



**HAL**  
open science

## **Proximate molecular quasar absorbers: Chemical enrichment and kinematics of the neutral gas**

P Noterdaeme, S Balashev, R Cuellar, J.-K Krogager, F. Combes, A de Cia, N Gupta, C Ledoux, S López, R Srianand

### ► **To cite this version:**

P Noterdaeme, S Balashev, R Cuellar, J.-K Krogager, F. Combes, et al.. Proximate molecular quasar absorbers: Chemical enrichment and kinematics of the neutral gas. *Astronomy & Astrophysics - A&A*, 2023, 673, pp.A89. <10.1051/0004-6361/202245554>. <hal-04038915>

**HAL Id: hal-04038915**

**<https://hal.science/hal-04038915v1>**

Submitted on 18 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.








L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# Proximate molecular quasar absorbers

## Chemical enrichment and kinematics of the neutral gas<sup>★</sup>

P. Noterdaeme<sup>1,2</sup>, S. Balashev<sup>3</sup>, R. Cuellar<sup>1,4</sup>, J.-K. Krogager<sup>5</sup>, F. Combes<sup>6</sup>, A. De Cia<sup>7</sup>, N. Gupta<sup>8</sup>,  
C. Ledoux<sup>9</sup>, S. López<sup>4</sup>, and R. Srianand<sup>8</sup>

<sup>1</sup> French-Chilean Laboratory for Astronomy, IRL 3386, CNRS and U. de Chile, Casilla 36-D, Santiago, Chile

<sup>2</sup> Institut d'Astrophysique de Paris, CNRS-SU, UMR 7095, 98bis bd Arago, 75014 Paris, France  
e-mail: noterdaeme@iap.fr

<sup>3</sup> Ioffe Institute, Polytekhnicheskaya 26, 194021 Saint-Petersburg, Russia

<sup>4</sup> Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile

<sup>5</sup> Centre de Recherche Astrophysique de Lyon, UMR5574, U. Lyon 1, ENS de Lyon, CNRS, 69230 Saint-Genis-Laval, France

<sup>6</sup> Observatoire de Paris, LERMA, Collège de France, CNRS, PSL University, Sorbonne University, 75014 Paris, France

<sup>7</sup> Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland

<sup>8</sup> Inter-University Centre for Astronomy and Astrophysics, Pune University Campus, Ganeshkhind, Pune 411007, India

<sup>9</sup> European Southern Observatory, Alonso de Córdova 3107, Vitacura, Casilla 19001, Santiago, Chile

Received 25 November 2022 / Accepted 25 February 2023

### ABSTRACT

Proximate molecular quasar absorbers (PH<sub>2</sub>) are an intriguing population of absorption systems that was recently uncovered through strong H<sub>2</sub> absorption at a small velocity separation from the background quasars. We performed a multi-wavelength spectroscopic follow-up of 13 such systems with VLT/X-shooter. Here, we present the observations and study the overall chemical enrichment measured from the H I, H<sub>2</sub>, and metal lines. We combined this with an investigation of the neutral gas kinematics with respect to the quasar host.

We find gas-phase metallicities in the range 2% to 40% of the solar value, that is, in the upper-half range of H I-selected proximate damped Lyman- $\alpha$  systems, but similar to what is seen in intervening H<sub>2</sub>-bearing systems. This is likely driven by similar selection effects that play against the detection of most metal- and molecule-rich systems in absorption. Differences are seen in the abundance of dust (from [Zn/Fe]) and its depletion pattern when compared to intervening systems, however, possibly indicating different dust production or destruction close to the active galactic nucleus. We also note the almost ubiquitous presence of a high-ionisation phase traced by N V in proximate systems. In spite of the hard UV field from the quasars, we found no strong overall deficit of neutral argon, at least when compared to intervening damped Lyman- $\alpha$  systems. The reason likely is that argon is mostly neutral in the H<sub>2</sub> phase, which accounts for a large fraction of the total amount of metals. We measured the quasar systemic redshifts through emission lines from both ionised gas and CO(3–2) emission, the latter being detected in all six cases for which we obtained 3 mm data from complementary NOEMA observations. For the first time, we observe a trend between the line-of-sight velocity with respect to systemic redshift and metallicity of the absorbing gas. This suggests that high-metallicity neutral and molecular gas is more likely to be located in outflows, while low-metallicity gas could be more clustered in velocity space around the quasar host, possibly with an infalling component.

**Key words.** galaxies: active – galaxies: evolution – quasars: general – quasars: absorption lines – quasars: emission lines

## 1. Introduction

The co-evolution of active galactic nuclei (AGNs) and their host galaxies is regulated by galaxy processes, including secular evolution, the inflow rate of gas (from the intergalactic medium to the interstellar medium and ultimately to the supermassive black hole (SMBH)), and merging history (see Hopkins et al. 2008). The apparent properties of AGNs may further depend on the orientation and the evolutionary stage at which the observations are taken (e.g. Urry & Padovani 1995). This constitutes an entire field of modern astrophysics, motivating large efforts from the community, on both theoretical and observational grounds. Because of the large diversity of observed properties, it is common to select objects based on a few of their many characteristics

(e.g. redshift, brightness, colour, radio-loudness, and variability) and study the relations between other marginalised properties. Each of these studies then brings a different and very valuable enlightening angle to the overall picture.

As cold gas constitutes the fuel for both star-formation and the growth of the SMBHs, many works have focused on detecting the dense molecular component of the brightest AGNs (quasars) through CO emission (e.g. Omont et al. 1996; Weiß et al. 2007; Wang et al. 2016) in order to explore the links between star formation, total molecular mass in the host, and growth and activity of the SMBH (e.g. Bischetti et al. 2017).

On the other hand, absorption studies enable detailed investigations of the gas properties along the line of sight to the central engine over a wide range of densities. For example, ionised winds powered by the accretion disc can be observed as broad absorption line systems (BALs) in roughly 15% of quasars, and their velocities range from a few thousand km s<sup>-1</sup> up to

<sup>★</sup> Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programmes 103.B-0260(A) and 105.203L.001.

$\sim 0.3 c$  (Hamann et al. 2018). Damped Lyman- $\alpha$  systems (DLAs) are absorption systems whose HI column density ( $N(\text{HI}) > 2 \times 10^{20} \text{ cm}^{-2}$ ) is high enough for them to be self-shielded from ionising radiation. They are frequently seen along quasar lines of sight (e.g. Prochaska et al. 2005; Noterdaeme et al. 2009). While most DLAs are unrelated to the quasar environment and are only intercepted by chance (called intervening DLAs), some DLAs are found at roughly the same redshift as the background quasar and are hence called proximate DLAs (PDLAs), where ‘proximate’ refers to proximity in velocity space (typically  $< a \text{ few } 10^3 \text{ km s}^{-1}$ ). Because the redshift path on which PDLAs are identified is much shorter than that available for intervening DLAs, the former are much rarer than the latter. Nonetheless, Prochaska et al. (2008) showed that the incidence of PDLAs (per unit redshift) is higher than expected from the intervening statistics, meaning that they are likely associated with a clustered environment around the quasar.

Ellison et al. (2010) performed the first systematic study of PDLAs at high resolution and concluded that at least a fraction of them exhibit different characteristics relative to the intervening population. They suggested, however, that PDLAs are probably not related to the quasar host, but preferably sample more massive galaxies in the highly clustered region around the quasar. Interestingly, Ly- $\alpha$  emission has been relatively frequently found in the core of several PDLAs (e.g. Møller & Warren 1993; Leibundgut & Robertson 1999; Ellison et al. 2002), when this is more rarely seen among the several  $10^4$  known intervening DLAs (but see Fynbo et al. 2010; Noterdaeme et al. 2012; Ranjan et al. 2018 for some example of direct detection, and Rahmani et al. 2010; Noterdaeme et al. 2014; Joshi et al. 2017 for statistical detection through stacking). Finley et al. (2013) and Fathivavsari et al. (2018) further identified a population of PDLAs with Ly- $\alpha$  emission spanning a range of strengths, up to the point at which the HI Ly- $\alpha$  absorption becomes unseen, but the presence of neutral gas is confirmed by strong low-ionisation metal lines. These emissions probably do not arise from star formation, but rather from Ly- $\alpha$  from the AGN that is not fully covered by the absorber. Because Ly- $\alpha$  photons scatter easily and up to large distances around quasars (e.g. Arrigoni Battaia et al. 2019; Borisova et al. 2016), it is not straightforward to conclude about the location, scale, and geometry of the absorbing gas. However, signatures of high excitation of the gas suggest an origin at a rather small distance from the AGN, at least for the most extreme cases.

Noterdaeme et al. (2019, hereafter Paper I) presented a new approach that focused on the presence of strong molecular hydrogen absorption directly detected at the quasar redshift in Sloan Digital Sky Survey (SDSS) spectra, without prior on the presence of HI or metal lines. Not only is H<sub>2</sub> sensitive to the physical and chemical conditions in the cold gas, but this approach also differs from studies based on CO molecular emission, for example, in the sense that no prior is made on the overall mass of dense molecular gas in the galaxy host, but rather on the interception of more diffuse cold gas by the quasar line of sight. The strong excess of proximate H<sub>2</sub> absorbers with respect to intervening statistics shows that most of these systems must be associated with the quasar environment, although it is not clear whether the absorbers arise from the quasar host or if the excess is due to strong clustering of galaxies around the quasar.

To shed light on this issue, we embarked on a follow-up study with the multi-wavelength medium-resolution spectrograph X-shooter on the Very Large Telescope (VLT). A detailed study of a first system revealed a likely origin in outflowing material intercepted by the line of sight to the

central engine (Noterdaeme et al. 2021a, Paper II). Remarkably strong and wide CO(3–2) emission as well as 3 mm continuum emission was detected in the same optically bright (moderately reddened) system through NOEMA observations (Noterdaeme et al. 2021b, Paper III). All this suggests that the nucleus is seen through a channel of lower density, but where some of the molecular content is carried out by outflows when the host still contains large amounts of dust and molecular gas with wide kinematics. This may be indicative of a post-merger system.

In this work, we focus on the overall chemical enrichment and kinematics of the gas of our follow-up sample. In Sect. 2 we present the sample, observations, data reduction, and post-processing. In Sect. 3 we present the column density and dust extinction measurements. In Sect. 4 we present the emission redshifts and absorption line kinematics. We discuss our findings and compare our results with other PDLAs as well as strong intervening H<sub>2</sub> systems in Sect. 5. We conclude in Sect. 6.

## 2. Sample, observations, and data reduction

### 2.1. Sample and observations

The sample studied here is entirely drawn from Paper I. In the statistical sample of 50 quasars with high-confidence proximate molecular absorbers, 22 have  $\delta < 20^\circ$ , that is, they can be reached from the southern hemisphere. We obtained VLT/X-shooter data for 13 of them: 11 based on their observability during the allocated periods of this specific programme (P103 and P105, but because of delays related to the observatory shutdown during the Covid-19 pandemic, the observations were actually completed in P109), plus the quasars SDSS J013644.02+044039.1 and SDSS J085859.67+174925.1 for which we used data collected in P94 as part of another programme (Balashev et al. 2019). We note that SDSS J123602.11+001024.5 has already been observed by Balashev et al. (2019). This quasar was then re-observed under better conditions and from various position angles through the availability of the X-shooter atmospheric dispersion corrector. We use the new data here. We finally note that because the quasars in this sample were selected from the statistical sample based upon their observability alone (and not based on any other property), they are likely a fairly good representation of the latter.

X-shooter (Vernet et al. 2011) simultaneously covers the whole optical to near-IR (NIR) wavelength range from 0.3 to 2.5  $\mu\text{m}$ , splitting the incoming light beam into three spectrographs (arms): UVB, ranging from 0.3 to 0.6  $\mu\text{m}$ ; VIS, ranging from 0.6 to 1.0  $\mu\text{m}$ ; and NIR, ranging from 1.0 to 2.5  $\mu\text{m}$ . All observing blocks (OBs) consisted of 2980 s, 3010 s, and  $5 \times 600$  s exposures for the UVB, VIS, and NIR arms obtained in stare mode. We used 1.0 and 0.9'' wide slits in the UVB and VIS, with slow read-out mode and  $1 \times 2$  pixel binning. The NIR observations were performed with 1.2'' slits for the two quasars observed in P103 (SDSS J001514.82+184212.3 and SDSS J232506.62+153929.3). The remaining observations were performed with a narrower NIR slit (0.9 arcsec) and a K-band blocking filter to avoid stray light in the instrument, effectively cutting out all wavelengths longer than 2.08  $\mu\text{m}$  for all remaining quasars but SDSS J013644.02+044039.1 and SDSS J085859.67+174925.1 (in P94), for which the K-band filter was not yet available. We executed two to five OBs for each quasar at different position angles. Most of the observations were obtained under good seeing, somewhat sharper

**Table 1.** Quasar sample and log of X-shooter observations

Quasar (SDSS) <sup>(a)</sup>	$z$	Observing dates	Seeing <sup>(b)</sup> (arcsec)	Airmass	Ref.
J001514.82+184212.3	2.63	29/07, 31/08, 31/08, 28/09 2019	1.67, 0.75, 0.65, 0.53	1.48, 1.39, 1.41, 1.42	(1)
J001930.55–013708.4	2.53	06/11, 07/11, 08/11 2021	0.91, 0.72, 0.71	1.30, 1.26, 1.16	(2)
J005917.64+112407.7	3.03	09/11, 09/11, 09/11 2021	0.77, 0.80, 1.14	1.24, 1.30, 1.47	(2)
J012555.11–012925.0	2.66	21/07, 01/09, 01/09, 11/10, 25/11 2021	1.14, 1.00, 0.98, 1.22, 0.92	1.11, 1.10, 1.17, 1.10, 1.17	(3)
J013644.02+044039.1	2.78	27/11, 28/11 2014	1.15, 1.15	1.16, 1.26	(4)
J085859.67+174925.1	2.65	24/12 2014, 15/02 2015	0.73, 1.38	1.36, 1.37	(4)
J123602.11+001024.5	3.03	07/02, 07/02, 07/02 2022	0.73, 0.83, 0.87	1.20, 1.12, 1.11	(2)
J124829.51+063925.6	2.53	08/02, 07/03, 07/03 2022	1.21, 0.61, 0.65	1.19, 1.18, 1.19	(2)
J125917.31+030922.5	3.23	06/02, 06/02 2022	0.79, 0.93	1.33, 1.43	(2)
J133111.41+020609.0	2.92	08/03, 29/05, 29/05 2022	0.80, 0.84, 0.98	1.25, 1.17, 1.30	(2)
J135808.94+141053.2	2.89	25/02, 01/03, 07/03 2022	1.01, 0.79, 0.68	1.29, 1.29, 1.30	(2)
J222807.36–022117.1	2.77	31/08, 15/09, 12/10, 26/10 2021	0.87, 0.71, 0.84, 0.85	1.10, 1.12, 1.09, 1.09	(2)
J232506.62+153929.3	2.62	29/08, 31/08, 30/09 2019	0.84, 0.77, 0.67	1.33, 1.33, 1.34	(2)

**Notes.** <sup>(a)</sup>The actual object designation following the rules from the International Astronomy Union, includes the SDSS preceding the J2000 coordinates sequence provided in this column. <sup>(b)</sup>Delivered seeing corrected for airmass, as measured at the telescope by the Shack-Hartmann wavefront sensor (fits keyword HIERARCH ESO TEL IA FWHM).

