are only detected in the high ions C IV and Si IV. Note that the orange Voigt profile fits have been vertically shifted for clarity. The grey shading shows the 1σ uncertainty in the flux. For Fe II λ 1608 and Al II λ 1670, the column densities used for the Voigt profile fit correspond to a 2σ upper limit derived using the apparent optical depth method. For Si II λ 1260, in the top-left panel, the Voigt profile at $+25\,\mathrm{km\,s^{-1}}$ corresponds to the Fe II λ 1260 transition associated with this absorber.

4 ABSORPTION LINE AND PHOTOIONIZATION ANALYSIS OF THE NEW NEAR-PRISTINE ABSORBERS

4.1 LLS0344

4.1.1 Metal-line column densities

LLS0344 was first identified in Prochaska et al. (2015) towards the $z_{\mathrm{em}} = 3.957$ quasar SDSS J034402.80−065300.0 (hereafter J0344−0653), based on absorption features at redshift $z_{\mathrm{abs}} = 3.843$, using a MIKE spectrum. At this redshift $N_{\mathrm{H\,I}}$ was estimated at $\log_{10}(N_{\mathrm{H\,I}}/\mathrm{cm}^{-2}) = 19.55 \pm 0.15$ along with $b = 30\,\mathrm{km\,s^{-1}}$. These values were determined by visual inspection of theoretical Voigt profiles superimposed on the Ly α line. At $z_{\mathrm{abs}} = 3.843$ they identified C II λ 1334, C IV λλ1548/1550, Si II λ 1260, Si IV λλ1393/1402, O I λ 1302, and obtained column density upper limits for Fe II λ 1608 and Al II λ 1670. The column densities were measured using the apparent optical depth method (Savage & Sembach 1991).

Guided by the initial results above, we search the HIRES spectrum for metal absorption lines and we list the detections and upper limits in Table 3. Fig. 1 depicts the strongest (highest oscillator strength) transitions of the most abundant metal species expected at $z_{\mathrm{abs}} = 3.843$. For a detection, the Voigt profile fit depicted by a solid blue line corresponds to the combination of the profiles of

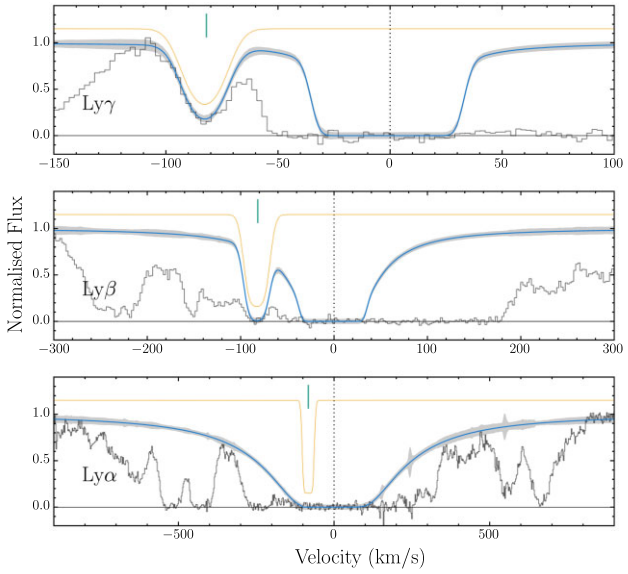

Figure 2. Continuum-normalized flux for Lyman series absorption lines of LLS0344 in the J0344−0653 HIRES spectrum (black histogram). The zero velocity redshift is set at $z_{\rm abs} = 3.843$. The blue solid lines correspond to our fiducial H I model with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.23 \pm 0.10$, $z_{\rm abs} = 3.843$ and $b = 10\,{\rm km\,s}^{-1}$. The solid orange line in each panel represents the Voigt profile fit of a D I line with $\log_{10}(N_{\rm D\,I}/{\rm cm}^{-2}) = 14.63 \pm 0.10$ and $b = 10\,{\rm km\,s}^{-1}$. It contributes to the total H I absorption profile and the vertical green ticks indicate its position, at the $-82\,{\rm km\,s}^{-1}$ isotopic shift from the associated H I line. The grey shading shows the 1σ uncertainty in the flux.

each velocity component (solid orange lines). For an upper limit, the profile illustrates the suitability of the 2σ upper limit derived through the use of the apparent optical depth method.

The absorption profiles in Fig. 1 show that the majority of the total optical depth in each ion is concentrated in three velocity components between -20 and $+20\,{\rm km\,s}^{-1}$ (red tick marks) for both the low ions (C II and Si II) and high ions (C IV and Si IV). All three components can be fitted simultaneously, with the same redshifts and b parameters, in all detected transitions, although the weakest, broadest component is not statistically required in the fit to Si II λ 1260. There are additional velocity components detected over a larger velocity extent in C IV and Si IV (cyan ticks). However, these are minor contributors to the total column densities in those ions (≈11 per cent for Si IV, ≈36 per cent for C IV). Given these observations, we assume for simplicity in our photoionization analysis that all the metal ion column density arises in the same phase as the H I detected in this system. Note that if, in reality, the high ion gas does not contribute to the H I column density, our simple assumption will result in a slight overestimate of the metallicity. That is, our assumption is conservative in that it will not result in a spurious near-pristine identification.

4.1.2 *H I column density*

Our fiducial H I model is depicted in Fig. 2 with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.23 \pm 0.10$ and $b = 10\,{\rm km\,s}^{-1}$. Having established the redshift of LLS0344, we checked the corresponding Lyman series lines to assess $N_{\rm H\,I}$. The model of Prochaska et al. (2015) with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.55 \pm 0.15$ appeared reasonable when compared with our HIRES spectrum. However, as Fig. 2 shows, there appears to be additional, lower column density Lyman absorption lines at almost all velocities across these profiles. This makes the presence, and extent, of the damping wings difficult to

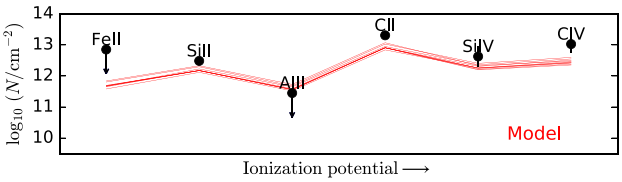

Figure 3. Minimal photoionization model of LLS0344. The observed column densities in Table 3 are shown as data points, with 1σ uncertainties and 2σ upper limits, and the red lines represent CLOUDY column density model values for 20 MCMC samples drawn at random.

discern. It may also appear possible that $N_{\rm H\,I}$ is substantially lower than Prochaska et al. (2015)'s estimate, which would weaken the evidence that this system is a near-pristine LLS.

However, the detection of D I at the expected $-82\,{\rm km\,s}^{-1}$ isotopic shift from the associated H I Ly γ line provides an alternative, more reliable estimate of $N_{\rm H\,I}$ in this system. The top panel of Fig. 2 shows this D I detection and that there is coincident, saturated absorption at the expected position of the Ly β D I line as well. This indicates that the majority – if not all – of the absorption at $-82\,{\rm km\,s}^{-1}$ in Ly γ is most likely due to D I. Indeed, Fig. 2 shows a single Voigt profile overlaying the data for these two D I lines, at the same redshift as the H I line, with $N_{\rm D\,I}$ and b adjusted in VPFIT to visually match the data. This yields $\log_{10}(N_{\rm D\,I}/{\rm cm}^{-2}) = 14.63 \pm 0.10$ and $b = 10\,{\rm km\,s}^{-1}$. Unfortunately, the higher order D I lines are severely blended with Lyman forest absorption so they could not be used to further corroborate the D I analysis. Nevertheless, using the D/H ratio measurement of Cooke et al. (2014), $\log_{10}(N_{\rm D\,I}/{\rm cm}^{-2}) = -4.597 \pm 0.006$, our $N_{\rm D\,I}$ estimate implies that $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.23 \pm 0.10$. As the bottom panel of Fig. 2 shows, overlaying a Voigt profile with this column density on the data (with the same b parameter of $10\,{\rm km\,s}^{-1}$ as for D I) provides a self-consistent result: the model does not fall below the limits imposed by the flux peaks at -350, -320, -150, $+360$, and $+470\,{\rm km\,s}^{-1}$. The uncertainty of 0.1 dex in $N_{\rm D\,I}$ (and therefore $N_{\rm H\,I}$) was determined visually to approximately match the scatter in the HIRES flux values and from the continuum placement. The latter is uncertain by as much as ±15–20 per cent and changing the continuum level by this amount requires an adjustment of the $N_{\rm D\,I}$ of ∼0.05 dex, well within the 0.1 dex uncertainty ascribed to $N_{\rm H\,I}$.

4.1.3 *Photoionization modelling results*

Following Section 3.2, we start the photoionization modelling of LLS0344 with the coarse CLOUDY grid and a 'minimal model': a minimal set of free parameters in the MCMC analysis to match the observing ionic column densities in Table 3. In this case, the minimal model considers that the singly and triply ionized species are produced by the same phase. Fig. 3 shows CLOUDY's predictions for the metal column densities. The main result derived by the MCMC sampling algorithm is the metallicity estimate of [Si/H] = -3.29 ± 0.26. However, the predicted values are larger than the measured ones for Al II, and smaller for C II, C IV, and Si IV. Retaining the assumption that these ions are produced by the same phase, these mismatches may correspond to a deviation from a solar abundance pattern for C and Al and/or uncertainties related to the shape of the HM12 UV background. Reducing the difference between the measured column densities and the predictions of CLOUDY is interesting as it may provide a measurement of [C/Si] and an upper limit on [Al/Si].

Given this low initial metallicity estimate and the possibility of non-solar abundance patterns for C and Al, we then follow Section 3.2

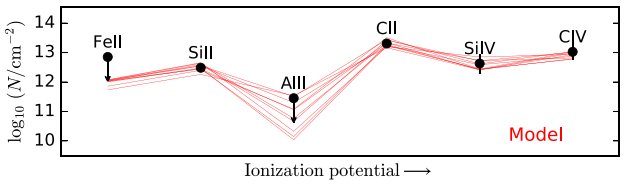

Figure 4. Same as Fig. 3 but for the fiducial photoionization model of LLS0344.

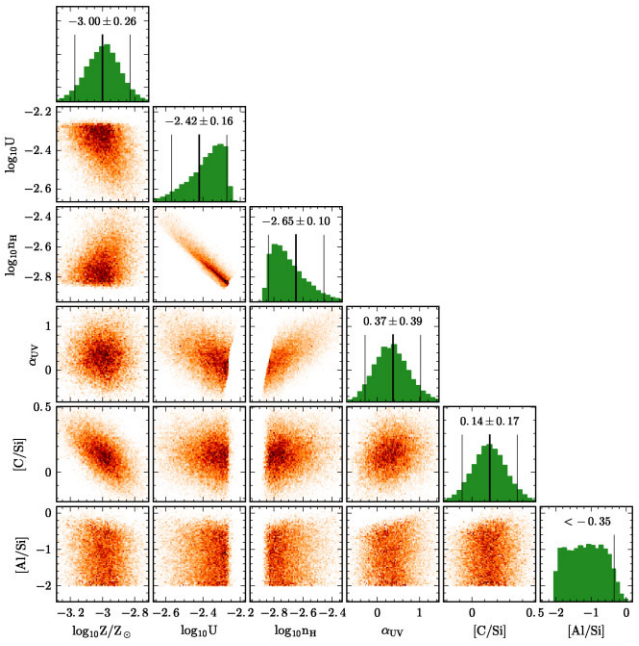

Figure 5. Fiducial photoionization model results for LLS0344. Each panel shows the MCMC distributions of parameter pairs involving metallicity ($Z/Z_{\odot}$), the ionization parameter (U), the volume density of hydrogen ($n_{\rm H}$), the [C/Si] and [Al/Si] ratios, and the slope of the UV background ($\alpha_{\rm UV}$). The panel at the top of each column shows the distribution of the corresponding parameter, together with its mean value or limit (95 per cent confidence).

and generate a finer grid of CLOUDY's models for LLS0344. The grid covers $-4.2 < \log_{10}(n_{\rm H}/{\rm cm}^{-3}) < -1$, $18.5 < \log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) < 19.5$, and $-6 <$ [Si/H] < -2, all with respective steps of 0.2 dex. As stated earlier, we introduce a free parameter $-2.50 < \alpha_{\rm UV} < 1.50$ with steps of 0.4 dex to account for uncertainties related to the shape of the HM12 UV background. We also apply a Gaussian prior on $\alpha_{\rm UV}$, centred on 0 with a standard deviation of $\sigma = 0.5$. Fig. 4 shows CLOUDY's predictions compared to the measured values. The distributions of [Si/H], $n_{\rm H}$, $\alpha_{\rm UV}$, [C/Si], and [Al/Si] derived by the MCMC sampling algorithm are shown in Fig. 5 for the fiducial model. For completeness, and comparison with models in other works, we also include the distribution of the ionization parameter, U, defined as the ratio of the ionising photon density to $n_{\rm H}$.

The main results from the fiducial model (Fig. 5) are a metallicity measurement of [Si/H] $= -3.00 \pm 0.26$, a carbon-to-silicon abundance ratio of [C/Si] $= 0.14 \pm 0.17$, and a upper limit on the aluminium-to-silicon abundance ratio of [Al/Si] ≤ -0.35. As opposed to the minimal model (Fig. 3), there are no mismatches between the CLOUDY predictions and measured column densities in Fig. 4, so it appears that this is a more reliable model. While the supersolar value of [C/Si] has only very marginal significance, the sub-solar upper limit on [Al/Si] is robust to different assumptions

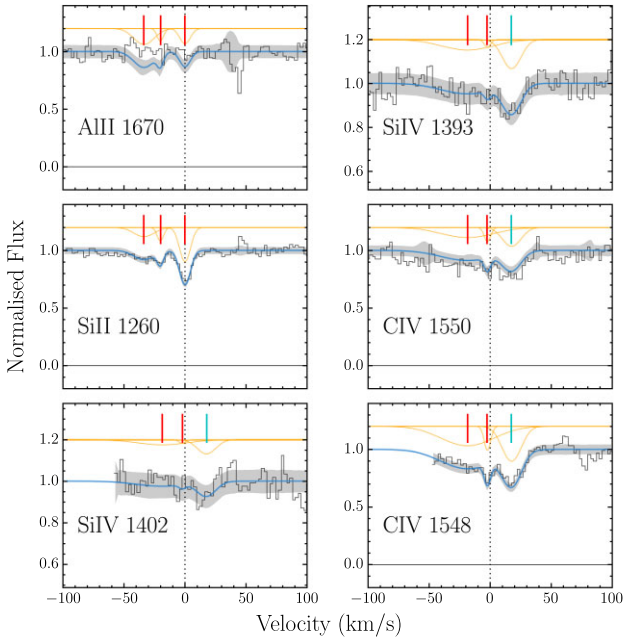

Figure 6. Same as Fig. 1 but for LLS1153 centred at $z_{\rm abs} = 4.038$ towards J1153+1011. For Al II λ 1670, the profile represents the 2σ upper limit on its total column density.

about the photoionization modelling. For example, assuming that the high ions (C IV and Si IV) are not produced by the same phase as the low ions (C II, Si II, Fe II, and Al II) and an HM12 UV background, [Al/Si] ≤ -0.14. Further changing the UV background to 'HM05' within CLOUDY, a revised version of that originally published by Haardt & Madau (1996), has little impact: we obtain [Al/Si] ≤ -0.18. We discuss how these values for [C/Si] and [Al/Si] can be compared to the nucleosynthetic yields of PopIII stars in Section 5.1.

4.2 LLS1153

4.2.1 Metal-line column densities

LLS1153 was first identified in Cooper et al. (2015) towards the $z_{\rm em} = 4.127$ quasar SDSS J115321.68+101112.9 (hereafter J1153+1011), based on absorption features at redshift $z_{\rm abs} = 4.038$, using a Magellan/MagE spectrum. At this redshift, rather than explicitly estimating $N_{\rm H\,I}$, they established a plausible range of values using the same approach explained in Section 3.1: $17.7 < \log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) < 19.0$. The upper limit derives from the lack of apparent damping wings for Ly α, while the lack of flux bluewards of the Lyman limit imposes the lower limit. At $z_{\rm abs} = 4.038$, they identified C IV λλ1548/1550, Si IV λλ1393/1402, and obtained column density upper limits for Si II λ 1260, λ1304, and λ1526.

Guided by the initial results above, we searched the HIRES spectrum for metal absorption lines and we list the detections and non-detections (upper limits) in the second column of Table 3. Fig. 6 depicts the strongest (highest oscillator strength) transitions of the most abundant metal species expected at $z_{\rm abs} = 4.038$. Fig. 6 clearly shows that most of the column density of highly ionized species resides in a component indicated by a cyan vertical tick, but that this is not present in the singly ionized species. Based on this difference, in our photoionization modelling we assume that the low and high ions arise in different phases of LLS1153, and that only the low ions are associated with the H I detected in the Lyman lines. Below, we

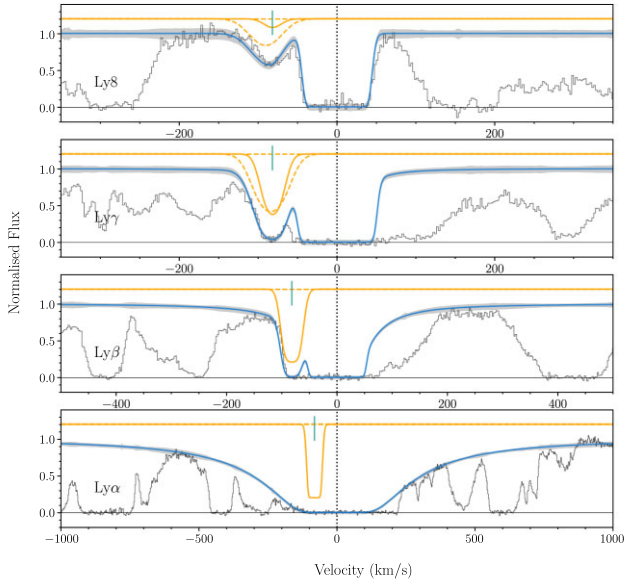

Figure 7. Same as Fig. 2 but for LLS1153 centred at $z_{\rm abs} = 4.038$ towards J1153+1011. The model (blue line) has $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.40 \pm 0.10$ and $b = 15\,{\rm km\,s}^{-1}$ derived from the deuterium absorption (solid orange line) at $-82\,{\rm km\,s}^{-1}$, with $\log_{10}(N_{\rm D\,I}/{\rm cm}^{-2}) = 14.80 \pm 0.10$ (same b). The orange dashed lines in the top two panels represent H I blends set at $z = 3.0289$ for Ly γ and 2.8243 for Ly-8.

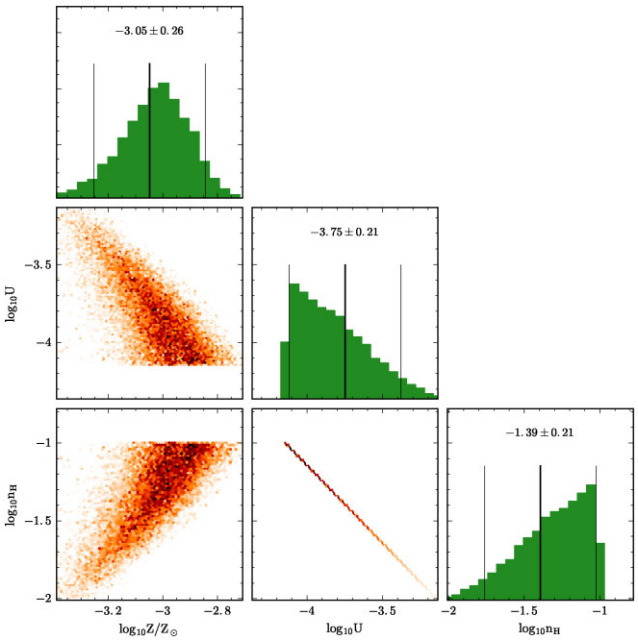

Figure 8. Same as Fig. 5 but for the fiducial photoionization model of LLS1153.

check the impact of including the triply ionized species on LLS1153's metallicity. In that case, the C IV and Si IV absorption in the two velocity components which are (seemingly) in common with the low ions are included in the model.

4.2.2 *H I column density*

Our fiducial H I model is depicted in Fig. 7 with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.40 \pm 0.10$ and $b = 15\,{\rm km\,s}^{-1}$. Having established the redshift of LLS1153 from its metal lines, we checked the corresponding Lyman series lines to assess its $N_{\rm H\,I}$. With a hint of damping wings apparent for the Ly α line in the third panel of Fig. 7, it should be expected that $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \geq 19$. However, similar to LLS0344, the detection of D I at the expected $-82\,{\rm km\,s}^{-1}$ isotopic shift from the associated H I Ly β, Ly γ, and Ly-8 lines provides a more reliable estimate of $N_{\rm H\,I}$. We emphasize again that such a method only provides estimates of the column density and width of the D I lines and not a reliable [D/H] value; indeed, possibly several additional H I blends would need to be comprehensively modelled to determine if a very precise fit of the D I lines was even possible.

The D I detection can be seen in the top panel of Fig. 7, at the expected $-82\,{\rm km\,s}^{-1}$ isotopic shift from the associated H I Ly-8 line. There are coincident, saturated absorption features at the same velocity in the Ly β and Ly γ lines. For both Ly-8 and Ly γ, part of the absorption at $-82\,{\rm km\,s}^{-1}$ is also due to unrelated H I blends. Thus, in Fig. 7, the two top panels each show the Voigt profile of an H I blend to account for some of this additional absorption: $z_{\rm abs} = 3.0289$ for Ly γ and 2.8243 for Ly-8. The third panel shows a single Voigt profile for D I Ly β. An exact, detailed fit to the data is not sought in this case, particularly for the blends, as the main uncertainty in $N_{\rm D\,I}$ stems from the continuum level, amounting to $\sim$0.10 dex (using the same approach as for LLS0344 in Section 4.1). Adjusting $N_{\rm D\,I}$, b, and the blend parameters, within VPFIT to visually match the data yields $\log_{10}(N_{\rm D\,I}/{\rm cm}^{-2}) = 14.80 \pm 0.10$ and $b = 15\,{\rm km\,s}^{-1}$. Using Cooke et al. (2014)'s average measured [D/H] ratio, our $N_{\rm D\,I}$ estimate implies that $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.40 \pm 0.10$ and $b = 15\,{\rm km\,s}^{-1}$. As can be appreciated from Fig. 7, this H I model is reasonable as it does not produce too much absorption at (i) the transmission peaks in Ly α at -600 and $400\,{\rm km\,s}^{-1}$; (ii) the sharp blue wing of the D I Ly β $-100\,{\rm km\,s}^{-1}$; (iii) the blue wing of D I at $-100\,{\rm km\,s}^{-1}$ and sharp rise in flux from zero at $-50\,{\rm km\,s}^{-1}$ in Ly γ; and (iv) the sharp rises in flux at -110 and $-40\,{\rm km\,s}^{-1}$ in Ly-8 and the transmission peak between them.

4.2.3 *Photoionization modelling results*

We start the modelling of LLS1153 on the coarse CLOUDY grids with minimal inputs to the photoionization model. We use the column densities listed in the second column of Table 3 and consider the bulk of H I to be associated with the singly ionized species. The result is an initial metallicity estimate from the minimal model of [Si/H] $= -3.15 \pm 0.26$.

Given this low initial metallicity estimate, we generate a finer grid of CLOUDY models for LLS1153. However, given that we only have a column density measurement of one metal ion (Si II), we do not consider varying the UV background slope or abundance ratios; in this case, the fiducial model has the same assumptions and inputs as the minimal model above. The distributions of $Z/Z_\odot$, U, and $n_{\rm H}$ derived by the MCMC sampling algorithm are shown in Fig. 8. The main result from the fiducial model is a metallicity measurement of [Si/H] $= -3.05 \pm 0.26$ (95 per cent-confidence). This is 0.1 dex higher than the initial estimate derived with the same model assumptions but a courser parameter sampling grid; even though this is well within the 95 per cent-confidence interval, it highlights the importance of using the finer grids.

So far, the high ions (C IV and Si IV) were not considered to be part of the same phase as the low ions and H I. However, from the metal-line velocity structures apparent in Fig. 6, it appears possible that the C IV and Si IV absorption between -40 and $+10\,{\rm km\,s}^{-1}$ (marked by red ticks) arises in the two fitted low-ion components at -20

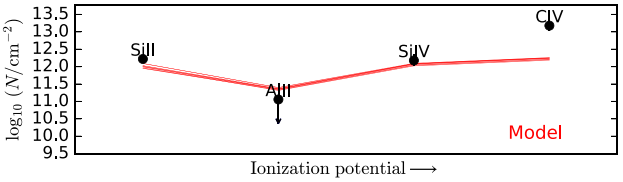

Figure 9. Same as Fig. 3 but for a test model of LLS1153 which includes the column densities of C IV and Si IV in the same phase as the singly ionized gas.

and 0 km s^{-1}. We therefore consider a test model for LLS1153 in which the two bluest (red tick) components are in the same phase as the low ions and are associated with the H I absorption shown in Fig. 7. Fig. 9 compares the metallicity measurements with a sample of CLOUDY's predictions from the MCMC algorithm for this test model. The metallicity measurement shifts to [Si/H] $= -3.58 \pm 0.26$, significantly below our fiducial metallicity, showing that the latter is robustly below [Si/H] $= -3$. However, this test model clearly fails to match the measured column densities of the high ions, similar to the case of LLS0344's minimal model. Indeed, $N_{\mathrm{C\,IV}}$ is more than 1 dex higher than the CLOUDY predictions, and causes a significant tension between the model and the upper limit on the Al II column density.

It is therefore interesting to consider a further test model, similar to the fiducial model of LLS0344, in which we also allow UV background slope variations and non-solar abundance ratios for C and Al. We use the same Gaussian prior on α_{UV} introduced for LLS0344; i.e. $\sigma = 0.5$ centred on 0. The resulting metallicity is [Si/H] $= -3.28 \pm 0.26$, $\alpha_{\mathrm{UV}} = 0.13 \pm 0.48$ and abundance ratios [C/Si] $= 0.89 \pm 0.17$ and [Al/Si] ≤ -0.01. The latter is consistent with the solar value and α_{UV} indicates consistency with the HM12 background. However, the low metallicity in this test model, combined with the overabundance of C, appears interesting at first. Nevertheless, we emphasize that the entire basis of this test model was to consider the same phase for the triply and singly ionized species, just to check the robustness of our fiducial metallicity [Si/H] ≤ -3. Furthermore, only C IV constrains the large value of [C/Si], i.e. it is simply a reflection of the assumption that C IV arises in the same phase as the singly-ionized metals. In these regards, the fiducial model of LLS1153 represents the simplest, most robust set of assumptions, and it produces the highest (i.e. most conservative) metallicity measurement.

4.3 LLS1156

4.3.1 Metal lines

LLS1156 was first identified by Prochaska et al. (2015) towards the $z_{\mathrm{em}} = 3.110$ quasar SDSS J115659.59+551308.1 (hereafter J1156+5513) based on absorption features at redshift $z_{\mathrm{abs}} = 2.616$ in a Keck/ESI spectrum. At this redshift $N_{\mathrm{H\,I}}$ was estimated at $\log_{10}(N_{\mathrm{H\,I}}/\mathrm{cm}^{-2}) = 19.10 \pm 0.30$ along with $b = 30$ km s^{-1}. This value of $N_{\mathrm{H\,I}}$ was determined by visual inspection of theoretical Voigt profiles superimposed on the Ly α line, which has clear enough damping wings to provide reasonable constraints, even at ESI's low resolution of $R = 8000$ (FWHM ~ 37 km s^{-1}). At $z_{\mathrm{abs}} = 2.616$, Prochaska et al. (2015) identified Al II λ 1670, C IV $\lambda\lambda$1548/1550 , and Si IV λ 1393, and obtained column density upper limits for Si II λ 1526, Fe II λ 2344, and Al III λ 1854.