**References.** (1) Noterdaeme et al. (2021a); (2) this work; (3) Balashev et al. (in prep); (4) Balashev et al. (2019).

than the slit widths at the corresponding wavelengths, resulting in higher than nominal spectral resolution. The observing log is given in Table 1. For convenience, we refer to these objects using a short version of their name in the following: SDSS Jhhmmss.ss+ddmmss.s is abbreviated as Jhhmm+ddmm (e.g. J0015+1842 instead of SDSS J001514.82+184212.3).

## 2.2. Data reduction and telluric corrections

All X-shooter spectra were processed using the official esorex pipeline version 3.5.3 (Modigliani et al. 2010) for stare-mode reduction, maintaining the default parameters and including the correction for instrument flexure (recipe xsh\_flexcomp). The raw images were pre-processed using the Python package astroscrappy (a reimplementation of the L.A.Cosmic algorithm by van Dokkum 2001) to interpolate over pixels contaminated by cosmic-ray hits using the neighbouring pixels. Each arm of each OB was reduced separately using the standard star observed closest in time in order to flux calibrate the given spectrum. The 1D spectra were then extracted using optimal extraction (Horne 1986).

Next, we used the ESO tool molecfit v.4.2.2a.1 (Smette et al. 2015) to correct each individual 1D VIS and NIR spectrum for the absorption lines produced by the Earth’s atmosphere. In the VIS, the main telluric absorption features are due to O<sub>2</sub> and H<sub>2</sub>O, while the NIR is strongly affected by H<sub>2</sub>O and CO<sub>2</sub> bands. Following the recommendation from the molecfit manual, the correction was made by choosing a few spectral regions with clear but not strongly saturated atmospheric lines for each quasar, typically, 0.68–0.69, 0.71–0.73, 0.76–0.77, 0.81–0.82, and 0.91–0.91  $\mu\text{m}$  in the VIS. In the NIR, we had to adjust the selected spectral regions for each quasar more strongly, depending on the strength of absorption bands and on the location of quasar emission lines. During this process, we also masked deviant pixels (e.g. cosmic remnants) as well as astrophysical lines.

The best-fit model was used to calculate the overall atmospheric transmission and correct the individual VIS and NIR 1D exposures accordingly. The telluric-corrected spectra were then shifted to vacuum-barycentric wavelengths, rebinned to a common wavelength grid, and combined using an inverse-variance weighting. The same vacuum-barycentric correction and combi-

nation of 1D spectra was applied to UV data, but without telluric correction.

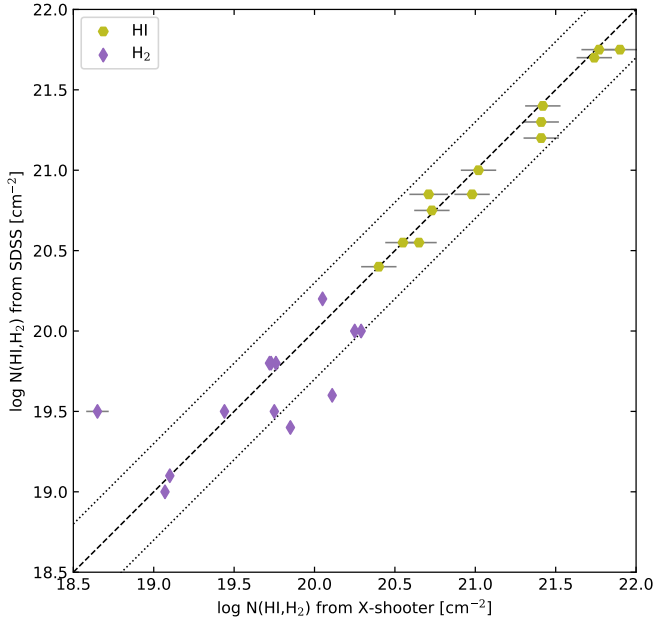
We note that molecfit also provides the actual spectral resolution, which is a free parameter of the transmission model. As expected from a seeing sharper than the slit widths for most of the observations, the resolving power is typically 25% higher than the nominal values in the documentation ( $R_{\text{nom}} = 5400$  (UV) and 8900 (VIS)). The average spectral resolution obtained from individual VIS spectra is then adopted when fitting the lines. In the UV, the spectral resolution is first estimated by scaling the nominal UV value by the observed-to-nominal VIS resolution ratio. We verified that the values broadly agree with those expected for a slit width that would correspond to the observed seeing. It was then adjusted during the fitting process if necessary. While the exact values have little influence on the metal column densities, we provide the adopted values of the resolutions for the combined spectra along with the best-fit model parameters in Appendix B to facilitate reproducibility.

## 2.3. NOEMA observations

While the present work focuses on X-shooter data, we also obtained observations with the NOEMA and the PolyFIX correlator in the 3 mm band between May 2020 and May 2021 for six out of the 13 objects presented here. The central frequency was chosen in order to cover the expected position of the CO(3–2) emission line, which we detected in all six cases. A description of data handling and analysis of NOEMA observations (including other quasars from the same parent sample that were not observed with X-shooter) will be presented in a future paper, but we already use the redshifts from the detected CO(3–2) emission lines in the discussion (Sect. 5). The Gaussian fits to the CO(3–2) emission lines are presented in Appendix F.

## 3. Column-density measurements

We analysed the absorption lines through standard, simultaneous, multi-component Voigt-profile fitting using vpf it v.12.3 (Carswell & Webb 2014) to measure the column densities, Doppler parameters, and redshifts. The quasar continuum was



**Fig. 1.** Comparison of H<sub>2</sub> (purple diamonds) and HI (green hexagons) column densities obtained using X-shooter and SDSS spectra. The dashed line depicts the one-to-one relation, and the dotted lines show  $\pm 0.3$  dex around it.

evaluated using spline functions, with nodes constrained by the unabsorbed regions and adjusted over strong lines in an iterative way when necessary. In order to best take the blending between the metal, H<sub>2</sub>, and HI lines into account, the final absorption models were obtained by fitting all lines simultaneously. However, for the sake of presentation, we describe the analysis separately for HI, H<sub>2</sub>, and metals. The results from the fits are presented in the appendix, and the main derived properties are summarised in Table 2.

### 3.1. Total HI and H<sub>2</sub> column densities

The total HI column densities were mostly constrained from the saturated core of the Ly- $\alpha$  line, which is little sensitive to the exact placement of the continuum on top of the Ly- $\alpha$  emission line of the quasars. Based on the wide wavelength coverage of X-shooter and the relatively high redshifts of quasars (implied by the selection upon H<sub>2</sub> in SDSS), all lines from the HI Lyman series are covered by our spectra, allowing us to constrain the HI column density further. While the formal statistical errors from fitting the lines are very small (typically, 0.01–0.02 dex), we note that the continuum placement as well as the choice of fitting regions are necessarily somewhat subjective and might introduce some systematic errors. When we allow for continuum variation using Chebychev polynomials, the uncertainties are smaller than 0.1 dex for similar DLA observations with X-shooter (e.g. Balashev et al. 2019; Ranjan et al. 2020). We hence added a conservative 0.1 dex uncertainty to the error budget.

In all the systems, we confirm strong H<sub>2</sub> absorption lines from the Lyman and Werner bands. These are typically detected from rotational level  $J = 0$  to  $J \sim 6$ –7 in our X-shooter spectra. The high column densities imply damped lines in at least the first few rotational levels and hence accurate column densities because lines in this regime do not depend on the Doppler parameter  $b$ . We also took into account and fit the detected high rotational levels, assuming the same Doppler parameter and red-

shift for all H<sub>2</sub> lines. However, the derived column densities for the high- $J$  levels are likely to be more sensitive to the assumptions on the Doppler parameters. The  $b$  values for H<sub>2</sub> lines have been found to depend on the rotational level in several cases (see e.g. Noterdaeme et al. 2007; Balashev et al. 2009). A more detailed study and exploration of systematics is hence needed in order to use these lines to constrain the excitation of the gas and the prevailing physical conditions. This is postponed to a future work, while we focus here on the total column densities, which are dominated by the low rotational levels. The fits to the H<sub>2</sub> absorption bands are presented in Appendix A.

In Fig. 1 we compare the total H<sub>2</sub> and HI column densities obtained with X-shooter with those derived from the SDSS. In the case of HI the agreement is very good and always better than 0.3 dex. For H<sub>2</sub>, for which the involved columns are much lower, the agreement remains remarkable (except for the lowest column density system, where the difference is 0.85 dex), in particular because the SDSS values were obtained by visually scaling and matching a template (i.e. assuming single component with fixed Doppler parameter and excitation temperature) to the low-resolution, low S/N SDSS spectra (see Paper I). This agreement is similar to what was found by Balashev et al. (2019) for intervening H<sub>2</sub> systems.

### 3.2. Low-ionisation metal species

In this section, we focus on the main low-ionisation metal lines that provide reliable column densities, that is, lines that have multiplets, are not strongly saturated, and do not lie in a region that is strongly affected by telluric lines. For example, we did not attempt to fit the only singly-ionised and very strong carbon line C II  $\lambda$ 1334. We fitted the low-ionisation metal lines assuming a common velocity structure with turbulent broadening (i.e. tied redshifts and Doppler parameters). We first identified the velocity components using the strong Fe II and Si II lines in the visual, which also has higher spectral resolution than UVB, and obtained a first fit, including weaker lines. We then added other metal species and adjusted the fitting regions and the number of components whenever necessary to obtain the final fit.

During this process, we also simultaneously fitted some other intervening systems that possess lines blended with that of the proximate systems, or masked the corresponding blended regions. We also note that Zn II  $\lambda$ 2026 is systematically blended with Mg I  $\lambda$ 2026, which we took into account assuming the same components, although Mg I is not discussed in this work. Whenever possible, we also used Mg I  $\lambda$ 1827 and/or Mg I  $\lambda$ 2852 to further constrain the (small) contribution of Mg I to the 2026 profile. We also took into account partial blending of Zn II with Cr II at 2062 Å. Finally, the analysis of neutral carbon (including fine-structure lines), neutral chlorine, and excited fine-structure lines (from C II, Si II and O I) is postponed to a future work because we focus here on chemical enrichment, to which their contribution is negligible. We already note the clear detection of Si II\* towards J0015+1842, J0019–0137, J0059+1124, J0125–0129, J1259+0309, J1358+1410, J2228–0221, and J2325+1539. The fit to the different metal lines is presented in Appendix B.

We note that the metal column densities obtained at medium spectral resolution should always be considered with some caution because the modelling of the velocity profile mostly implied fewer but broader ( $b$  frequently as high as 20 km s<sup>–1</sup>) components than typically observed in high-resolution spectra, where  $b < 10$  km s<sup>–1</sup>. Therefore, in spite of our best efforts to model the profiles and use lines with a range of strengths, unresolved line saturation may still affect the column density estimates for

**Table 2.** Main properties of the proximate absorption systems

Quasar	$\log N(\text{H I})$ ( $\text{cm}^{-2}$ )	$\log N(\text{H}_2)$ ( $\text{cm}^{-2}$ )	[Zn/H]	[S/H]	[Si/H]	[Fe/H]	$A_V$ <sup>(a)</sup> (mag)	$\Delta v_{90}$ <sup>(b)</sup> ( $\text{km s}^{-1}$ )
J0015+1842	20.71±0.12	19.73 ± 0.02	-0.41±0.13	-0.54±0.14	-0.78±0.12	-1.79±0.12	0.40±0.10	450
J0019-0137	21.02±0.11	20.29 ± 0.01	-0.90±0.12	-1.00±0.11	-0.96±0.12	-2.03±0.11	0.08±0.01	190
J0059+1124	21.77±0.11	19.07 ± 0.02	-1.40±0.15	-1.37±0.16	-1.03±0.34	-2.02±0.13	0.04±0.02	170
J0125-0129	21.90±0.15	20.05 ± 0.01	-0.96±0.15	-0.96±0.16	-0.92±0.16	-1.66±0.15	0.19±0.01	430
J0136+0440	20.73±0.11	18.65 ± 0.07	–	-0.59±0.11	-0.80±0.12	-1.24±0.11	0.04±0.01 <sup>(c)</sup>	190
J0858+1749	20.40±0.11	19.72 ± 0.02	-0.52±0.12	-0.57±0.12	-0.96±0.11	-2.26±0.11	0.12±0.01 <sup>(c)</sup>	190
J1236+0010	20.65±0.11	19.75 ± 0.01	-0.55±0.13	-0.48±0.11	-0.71±0.11	-1.70±0.11	0.07±0.02	400
J1248+0639	20.55±0.11	19.76 ± 0.01	-0.55±0.16	-0.48±0.19	-0.89±0.12	-1.76±0.11	-0.05±0.02	240
J1259+0309	21.42±0.11	19.10 ± 0.01	-1.57±0.12	-1.39±0.11	-1.33±0.11	-1.85±0.11	-0.05±0.01	90
J1331+0206	21.41±0.11	20.11 ± 0.01	-1.56±0.14	-1.67±0.12	-1.61±0.12	-1.94±0.12	0.04±0.02	80
J1358+1410	21.19±0.11	19.85 ± 0.02	-0.96±0.14	-0.96±0.11	-0.87±0.11	-1.76±0.11	0.06±0.03	170
J2228-0221	20.98±0.11	19.44 ± 0.02	-0.52±0.11	-0.54±0.11	-1.00±0.11	-1.94±0.11	0.24±0.03	110
J2325+1539	21.74±0.11	20.25 ± 0.01	-1.15±0.11	-1.03±0.11	-1.14±0.11	-1.70±0.11	0.16±0.03	240

**Notes.** <sup>(a)</sup>Statistical error only from the dispersion between measurements. The systematic error on  $A_V$  is about 0.07 mag (see text). <sup>(b)</sup>Measured from the low-ion metal lines (see text). <sup>(c)</sup>Balashhev et al. (2019) quoted  $A_V = 0.16$  and  $0.29$  for J0136+0440 and J0858+1749, respectively. However, these values were obtained with a varying intrinsic power law, as is usually done in attempt to separate the intrinsic slope from that due to an intervening absorber. Here, as for other quasars in our sample, we consistently re-measured  $A_V$  using a fixed power law to quantify the amount of dust in the overall systems.

some of the species more than the formal uncertainty suggests. Metallicities were derived using weak Zn II lines in all but one case, where the total column densities are less sensitive to the exact velocity structure decomposition.

### 3.3. Dust

We quantified the amount of dust in the absorbers by measuring the extinction of the quasar, following a template-matching procedure as was used in many works (e.g. Srianand et al. 2008; Ma et al. 2015; Krogager et al. 2016; Zhang et al. 2015). In order to measure the dust reddening using the full wavelength range provided by X-shooter, we combined the spectra from all three arms to obtain a joint spectrum for each quasar. The combination was performed using an inverse-variance weighting in the overlapping regions. During the combination, all three arms were interpolated onto a common wavelength grid using a fixed sampling of  $0.2 \text{ \AA}$ . The joint spectra were finally corrected for Galactic extinction using the map by Schlafly & Finkbeiner (2011).

The individual joint spectra were then fitted using the quasar template by Selsing et al. (2016) assuming a fixed extinction law at the redshift of the quasar. This implicitly assumes that no other intervening absorber causes the reddening. This assumption is reasonable because the proximate systems likely contain most of the dust present along the line of sight, as indicated by  $\text{H}_2$  and the low probability of intercepting another rare dust-rich absorber by chance. We did not find signs for additional metal-rich intervening absorbers in any of the quasar spectra. Since we observe no evidence of  $2175 \text{ \AA}$  extinction features, we used the extinction law inferred for the Small Magellanic Cloud (Gordon et al. 2003). To avoid biasing the fit, we excluded regions of the spectra around the strong, broad emission lines. We furthermore exclude regions that are affected by strong telluric absorption.

For each target, we derived the dust extinction as the average of the  $A_V$  values for the individual observations. The statistical uncertainty was determined as the standard deviation of the individual measurements. This uncertainty includes any uncertainty

on the flux calibration of the different spectra. Lastly, from previous work, we assigned a systematic uncertainty of 0.07 mag (Noterdaeme et al. 2017) due to any possible intrinsic mismatch of the template and the given target (i.e. accounting for scatter in the intrinsic spectral energy distribution of quasars). The results are summarised in the penultimate column of Table 2, and the fits shown in Appendix C.

## 4. Absorption line kinematics and systemic redshifts

### 4.1. Absorption line kinematics

We quantified the kinematics of the neutral absorbing gas using the velocity width as defined by Prochaska & Wolfe (1997):  $\Delta v_{90} = c[\lambda_{95} - \lambda_5]/\lambda_{50}$ , where  $\lambda_5$ ,  $\lambda_{50}$ , and  $\lambda_{95}$  are the wavelengths corresponding to the 5, 50 and 95 percentiles of the apparent optical depth distribution. For each system, we followed the criteria from Ledoux et al. (2006) and selected a few low-ionisation absorption lines that are not too strongly saturated (it would otherwise be impossible to measure the apparent optical depth) nor too weak (in which case, part of the gas would remain untraced). We then checked for consistency between results. The  $\Delta v_{90}$  for each PDLA is shown in Appendix D and provided in the last column of Table 2.

We note that while the smearing of the lines at medium spectral resolution might result in a small overestimation of  $\Delta v_{90}$  in principle (Prochaska et al. 2008; Arabalmani et al. 2015), this is mostly true for low-velocity widths and only becomes an issue when statistics of samples are compared that were taken at different spectral resolutions. Moreover, the spectral resolution and  $\Delta v$  are high enough for the corrections to be negligible here, although, strictly speaking, the lowest values (for J1331+0206 and J1259+0309) might be considered as upper limits.

### 4.2. Quasar redshifts

A precise measurement of the quasar redshift is then required to study the motion of the absorbing gas with respect to the

**Table 3.** Measurements of the quasar emission line redshift.

Quasar	Redshift									
	[O II]	[O III]	Mg II	C III]		He II		Template	$z'_{\text{sys}}$	CO(3–2)
				Meas.	Corr.	Meas.	Corr.			
J0015+1842	2.6310	2.6309	–	–	–	–	–	–	2.6310	2.6285
J0019–0137	–	2.5217	2.5276	2.5185	2.5212	–	–	2.5253	2.5217	–
J0059+1124	–	3.0369	–	3.0338	3.0369	3.0353	3.0406	3.0395	3.0369	3.0378
J0125–0129	–	2.6534	–	2.6466	2.6494	–	–	2.6540	2.6534	2.6618
J0136+0440	–	–	2.7905	2.7820	2.7849	2.7800	2.7814	2.7846	2.7849	–
J0858+1749	–	–	2.6643	2.6450	2.6478	–	–	2.6589	–	–
J1236+0010	–	–	3.0284	3.0208	3.0239	–	–	3.0318	3.0284	–
J1248+0639	–	–	2.5273	2.5206	2.5233	–	–	2.5221	2.5273	–
J1259+0309	–	–	3.2365	3.2277	3.2309	–	–	3.2369	3.2365	3.2405
J1331+0206	2.9179	2.9147	–	2.9099	2.9129	2.9100	2.9155	2.9153	2.9147	2.9168
J1358+1410	2.8992	–	2.8926	2.8971	2.9001	–	–	2.8996	2.8926	–
J2228–0221.	–	–	–	2.7676	2.7705	–	–	2.7763	2.7705	2.7780
J2325+1539	–	–	–	2.6044	2.6072	–	–	2.6051	2.6072	–

central engine along the line of sight, and more generally, to investigate dynamical processes in the circum-nuclear region and/or in the host galaxy. Velocity shifts between the different emission lines are seen in quasar spectra, reflecting the diverse kinematics of the emitting regions (e.g. Gaskell 1982; Tytler & Fan 1992; Vanden Berk et al. 2001; Shen et al. 2016). Overall, it was found that high-ionisation lines tend to be strongly blueshifted with respect to the systemic redshift ( $z_{\text{sys}}$ ), while low-ionisation lines provide a better means to measure  $z_{\text{sys}}$ . The most reliable of these lines are the oxygen lines.

We hence first focused on the [O II] and [O III] lines, which are redshifted in the NIR. Unfortunately, telluric absorption bands frequently affect these lines, and in spite of our best efforts to correct for the atmospheric absorption, it is not always possible to recover the unabsorbed emission line properly. When the lines were reasonably well visible, we modelled the region covering H $\beta$  and [O III] using a simple linear continuum, to which we added Gaussian lines (with more than one component whenever necessary), but we tied the parameters of the two lines of the [O III] doublet together: the same velocity, the same width, and a flux ratio of 3. [O II] was fitted independently. In order to remove spikes, skyline, or cosmic-ray residuals as well as absorption lines, we iteratively fitted the data rejecting deviant pixels (at more than  $3\sigma$ ) at each iteration until convergence (i.e. until no more pixels were rejected). We then visually compared the derived model with the rebinned data obtained from median averaging in bins of typically 25 pixels. Unfortunately, for two [O II] lines, the sky line residuals are close to the peak, and the [O II] redshifts may therefore not be as reliable as expected. This effect is mitigated for [O III], which is generally stronger and has two lines.

In principle, Mg II is the following choice at shorter wavelengths (Shen et al. 2016). We note that owing to our selection on the presence of a proximate absorber, the line is affected by strong Mg II absorption, which we masked during the fitting process. We then also considered the strong C III] and the weaker He II lines at 1908 and 1640 Å rest frame, respectively, which are free from strong absorption. We followed the prescription from Shen et al. (2016) and measured the line redshifts from the peak of the emission, obtained from a model that reproduced the data. Multiple Gaussians provided convenient models through their flexibility and the relatively few free parameters.

The redshifts obtained this way are given in Table 3 using the same rest-frame wavelengths as in Shen et al. (2016). In the case of [O II], [O III], and Mg II, the authors found small systematic shifts ( $\sim 8$ ,  $-45$ , and  $-57$  km s $^{-1}$ , respectively) compared to the systemic value, which they determined from Ca II absorption. This is less than our measurement uncertainties, therefore we did not attempt any correction here. In the case of the C III] complex (which includes Si III] and Al III], they found a luminosity-independent systematic shift of  $-229$  km s $^{-1}$ , and in the case of He II the systematic shift depends on luminosity. Here we also followed the prescription from Shen et al. (2016) and applied the correction accordingly (see their Eq. (1) and best-fit parameters in their Table 2). We note that these empirical shifts include any possible physical shifts of the emitting regions, plus the fact that several species contribute to the complex. We also provide the corrected redshifts in Table 3 in an attempt to remove the average trends seen in quasars. As the average shifts from Shen et al. (2016) were computed based on low-redshift quasars observed in the SDSS, we finally performed a last test by matching the quasar composite from Selsing et al. (2016) with our data in a region of the visual spectrum covering the range 1620–2450 Å rest-frame, that is, covering the features from He II to [Ne IV]. This range is slightly reduced in a few cases to avoid poor-quality regions, generally at long wavelengths. Absorption lines, bad pixels, and sky-line residuals were masked during the fitting process. The only four free parameters are the scaling value, an additional continuum (power-law+constant), and the redshift.

The different measurements (see Appendix E) then allowed us to cross-check the emission redshift and determine the best estimate based on ionised lines (which we call  $z'_{\text{sys}}$ ), as given in the penultimate column of Table 3. In general, we prioritised the lines following the prescription from Shen et al. (2016), although in some cases, we preferred a second-choice line with high S/N over a line with more uncertain measurement. This was the case for J0136+0440 and J0858+1749, where we preferred the corrected C III] over Mg II, and J1331+0206 where we preferred Mg II over [O II]. At the end, He II and template measurement were not used, but they are provided for comparison and to give more confidence to the measurement. Finally, in the last column of this table, we indicate the redshift of the CO(3–2) emission line derived from fitting the 3 mm NOEMA data with single Gaussian profiles. This is expected to be an even better

indicator of the quasar host systemic redshift because it traces the bulk of the molecular gas (Wang et al. 2016). The corresponding fits are shown in Appendix F. Interestingly, the CO-based redshifts appear to be systematically higher by about  $250 \text{ km s}^{-1}$  (average based on the six systems for which CO(3–2) measurements are available) with respect to the systemic redshifts based on rest-frame UV/optical transitions. This is discussed in Sect. 5.3.

## 5. Results and discussions

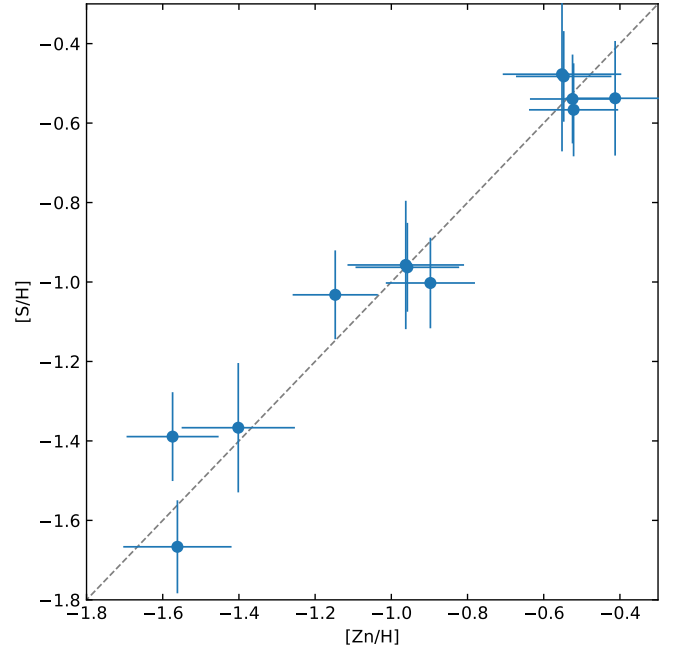
### 5.1. Chemical enrichment

In the following, we discuss the chemical properties of our sample based on the measured column densities, and we further compare them with other samples of absorption systems: (1) Proximate H I -selected DLAs observed at high spectral resolution by Ellison et al. (2010); (2) a compilation of intervening DLAs (De Cia et al. 2016) with robust abundance measurements, and (3) a compilation of H<sub>2</sub> systems from Balashev et al. (2019) that was confirmed from follow-up observations. This compilation contains H<sub>2</sub>-bearing systems identified with a number of different selections: in regular DLA systems (Noterdaeme et al. 2008; Balashev et al. 2014, 2017), extreme H I absorbers (Noterdaeme et al. 2015; Ranjan et al. 2020), and C I-selected systems (Ledoux et al. 2015; Noterdaeme et al. 2018). A discussion of the selection function is presented in Balashev et al. (2019), and therefore, we here only consider the sample as a whole, but keep in mind that its distribution in parameter space is a sum of several selections. We removed some systems that have  $z_{\text{abs}} \approx z_{\text{QSO}}$  and might be considered as proximate from the intervening samples. In order to avoid possible evolutionary effects in our comparison, we also restricted these samples to  $z > 2$ . Finally, we removed known H<sub>2</sub> systems from the sample by De Cia et al. (2016) to avoid duplicates with Balashev et al. (2019).

We used the standard notation of the gas-phase abundances of the different species X expressed relative to the solar values as  $[X/H] = \log(N(X)/N(\text{H}_{\text{tot}})) - \log(X/H)_{\odot}$ , and where we took the solar abundances from Asplund et al. (2021) following the recommendations of Lodders (2003) on whether to take photospheric, meteoritic, or average values (see e.g. Table 1 of Konstantopoulou et al. 2022). We assumed, as is common, that ionisation corrections are negligible for DLAs because of the self-shielding from  $>13.6 \text{ eV}$  photons, that is, that the considered species are in the main ionisation stage in the neutral gas (e.g.  $N(\text{Zn II}) = N(\text{Zn})$ ). We included H<sub>2</sub> in the total hydrogen column density, however, that is,  $N(\text{H}) = N(\text{H I}) + 2N(\text{H}_2)$ . We recall that we considered the gas-phase abundances only when quoting abundances as  $[X/H]$ . In other words, we did not attempt any correction for dust depletion. We further investigate this effect by considering volatile and refractory elements in the following sections. A summary of the total abundances of the main species of interest is provided in Table 2.