Guided by the initial results above, we searched our HIRES spectrum for metal absorption lines and we list the detections and non-detections (upper limits) in the third column of Table 3. Fig. 10

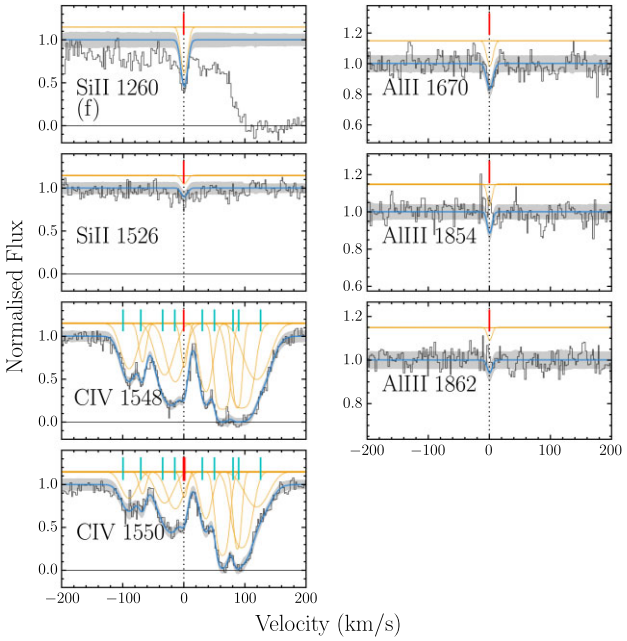

Figure 10. Same as Fig. 1 but for LLS1156 centred at $z_{\mathrm{abs}} = 2.616$ towards J1156+5513. The '(f)' label (in the Si II λ 1260 panel here) indicates that a transition falls in the Lyman forest of the quasar. For Al III $\lambda\lambda$1854/1862, the column densities used for the Voigt profile fits correspond to a 2σ upper limit on the column density.

depicts the strongest transitions of the most abundant metal species expected at $z_{\mathrm{abs}} = 2.616$. Fig. 10 clearly shows that the high-ion (C IV) absorption has a broad, complex velocity structure, which is very different to the very simple, apparently single-component low-ion absorption (Si II and Al II). Si II λ 1260 falls in the Lyman forest of J1156+5513 and sits amongst an H I structure which may affect our column density estimate. Indeed, some of the absorption we attribute to the H I structure could, in fact, be additional Si II λ 1260 absorption; our column density may be a lower limit in that case. The weak detection of Al II λ 1670 will also allow us to constrain the [Al/Si] ratio in Section 4.3.3.

4.3.2 H I column density

Our fiducial H I model is depicted in Fig. 11 (solid blue line) with $\log_{10}(N_{\mathrm{H\,I}}/\mathrm{cm}^{-2}) = 19.30 \pm 0.10$ and $b = 15$ km s^{-1}. Having established the redshift of LLS1156 from the low-ionization metal lines, we checked the corresponding Lyman series lines to assess $N_{\mathrm{H\,I}}$. LLS1156's redshift value is low compared to the rest of our LLS sample, so we only have access to its associated Ly α line. However, its damping wings are clearly seen in Fig. 11 and they directly constrain the neutral hydrogen column density to be $\log_{10}(N_{\mathrm{H\,I}}/\mathrm{cm}^{-2}) = 19.30 \pm 0.10$. We also use $b = 15$ km s^{-1} in our model but note that the precise value does not have an impact on the column density estimate. The uncertainty on $N_{\mathrm{H\,I}}$ was determined from the flux scatter in the HIRES spectrum and the uncertainties in the continuum placement, the latter being the dominant factor.

4.3.3 Photoionization modelling results

We first construct a minimal photoionization model of LLS1156 using the column densities listed in the third column of Table 3. Following the discussion above (Section 4.3.1), we consider the bulk of H I to

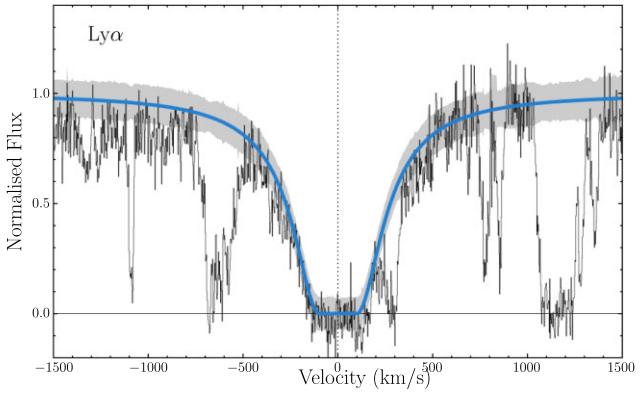

Figure 11. Same as Fig. 2 but for the Lyα line of LLS1156 centred at $z_{\rm abs} = 2.616$ towards J1156+5513. The model (blue solid line) has $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.30 \pm 0.10$ and $b = 15\,{\rm km\,s}^{-1}$.

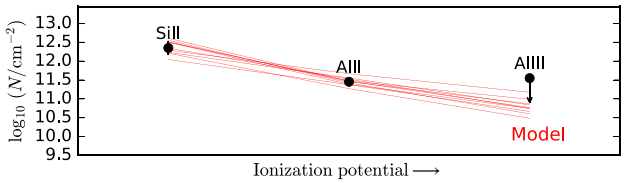

Figure 12. Same as Fig. 3 but for the fiducial model of LLS1156.

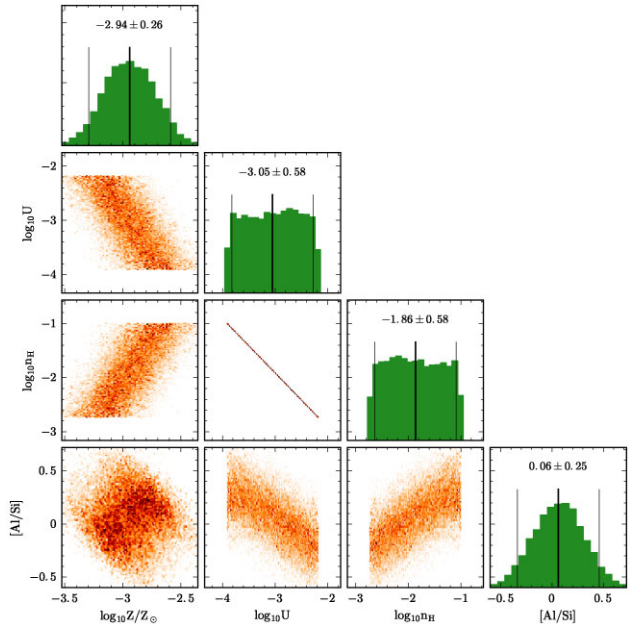

Figure 13. Same as Fig. 5 but for the fiducial photoionization model of LLS1156.

be associated with the singly ionized metal lines (Si II and Al II). Comparison of CLOUDY's predictions with the measured column densities reveals no significant discrepancies, therefore hinting that [Al/Si] must be very close to the solar value. The initial metallicity estimate from this minimal model is [Si/H] $= -2.93 \pm 0.26$. While strictly higher than -3, this is still consistent with it being a 'near-pristine' system compared to the higher metallicity LLSs described in Appendix A where [Si/H] $\gtrsim -2.5$, so it warrants further investigation.

We therefore generated a finer grid of CLOUDY's models for LLS1156, with a fiducial model in which [Al/Si] is also a free parameter. However, given that even the minimal model above closely matched the observed column densities, and that we do not detect higher ions (Al III), we do not include the possibility of varying the UV background slope in the fiducial model. Fig. 12 shows CLOUDY's predictions compared to the measured column densities and Fig. 13 shows the distributions of $Z/Z_\odot$, U, $n_{\rm H}$, and [Al/Si] from the MCMC sampling algorithm. The fiducial metallicity measurement is [Si/H] $= -2.94 \pm 0.26$, almost identical to the minimal model value. The fiducial model also confirms a near-solar [Al/Si] ratio of 0.06 ± 0.25. Unfortunately, this does not provide an important constraint on LLS1156's chemical abundance pattern; i.e. this single abundance ratio does not enable a detailed comparison with models of pristine gas enrichment.

5 DISCUSSION

In Section 4, we reported the discovery of three new near-pristine LLSs (LLS0344, LLS1153, and LLS1156) with metallicity measurements of, respectively, [Si/H] $= -3.00$, -3.05, and -2.94, all with 95 per cent confidence uncertainties of 0.26 dex (dominated by the uncertainties in $N_{\rm H\,I}$ which are estimated to be 0.1 dex in all cases). Fig. 14 summarizes the metallicity distribution at $z > 2$ of LLSs and DLAs in the literature (updated from Robert et al. 2019). It appears that the three new near-pristine LLSs reported here are separated from the bulk of the significant and representative portion of the LLS population shown on the figure. Indeed, the majority of LLSs generally have metallicities $\gtrsim$0.5 dex higher than those of the three 'near-pristine' LLSs. It is important to note that the range of LLS metallicities in each redshift bin (red points), taken directly from Fumagalli et al. (2016a), represent the composite probability distribution from the constituent absorbers. The red error bars in Fig. 14 indicate the 25–75 per cent range of these composite distributions, and this extends almost down to $\log_{10}(Z/Z_\odot) = -3$ for the $z = 3$–4 bin. However, the low-metallicity tails of the distributions for individual LLSs in this bin are particularly broad, with no real lower limit, because a larger fraction of them ($\sim$30 per cent) have poorly constrained H I column densities in the range $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \approx 17.9$–$18.9$ (fig. 1 of Fumagalli et al. 2016a, i.e. damping wings were not detected but no flux was detected bluewards of the Lyman limit). Indeed, no individual LLS was measured to have $\log_{10}(Z/Z_\odot) \lesssim -3$ in this sample. Furthermore, our survey targeted six of these absorbers, which had metallicity upper limits of ≈ -3 from that study, but we found that three actually have considerably higher metallicities ([Si/H] $\gtrsim -2.5$; see Appendix A). That is, the binned LLS points in Fig. 14 are biased to low metallicities, particularly the lower error bars and especially for $z = 3$–4, and include the LLSs we targeted, which all tend to artificially reduce the apparent gap between the near-pristine systems and the higher metallicity bulk of the LLS sample.

Regarding the metallicity distribution of LLSs at $z_{\rm abs} \geq 2$, it appears to be broadly unimodal, with a peak at [Si/H] ~ -2. This was found in both Lehner et al. (2016) and Fumagalli et al. (2016a), who also derived probabilities for finding LLSs with [Si/H] < -3 of $\sim$10 and $\sim$3–18 per cent, respectively, at $2 \leq z_{\rm abs} \leq 4.5$. In Robert et al. (2019), we estimated that LLSs in this redshift range with metallicity [Si/H] < -4 constitute $\sim$1.6 per cent of the population, drawing from an effective sample of $\sim$191 LLSs without strong metallicity biases. With the discovery of three new near-pristine LLSs, we now extend this simple estimate to 8 out of 191 systems with $\log_{10}(Z/Z_\odot) \leq -3$, a proportion of $\sim$4 per cent. This is substantially lower than

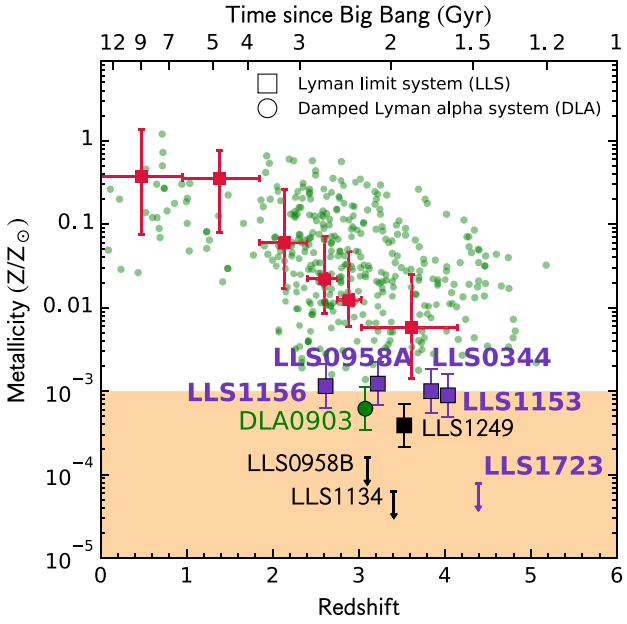

Figure 14. Summary of the metallicity distribution of LLSs (squares) and DLAs (circles) in the literature, showing the three new near-pristine systems discovered in this work: LLS0344, LLS1153, and LLS1156, shown in purple. Our remeasurement of LLS0958A's metallicity is also shown in purple (0.44 dex higher than in Lehner et al. 2016). The lowest metallicity measurement for a LLS from Crighton et al. (2016), LLS1249, is shown in black. All these near-pristine LLS metallicities are plotted with 95 per cent confidence error bars. The three known, apparently metal-free LLSs are indicated with arrows (2σ upper limits): LLS0958B and LLS1134 from Fumagalli et al. (2011a) in black, and LLS1723 from Robert et al. (2019) in purple. The red squares and error bars represent the LLS sample of Fumagalli et al. (2016a) and show the median and 25–75 per cent range of the composite posterior probability density function in redshift bins containing at least 25 LLSs each. The green circles are the DLA samples of Cooke et al. (2011b), Rafelski et al. (2012) and Jorgenson, Murphy & Thompson (2013) [while more recent, dust-corrected DLA metallicity samples are available in De Cia et al. (2018) and Péroux & Howk (2020), dust bias at metallicities $\lesssim -2$ is negligible]. The lowest metallicity measurement ([Si/H]) for a DLA from Cooke, Pettini & Steidel (2017), DLA0903, is shown in green. The shaded orange region shows the expected metallicity range for gas enriched by PopIII supernovae from the simulations of Wise et al. (2012).

the ∼10 per cent and ∼18 per cent upper bounds of Lehner et al. (2016) and Fumagalli et al. (2016a). And, while the latter's lower bound of ∼3 per cent is close to our ∼4 per cent estimate, we must recall that theirs includes the low-metallicity tails of the probability distributions for many LLSs, none of which actually had measured metallicities 3 dex (or more) lower than the solar value. That is, our new estimate appears to indicate that very metal-poor LLSs are rarer than previously estimated. Nevertheless, we note that our sample of eight candidates was deliberately selected to have higher $N_{\rm H\,I}$ values, on average, than these previous studies, so it remains possible that a smaller fraction of higher $N_{\rm H\,I}$ LLSs have $\log_{10}(Z/Z_\odot) \lesssim -3$ than lower $N_{\rm H\,I}$ systems. Naively, we would also expect that the cosmological abundance of $\log_{10}(Z/Z_\odot) \lesssim -3$ LLS may evolve with redshift.