#### 5.1.1. Overall metallicity from volatile species: Zinc and sulphur

Zinc and sulphur are generally considered as the best reference elements for the intrinsic metallicity (i.e.  $\log Z/Z_{\odot} \approx [\text{Zn}/\text{H}] \approx [\text{S}/\text{H}]$ ) because they are volatile and hence little depleted onto dust grains (see e.g. Welty et al. 1999; De Cia et al. 2016, among many other works). We confirm that for all PDLAs where both zinc and sulphur abundances can be measured, the agreement between both values is good, with an average measured sulphur-



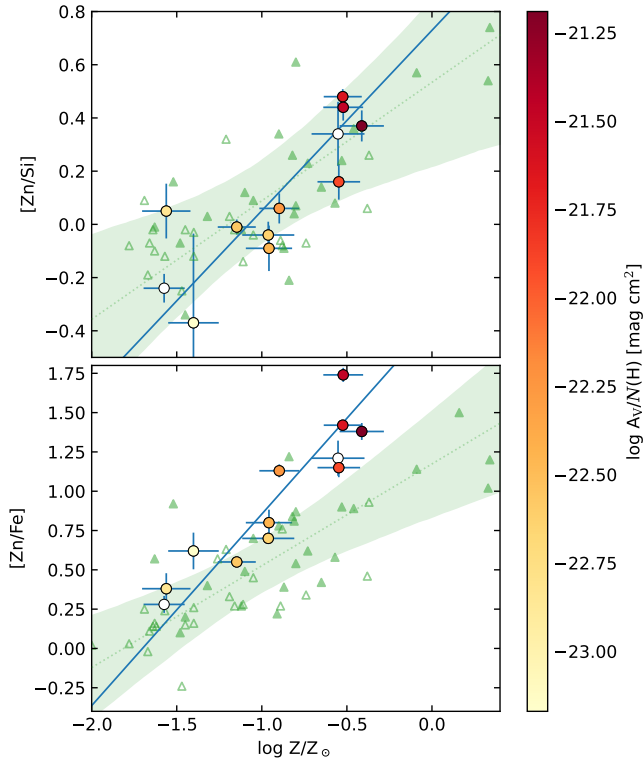
**Fig. 2.** Observed abundance of sulphur against that of zinc in proximate H<sub>2</sub> systems.

to-zinc abundance ratio  $[\text{S}/\text{Zn}] \approx 0$  and dispersion 0.1 dex (see Fig. 2). We note that, in principle, this does not exclude depletion of both species to similar amounts. However, we would expect any depletion to start differing at higher  $[\text{Zn}/\text{H}]$ , which is not apparent. We do not see any trend of  $[\text{S}/\text{Zn}]$  with  $[\text{Zn}/\text{Fe}]$  either. In any case, for a comparison with other measurements in the literature, we continue to assume  $\log Z/Z_{\odot} \equiv [\text{Zn}/\text{H}]$  in the following. For the only system without a Zn measurement (J0136+0440, where the corresponding lines are too heavily blended with atmospheric lines), we hence used sulphur instead. Overall, we found that the metallicities in proximate H<sub>2</sub> systems span a range from  $\log Z/Z_{\odot} \approx -1.6$  to  $-0.4$ . This corresponds to the upper half range of H I-selected proximate DLAs, as investigated by Ellison et al. (2010), but is similar to the metallicity of intervening H<sub>2</sub>-bearing DLA systems (e.g. Petitjean et al. 2006; Balashev et al. 2019).

#### 5.1.2. Dust depletion from relative abundances

Silicon and iron are refractory elements known to be respectively mildly and strongly depleted onto dust grains, and their gas-phase abundances are typically lower than those of zinc and sulphur. In Fig. 3 we show the abundances ratios  $[\text{Zn}/\text{Fe}]$  and  $[\text{Zn}/\text{Si}]$  as a function of metallicity. These ratios clearly increase with increasing metallicity, as has been observed in a number of works on intervening DLAs (e.g. Ledoux et al. 2003; Noterdaeme et al. 2008; De Cia et al. 2016; Balashev et al. 2019). This can be attributed to the increasing abundance of dust with increasing metallicity.

Iron appears to be significantly depleted even down to the lowest metallicities, while  $[\text{Si}/\text{Zn}]$  becomes positive only for metallicities higher than one-tenth solar. Interestingly,  $[\text{Zn}/\text{Fe}]$  in proximate H<sub>2</sub> systems is clearly higher for a given metallicity than the values seen in regular intervening DLAs, suggesting a higher depletion of iron in proximate systems. This remains true when proximate H<sub>2</sub> absorbers are compared only to intervening



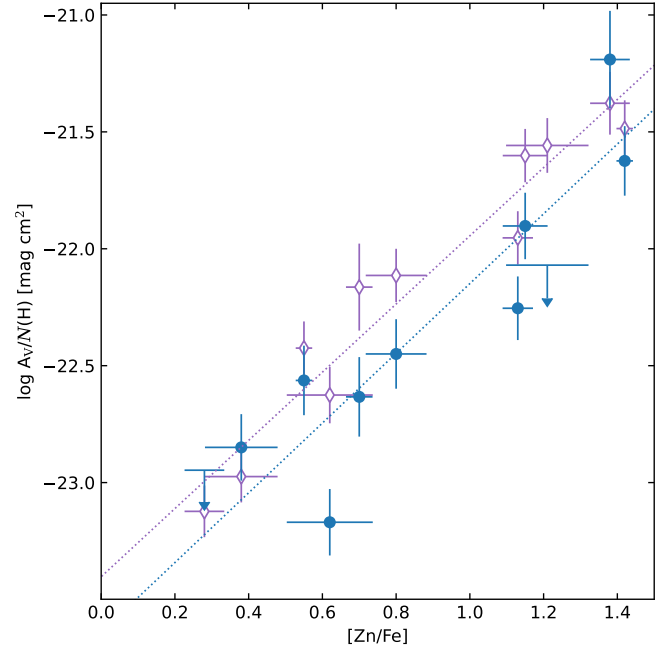
**Fig. 3.** Zinc-to-iron (bottom) and zinc-to-silicon (top) ratio for our sample of proximate  $H_2$  systems (circles with error bars). The colour scale provides the value of  $A_V/N(H)$ , which is a proxy of the dust-to-gas ratio. For two systems, the derived  $A_V$  values are negative (hence unphysical) and the corresponding points are empty. Green triangles represent measurements in intervening DLAs (De Cia et al. 2016), and filled triangles correspond to known  $H_2$  intervening systems (Balashev et al. 2019). The solid blue lines show the best linear fit to the value for the proximate  $H_2$  systems, taking uncertainties on both axes into account. The green lines and shadowed region show the best linear fit and associated  $1\sigma$  uncertainty for the overall sample of intervening DLAs.

$H_2$  ones, even if the latter tend to have high depletion among the overall (mostly non- $H_2$  bearing) intervening population at a given metallicity. It is also interesting to see a lower dispersion for proximate systems around the main trend (dotted blue line), likely suggesting a more homogeneous population.

On the other hand, the difference between proximate  $H_2$  and intervening systems (regardless of their  $H_2$  content) is less clear in the case of  $[Zn/Si]$ . Since silicon is an  $\alpha$ -element predominantly produced in massive stars, it is possible that a stronger intrinsic depletion (with respect to intervening systems) is compensated for by recent star formation, or that the depletion sequence for Si and Fe differs as well.

In short, the high depletion (from  $[Zn/Fe]$ ) at a given metallicity suggests a mechanism for build-up and/or destruction of dust grains that is similar in the population of proximate  $H_2$  systems, but different from that seen in intervening DLAs. Any correction to account for some depletion of zinc and to recover the intrinsic metallicity could then also be different for proximate systems. We note that if the same depletion sequences as in intervening systems were assumed, the corrected metallicities, following De Cia et al. (2016), would be a factor two higher on average.

The abundance of dust can independently be quantified using the extinction per hydrogen atom,  $A_V/N(H)$ . While the 0.07 mag uncertainty on  $A_V$  measurements remains quite large compared

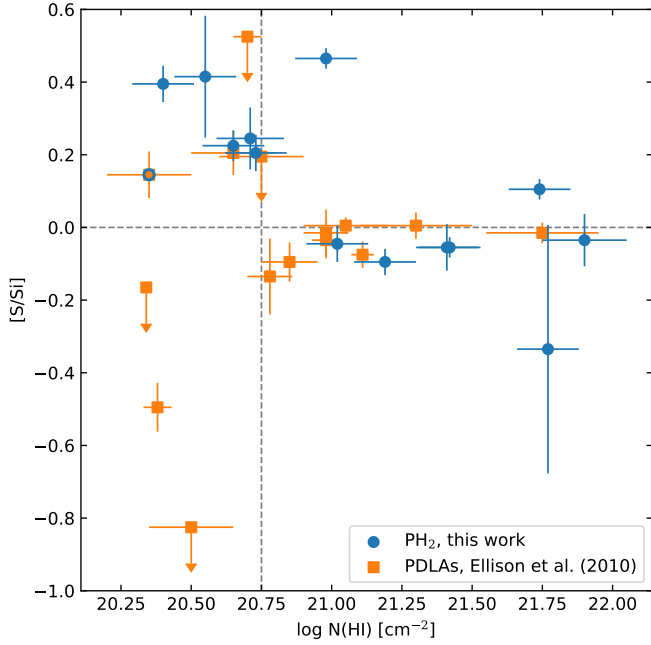


**Fig. 4.** Dust-to-gas ratio measured from the continuum spectrum (filled circles) and derived from the metal column densities (open diamonds), following De Cia et al. 2016) as a function of depletion factor. The blue (purple) dotted line corresponds to a linear fit to  $([Zn/Fe], \log A_V/N(H))$  data points, where  $A_V$  is based on continuum fitting (metal column density).

to their values, there is a clear tendency for higher  $A_V/N(H)$  at higher  $[Zn/Fe]$ ,  $[Zn/Si]$  and metallicities, strengthening a dust-depletion origin for the observed underabundance of iron and silicon. The depletion and value of  $A_V/N(H)$  strongly correlate with each other with a Pearson coefficient  $r = 0.9$  and a probability of chance coincidence of 0.07%; see Fig. 4. The low dispersion around the approximate linear relation ( $\log A_V/N(H) \approx 1.3[Zn/Fe] - 23.5$ ) indicates that the quasar template matching provides reliable estimates, at least relative to each other. Interestingly, it is also possible to indirectly derive the extinction from the gas-phase abundances, assuming some scaling with respect to the Galactic values. We represent this in Fig. 4 as well, based on calculations following De Cia et al. (2016). These values appear to agree relatively well with the reddening-based measurements, but the former are higher by 0.25 mag on average than the latter. This may be another indication for a different type of dust close to the quasar. However, a comparison of extinction derived from gas-phase abundances with that directly obtained from the spectral slope should be performed first in intervening systems to validate the prescription by De Cia et al. (2016) at this precision level.

### 5.1.3. Sulphur-to-silicon ratio in proximate absorption systems

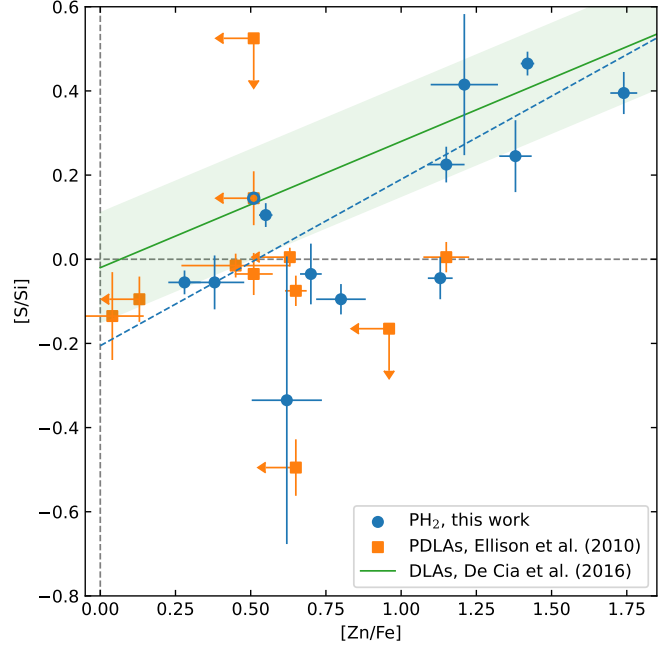
The  $[S/Si]$  ratio is also an interesting quantity because both species are  $\alpha$  elements and their intrinsic abundances are expected to follow each other. According to Ellison et al. (2010), based on modelling by Rix et al. (2007), the observed negative  $[S/Si]$  values at low HI column densities in their sample of PDLAs is due to ionisation by a hard spectrum. We compare this ratio in our sample of proximate systems with those from Ellison et al. (2010) as a function of the HI column den-



**Fig. 5.** Sulphur-to-silicon abundance ratio vs H I column density. The horizontal dashed line marks the solar ratio, and the vertical line shows the  $N(\text{H I})$  limit below which [Ellison et al. \(2010\)](#) observed lower-than-solar abundances in PDLAs (see text). While not in our sample of proximate  $\text{H}_2$ , one PDLA from [Ellison et al. \(2010\)](#), namely J2340-00, also has strong  $\text{H}_2$  and is marked with a blue circle.

sity in Fig. 5. In contrast to what these authors observed in PDLAs, we observe no drop in  $[\text{S}/\text{Si}]$  below  $\log N(\text{H I}) = 20.75$  in our strong  $\text{H}_2$ -bearing PDLAs. While we did not probe H I columns as small as these authors, it is interesting to note that both PDLAs and  $\text{PH}_2$  have close to solar ratios at high column densities  $\log N(\text{H I}) > 21$ . In contrast with the PDLAs from [Ellison et al. \(2010\)](#), we see a tendency for increasing  $[\text{S}/\text{Si}]$  with decreasing H I column, which is opposite to what is expected if these ionisation corrections are at play, but rather reflects an increasing depletion with decreasing H I. This can be due to a bias against systems with high column densities of dust that would either be excluded from the parent quasar sample ([Ledoux et al. 2009](#); [Augustin et al. 2018](#)) or result in a low S/N of the corresponding spectra. It can also be due to a higher abundance of metals or dust required for the presence of  $\text{H}_2$  at low H I column densities (e.g. [Bialy et al. 2017](#)), thereby increasing the dust shielding,  $\text{H}_2$  formation rate, and gas cooling through atomic lines. In the low  $N(\text{H I})$  regime, we would qualify the observed  $[\text{S}/\text{Si}]$  in (non- $\text{H}_2$ ) PDLAs as highly dispersed and possibly interpret this as a wide range of ionisation.