While the above results cannot be considered reliable estimates of the cosmological abundance of these systems, they nevertheless may still provide interesting insights. Indeed, the real shape of the LLS metallicity distribution at $z \geq 2$ is currently unclear. After all, we have been able to find four LLSs with $\log_{10}(Z/Z_\odot) \leq -3$ non-serendipitously in this paper and Robert et al. (2019). If, as we estimate, they are apparently very rare, they may simply be examples of the very low metallicity tail of a broad unimodal metallicity distribution. On the other hand, if they are significantly more common than we estimate (due to unknown biases), they may instead constitute a second mode of the metallicity distribution, possibly with a different origin to the higher metallicity LLSs. Clarifying this issue would help infer the origins of very metal-poor LLSs. Measuring the redshift evolution of the cosmological abundance of $\log_{10}(Z/Z_\odot) \leq -3$ LLSs, and its dependence of $N_{\rm H\,I}$, would be particularly important in this context. In Section 5.1, we consider the possible origin scenarios for the near-pristine LLSs in a similar way to LLS1723 in Robert et al. (2019). The fact that our new near-pristine LLSs have detected metal lines, unlike LLS1723, leaves the question of their enrichment histories more open. Then, to refine the current probability estimates of finding very metal-poor LLSs, we discuss in Section 5.2 how we can better identify very metal-poor candidates in future searches.

5.1 Possible origins of the three new near-pristine LLSs

The abundance pattern of a given absorption system should theoretically exhibit the chemical fingerprint left by the contributing stars that ended their life as SNe. Establishing its chemical profile offers the possibility of investigating the properties of the progenitor stars such as their number, mass distribution and SN explosion energies. Perhaps more importantly, it would determine in which stellar generation the progenitor stars were formed. If the progenitors could be confidently identified as being PopIII stars, this would represent an important, relatively direct observational probe of the properties of the first stars. A similar approach has been employed for very metal-poor local stars in the Milky Way (e.g. Christlieb et al. 2002; Keller et al. 2014) which may have been created from the remnants of PopIII stars. However, the interpretation of stellar abundances is neither direct nor simple (e.g. Asplund 2005; Beers & Christlieb 2005). For instance, the abundance pattern will change substantially over time as the interior is polluted by a star's environment and its internal nuclear burning. Compared to very metal-poor local stars, the large reservoirs of hydrogen in LLSs and DLAs may be less complex to model and could, in principle, have abundance patterns subject to fewer sources of metal enrichment. So far, this approach has been mostly applied to DLAs (e.g. Pettini et al. 2008; Penprase et al. 2010; Cooke et al. 2011a), and offers promising results in some cases (e.g. Cooke, Pettini & Jorgenson 2015; Cooke et al. 2017). However, as can be seen from Fig. 14, LLSs appear to be more metal-poor than DLAs, on average, at $z > 2$. This may indicate that LLSs have a greater chance of pollution by fewer generations of stars – possibly PopIII stars – or even remaining metal-free as has been suggested for LLS1134 and LLS0958B by Fumagalli et al. (2011a) and LLS1723 by Robert et al. (2019).

Generally speaking, the enrichment level of a gas cloud through the explosion of SNe depends on many properties, such as the form of the initial mass function (IMF) of the first stars, their explosion energies, and the level of mixing between stellar layers. Therefore, the fate of these remnants and their interactions with their environment spans many orders of magnitude in length-scale: from sub-pc to Mpc. This is a complex process and key point for numerical simulations if they are to provide clues for the enrichment history of LLSs. In Robert et al. (2019), we discussed the possible origins of LLS1723, an apparently metal-free system. Briefly, it may be a cold gas stream accreting towards a galaxy, or it may represent a high-density portion of the intergalactic medium containing either pristine gas (unpolluted by stellar debris for 1.4 Gyr after the big bang) or the remnants of low-energy

supernovae from (likely low-mass) PopIII stars. For the three new near-pristine LLSs, we may reject the pristine gas possibility because we clearly detected metal absorption. The other two possibilities – cold streams and PopIII remnants – are considered further below.

Before discussing the different possible origin scenarios for the three new near-pristine LLSs, we note that, before this work, only two others had been found. Crighton et al. (2016) and Lehner et al. (2016) analysed the properties of LLS1249 and LLS0958A, respectively, both with metallicities[3] $\log_{10}(Z/Z_{\odot}) \sim -3.4$. For both, only a [C/Si] ratio could be determined, and therefore distinguishing between PopIII and PopII/I scenarios for the progenitor star(s) was not possible. One of the main motivations for our survey was to find more such systems, but with higher column densities, to increase the possibility of establishing more than a single abundance ratio. However, only two of the three new near-pristine systems provided a potentially useful ratio: a $[C/Si] = 0.14 \pm 0.17$, plus an upper limit on $[Al/Si] \leq -0.35$ (95 per cent confidence) for LLS0344, and $[Al/Si] = 0.06 \pm 0.25$ for LLS1156. Unfortunately, these are again not enough to single out a particular nucleosynthetic model. Indeed, using the online fitting tool STARFIT[4] (Frebel et al. 2019) we could not conclude which model best-fitted our measured ratios in LLS0344: enrichment coming from a 140 to 260 $M_{\odot}$ PopIII progenitor, a $\leq 100\,M_{\odot}$ PopIII progenitor, or a PopII/I progenitor gave equally acceptable fits. Therefore, without more detailed abundance ratio constraints coming from their weak metal lines, the possible origins for the three new near-pristine LLSs becomes similar to that discussed for the apparently metal-free LLS1723, though with their higher metallicities taken into account.

A first possible origin scenario for the three new near-pristine LLSs (LLS0344, LLS1153, LLS1156) is that they could be intergalactic clouds that have been enriched by PopIII stars. Indeed, the fact that these three LLSs have $\log_{10}(Z/Z_{\odot}) \leq -3$ is reminiscent of the metallicity floor found in different numerical simulations (Tornatore, Ferrara & Schneider 2007; Maio et al. 2011; Wise et al. 2012; Jaacks et al. 2018). In Wise et al. (2012), a single pair-instability supernova in the mass range 140–260 $M_{\odot}$ can enrich its host halo and surrounding ∼5 kpc to such a metallicity as early as $z = 13$–16. Such a scenario would leave near-pristine LLSs as isolated portions of the IGM that have not been polluted any further by $z \sim 3$. It may also be possible for intergalactic LLSs to have been enriched to a similar level by less massive PopIII stars. A less massive progenitor, with a typical mass $\leq 40\,M_{\odot}$, means less metallic ejecta to enrich the surrounding IGM. Therefore, the enrichment level necessary to form PopII/I stars would not be reached immediately after the death of these less massive PopIII stars, i.e. in contrast to a single very massive PopIII star, many $\leq 40\,M_{\odot}$ PopII progenitors would be needed to reach an enrichment level of $\log_{10}(Z/Z_{\odot}) \sim -3$.

Similar conclusions were reached by Welsh, Cooke & Fumagalli (2019) who consider the possible origins of the 11 most metal-poor DLAs known. DLA0903, represented in Fig. 14, is the most metal-poor DLA from this sample, by a considerable margin, with $\log_{10}(Z/Z_{\odot}) = -3.21 \pm 0.05$ (95 per cent-confidence), which is very similar to the near-pristine LLSs. One of Welsh et al.'s main conclusions is that the 'typical' metal-poor DLA has been enriched by ≤ 72 massive stars with masses $\leq 40\,M_{\odot}$. Moreover, the simulations of Wise et al. (2012) suggest that haloes in which one or many of these low-mass PopIII stars explodes would tend to keep the metal ejecta within them. Such self-enrichment would allow reaching the metallicity level necessary to form PopII/I stars. It would then be interesting to consider how portions of the surrounding IGM could end up enriched by low-mass PopIII remnants, and stay isolated until $z \sim 3$. Whalen et al. (2008) considered the evolution of different types of PopIII SNe for haloes of different masses. For 15–40 $M_{\odot}$ stars, their explosions can disrupt haloes with $\leq 10^7\,M_{\odot}$ but not more massive haloes where the ejecta stays contained and self-enrichment can occur. From the same work, PopIII stars with masses in the range 140–260 $M_{\odot}$ easily disrupt haloes more massive than $10^7\,M_{\odot}$ when exploding as pair-instability supernovae, and it can be approximated that any PopIII supernova explosion can disrupt a halo less massive than $10^7\,M_{\odot}$. Therefore, to have a portion of the IGM polluted by the remnants of $\leq 40\,M_{\odot}$ PopIII stars, the progenitor stars' birth site would need to be within a small halo with $\leq 10^7\,M_{\odot}$.

While still assuming that the near-pristine LLSs arise in the intergalactic medium, we can instead consider a PopII/I enrichment scenario. However, following a similar argument about the apparently metal-free LLS1723 in Robert et al. (2019), this seems unlikely. First, it would require prior enrichment by PopIII stars which, as already discussed, already raises the metallicity of pristine gas to $\log_{10}(Z/Z_{\odot}) \sim -3$. The additional enrichment from one or more PopII/I supernovae will then increase the metallicity significantly above what is observed in the near-pristine LLSs. Furthermore, because most PopII/I stars would be less massive than PopIII stars, the PopII/I supernovae ejecta is more likely to be contained within the host halo. That is, the halo would self-enrich rather than lead to pollution of the IGM with $\log_{10}(Z/Z_{\odot}) \sim -3$.

Given that the hydrogen column densities for the new three near-pristine absorbers are all $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) > 19$, it may be more likely that they arise in a circumgalactic region, possibly as part of a cold stream of gas. This scenario is a well-known prediction of cosmological simulations where the streams of gas being accreted into galactic haloes have column densities consistent with LLSs (e.g. Faucher-Giguère & Kereš 2011; Fumagalli et al. 2011b; van de Voort et al. 2012). None the less, the origin of their metal enrichment is unclear and may arise in several different ways. Indeed, they could have been enriched by PopIII or PopII/I remnants before the redshift at which they are observed ($z_{\rm abs}$). The gas would have then remained unaffected, without further metal pollution, and become part of a cold stream observed at $z_{\rm abs}$. On the other hand, the LLSs could have been polluted for the first time at $z_{\rm abs}$ by metals with a PopII/I origin, coming from a recent SN event in the galaxy/galaxies they surround. Indeed, as pointed out above with the work of Whalen et al. (2008), while remnants of relatively low-mass PopII/I stars are unlikely to be able to reach the IGM, it is possible they may reach a circumgalactic cold stream of gas close to the host-galaxies. The added enrichment from PopII/I explosions would likely exceed or stay close to $\log_{10}(Z/Z_{\odot}) \sim -3$, being consistent with the metallicities derived for the three new near-pristine LLSs discovered in Section 4. In the absence of detailed chemical profiles (i.e. several different abundance ratios) to compare to predictions of theoretical nucleosynthetic yields, both the above options are possible for the source of enrichment in the circumgalactic gas stream scenario.

Finally, we note that, with the advent of optical integral field spectrographs such as Keck/KCWI and VLT/MUSE, it is now possible to observe the galactic environment of metal-poor LLSs relatively efficiently. This is a promising approach that is likely to help better understand the origins of these systems. For instance,

[3]In our reassessment of LLS0958A's metallicity in Appendix B, we find $\log_{10}(Z/Z_{\odot}) \approx -2.9$. Our estimate differs from Lehner et al. (2016) as we include metal absorption over the full velocity structure. However, this does not change the fact that LLS0958A may be in an isolated, intergalactic-like environment, as discussed further below.

[4]See https://2sn.org/starfit/.

Fumagalli et al. (2016b) used ∼5 h of VLT/MUSE observations to map the Ly α-emitting galaxy environment down to luminosities $L_{\mathrm{Ly}\,\alpha} \geq 3 \times 10^{41}$ erg s^{-1} ($0.1L^{*}_{\mathrm{Ly}\alpha}$) in a ∼160 kpc radius around LLS0958A and LLS0958B, with metallicities $\log_{10}(Z/Z_{\odot}) = -3.35 \pm 0.05$ (though see footnote 3) and ≤ -3.8, respectively. Interestingly, this work reports two different environments for these possibly near-pristine and apparently metal free LLSs. On the one hand, the near-pristine LLS0958A shows no nearby galaxies while the apparently metal-free LLS0958B is surrounded by five Ly α emitters at its redshift with impact parameters in the range 77–312 kpc. Three of the emitters appear aligned in projection and may indicate a filamentary intergalactic structure. Interestingly, LLS0958A may then correspond to the case of an isolated, enriched portion of the IGM as we described above, while LLS0958B may be akin to a portion of a circumgalactic cold stream that has remained free of metal enrichment. By contrast, recently Lofthouse et al. (2020) studied the galactic environment of the first near-pristine system, LLS1249, with MUSE and discovered three Ly α-emitting galaxies within 185 kpc (projected) of the absorber.