To investigate the effect of differential dust depletion further, we plot  $[\text{S}/\text{Si}]$  as a function of  $[\text{Zn}/\text{Fe}]$  in Fig. 6. Here, the correlation between the two abundance ratio is apparent in the case of proximate  $\text{H}_2$  systems (Pearson correlation coefficient of 0.76 with a 0.4% probability to be due to chance coincidence). This favours depletion of silicon as a more likely origin for the high  $[\text{S}/\text{Si}]$  values. Fitting a linear relation to the proximate  $\text{H}_2$  data that takes uncertainties on both axes into account, we obtain  $[\text{S}/\text{Si}] \approx 0.39(\pm 0.08) \times [\text{Zn}/\text{Fe}] - 0.21(\pm 0.09)$ . A correlation between the two abundances ratio has also been observed in intervening DLAs (Pearson correlation  $r = 0.7$ ), for which we show the best-fit linear relation obtained by [De Cia et al. \(2016\)](#) together with the associated intrinsic



**Fig. 6.** Sulphur-to-silicon vs zinc-to-iron abundance ratios in proximate  $\text{H}_2$  and DLA systems. The green line and shaded area represent the best fit for intervening DLAs by [De Cia et al. \(2016\)](#) and the  $1\sigma$  scatter around this relation. The dashed blue line is a linear fit to the proximate  $\text{H}_2$  data. Dashed vertical and horizontal lines mark solar values. Symbols and colours for data points are as in Fig. 5.

dispersion<sup>1</sup>. However, proximate systems differ from intervening DLAs in the sense that they have  $[\text{S}/\text{Si}] \approx 0$  for a wider range of  $[\text{Zn}/\text{Fe}]$  values. This is observed for PDLAs regardless of whether they contain  $\text{H}_2$  when the intercept for intervening DLAs is almost zero.

In summary, the  $\text{PH}_2$  systems with high  $[\text{S}/\text{Si}]$  values appear to behave like intervening DLAs. The five systems with highest values also have low H I column densities, which is naturally explained if high-H I high-depletion systems are missing due to dust bias. At low  $[\text{Zn}/\text{Fe}]$  (and high  $N(\text{H I})$ ), the  $[\text{S}/\text{Si}]$  values tend to differ between intervening and proximate systems, the latter having typically lower observed  $[\text{S}/\text{Si}]$  from S II and Si II closer to solar. We conclude that the main driver of the  $[\text{S}/\text{Si}]$  ratio is depletion onto dust, with a possible different depletion sequence than that seen for intervening DLAs, again suggesting a different dust type proximate to the quasar. However, ionisation effects might still be at play, meaning that part of these low-ionisation metals might arise from ionised gas. Larger samples of proximate DLAs (with and without  $\text{H}_2$ ) are necessary to investigate this further.

#### 5.1.4. Argon as a probe of a hard UV ionising field

Argon can provide an additional diagnostic on the ionising field. Owing to its high first-ionisation energy (15.76 eV), argon is expected to be mostly neutral in DLAs, which are self-shielded against H I-ionising photons. While argon is predominantly ionised by cosmic rays in the local ISM, high-energy photons significantly above 13.6 eV can still penetrate the cloud and ionise argon, which further has a low recombination to ion-

<sup>1</sup> The differences in adopted solar values between this work and [De Cia et al. \(2016\)](#) only affect the  $[\text{Zn}/\text{Fe}]$  ratio by 0.005 dex and hence can safely be ignored here.

isation rate in diffuse atomic medium (see e.g. Sofia & Jenkins 1998). Zafar et al. (2014a) found a deficiency of argon by about 0.4 dex (with standard deviation about 0.2 dex) with respect to other  $\alpha$ -elements in intervening DLAs, with a weak dependence on HI column density. They concluded that because argon is extremely volatile (Lodders 2003), the argon deficit is likely dominated by ionisation from the quasar metagalactic radiation, with local modulation by HI column inside the absorbing galaxies.

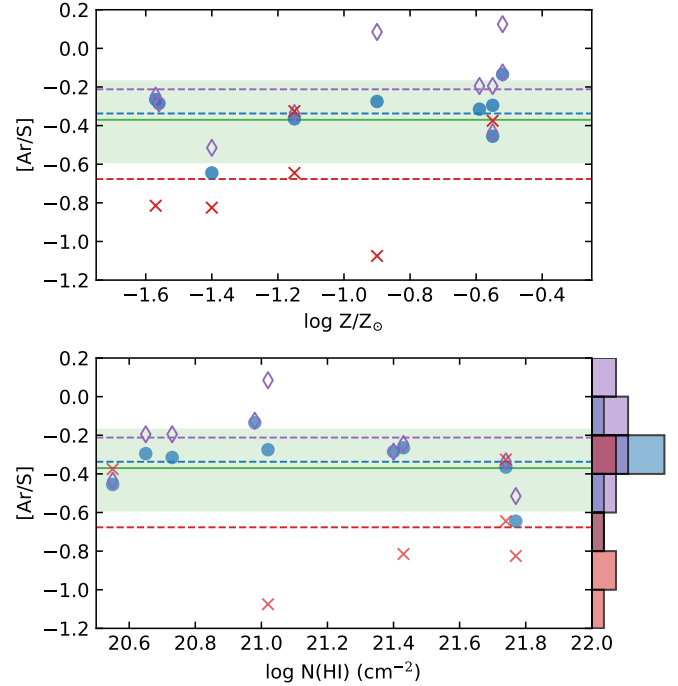
We constrained the Ar I column densities in our sample of proximate H<sub>2</sub> absorbers using the two available absorption lines (at 1048 and 1066 Å). Because these fall into the Lyman- $\alpha$  forest and can easily be blended with H<sub>2</sub>/HD lines, we used a slightly different procedure: We fitted the Ar I lines by fixing the redshifts and Doppler parameters to that previously obtained for other metal metal lines and fitting for the column densities. We included the H<sub>2</sub> absorption profile constrained by the independent fit (see Sect. 3.1) and avoided blends with the Ly- $\alpha$  forest as best as possible. In a few cases, however, we were unable to obtain a meaningful constraint. The fit to the argon lines is shown in Fig. B.10.

The derived [Ar/S] (obtained from Ar I and S II) values are shown as a function of the total HI column densities and overall metallicities in Fig. 7. First, we do not see any dependence of [Ar/S] on  $N(\text{HI})$ , in agreement with the expected lack of shielding by neutral hydrogen. We do not see any dependence on overall metallicity either, in contrast to the possible negative (but not easily explained) trend seen by Zafar et al. (2014a). The lack of a dependence on overall metallicity is in line with the expected non-depletion of argon onto dust. Second, [Ar/S] is found to be very similar (-0.35 dex on average) to that seen in intervening DLAs, which could be surprising given the proximity of the quasar. However, our measurements can still be affected by Ly- $\alpha$  interlopers so that the Ar I columns should in principle be considered as upper limits in a conservative approach<sup>2</sup>, but more importantly, the metal column density is generally dominated by the H<sub>2</sub>-bearing component, where we observe a lower deficit of argon (-0.2 dex) with respect to components without detectable H<sub>2</sub> (-0.7 dex). In Fig. B.10, we also show the expected Ar I profiles obtained assuming a solar [Ar/S] ratio. This over-predicts the absorption in most cases, but this over-prediction also appears to be less severe at the velocity of H<sub>2</sub> components.

The relatively low deficit of argon in H<sub>2</sub> components is most likely a result from an efficient conversion of Ar<sup>+</sup> into neutral argon through chemical reactions with H<sub>2</sub> (Ar<sup>+</sup> + H<sub>2</sub> → ArH<sup>+</sup> + H followed by ArH<sup>+</sup> + H<sub>2</sub> → Ar + H<sub>2</sub><sup>+</sup> (see Duley 1980; Schilke et al. 2014; Neufeld & Wolfire 2016). The situation could be similar to that with neutral chlorine (Balashev et al. 2015) and all argon being in neutral form in the presence of H<sub>2</sub>.

In the regions where hydrogen is predominantly atomic, the ionisation fraction of argon should in turn be very dependent on the ionising field. At higher  $I_{\text{UV}}$ , the photo-ionisation rate indeed increases while the H<sub>2</sub> abundance decreases, impeding the molecular-ion recombination route. Hence the ionisation fraction of argon should rise at least proportionally to  $I_{\text{UV}}$ . Therefore, the systematically lower abundance in non-H<sub>2</sub>-bearing components confirms that the medium is likely exposed to a significantly enhanced (and hard) radiation field and/or cosmic-ray ionisation. While quantitative statements would require a full

<sup>2</sup> This should be less of an issue at high spectral resolution, as in the work by Zafar et al. (2014a), which was based on UVES data.

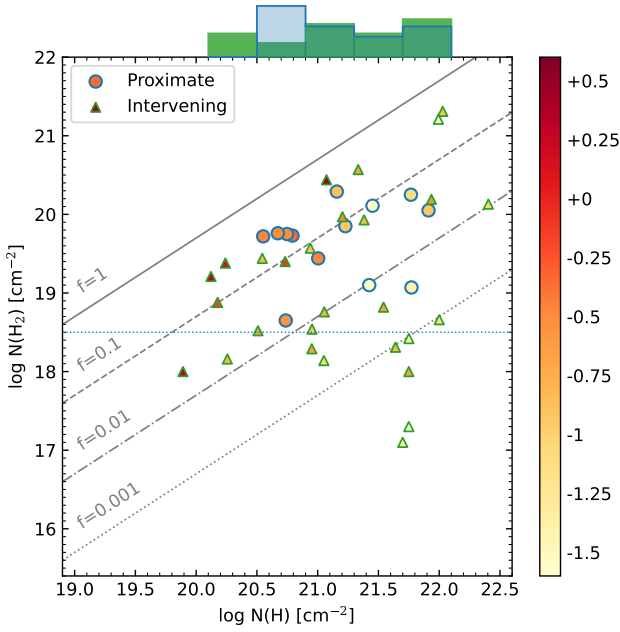


**Fig. 7.** Observed [Ar/S] ratio (based on Ar I and S II) as a function of overall metallicity (top) and HI column density (bottom) in our sample of proximate systems. Blue circles correspond to the overall value in each system, and purple diamonds and red crosses show the H<sub>2</sub> and non-H<sub>2</sub> components, respectively. The dashed lines correspond to the corresponding mean values. The distribution is shown in the side panel. The green line corresponds to the mean [Ar/S] value (corrected to the same solar reference) and dispersion for intervening DLAs (Zafar et al. 2014b). Even though we tried to take blends into account, all points should be considered as upper limits in principle.

modelling of the observed column densities and a better-assessed association between H<sub>2</sub> and metal components (i.e. at high spectral resolution), which is beyond the scope of this paper, we can estimate that the -0.2 dex deficit of argon means that about half of the metal column seen at the velocity of H<sub>2</sub> components could come from their (H<sub>2</sub>-free) HI-envelope where argon would be mostly in ionised form.

## 5.2. H<sub>2</sub> in proximate versus intervening systems

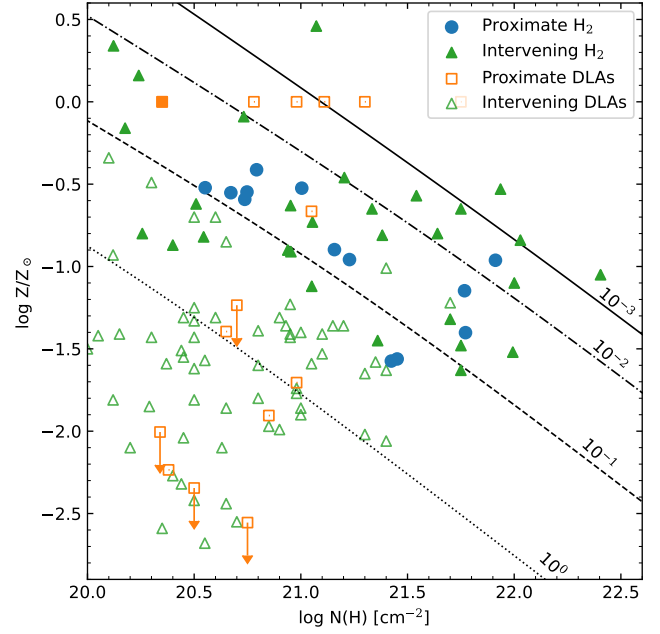
We compare proximate and intervening systems according to their H<sub>2</sub> and total hydrogen column densities in Fig. 8. The absence of proximate systems at low H<sub>2</sub> columns is naturally due to our selection based on the presence of strong H<sub>2</sub> Lyman-Werner bands (Paper I), and all our systems have  $\log N(\text{H}_2) > 18.5$ . Proximate H<sub>2</sub> were in turn selected independently of  $N(\text{HI})$ , but as this figure shows, they are limited to  $\log N(\text{H}) > 20.5$  when intervening systems can have lower  $N(\text{H})$  even at similar H<sub>2</sub> column densities (see the distributions at the top of this figure). However, this is driven by a few very metal-rich intervening absorbers (as can be seen from the colour coding), two of them following a C I-based selection, skewed towards very high metallicities. Metallicities tend to increase with decreasing  $N(\text{H})$  and increasing  $N(\text{H}_2)$ . This might naturally be explained by the  $1/Z$  dependence of the critical column density for H<sub>2</sub> to form (e.g. Krumholz et al. 2008), but also to a bias against high-metallicity systems at high- $N(\text{HI})$  due to stronger dust extinction.