While it is promising that the different origin scenarios for metal-poor LLSs described in this section and in Robert et al. (2019) can be observed, the initial results above from Fumagalli et al. (2016b) and Lofthouse et al. (2020) above highlight the complexity of metal enrichment at $z > 2$. Indeed, the 'naive' expectation that the most metal-poor LLSs should reside in the least galaxy-rich environment, similar to the IGM, appears to be contradicted by the examples of LLS0958A and B, but perhaps confirmed in the case of LLS1249. However, it is important to consider that LLS0958A may still arise close to a galaxy, or galaxies, which simply fall below to detection limits of the VLT/MUSE observations (Fumagalli et al. 2016b). In any case, it is clear that to refine our discussion of the nature of metal-poor LLSs, building a larger, statistical sample is much needed (ideally, combining absorption and integral field spectroscopy). In that regard, we discuss in Section 5.2 how our strategy for identifying new metal-poor LLSs can be improved.

5.2 Strategies for identifying new metal-poor LLSs

In this work, we targeted with Keck/HIRES eight LLSs without existing high-SNR, high-resolution spectra. These systems were selected because their metallicity upper limits were found to be $\log_{10}(Z/Z_{\odot}) \leq -3$ by Cooper et al. (2015) or Fumagalli et al. (2016a). Nevertheless, four LLSs turned out to be more metal rich than expected, with $\log_{10}(Z/Z_{\odot}) \gtrsim -2.5$ for three of them (LLS0952, LLS1304, LLS2241) and one not a LLS at all (LLS0106) – see Appendix A. We do not discuss their possible origins since they have metallicities not too different from the vast majority of LLSs already studied in the literature (e.g. Fumagalli et al. 2016a; Lehner et al. 2016; Prochaska et al. 2017). None the less, as emphasized above, new observations of metal-poor LLSs are clearly needed, given the variety of possible origin cases for the (so-far) small sample of systems with a robustly measured low metallicity. Therefore, we consider the four rejected LLSs to help establish a better strategy for identifying future new metal-poor systems.

Among the four rejected LLSs, the case of LLS0106 is the simplest. At redshift $z_{\mathrm{abs}} = 4.172$, Prochaska et al. (2015) estimated $\log_{10}(N_{\mathrm{H\,I}}/\mathrm{cm}^{-2}) = 19.05 \pm 0.20$ (1σ) based on visual inspection of theoretical Voigt profiles superimposed on the Ly α line. When selecting our metal-poor candidates, we confirmed these values, and the non-detections of strong metal lines, using the same MIKE spectrum as Prochaska et al. (2015). However, some remaining flux bluewards of LLS0106's Lyman limit can, in fact, be observed in this spectrum, implying an upper limit on $N_{\mathrm{H\,I}}$ of $\log_{10}(N_{\mathrm{H\,I}}/\mathrm{cm}^{-2}) \leq 17$. Even with no secure metal-line detection, the very low $N_{\mathrm{H\,I}}$ would only provide a very high – and, in this context, not useful – upper limit on the metallicity of this LLS, so it would not have been selected for Keck/HIRES follow-up. This illustrates the point that establishing the different properties of LLSs in large surveys is, currently, still very dependent on manual, human interaction which, occasionally, is prone to error. More systematic checks of key observables (in this case the existence of detectable flux bluewards of the Lyman limit) should be undertaken to avoid similar, erroneous LLS identifications.

Low spectral resolution can also lead to underestimating the metallicity of an absorber. Moderate resolution (~ 40 km s^{-1}) spectra, such as those used in Cooper et al. (2015) and Prochaska et al. (2015), may not allow metal column density measurements accurate enough, in some cases, to be sure a system is truly a very metal-poor candidate. For the four rejected LLSs in Appendix A, we note that for important transitions such as Si II λ 1260 or C II λ 1334 only upper limits on their column densities were provided. Clearly, it is not possible to unambiguously detect weak, narrow ($b \sim 5$ km s^{-1}) metal lines in such low-to-medium resolution spectra. We therefore recommend careful assessment of whether the column density measurements and upper limits derived using such spectra confidently constrain the metallicity to $\log_{10}(Z/Z_{\odot}) \leq -3$ in each individual absorber.

On a similar note, more detailed information about the metallicity upper limit should be used to select metal-poor candidate LLSs. Our candidates in Table 1 were selected using estimates of their metallicity upper limits from Lehner et al. (2016) and Cooper et al. (2015). These were derived with a method similar to that described in Section 4, i.e. an MCMC sampling to compare observed metal column densities to the predictions of a grid of CLOUDY photoionization models. In general (except for LLS1304), the resulting metallicity posterior distribution was very broad for each candidate, so the metallicity upper limit was poorly defined. Also, as stated above, the low-to-medium resolution spectra used in these studies mostly provided upper limits on the metal column densities, and this strongly impacted the shape of the metallicity posterior distribution. Therefore, to more securely select very metal-poor candidates from low-to-medium resolution spectra, we recommend running a MCMC sampling of a full suite of photoionization models which cover all conditions found in known LLSs (e.g. Lehner et al. 2016; Fumagalli et al. 2016a). Follow-up spectroscopy could then only be conducted if its metallicity distribution shows, with good confidence, an upper-limit consistent with $\log_{10}(Z/Z_{\odot}) \leq -3$.

Finally, it is also important that the wavelength coverage of an LLS survey is tailored not only to the redshift range, but also the column density range targeted. For systems with $\log_{10}(N_{\mathrm{H\,I}}/\mathrm{cm}^{-2}) \leq 19.0$, the high-order Lyman series and Lyman limit must be covered to ensure the best constraints on $N_{\mathrm{H\,I}}$, while the Ly α line should be sufficient to cover for higher $N_{\mathrm{H\,I}}$ systems because its damping wings strongly constrain $N_{\mathrm{H\,I}}$. Clearly, the SNRs required in the spectral regions of these features also needs careful consideration. The SNR and wavelength coverage requirements for an LLS survey are also impacted by the H I Ly α forest. Indeed, depending on an LLS's redshift, many important metal lines can fall inside the forest. Therefore, H I blends might prevent estimation of a metal transition's column density at low-to-medium resolution. Furthermore, while high resolution can allow deblending a narrow metal line from a H I forest feature, it can also fail depending on how broad and saturated the H I feature is. To reduce this problem, observing a wavelength range that allows different metal lines to be observed simultaneously is recommended. Obviously, one can select metal transitions that are redwards of the Ly α forest so they are less likely to be contaminated

by unrelated features. While it is common practice to use the strongest available transition of an ion to measure its column density, the weaker ones can also be useful. They have some chance of being less affected by H I forest blends but may also be redwards of it. None the less, for transitions that are much weaker than their ion's strongest one (e.g. Si II λ 1526 and λ1260, respectively), it might not be feasible to obtain a detection or a useful upper limit when $\log_{10}(Z/Z_{\odot}) < -3$. For example, detecting Si II λ 1526 in LLS1304 would require SNR $\sim$ 300 at 6620 Å with MagE ($\sim$187 for ESI). Such high SNRs are clearly not reachable in most surveys. For comparison, only SNR $\sim$ 40 would be required for HIRES.

To summarize the discussion above, the following are important building blocks for a survey to identify very metal-poor candidates with low-to-medium resolution spectra:

(i) Resolution high enough, for the SNR targeted, to allow column density measurements precise enough (and low enough upper limits for metals) for a maximum metallicity of $\log_{10}(Z/Z_{\odot}) \leq -3$ to be determined through photoionization modelling. What constitutes 'enough' for these various survey parameters needs to be carefully modelled, with simulated spectra, when initially considering a survey, depending on the available telescopes and spectrographs.
(ii) MCMC sampling of a full suite of photoionization models of very metal-poor candidates, and to conduct a high-resolution spectroscopic follow-up of the candidates whose metallicity distribution shows an upper limit consistent with $\log_{10}(Z/Z_{\odot}) \leq -3$ at high confidence.
(iii) Regardless of the spectral quality of the data, a reliable estimate of $N_{\rm H\,I}$ is important.

Regarding the high-resolution follow-up observing strategy, we recommend

(i) wavelength coverage of the strong metal transitions (e.g. Si II λ 1260) and of weaker representatives of the same ions;
(ii) wavelength coverage of at least the Lyman limit and Ly α line for LLSs with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \leq 19.0$ and $\geq$19.0, respectively;
(iii) high enough SNR and resolving power for these regions (SNR $\geq$25–30 and $R \sim 35\,000$) to robustly measure the column density of absorption lines and resolve important velocity structure.

As a final note, we add that it is also important to apply consistent methodologies to establish a large sample of near-pristine absorbers. In Section 5, we mentioned that only two other near-pristine LLSs had previously been found: LLS1249 and LLS0958A by Crighton et al. (2016) and Lehner et al. (2016), respectively, who measured their metallicities to be $\log_{10}(Z/Z_{\odot}) \approx -3.4$. Due to the availability of a high-resolution HIRES spectrum of LLS0958A in the KODIAQ sample of O'Meara et al. (2015), we conducted a reassessment of its properties, particularly its metallicity, with the same methodology we applied to the other eight LLSs described in this work. Details are provided in Appendix B. We find $\log_{10}(Z/Z_{\odot}) = -2.91 \pm 0.26$ for LLS0958A, which is substantially higher than Lehner et al.'s (2016) value of -3.35 ± 0.05. However, LLS0958A's low metallicity is still intriguing given that it could be in an isolated, intergalactic-like environment in Fumagalli et al. (2016b), as we discussed in Section 5.1.

6 CONCLUSIONS

In this work, we described a dedicated search for near-pristine LLSs in which we selected a total of eight LLSs from large, low-to-medium resolution spectroscopic surveys (Cooper et al. 2015; Prochaska et al. 2015). The eight candidates exhibited no or only weakly detected metal absorption lines, suggesting a very low metallicity, $\log_{10}(Z/Z_{\odot}) < -3$, derived in the photoionization modelling of Cooper et al. (2015) and Fumagalli et al. (2016a). To confirm their very metal-poor nature, we observed them in a Keck/HIRES high-resolution spectroscopy campaign from 2016 to 2017. One of these eight LLSs, LLS1723, did not display metal lines in the follow-up spectra, as presented in Robert et al. (2019). All the remaining seven LLSs, presented in this work, exhibited detectable, albeit weak, metal absorption lines in the Keck/HIRES spectra.

To make more accurate metallicity measurements for these seven near-pristine candidates, we first derived metal and hydrogen column densities from the Keck/HIRES spectra using multicomponent Voigt profile fits. These were compared to a grid of CLOUDY photoionization models, sampled via an MCMC algorithm, to measure the metallicities. This confirmed the discovery of three new near-pristine LLSs – LLS0344, LLS1153, and LLS1156 – with final metallicities [Si/H] $= \log_{10}(Z/Z_{\odot}) = -3.00 \pm 0.26$, -3.05 ± 0.26, and -2.94 ± 0.26 (95 per cent confidence), respectively. Unfortunately, only two of these systems exhibited more than one detected metal ion species: LLS0334's [C/Si] ratio is consistent with solar, while its [Al/Si] ratio is constrained to be less than 0.14 dex below solar; for LLS1156, the [Al/Si] ratio is consistent with solar. These constraints on the nucleosynthetic profile of these LLSs are not detailed enough to distinguish between PopIII and later population progenitor models. Even the low [Al/Si] ratio for LLS0344 did not strongly constrain the progenitor mass within the PopIII scenarios explored. For the remaining four LLSs, we found that the previous metallicity estimates, prior to this work, were underestimated; our Keck/HIRES spectra and photoionization modelling strongly suggest metallicities $\log_{10}(Z/Z_{\odot}) \gtrsim -2.5$ (see Appendix A).

With the discovery of three new systems via our dedicated search, and the first two serendipitous discoveries (LLS1249 by Crighton et al. 2016 and LLS0958A by Lehner et al. 2016, noting our reassessment of the latter in Appendix B), there are now five LLSs with metal detections known to be consistent with $\log_{10}(Z/Z_{\odot}) \leq -3$. Fig. 14 summarizes their metallicities in the context of the general LLS population. The near-pristine systems have clearly lower metallicities than the vast majority of LLSs and are relatively rare. Given the samples from which these five systems and the three apparently metal-free LLSs (Fumagalli et al. 2011a; Robert et al. 2019) were drawn, we estimated that systems with $\log_{10}(Z/Z_{\odot}) \leq -3$ constitute only $\sim$4 per cent of LLSs. However, it is not clear from the metallicity distribution alone whether they are truly a separate population, with a different origin, to the higher metallicity LLSs. We considered the possible origins for near-pristine LLSs, based on results from existing simulations in the literature, in Section 5.1. Systems arising in the high-density regions of the IGM (i.e. not near galaxies) may well be the chemical remnants of PopIII stars with no contamination from later generations. However, if the near-pristine LLSs arise in circumgalactic environments, as seems most likely for the new systems, given their higher column densities [i.e. $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) > 19$], then their metallicities are consistent with being polluted by PopIII or PopII/I stars at very early times, or by PopII/I supernovae from the nearby galaxy/galaxies, at or just prior to the epoch at which we observe them. Therefore, the origin of the three new near-pristine systems remains unclear.

Simulations may be able to help clarify the origins of near-pristine LLSs. The major challenge in this context is to resolve very low-density gas. Peeples et al. (2019) made concerted progress in this direction for circumgalactic gas, but simulating the IGM at high enough mass resolution will be substantially more difficult.