**Fig. 8.** Column density of molecular hydrogen vs total neutral hydrogen column density  $N(\text{H}) = N(\text{HI}) + 2N(\text{H}_2)$  for our sample of proximate  $\text{H}_2$  absorbers and  $z > 2$  intervening strong  $\text{H}_2$ -bearing DLAs from Balashev et al. (2019). For a fair comparison, we removed systems that may potentially be proximate, and kept only systems with a conservative cut of  $c(z_{\text{abs}} - z_{\text{QSO}})/(1 + z_{\text{QSO}}) < -10000 \text{ km s}^{-1}$ . Our selection of proximate  $\text{H}_2$  systems based on strong  $\text{H}_2$  lines directly seen in the low-resolution SDSS spectra naturally skews our sample to high  $\text{H}_2$  columns. The horizontal blue line shows the approximate and empirical lower limit of at  $\log N(\text{H}_2) > 18.5$  for the present sample. The top histograms show the  $N(\text{H})$  distributions for the proximate (blue) and intervening (green) systems with  $N(\text{H}_2) > 18.5$ . Finally, the different lines correspond to different averaged molecular fractions  $f_{\text{H}_2} \equiv 2N(\text{H}_2)/(2N(\text{H}_2) + N(\text{HI}))$ .

This is better visible in Fig. 9, in which the location  $\text{H}_2$  systems is shown in the  $N(\text{H})$ - $Z$  plane. The different lines show the ratio of UV intensity to number density as a function of the HI column density in the envelope of a single  $\text{H}_2$  cloud illuminated on one side, following Eq. (2) of Paper I, based on theoretical calculations by Bialy et al. (2017) and where we used  $Z/Z_{\odot}$  as a proxy for  $\sigma_g$ , the grain Lyman-Werner photon absorption cross section per hydrogen nucleon normalised to the fiducial Galactic value. We note that these calculations do not provide the column density of  $\text{H}_2$ , which just continues to build up as more gas would be added to a given cloud. In practice, the total column is dominated by HI, so that the plot against  $N(\text{HI})$  only is almost identical. Three-quarter (9 out of 12) of the proximate  $\text{H}_2$  systems are located between the  $I_{\text{UV}}/n = 10^{-2}$  and  $10^{-1} \text{ cm}^3$  lines, suggesting similar physical conditions. We caution, however, that the measured HI column corresponds to that integrated over the profiles, while the calculations provide the HI column that participates in the shielding of  $\text{H}_2$  clouds. Any atomic gas, with some amount of dust, that is located upstream along the line of sight will still participate in shielding  $\text{H}_2$  from the quasar field.

Determining the physical conditions from population levels is thus necessary to constrain the distances of the cloud from the AGNs. Owing to the different selections, intervening  $\text{H}_2$  systems in turn tend to have a wider spread of values. We also note that the lines of constant  $I_{\text{UV}}/n$  closely follow the constant metal column density (for volatile species), and hence almost



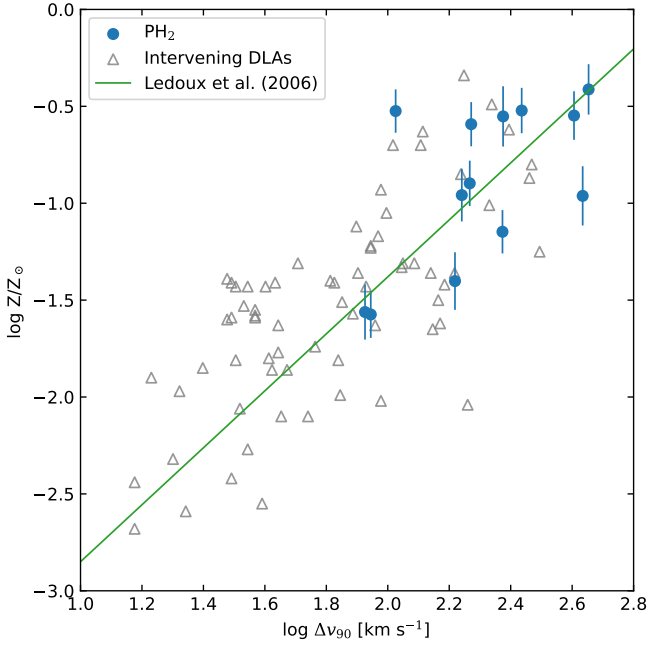
**Fig. 9.** Metallicity vs total hydrogen column density. The different lines correspond to the hydrogen column density of a single  $\text{H}_2$  cloud illuminated by a UV field for different UV to density ratios  $I_{\text{UV}}/n$  as indicated above each line, with  $I_{\text{UV}}$  in units of the Draine field and  $n$  in  $\text{cm}^{-3}$ . Filled symbols correspond to  $\text{H}_2$ -bearing systems. Proximate  $\text{H}_2$  systems are from this work, proximate DLAs from Ellison et al. (2010), and intervening systems from Noterdaeme et al. (2008) and Balashev et al. (2019).

constant extinction (e.g. Vladilo et al. 2006). In other words, the conditions that facilitate the presence of  $\text{H}_2$  are also those that can extinguish the background sources to the point that may be excluded from the optically selected samples. However, since most of the intervening systems were also detected towards quasars from the SDSS, it is likely that more reddened quasars with more metal-rich proximate  $\text{H}_2$  absorbers are already present in the SDSS database, in particular because these systems could afford lower densities to survive, but the S/N achieved in the blue prevented their detection directly from  $\text{H}_2$  lines.

We also show in Fig. 9 the location of proximate and intervening DLAs without strong  $\text{H}_2$  as open symbols. In the case of intervening non- $\text{H}_2$  systems, we used the sample from Noterdaeme et al. (2008), for which  $\text{H}_2$  has been systematically assessed. The fact that no  $\text{H}_2$  is detected, even at high metallicity and HI column density, means that the corresponding system does not fulfil the requirements for  $I_{\text{UV}}/n$ . For example, for a UV field about one-tenth of the Draine field, these systems would have low number densities  $n < 1 \text{ cm}^{-3}$  (which is typical for the warm neutral phase), which is too low to build up molecular hydrogen. Alternatively, their total HI column density is spread over many different clouds. We finally note that the above implicitly assumes static equilibrium, which may not necessarily be the case, in particular if  $\text{H}_2$  is found in outflowing gas. For  $n = 100 \text{ cm}^{-3}$ , the characteristic timescale for  $\text{H}_2$  formation indeed is  $\sim 10^7 \text{ yr}$ , which is then comparable to the dynamical timescales.

### 5.3. Kinematics

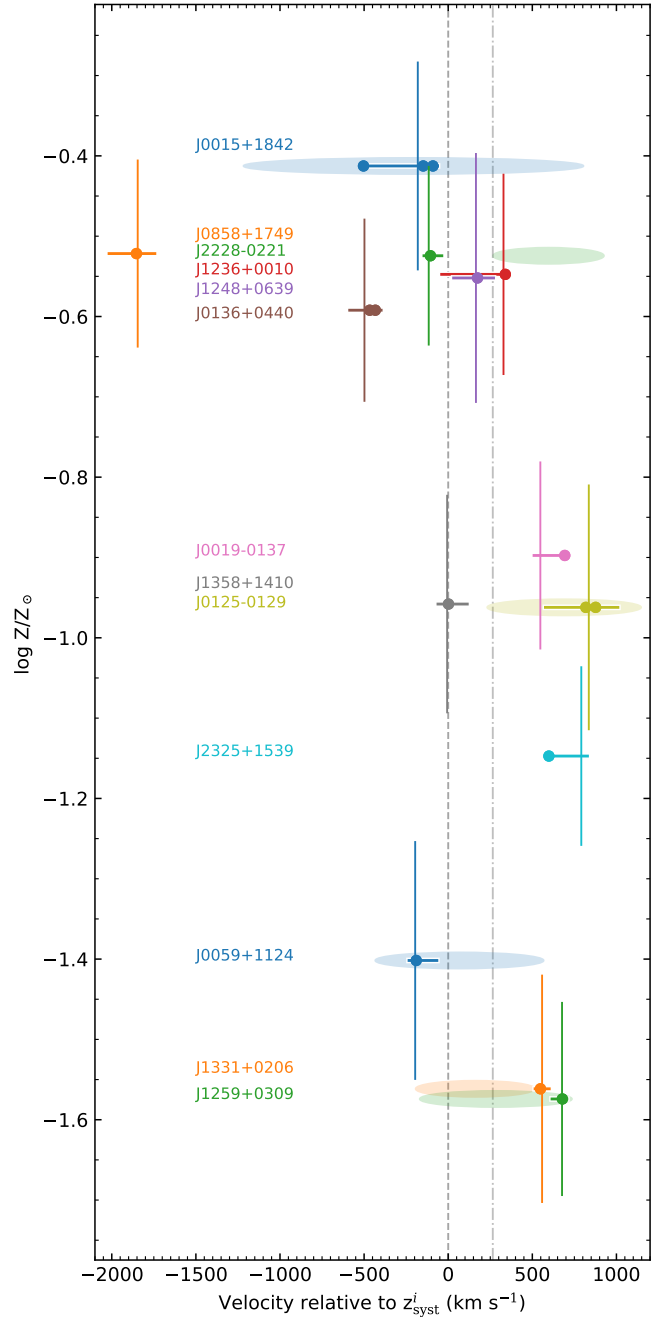
The velocity extent of low-ionisation metal lines  $\Delta v_{90}$  is considered as a good way to quantify the kinematics of the neutral gas



**Fig. 10.** Metallicity vs velocity extent of the low-ionisation metal lines for our sample of proximate H<sub>2</sub> systems and a sample of intervening DLAs (mostly from Ledoux et al. 2006 with some additions by Noterdaeme et al. 2008). The green line shows the best-fit relation for high-*z* intervening DLAs by Ledoux et al. (2006).

(because this information is not available from the HI lines). In the case of intervening DLAs, Ledoux et al. (2006) first observed a correlation between this velocity width and the gas metallicity, suggesting that this reflects an underlying relation between mass and metallicity because the motion of low-ionisation lines is expected to be governed by gravity. This has been discussed in a number of subsequent works (Møller et al. 2013; Neeleman et al. 2013; Christensen et al. 2019), with attempts to connect this to the mass-metallicity relation seen in emission (e.g. Møller & Christensen 2020). While the assumption that  $\Delta v_{90}$  primarily depends on the galaxy mass seems to be consolidated at least in a statistical sense, there is evidence that outflows, tidal streams, and the environment in general can affect the velocity spread in a more complex way (e.g. Zou et al. 2018; Nielsen & Kacprzak 2022). Because of the presence of the quasar, the situation might be expected to be even more complex in proximate systems. For example, the case of J0015+1842 is strongly suggestive of outflowing material (Paper II). Surprisingly enough, the distribution of  $\Delta v_{90}$  in proximate H<sub>2</sub> systems follows the trend, and dispersion seen in intervening DLAs, although it is limited to high values, as expected from the presence of H<sub>2</sub> (Noterdaeme et al. 2008). Overall, the velocity spread of the neutral gas in proximate H<sub>2</sub> systems suggests an origin in massive galaxies, but without a clear indication of perturbed kinematics, or at least, not more perturbed than similar high-metallicity systems in the intervening population (see Fig. 10).

In order to investigate now the kinematics of the gas with respect to the quasar, we represent in Fig. 11 the kinematics of the neutral gas as a function of its metallicity. There is an apparent trend for high-metallicity systems to be around the quasar systemic redshift as obtained from the ionised lines, while low-metallicity systems tend to be redshifted. The Pearson correlation coefficient between  $v_{50} = c(z_{\text{QSO}} - z_{50})/z_{\text{QSO}}$  (i.e. the velocity of the centroid of the absorption profile), and the gas



**Fig. 11.** Summary of absorbing gas metallicity and kinematics. The differently coloured segments represent the  $\Delta v_{90}$  regions, and dots indicate the location of H<sub>2</sub> components. The zero of the velocity scale is set to the best estimate of the systemic redshift  $z_{\text{sys}}$  from ionised lines (Sect. 4.2). Positive velocities mean here  $z > z_{\text{sys}}^i$ . Ellipses are shown at the location of the CO(3-2) emission whenever observed, and the major axis is equal to the line FWHM.  $z_{\text{CO}}$  is shifted with respect to  $z_{\text{sys}}^i$  by about 250 km s<sup>-1</sup> on average (dash-dotted line).

metallicity is  $r = -0.50$  with a  $p$ -value 0.08. In other words, the probability is only about 8% that this anti-correlation is due to chance alone. One system, towards J0858+1749, is significantly more blueshifted than the rest of the sample. This does not drive the correlation, however, as removing it from the correlation test only changes the Pearson values to  $r = -0.49$ ,  $p = 0.10$ . The  $\sim 2000$  km s<sup>-1</sup> blueshift of this system makes it intermediate between what could be considered as clearly intervening or

proximate. For example, at this velocity, the incidence of strong H<sub>2</sub> absorbers already reaches that of intervening absorbers (see Fig. 3, [Noterdaeme et al. 2019](#)). The derived UV flux in this system ([Balashev et al. 2019](#)) suggests that this system is located farther from the quasar, at a distance of at least several hundred kiloparsec. To account for the statistical uncertainties on metallicity and on the quasar systemic redshift, we calculated the Pearson ( $r, p$ ) values for 10 000 random samples, for which the metallicity of each system was taken using a normal distribution around the best value with the corresponding uncertainties. Similarly, the velocity was chosen using a normal distribution around the measured value, with  $\sigma$  set to the intrinsic velocity dispersion of  $z_{\text{QSO}}$  provided by [Shen et al. \(2016\)](#); i.e. 46, 56, 205, and 233 km s<sup>-1</sup> for  $z_{\text{QSO}}$  based on [O II], [O III], Mg II and C III], respectively). From this exercise, we find an approximate normal distribution of  $r$  values, giving  $r = -0.47 \pm 0.09$  ( $r = -0.44 \pm 0.11$  ignoring J0858+1749). The  $p$ -values have a median of 0.11 (0.14), but with some tail towards higher values. In short, this means that the trend remains even when statistical uncertainties on metallicity and emission redshift measurements are taken into account.

In principle, the projected line-of-sight velocity towards the quasar for low-metallicity systems could be interpreted as neutral gas infalling onto the quasar host galaxy. However, it is unlikely that gas would feed the quasar directly, and we may instead expect any feeding gas to orbit the quasar host galaxy without a significant line-of-sight velocity component. It is possible, however, that gas orbits the galaxy on non-circular orbits after a galaxy companion interaction, where the outer neutral gas returns as a fountain after having been dragged out in a tidal tail.