Testing the predictions of these simulations will also require a larger survey for near-pristine and apparently metal-free LLSs, with well-controlled selection effects. For example, a basic test would be to compare the redshift and column density dependence of the number density of such LLSs with simulations of different origin scenarios. Similarly, discovering near-pristine LLSs with a suitably large variety of detected metal species to allow detailed nucleosynthetic modelling will also likely require a significantly larger sample. Our success in discovering three near-pristine LLSs in a targeted way opens the possibility for creating a well-defined, larger statistical sample in the future to conduct such tests. We recommended observational strategies for future surveys in Section 5.2.

ACKNOWLEDGEMENTS

We thank Ryan Cooke for assisting with aspects of the data reduction. PFR acknowledges support through a Swinburne University Postgraduate Research Award (SUPRA) PhD scholarship and travel support through the International Telescopes Support Office. MTM thanks the Australian Research Council for *Discovery Project* grant DP130100568 that supported this work. MF acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 757535) and funding support by Fondazione Cariplo, grant no. 2018-2329.

The data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. Australian access to the W. M. Keck Observatory has been made available through Astronomy Australia Limited via the Australian Government's National Collaborative Research Infrastructure Strategy, via the Department of Education and Training, and an Australian Government astronomy research infrastructure grant, via the Department of Industry, Innovation and Science. This research has made use of the KOA, which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration. Our analysis made use of ASTROPY (Astropy Collaboration 2013), MATPLOTLIB (Hunter 2007), and BARAK (https://github.com/nhmc/Barak).

DATA AVAILABILITY

All raw exposures of all quasars studied in this work are publicly available in the KOA. The final, combined spectrum of each quasar is available on request to the corresponding author (MTM).

REFERENCES

Asplund M., 2005, ARA&A, 43, 481
Astropy Collaboration, 2013, A&A, 558, A33
Beers T. C., Christlieb N., 2005, ARA&A, 43, 531
Bromm V., Larson R. B., 2004, ARA&A, 42, 79
Bromm V., Yoshida N., Hernquist L., McKee C. F., 2009, Nature, 459, 49
Carswell R. F., Webb J. K., 2014, Astrophysics Source Code Library, record ascl:1408.015
Cen R., Riquelme M. A., 2008, ApJ, 674, 644
Christlieb N. et al., 2002, Nature, 419, 904
Cooke R. J., Madau P., 2014, ApJ, 791, 116
Cooke R., Pettini M., Steidel C. C., Rudie G. C., Jorgenson R. A., 2011a, MNRAS, 412, 1047
Cooke R., Pettini M., Steidel C. C., Rudie G. C., Nissen P. E., 2011b, MNRAS, 417, 1534
Cooke R. J., Pettini M., Jorgenson R. A., Murphy M. T., Steidel C. C., 2014, ApJ, 781, 31
Cooke R. J., Pettini M., Jorgenson R. A., 2015, ApJ, 800, 12
Cooke R. J., Pettini M., Steidel C. C., 2017, MNRAS, 467, 802
Cooke R. J., Pettini M., Steidel C. C., 2018, ApJ, 855, 102
Cooper T. J., Simcoe R. A., Cooksey K. L., O'Meara J. M., Torrey P., 2015, ApJ, 812, 58
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
Crighton N. H. M., O'Meara J. M., Murphy M. T., 2016, MNRAS, 457, L44
De Cia A., Ledoux C., Petitjean P., Savaglio S., 2018, A&A, 611, A76
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Dekel A. et al., 2009, Nature, 457, 451
Faucher-Giguère C.-A., Kereš D., 2011, MNRAS, 412, L118
Ferland G. J. et al., 2013, Rev. Mex. Astron. Astrofis., 49, 137
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Frebel A. et al., 2005, Nature, 434, 871
Frebel A., Christlieb N., Norris J. E., Thom C., Beers T. C., Rhee J., 2007, ApJ, 660, L117
Frebel A., Ji A. P., Ezzeddine R., Hansen T. T., Chiti A., Thompson I. B., Merle T., 2019, ApJ, 871, 146
Fumagalli M., O'Meara J. M., Prochaska J. X., 2011a, Science, 334, 1245
Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011b, MNRAS, 418, 1796
Fumagalli M., O'Meara J. M., Prochaska J. X., Worseck G., 2013, ApJ, 775, 78
Fumagalli M., O'Meara J. M., Prochaska J. X., 2016a, MNRAS, 455, 4100
Fumagalli M., Cantalupo S., Dekel A., Morris S. L., O'Meara J. M., Prochaska J. X., Theuns T., 2016b, MNRAS, 462, 1978
Haardt F., Madau P., 1996, ApJ, 461, 20
Haardt F., Madau P., 2012, ApJ, 746, 125
Heger A., Woosley S. E., 2010, ApJ, 724, 341
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jaacks J., Thompson R., Finkelstein S. L., Bromm V., 2018, MNRAS, 475, 4396
Jorgenson R. A., Murphy M. T., Thompson R., 2013, MNRAS, 435, 482
Keller S. C. et al., 2014, Nature, 506, 463
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Lehner N. et al., 2013, ApJ, 770, 138
Lehner N., O'Meara J. M., Howk J. C., Prochaska J. X., Fumagalli M., 2016, ApJ, 833, 283
Lehner N., Wotta C. B., Howk J. C., O'Meara J. M., Oppenheimer B. D., Cooksey K. L., 2019, ApJ, 887, 5
Liu B., Bromm V., 2020, MNRAS, 497, 2839
Lofthouse E. K. et al., 2020, MNRAS, 491, 2057
Maio U., Khochfar S., Johnson J. L., Ciardi B., 2011, MNRAS, 414, 1145
Murphy M. T., 2016, UVES_popler: POst PipeLine Echelle Reduction software , https://doi.org/10.5281/zenodo.1297190
Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2019, MNRAS, 482, 3458
O'Meara J. M., Prochaska J. X., Worseck G., Chen H.-W., Madau P., 2013, ApJ, 765, 137
O'Meara J. M. et al., 2015, AJ, 150, 111
Oke J. B., Gunn J. E., 1982, PASP, 94, 586
Peeples M. S. 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., Zych B. J., Steidel C. C., Chaffee F. H., 2008, MNRAS, 385, 2011
Pointon S. K., Kacprzak G. G., Nielsen N. M., Muzahid S., Murphy M. T., Churchill C. W., Charlton J. C., 2019, ApJ, 883, 78
Prochaska J. X., O'Meara J. M., Worseck G., 2010, ApJ, 718, 392

Prochaska J. X., O'Meara J. M., Fumagalli M., Bernstein R. A., Burles S. M., 2015, ApJS, 221, 2
Prochaska J. X. et al., 2017, ApJ, 837, 169
Rafelski M., Wolfe A. M., Prochaska J. X., Neeleman M., Mendez A. J., 2012, ApJ, 755, 89
Robert P. F., Murphy M. T., O'Meara J. M., Crighton N. H. M., Fumagalli M., 2019, MNRAS, 483, 2736
Savage B. D., Sembach K. R., 1991, ApJ, 379, 245
Steidel C. C., 1990, ApJS, 74, 37
Tornatore L., Ferrara A., Schneider R., 2007, MNRAS, 382, 945
van de Voort F., Schaye J., Altay G., Theuns T., 2012, MNRAS, 421, 2809
Welsh L., Cooke R., Fumagalli M., 2019, MNRAS, 487, 3363
Whalen D., van Veelen B., O'Shea B. W., Norman M. L., 2008, ApJ, 682, 49
Wise J. H., Turk M. J., Norman M. L., Abel T., 2012, ApJ, 745, 50
Wotta C. B., Lehner N., Howk J. C., O'Meara J. M., Prochaska J. X., 2016, ApJ, 831, 95
Wotta C. B., Lehner N., Howk J. C., O'Meara J. M., Oppenheimer B. D., Cooksey K. L., 2019, ApJ, 872, 81
Yong D. et al., 2013, ApJ, 762, 26

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

LLS0952_column_densities.pdf
LLS0952_parameter_distributions.pdf
LLS1304_parameter_distributions.pdf
LLS1304_column_densities.pdf
LLS2241_column_densities.pdf
LLS2241_parameter_distributions.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.

APPENDIX A: HIGHER METALLICITY LYMAN LIMIT SYSTEMS IN THE SAMPLE

As explained in Section 3.2, only three of the near-pristine candidates have measured metallicities consistent with $\log_{10}(Z/Z_\odot) \leq -3$: LLS0344, LLS1153, and LLS1156. Among the remaining four, LLS0106 has $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2})$ well below the normal LLS definition of 17.2, and the others (LLS0952, LLS1304, and LLS2241) are consistent with $\log_{10}(Z/Z_\odot) \gtrsim -2.5$. The analysis of their absorption features is described very briefly below for completeness. Plots of the photoionization results for each system (akin to Figs 4 and 5 for LLS0344) are available as supporting information online. Table A1 summarizes the metal and H I column density measurements and upper limits, and metallicity measurements for each system. These four systems were considered in Section 5.2 when discussing how to improve searches for very metal-poor candidates in future surveys. Finally, for LLS2241, we find that its H I column density is $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \approx 20.3$, consistent with the normal DLA definition, so this system may be better considered as a very low metallicity DLA rather than an LLS.

A1 LLS0106

LLS0106 was first identified by Prochaska et al. (2015) towards the $z_{\rm em} = 4.43$ quasar SDSS J010619.20+004823.3 (hereafter J0106+0048), based on absorption features at redshift $z_{\rm abs} = 4.172$, using a Magellan/MIKE spectrum. At this redshift, $N_{\rm H\,I}$ was estimated at $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.05 \pm 0.20$ along with $b = 30\,{\rm km\,s}^{-1}$ from the Ly α and Ly β lines. Prochaska et al. (2015) did not detect any metal lines, finding upper limits for C IV $\lambda\lambda$1548/1550, Si II λ 1304, Si IV $\lambda\lambda$1393/1402 , and O I λ 1302. In our HIRES spectrum, we tentatively detect Si II λ 1260, the strongest Si II line, in a broad Ly α forest feature and provide the corresponding column density Table A1. Fig. A1 depicts this Si II λ 1260 feature and the weaker (smaller oscillator strength) Si II λ 1304 and Si II λ 1526 transitions for comparison. The weaker Si II lines were also used to constrain N(Si) in the VPFIT modelling.

Our fiducial H I model is depicted in Fig. A2: $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 17$ and $b = 15\,{\rm km\,s}^{-1}$. Despite the previous estimate of $N_{\rm H\,I} \geq 19$ (Prochaska et al. 2015), we find that there is significant flux bluewards of the Lyman limit, as can be seen in Fig. A2. The nominal upper limit value on $N_{\rm H\,I}$ of $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \leq 17$ is therefore conservative in this case, demonstrating that this system is almost certainly not a true LLS considering the conventional threshold of $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) \geq 17.2$. Therefore, we do not proceed to attempt a metallicity measurement for LLS0106.

Table A1. Same as Table 3 but for the four LLSs found to have higher metallicities than near-pristine systems [$\log_{10}(Z/Z_\odot) \leq -3$].

Ion	$\log_{10}(N/{\rm cm}^{-2})$			
	LLS0106	LLS0952	LLS1304	LLS2241
Si II	11.75 ± 0.15	13.55 ± 0.04	12.55 ± 0.01	13.22 ± 0.05
Si IV	–	–	13.01 ± 0.02	–
C II	–	14.41 ± 0.17	13.30 ± 0.03	–
C IV	–	–	13.15 ± 0.06	–
Fe II	–	13.59 ± 0.05	–	–
Al II	–	–	–	12.08 ± 0.04
Al III	–	–	≤ 11.79	≤ 12.35
O I	–	14.99 ± 0.18	–	–
H I	≤ 17	19.80 ± 0.10	17.90 ± 0.10	20.3 ± 0.10
[Si/H]	[a]	-1.85 ± 0.26	-2.30 ± 0.26	-2.75 ± 0.26

[a] We did not calculate a metallicity for LLS0106 because its $N_{\rm H\,I}$ was below that of the normal definition for an LLS (see Section A1).

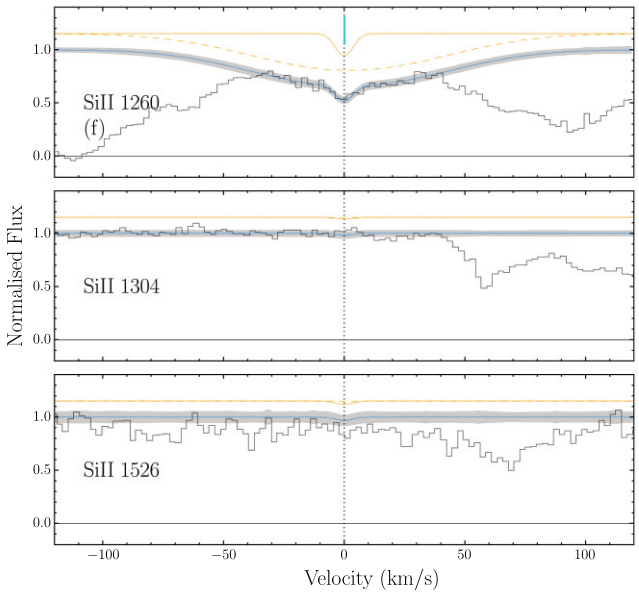

Figure A1. Same as Fig. 1 but for J0106+0048 and LLS0106 with the zero velocity redshift set at $z_{\rm abs} = 4.172$. For Si II λ 1304 and Si II λ 1526, the column densities used for the Voigt profile fits are the same used for Si II λ 1260, the strongest Si II transition.