Alternatively, the systemic redshift might be underestimated if the blueshifts of the emission lines are higher than observed by [Shen et al. \(2016\)](#). In their work, the authors used Ca II absorption as a reference to compute the (non)-shifts of [O II] and [O III] and then the shift from other lines (Mg II, C III], etc.) are computed with respect to the latter. The use of SDSS data (covering wavelengths shorter than 10 400 Å) implies that the computed shifts correspond to low- to intermediate-redshift quasars (which therefore are likely not particularly luminous). At high redshifts ( $z \sim 5-7$ ), [Schindler et al. \(2020\)](#) and [Eilers et al. \(2021\)](#) found higher Mg II blueshifts in luminous quasars, using precise systemic redshifts from [C II]158 $\mu\text{m}$  and/or CO emission line measurements from ALMA and NOEMA as reference. These lines arise from star formation and molecular gas, respectively, and hence should better represent the systemic redshift of the quasar host, which is also very likely to dominate other lower-mass galaxies in the quasar group. The sample presented here has redshifts  $z \sim 3$ , so that we may naively expect line shifts in between those obtained by [Shen et al. \(2016\)](#) and those obtained by [Schindler et al. \(2020\)](#) or [Eilers et al. \(2021\)](#). For half of our sample, we have CO(3-2) redshifts from our NOEMA observations, which we represent as ellipses in Fig. 11. The CO redshifts are about 250 km s<sup>-1</sup> higher on average than those obtained from the (corrected) ionised emission lines. If  $z_{\text{CO}}$  represents better estimates of the host galaxy redshift, then the absorbing gas now appears to have no particular line-of-sight velocity at low metallicities, while the outward velocities at high metallicities is reinforced. This maintains the anti-correlation to  $r \approx -0.6$ ,  $p \approx 0.2$  (now computed on six quasars). In this case, high-metallicity gas (above  $\sim 25\%$  solar) is more likely to represent gas enriched by intense star formation activity that is expelled through associated winds, as is clearly seen in the case of J0015+1842 ([Paper II](#)), which has the highest metallicity in our sample. This would be consistent with chemically enriched

outflowing gas coming from nuclear regions, in contrast to outer gas, due to the radial  $Z$ -gradient.

Determining the distances between the absorbing gas and the active nucleus by modelling the excitation of molecular and atomic species should help understanding the origin of the gas with different metallicities further. This study using the present sample is devoted to a future paper.

#### 5.4. AGN proximity from the presence of N V

While we focused on the neutral phase (and low ionization metal species) in this paper, we briefly discuss here the presence of N V absorption at the quasar redshift. Photo-ionisation of nitrogen to this stage requires an energy of 77.5 eV, meaning a hard ionisation spectrum because stellar light falls rapidly after 54 eV, when He II ionisation occurs. This makes N V one of the best diagnostics for the physical proximity of the AGN and distinguishes between intervening and associated system (e.g. [Perrotta et al. 2016](#)). [Fox et al. \(2009\)](#) found that N V is detected in only 13% of the intervening DLAs, but they also found a surprisingly similar fraction in PDLAs out to 5000 km s<sup>-1</sup>, suggesting that proximity in velocity does not necessarily mean physical proximity. 5000 km s<sup>-1</sup> in the Hubble flow corresponds to about 16 Mpc proper distance at  $z \sim 3$ . [Ellison et al. \(2010\)](#) found a tentative excess of N V detection rate (2 out of 7) at similar velocities in their PDLA sample, but we note that this fraction would increase if they had limited the statistics to smaller velocity separations from the quasar. Finally, using composite spectra of about one thousand C IV absorbers and interpreting their results with the help of ionisation modelling, [Perrotta et al. \(2018\)](#) confirmed that N V is only present in proximate systems and concluded that it is a good tool for identifying gas that is physically related to the quasar host. N V absorption is clearly seen at the quasar redshift for 9 out of our 13 proximate H<sub>2</sub> absorbers (see Fig. A.1). The non-detection of N V in the two systems with the lowest metallicity in our sample may just be due to column density effect rather than lower ionisation. Overall, the very high incidence of N V in our sample supports an origin in the close environment of the quasars. Ionisation modelling, together with comparative kinematic (N V is typically present over wider velocity ranges than low-ionisation species), should help us to understand the ionised component of H<sub>2</sub>-selected proximate systems.

## 6. Conclusion

The chemical enrichment and velocity dispersion of the neutral gas in strong proximate H<sub>2</sub> systems appears not to be significantly different from what is seen in intervening systems. While the metallicities in proximate H<sub>2</sub> systems are higher than the general population of intervening DLAs and are in the upper half when compared to the sample of proximate DLAs from [Ellison et al. \(2010\)](#), they are consistent with intervening systems that also have high H<sub>2</sub> content (see [Balashev et al. 2019](#)). The relative abundances of sulphur, silicon, iron, and zinc follow the expected trends due to increasing depletion of refractory elements with increasing metallicity (e.g. [De Cia et al. 2016](#)), indicating an increased abundance of dust with increasing chemical enrichment. This is also confirmed independently by the increasing quasar extinction per hydrogen atom with depletion and metallicity. We note, however, the higher depletion of iron in proximate systems compared to intervening systems at given metallicity, suggesting different dust production and/or destruction close to the AGN. The ubiquitous presence of N V absorption indeed indicates a strong and hard UV field, which also

explains the deficit of neutral argon in non-H<sub>2</sub>-bearing components. Argon in turn remains predominantly in neutral form in the presence of H<sub>2</sub>, aided by ion-molecule reactions.

The lack of proximate and intervening systems with both high-metallicity (or depletion) and hydrogen column density is likely due to a selection effect against highly reddened systems, which would be more severe than just exclusion from optically selected quasar samples. Systems like this may already be present in the SDSS quasar database, but not selected through H<sub>2</sub> absorption because of their low S/N in the blue. These systems would also be interesting to recover because they should be able to survive with relaxed densities for a given intensity of UV field (i.e. distance to the central engine). Upcoming surveys should also provide less biased samples of quasars and absorbers. For example, 4MOST (de Jong et al. 2019) will collect optical spectra of quasars selected on astrometry (The 4MOST-*Gaia* Purely Astrometric Quasar Sample, 4G-PAQS; PI: Krogager), infrared, or X-ray properties (Merloni et al. 2019). The higher spectral resolution and larger collecting area should push deeper into the dust-obscured population of absorbers. Similarly, the MeerKAT Absorption Line Survey (MALS; Gupta et al. 2016) also uncovers quasars in a dust-unbiased manner, owing to the radio+IR selection of the targets (see Krogager et al. 2018; Gupta et al. 2022a). A complementary view of proximate cold gas absorbers can then be done through associated 21 cm absorption. Furthermore, 21 cm absorption provides different (and multiple) lines of sight because radio-continuum emission regions are distinct from the region that produces the optical continuum (Gupta et al. 2022b).

The significant excess of proximate H<sub>2</sub> systems with respect to the expected incidence from the intervening statistics showed that most of the former are very likely to be associated with the quasar environment (quasar host or galaxies in the quasar group). On the other hand, the similar column densities of neutral and molecular hydrogen at similar metallicities when compared to intervening H<sub>2</sub>-bearing DLAs means that proximate H<sub>2</sub> should have similar UV-to-density ratios. This means that proximate H<sub>2</sub> systems should have a significantly higher density than their intervening analogues if they are located within the dominating influence of the quasar UV field (i.e. within the host galaxy, or possibly several hundred kiloparsec away). The higher densities if located close to the quasars would explain the higher fraction of systems with excited atomic fine-structure lines. Si II\* is indeed clearly detected in 8 out of the 13 systems presented here. Determining the physical conditions is necessary in that respect as well.

While our parent sample was selected by having  $z_{\text{abs}} \approx z_{\text{QSO}}$ , the systems might have been found within a wide velocity range around the quasar owing to the large uncertainties on the quasar emission redshifts in SDSS spectra, together with the low S/N and resolution, allowing H<sub>2</sub> detection over a relatively wide range. It is then almost surprising to find that all systems but one were found to have velocities significantly lower than 1000 km s<sup>-1</sup> from the quasar, after securing precise absorption and emission redshifts. In the absence of peculiar motions (i.e. pure Hubble flow), this would correspond to at most a few million parsec. If peculiar motions contribute significantly to the observed velocities, then this becomes an upper limit to the actual distance. Velocities of  $z_{\text{abs}} > z_{\text{QSO}}$  systems indicate the dispersion due to peculiar velocities because this clearly cannot be due to the Hubble flow. The kinematics are then all consistent with no Hubble-flow component at all (except for J0858+1749), suggesting that most systems are indeed associated with the quasar host or its group environment.

We observe a trend for different line-of-sight velocities of the absorbing gas with respect to the quasar as a function of metallicity. When we assume systemic redshifts from ionised lines, this would suggest that low-metallicity gas tends to move towards the quasar. However, it is more likely that the low-metallicity gas is more clustered (in velocity space) around the quasar, as corroborated by the use of CO-based redshifts, and that the high-metallicity gas more likely arises from outflowing winds. While this remains speculative for now, if it is confirmed in a large sample of systems, including non-H<sub>2</sub> systems, the relation between the motion of neutral gas with respect to the quasar host galaxy and/or the central engine and the chemical enrichment of this gas would open a new window for understanding quasar evolution.

Selection effects appear to be complex in proximate systems, however. For example, the classical selection of PDLAs based on saturated cores biases the samples against systems with leaking Lyman- $\alpha$  emission, which is observed in about half of our proximate H<sub>2</sub> systems. If, as expected, the amount of gas and its distribution around the quasar affects the ability of Lyman- $\alpha$  photons to scatter at large distances, this bias will also affect any conclusion on the nature of proximate systems.

We conclude that proximate systems have a strong potential for investigating many processes that drive the evolution of quasars. Determining the physical conditions and distances of gas clouds from the central engine is the next natural step in this matter.

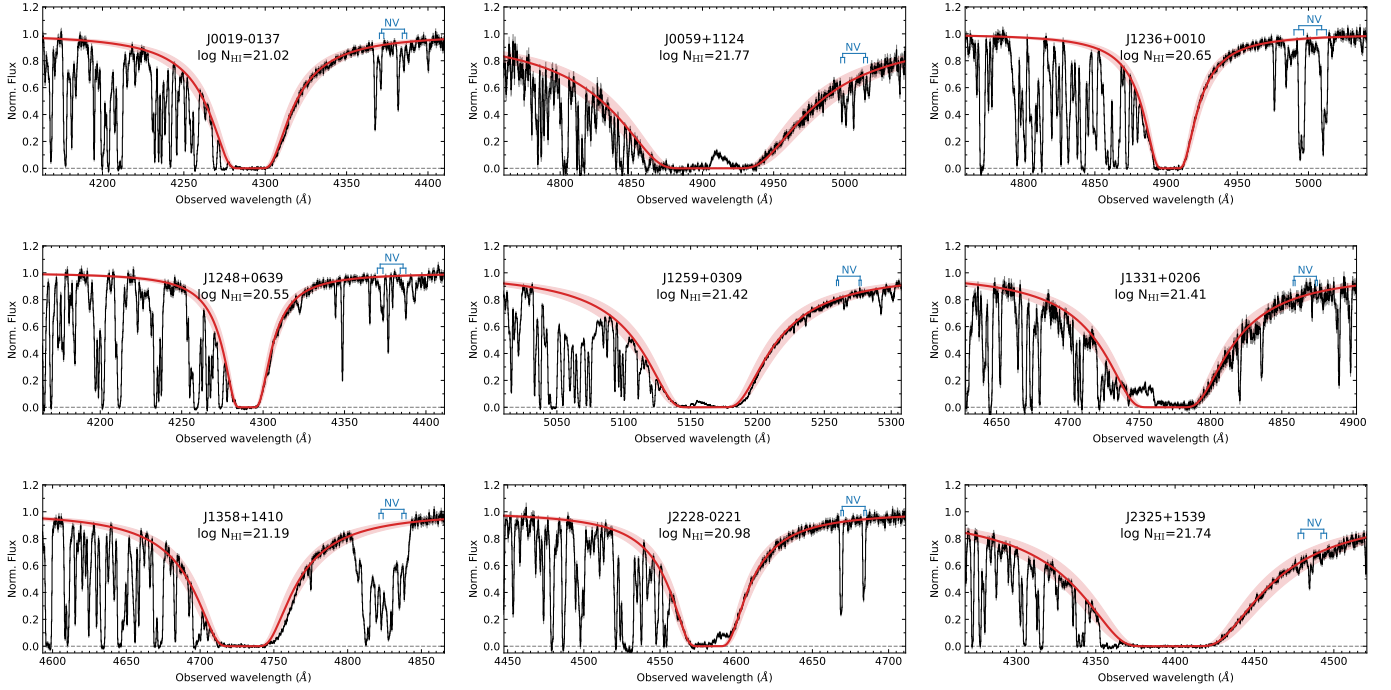
*Acknowledgements.* We thank the anonymous referee for useful comments and suggestions. The research leading to these results has received support from the French Agence Nationale de la Recherche under ANR grant 17-CE31-0011-01/project “HIH2”. R.C. gratefully acknowledges support from the French-Chilean Laboratory for Astronomy (IRL 3386). S.L. acknowledges support by FONDECYT grant 1191232.