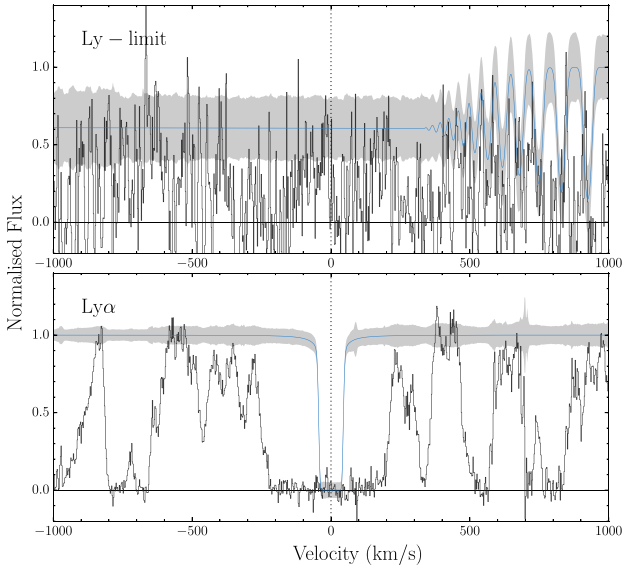

Figure A2. Same as Fig. 2 but for J0106+0048 and LLS0106 with the zero velocity redshift set at $z_{\rm abs} = 4.172$ and profile model (blue solid line) parameters $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 17$ and $b = 15\,{\rm km\,s}^{-1}$.

A2 LLS0952

LLS0952 was first identified by Prochaska et al. (2015) towards the $z_{\rm em} = 3.396$ quasar SDSS J095256.41+332939.0 (hereafter J0952+3329), based on absorption features at redshift $z_{\rm abs} = 3.212$, using a Keck/ESI spectrum. $N_{\rm H\,I}$ was estimated at $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.90 \pm 0.20$ (1σ) along with $b = 30\,{\rm km\,s}^{-1}$ from the Ly α line, the main constraint coming from its damping wings. Prochaska et al. (2015) reported detections of C II λ 1334, Si IV $\lambda\lambda$1393/1402 , and O I λ 1302, with only upper limits determined for other singly ionized species (Si II, Al II, Fe II) and higher ions (Al III and Si IV). Our Keck/HIRES spectrum provides strong detections of C II, O I, Si II, and even Fe II, with column density measurements listed in Table A1, but we do not detect the high ions (C IV and Si IV). Fig. A3 depicts the strongest transitions of the most abundant metal species in this LLS at $z_{\rm abs} = 3.213$, clearly showing a low-ionization phase for LLS0952: the different transitions depicted are aligned in velocity space and comprise the same, simple velocity structure. Therefore, for the fiducial model of LLS0952, all the transitions depicted are considered to be produced by the same phase as the H I absorption.

Our fiducial H I model is depicted in Fig. A4 with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.80 \pm 0.10$ and $b = 15\,{\rm km\,s}^{-1}$. Fig. A4 clearly shows that the Ly α transition has damping wings and will therefore strongly constrain the column density. The uncertainty in $N_{\rm H\,I}$ is determined in a similar way to that for LLS0344, taking into account the scatter in flux and uncertainty in the continuum placement.

A3 LLS1304

LLS1304 was first identified by Cooper et al. (2015) towards the $z_{\rm em} = 3.651$ quasar SDSS J130452.57+023924.8 (hereafter J1304+0239), based on absorption features at redshift $z_{\rm abs} = 3.3369$, using a Magellan/MagE spectrum. A plausible range of values for $N_{\rm H\,I}$ was estimated based on the absence of detected flux bluewards of the Lyman limit and the apparent lack of Ly α damping wings: $17.9 < \log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) < 18.7$. Cooper et al. (2015) reported detections of the high ions (C IV and Si IV) and non-detections (upper limits) of

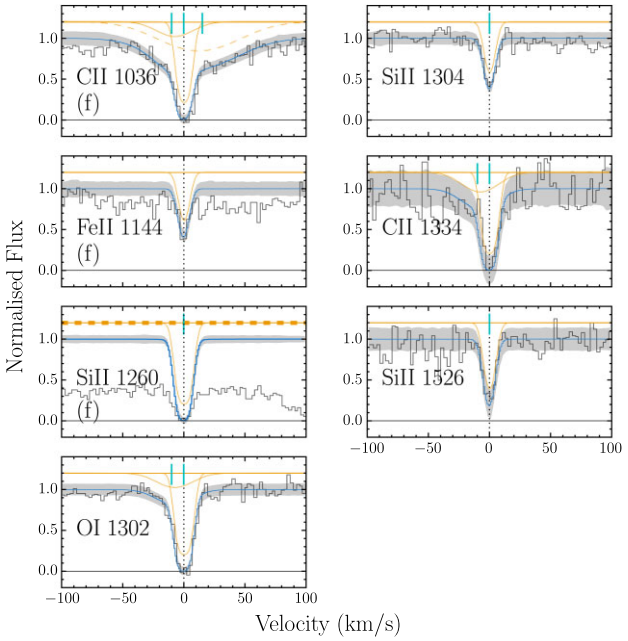

Figure A3. Same as Fig. 1 but for J0952+3329 and LLS0952 with the zero velocity redshift set at $z_{\rm abs} = 3.213$.

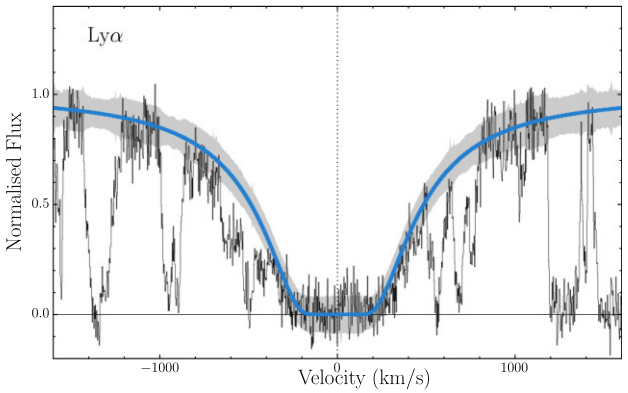

Figure A4. Same as Fig. 2 but for J0952+3329 and LLS0952 with the zero velocity redshift set at $z_{\rm abs} = 3.213$ and profile model (blue solid line) parameters $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 19.80 \pm 0.10$, and $b = 15\,{\rm km\,s}^{-1}$.

Si II and Al III. Our HIRES spectrum provided detections of two low ions (C II and Si II) and the same high ions, as shown in Fig. A5. Si II λ 1260 is in the Lyman forest of J1304+0239, between two H I blends, and this almost certainly biases the Si II column density estimate to higher values. We therefore treated it as an upper limit in the photoionization modelling. The low and high ions have very similar velocity structures, suggesting they may be part of the same phase. However, the metallicity obtained from the photoionization modelling does not depend strongly on this assumption.

Fig. A6 shows our fiducial H I model with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 17.90 \pm 0.10$ and $b = 20\,{\rm km\,s}^{-1}$. There is no apparent flux bluewards of the Lyman limit so we proceed by following Cooper et al. (2015) to establish a plausible $N_{\rm H\,I}$ range. The upper limit is defined by the Ly α and β lines, which show no apparent damping wings. The lower limit on $N_{\rm H\,I}$ (from the Lyman limit), $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 17.90 \pm 0.10$, provides the highest (i.e. most conservative) estimate of LLS1304's metallicity. The uncertainty on $N_{\rm H\,I}$ is determined by taking into

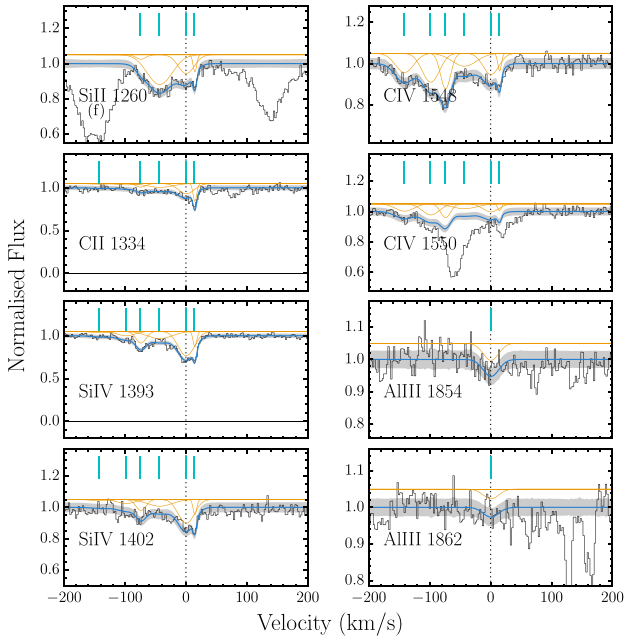

Figure A5. Same as Fig. 1 but for J1304+0239 and LLS1304 with the zero velocity redshift set at $z_{\rm abs} = 3.3369$. For Al III λλ1854/1862, the column densities used for the Voigt profile fits correspond to a 2σ upper limit derived using the apparent optical depth method. Note that C IV λ 1550 is blended with C IV λ 1548 absorption from a higher redshift $z_{\rm abs} = 3.3433$ system, the absorption profiles of which are not shown.

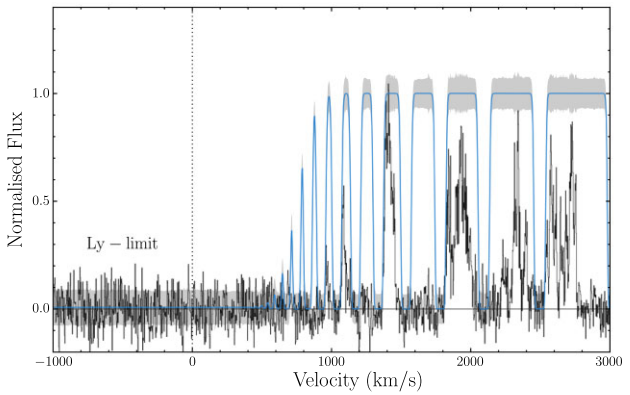

Figure A6. Same as Fig. 2 but for J1304+0239 and LLS1304 with the zero velocity redshift set at $z_{\rm abs} = 3.33695$ and profile model (blue solid line) parameters $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 17.90 \pm 0.10$, and $b = 20\,{\rm km\,s}^{-1}$.

account the scatter in flux in the HIRES spectrum, and the continuum placement.

A4 LLS2241

LLS2241 was first identified by Prochaska et al. (2015) towards the $z_{\rm em} = 4.470$ quasar SDSS J224147.70+135203.0 (hereafter J2241+1352), based on absorption features at redshift $z_{\rm abs} = 3.653\,393$, using a Keck/ESI spectrum. Damping wings in the Ly α profile indicated $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 20.20 \pm 0.20$ with $b = 30\,{\rm km\,s}^{-1}$. Prochaska et al. (2015) reported no metal detections, with upper limits determined for Si II, Al II, Fe II, Al III and C IV. We detect only low ions Si II and Al II in our Keck/HIRES spectrum, with an upper limit determined for Al III, as detailed in Table A1,

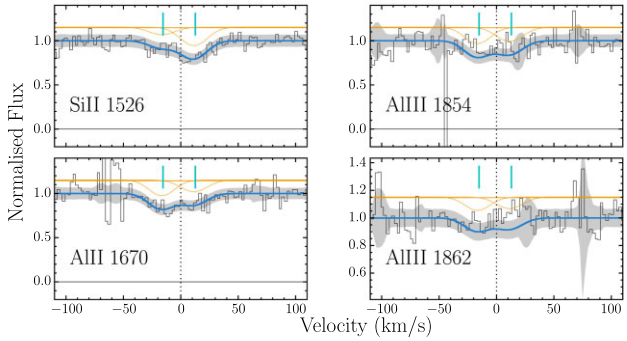

Figure A7. Same as Fig. 1 but for J2241 + 1352 and LLS2241 with the zero velocity redshift set at $z_{\rm abs} = 3.65492$. For Al III λλ1854/1862, the column densities used for the Voigt profile fits correspond to a 2σ upper limit derived using the apparent optical depth method.

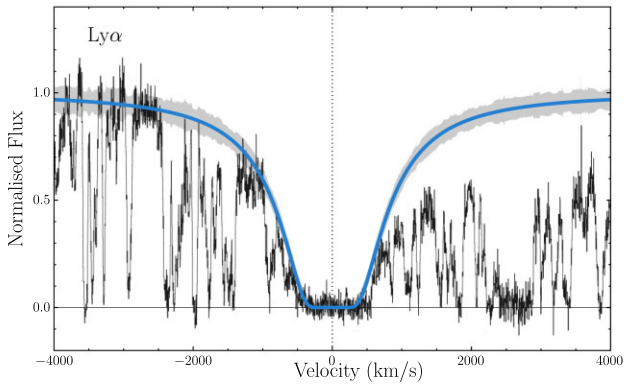

Figure A8. Same as Fig. 2 but for J2241 + 1352 and LLS2241 with the zero velocity redshift set at $z_{\rm abs} = 3.65492$ and profile model (blue solid line) parameters $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 20.30 \pm 0.10$, and $b = 30\,{\rm km\,s}^{-1}$.

and depicted in Fig. A7. C IV λλ1548/1550 falls in a region affected strongly by telluric absorption, while the wavelengths where Si IV λλ1393/1402 fall are heavily crowded with Ly α forest lines. We associate the detected singly-ionized metal species with the H I content to estimate the metallicity.

Fig. A8 depicts our fiducial H I model with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 20.30 \pm 0.10$ and $b = 30\,{\rm km\,s}^{-1}$. The Ly α damping wings are very prominent in LLS2241, and these strongly constrain $N_{\rm H\,I}$. The uncertainty on $N_{\rm H\,I}$ is determined by taking into account the scatter in flux in the HIRES spectrum and the continuum placement.

APPENDIX B: REASSESSMENT OF LLS0958A'S METALLICITY

Here we measure the metallicity of LLS0958A, which was previously identified as a near-pristine absorption system by Lehner et al. (2016), using the same approach as used for our other LLSs in this work.

B1 Metal-line column densities

LLS0958A was first identified by Steidel (1990) towards the $z_{\rm em} = 3.297$ quasar Q095852.3+120243 (hereafter Q0958+1202), based on absorption features at redshift $z_{\rm abs} = 3.2228$, using a Hale Telescope/Double Spectrograph (Oke & Gunn 1982) spectrum. High-resolution spectroscopy with HIRES was later obtained by Fumagalli et al. (2011a), establishing $z_{\rm abs} = 3.223\,194 \pm 0.000\,002$,

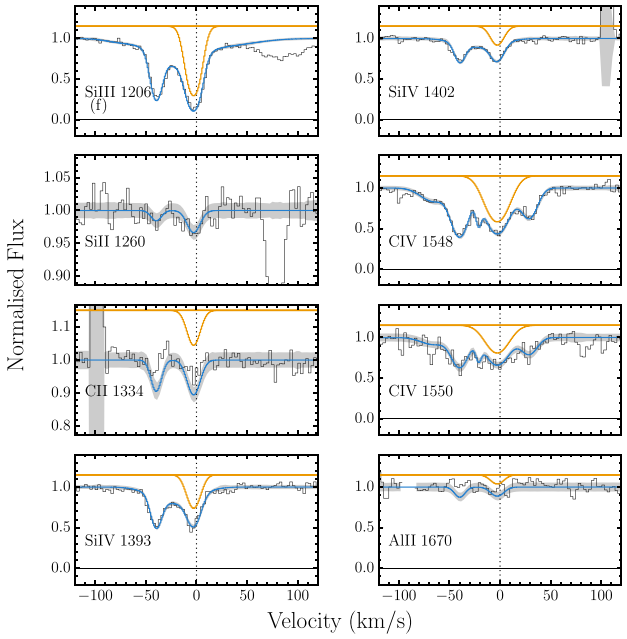

Figure B1. Similar to Fig. 2 but for Q0958+1202 and LLS0958A with the zero velocity redshift set at $z_{\rm abs} = 3.223\,194$. However, in this case the blue solid lines here represent our fit to the entire metal absorption observed, with total column densities provided in the first column of Table B1, while the orange lines show just the component at $0\,{\rm km\,s^{-1}}$ in our model, with column densities in the second column of Table B1. For Al II λ 1670, the column density used for the profile fit corresponds to a 2σ upper limit derived using the apparent optical depth method.

$\log_{10}(N_{\rm H\,I}/{\rm cm^{-2}}) = 17.36 \pm 0.05$ and $b = 20.4\,{\rm km\,s^{-1}}$. These values were estimated from the flux decrement at the Lyman limit. In the same spectrum, Lehner et al. (2016) detected Si II λ 1206 and 1260, and the high ion doublets (C IV and Si IV), and determined upper limits for C II λ 1334 and Al II λ 1670. The column densities were measured using the apparent optical depth method over a velocity range of -15 to $+30\,{\rm km\,s^{-1}}$ which included only the strongest metal absorption features; a range -15 to $+15\,{\rm km\,s^{-1}}$ was used to determine the upper limits for non-detections.

We confirmed the metal detections and non-detections of Lehner et al. (2016) in the HIRES spectrum of LLS0958A from the KODIAQ sample (O'Meara et al. 2015). Fig. B1 depicts the strongest transitions of the most abundant metal species at $z_{\rm abs} = 3.223\,194$, and we list the detections and upper limits in the first column of Table B1. Fig. B1 illustrates how our reassessment of LLS0958A's metallicity differs from Lehner et al.'s: consistent with our approach to other LLSs in this work, we included all of the detected metal absorption (blue curves); by contrast, Lehner et al. effectively included only the strongest metal component in their narrower velocity range. Given the lack of information about the hydrogen velocity structure offered by the Lyman series lines (i.e. they are too broad), we did not find evidence that the H I content could be associated with only one of the main metal-line components. Obviously, including all the metal absorption implies larger total metal column densities: the other main absorption feature, at $-40\,{\rm km\,s^{-1}}$ in Fig. B1, clearly contains a comparable column density as the main feature at $0\,{\rm km\,s^{-1}}$. This is evident in Table B1 that compares our measured column densities with those of Lehner et al. Our larger metal column densities imply a higher metallicity estimate for LLS0958A, as calculated below. Table B1 also provides the column density contained in the main component at $0\,{\rm km\,s^{-1}}$ in our model, as depicted by the orange profiles in Fig. B1. These values are much closer to the those derived from Lehner et al. using the apparent optical depth method.

Table B1. Same as Table 3 but for LLS0958A. For the fiducial model in the first column, the total column densities, from all fitted velocity components in Fig. B1, are listed. Those for the main component at $v = 0\,{\rm km\,s^{-1}}$ are provided in the second column for the best comparison with the apparent optical depth values in the third column, as derived by Lehner et al. (2016) using only the -15 to $+30\,{\rm km\,s^{-1}}$ velocity range. The last row provides the metallicity measurements inferred from the corresponding photoionization analyses.

Ion	$\log_{10}(N/{\rm cm^{-2}})$		
	Fiducial	Main component	Lehner et al.
Si II	11.40 ± 0.10	11.09 ± 0.18	11.26 ± 0.10
Si III	13.10 ± 0.12	12.77 ± 0.54	12.85 ± 0.01
Si IV	13.16 ± 0.01	12.65 ± 0.15	12.87 ± 0.01
C II	≤ 12.60	≤ 12.60	≤ 12.07
C IV	13.76 ± 0.20	13.32 ± 0.14	13.55 ± 0.01
Al II	≤ 11.40	≤ 11.40	≤ 11.11
H I		17.36 ± 0.03	
[Si/H]	-2.91 ± 0.26	-3.25 ± 0.26	-3.35 ± 0.05

Our fiducial photoionization model of LLS0958A assumes that only the low ions in Fig. B1 are associated with the bulk of the H I content. This is due to the apparent difference in velocity structure displayed by C IV compared to the low ions. We note that Lehner et al. assumed that both the low and high ions are in the same phase and are all associated with the H I content. Indeed, this may be justified because the Si IV velocity structure appears more similar to that of the low ions. Therefore, we test the effect of including the high ions in our photoionization modelling below (Section B3) and find the metallicity is fairly insensitive to this assumption.

B2 H I column density

Our fiducial H I model is depicted in Fig. B2 that shows the high-order Lyman series and limit for LLS0958A. We find that the $N_{\rm H\,I}$ value derived by Lehner et al. is entirely appropriate, even though we took a very different approach to determining the continuum in the Lyman limit region of the spectrum. To form their combined spectrum, Lehner et al. first fitted continua to the spectra from each echelle order in each exposure (as in O'Meara et al. 2015). The continuum-normalized orders were then averaged to form a normalized, combined spectrum. While this approach will work well for orders containing only a small number of narrow absorption features (e.g. redwards of the Ly α emission line), it is unlikely to work well for orders spanning the Lyman limit, like in Fig. B2. Our approach was the same as for the rest of this paper: i.e. reduce the original exposures with MAKEE and combine them with UVES_POPLER. We obtained all available exposures of Q0958+1202, and appropriate calibration exposures, from the Keck Observatory Archive (KOA).[5] UVES_POPLER scales individual order spectra to optimally match each other before averaging them and fitting a continuum to the final, combined spectrum. The result is the Lyman limit spectrum shown in Fig. B2 overlaid with the H I model of Lehner et al.: $\log_{10}(N_{\rm H\,I}/{\rm cm^{-2}}) = 17.36 \pm 0.05$ (1σ) and $b = 20.4\,{\rm km\,s^{-1}}$. Clearly, this model adequately matches our combined spectrum in this region, so we adopt this $N_{\rm H\,I}$ value without alteration.

[5] https://www2.keck.hawaii.edu/koa/public/koa.php

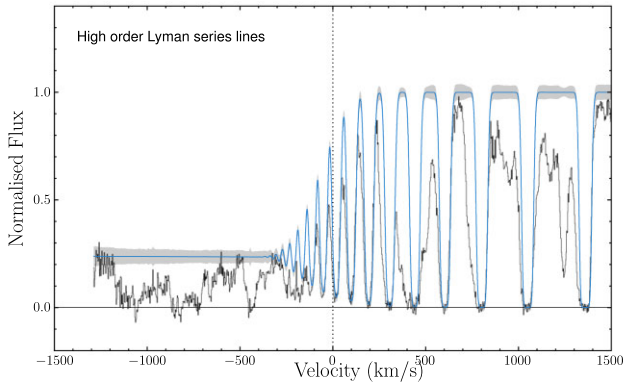

Figure B2. Continuum normalized flux of the Lyman limit and other high order Lyman series lines used to constrain $N_{\rm H\,I}$ for LLS0958A in the Q0958+1202 HIRES spectrum (black histogram). The zero velocity redshift is set at $z_{\rm abs} = 3.223\,194$ for Ly-19 (at 914.039 Å in the rest frame). The blue solid lines corresponds to Lehner et al. (2016)'s fiducial model of LLS0958A with $\log_{10}(N_{\rm H\,I}/{\rm cm}^{-2}) = 17.36 \pm 0.05, b = 20.4\,{\rm km\,s}^{-1}$ at $z_{\rm abs} = 3.223\,194$. The grey shading shows the 1σ uncertainty in the flux.

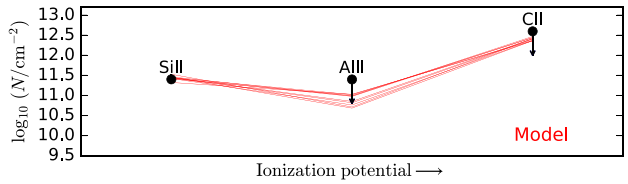

Figure B3. Same as Fig. 5 but for LLS0958A's fiducial model where only the singly ionized species listed in Table B1 are associated with the H I.

B3 Photoionization modelling results

We use the same photoionization modelling approach for LLS0958A as the other systems in this work (see Sections 3.2 and 4). As discussed above, our fiducial model assumes that the bulk of the H I is associated with the low-ions only: C II, Si II, Al II. As only Si II is regarded as a detection, we assume a non-tilted HM12 UV background (i.e. we set $\alpha_{\rm UV} = 0$) and solar values of [C/Si] and [Al/Si]. Fig. B3 shows the comparison between CLOUDY's predictions for the metal column densities and the values listed in the first column of Table B1. This simple fiducial model is clearly consistent with the data. The corresponding distributions of $Z/Z_\odot$, $n_{\rm H}$, U were derived by the MCMC sampling algorithm as usual, providing a metallicity measurement of $\log_{10}(Z/Z_\odot) = -2.91 \pm 0.26$ (95 per cent confidence). This is 0.44 dex higher than Lehner et al.'s value of -3.35 ± 0.05.

Fig. B4 compares the fiducial column density measurements with a CLOUDY model where all metal ions are assumed to trace the observed H I. In this case, there are clear mismatches between the data and model: the model cannot simultaneously fit the low and high ions (Si II/IV and C II/IV). This remains the case even when allowing a variable slope for the UV background and non-solar value for [C/Si]. In both cases, the metallicity remains ≥ -3: $\log_{10}(Z/Z_\odot) =$

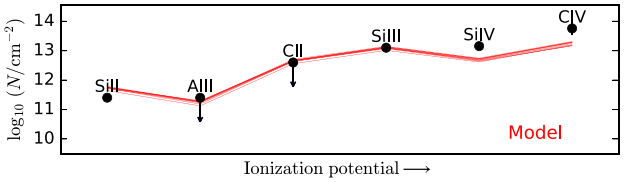

Figure B4. Same as Fig. B3 but where all metallic species listed in Table B1 are associated with the H I.

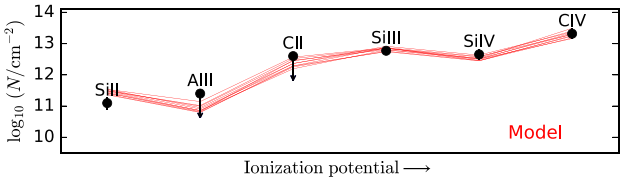

Figure B5. Same as Fig. B3 but using only the main metal component's column densities (second column of Table B1) and the same photoionization modelling assumptions as Lehner et al. (2016).

-2.89 ± 0.26 and -2.60 ± 0.26, respectively. The former, simpler case reflects most closely the assumptions made by Lehner et al. in their photoionization analysis. In this respect, it appears that the assumption about which ions trace the H I – low-ions only, or all ions – makes little difference to the metallicity estimate.

Instead, our metallicity estimate is 0.44 dex higher than Lehner et al.'s predominantly because we include all the metal absorption detected, rather than that found within a small velocity range. To illustrate that this is not adequately explained by differences in our modelling approaches, we also analysed LLS0958A replicating Lehner et al.'s assumptions as closely as possible. We used only the metal column densities from the main, $v = 0\,{\rm km\,s}^{-1}$ component in Fig. B1. These compare closely with those derived using the apparent optical depth method, in a restricted velocity range around that component, by Lehner et al., as can be seen in the second and third columns of Table B1. Again, we assume that all metal ions are associated with the H I and we use the same UV background in our CLOUDY model as Lehner et al., i.e. HM05 – a revised version of that originally published by Haardt & Madau (1996). Finally, a non-solar [C/Si] value is allowed as a free parameter. Fig. B5 compares this model with the metal column densities from the main component. There are some mismatches between the model and data; in particular, the model does not simultaneously reproduce the column densities of Si II, III, and IV. The distributions of $Z/Z_\odot$, $n_{\rm H}$, U, and [C/Si] were derived by the MCMC sampling algorithm in the same way as the other absorbers in this paper, with the most likely metallicity being $\log_{10}(Z/Z_\odot) = -3.25 \pm 0.26$. This is only 0.1 dex higher than Lehner et al.'s value of -3.35, demonstrating that differences in photoionization modelling are not responsible for the higher metallicity we find for LLS0958A.

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