## References

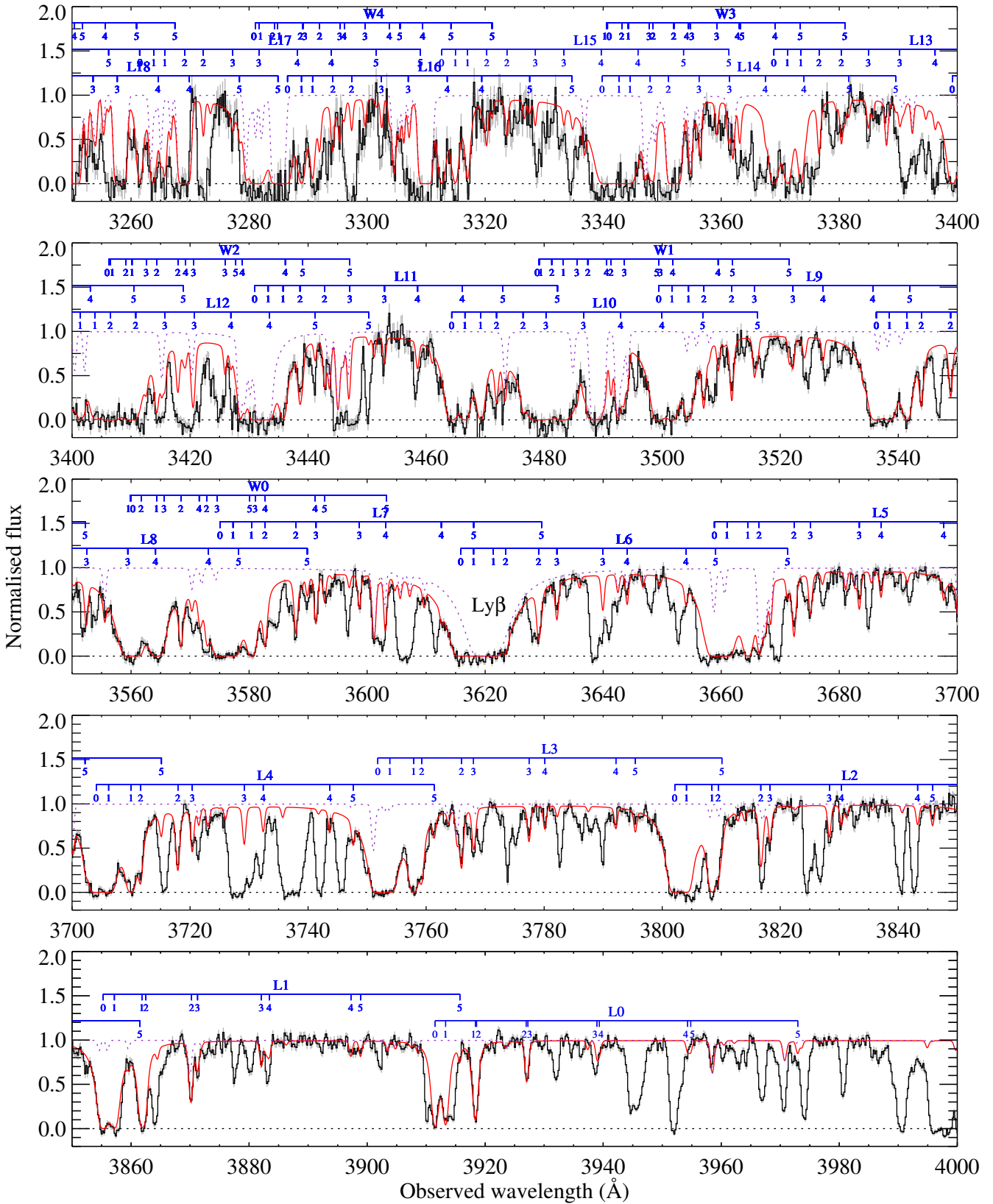
- Arabsalmani, M., Møller, P., Fynbo, J. P. U., et al. 2015, *MNRAS*, 446, 990
- Arrigoni Battaia, F., Hennawi, J. F., Prochaska, J. X., et al. 2019, *MNRAS*, 482, 3162
- Asplund, M., Amarsi, A. M., & Grevesse, N. 2021, *A&A*, 653, A141
- Augustin, R., Péroux, C., Møller, P., et al. 2018, *MNRAS*, 478, 3120
- Balashev, S. A., Varshalovich, D. A., & Ivanchik, A. V. 2009, *Astron. Lett.*, 35, 150
- Balashev, S. A., Klimenko, V. V., Ivanchik, A. V., et al. 2014, *MNRAS*, 440, 225
- Balashev, S. A., Noterdaeme, P., Klimenko, V. V., et al. 2015, *A&A*, 575, L8
- Balashev, S. A., Noterdaeme, P., Rahmani, H., et al. 2017, *MNRAS*, 470, 2890
- Balashev, S. A., Klimenko, V. V., Noterdaeme, P., et al. 2019, *MNRAS*, 490, 2668
- Bialy, S., Bihr, S., Beuther, H., Henning, T., & Sternberg, A. 2017, *ApJ*, 835, 126
- Bischetti, M., Piconcelli, E., Vietri, G., et al. 2017, *A&A*, 598, A122
- Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, *ARA&A*, 51, 207
- Borisova, E., Cantalupo, S., Lilly, S. J., et al. 2016, *ApJ*, 831, 39
- Carilli, C. L., & Walter, F. 2013, *ARA&A*, 51, 105
- Carswell, R. F., & Webb, J. K. 2014, Astrophysics Source Code Library [record ascl:1408.015]
- Christensen, L., Møller, P., Rhodin, N. H. P., Heintz, K. E., & Fynbo, J. P. U. 2019, *MNRAS*, 489, 2270
- De Cia, A., Ledoux, C., Mattsson, L., et al. 2016, *A&A*, 596, A97
- de Jong, R. S., Agertz, O., Berbel, A. A., et al. 2019, *The Messenger*, 175, 3
- Duley, W. W. 1980, *MNRAS*, 190, 683
- Eilers, A.-C., Hennawi, J. F., Davies, F. B., & Simcoe, R. A. 2021, *ApJ*, 917, 38
- Ellison, S. L., Yan, L., Hook, I. M., et al. 2002, *A&A*, 383, 91
- Ellison, S. L., Prochaska, J. X., Hennawi, J., et al. 2010, *MNRAS*, 406, 1435
- Fathivavari, H., Petitjean, P., Jamialahmadi, N., et al. 2018, *MNRAS*, 477, 5625
- Finley, H., Petitjean, P., Pâris, I., et al. 2013, *A&A*, 558, A111
- Fox, A. J., Prochaska, J. X., Ledoux, C., et al. 2009, *A&A*, 503, 731
- Fynbo, J. P. U., Laursen, P., Ledoux, C., et al. 2010, *MNRAS*, 408, 2128
- Gaskell, C. M. 1982, *ApJ*, 263, 79
- Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., & Wolff, M. J. 2003, *ApJ*, 594, 279

- Gupta, N., Srianand, R., Baan, W., et al. 2016, in *MeerKAT Science: On the Pathway to the SKA*, 14
- Gupta, N., Shukla, G., Srianand, R., et al. 2022a, *ApJ*, **929**, 108
- Gupta, N., Srianand, R., Momjian, E., et al. 2022b, *ApJ*, **927**, L24
- Hamann, F., Chartas, G., Reeves, J., & Nardini, E. 2018, *MNRAS*, **476**, 943
- Hopkins, A. M., McClure-Griffiths, N. M., & Gaensler, B. M. 2008, *ApJ*, **682**, L13
- Horne, K. 1986, *PASP*, **98**, 609
- Joshi, R., Srianand, R., Noterdaeme, P., & Petitjean, P. 2017, *MNRAS*, **465**, 701
- Konstantopoulou, C., De Cia, A., Krogager, J.-K., et al. 2022, *A&A*, **666**, A12
- Krogager, J. K., Fynbo, J. P. U., Noterdaeme, P., et al. 2016, *MNRAS*, **455**, 2698
- Krogager, J. K., Gupta, N., Noterdaeme, P., et al. 2018, *ApJS*, **235**, 10
- Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2008, *ApJ*, **689**, 865
- Ledoux, C., Petitjean, P., & Srianand, R. 2003, *MNRAS*, **346**, 209
- Ledoux, C., Petitjean, P., & Srianand, R. 2006, *ApJ*, **640**, L25
- Ledoux, C., Vreeswijk, P. M., Smette, A., et al. 2009, *A&A*, **506**, 661
- Ledoux, C., Noterdaeme, P., Petitjean, P., & Srianand, R. 2015, *A&A*, **580**, A8
- Leibundgut, B., & Robertson, J. G. 1999, *MNRAS*, **303**, 711
- Lodders, K. 2003, *ApJ*, **591**, 1220
- Ma, J., Caucal, P., Noterdaeme, P., et al. 2015, *MNRAS*, **454**, 1751
- Merloni, A., Alexander, D. A., Banerji, M., et al. 2019, *The Messenger*, **175**, 42
- Modigliani, A., Goldoni, P., Royer, F., et al. 2010, *Proc. SPIE*, **7737**, 773728
- Møller, P., & Christensen, L. 2020, *MNRAS*, **492**, 4805
- Møller, P., & Warren, S. J. 1993, *A&A*, **270**, 43
- Møller, P., Fynbo, J. P. U., Ledoux, C., & Nilsson, K. K. 2013, *MNRAS*, **430**, 2680
- Neeleman, M., Wolfe, A. M., Prochaska, J. X., & Rafelski, M. 2013, *ApJ*, **769**, 54
- Neufeld, D. A., & Wolfire, M. G. 2016, *ApJ*, **826**, 183
- Nielsen, N. M., Kacprzak, G. G., Sameer, S., et al. 2022, *MNRAS*, **514**, 6074
- Noterdaeme, P., Ledoux, C., Petitjean, P., et al. 2007, *A&A*, **474**, 393
- Noterdaeme, P., Petitjean, P., Ledoux, C., Srianand, R., & Ivanchik, A. 2008, *A&A*, **491**, 397
- Noterdaeme, P., Petitjean, P., Ledoux, C., & Srianand, R. 2009, *A&A*, **505**, 1087
- Noterdaeme, P., Laursen, P., Petitjean, P., et al. 2012, *A&A*, **540**, A63
- Noterdaeme, P., Petitjean, P., Pâris, I., et al. 2014, *A&A*, **566**, A24
- Noterdaeme, P., Petitjean, P., & Srianand, R. 2015, *A&A*, **578**, L5
- Noterdaeme, P., Krogager, J.-K., Balashev, S., et al. 2017, *A&A*, **597**, A82
- Noterdaeme, P., Ledoux, C., Zou, S., et al. 2018, *A&A*, **612**, A58
- Noterdaeme, P., Balashev, S., Krogager, J. K., et al. 2019, *A&A*, **627**, A32
- Noterdaeme, P., Balashev, S., Krogager, J. K., et al. 2021a, *A&A*, **646**, A108
- Noterdaeme, P., Balashev, S., Combes, F., et al. 2021b, *A&A*, **651**, A17
- Omont, A., Petitjean, P., Guilloteau, S., et al. 1996, *Nature*, **382**, 428
- Perrotta, S., D'Odorico, V., Prochaska, J. X., et al. 2016, *MNRAS*, **462**, 3285
- Perrotta, S., D'Odorico, V., Hamann, F., et al. 2018, *MNRAS*, **481**, 105
- Petitjean, P., Ledoux, C., Noterdaeme, P., & Srianand, R. 2006, *A&A*, **456**, L9
- Prochaska, J. X., & Wolfe, A. M. 1997, *ApJ*, **487**, 73
- Prochaska, J. X., Tripp, T. M., & Howk, J. C. 2005, *ApJ*, **620**, L39
- Prochaska, J. X., Chen, H.-W., Wolfe, A. M., Dessauges-Zavadsky, M., & Bloom, J. S. 2008, *ApJ*, **672**, 59
- Rahmani, H., Srianand, R., Noterdaeme, P., & Petitjean, P. 2010, *MNRAS*, **409**, L59
- Ranjan, A., Noterdaeme, P., Krogager, J. K., et al. 2018, *A&A*, **618**, A184
- Ranjan, A., Noterdaeme, P., Krogager, J. K., et al. 2020, *A&A*, **633**, A125
- Rix, S. A., Pettini, M., Steidel, C. C., et al. 2007, *ApJ*, **670**, 15
- Schilke, P., Neufeld, D. A., Müller, H. S. P., et al. 2014, *A&A*, **566**, A29
- Schindler, J.-T., Farina, E. P., Bañados, E., et al. 2020, *ApJ*, **905**, 51
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, **737**, 103
- Selsing, J., Fynbo, J. P. U., Christensen, L., & Krogager, J.-K. 2016, *A&A*, **585**, A87
- Shen, Y., Brandt, W. N., Richards, G. T., et al. 2016, *ApJ*, **831**, 7
- Smette, A., Sana, H., Noll, S., et al. 2015, *A&A*, **576**, A77
- Sofia, U. J., & Jenkins, E. B. 1998, *ApJ*, **499**, 951
- Srianand, R., Gupta, N., Petitjean, P., Noterdaeme, P., & Saikia, D. J. 2008, *MNRAS*, **391**, L69
- Tytler, D., & Fan, X.-M. 1992, *ApJS*, **79**, 1
- Urry, C. M., & Padovani, P. 1995, *PASP*, **107**, 803
- Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, *AJ*, **122**, 549
- van Dokkum, P. G. 2001, *PASP*, **113**, 1420
- Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, *A&A*, **536**, A105
- Vladilo, G., Centurión, M., Levshakov, S. A., et al. 2006, *A&A*, **454**, 151
- Wang, R., Wu, X.-B., Neri, R., et al. 2016, *ApJ*, **830**, 53
- Weiß, A., Downes, D., Neri, R., et al. 2007, *A&A*, **467**, 955
- Welty, D. E., Frisch, P. C., Sonneborn, G., & York, D. G. 1999, *ApJ*, **512**, 636
- Zafar, T., Vladilo, G., Péroux, C., et al. 2014a, *MNRAS*, **445**, 2093
- Zafar, T., Centurión, M., Péroux, C., et al. 2014b, *MNRAS*, **444**, 744
- Zhang, S., Ge, J., Jiang, P., et al. 2015, *ApJ*, **802**, 92
- Zou, S., Petitjean, P., Noterdaeme, P., et al. 2018, *A&A*, **616**, A158

## Appendix A: Fit to H<sub>I</sub> Ly- $\alpha$ and H<sub>2</sub> Lyman-Werner absorption bands



**Fig. A.1.** Damped H I-Ly- $\alpha$  lines for the proximate H<sub>2</sub> systems. The X-shooter data are shown in black, and the profile corresponding to best fit is shown in red. The red shadows represent the conservative range of  $N(\text{H I})$ , including continuum placement uncertainties. There is residual Ly- $\alpha$  emission in the core of several DLAs. We also mark the expected position of the NV  $\lambda\lambda$ 1238, 1242 doublet in blue. The tick marks represent the range seen in low-ionisation species (i.e.  $z_5$  and  $z_{95}$ ). It is here detected in all cases but J1259+0309 and J1331+0206.



**Fig. A.2.** Portion of an X-shooter spectrum of J0019–0137 featuring H<sub>2</sub> absorption bands. The solid red line shows the overall absorption model, and the contribution from non-H<sub>2</sub> lines (H I, metals and lines from intervening systems) is shown as a dashed profile. The blue segments connect lines from different rotational levels (shown here for convenience from J=0 to J=5) of a given band (indicated by the label above, 'L' for Lyman, 'W' for Werner, and the following number is the upper vibrational level).

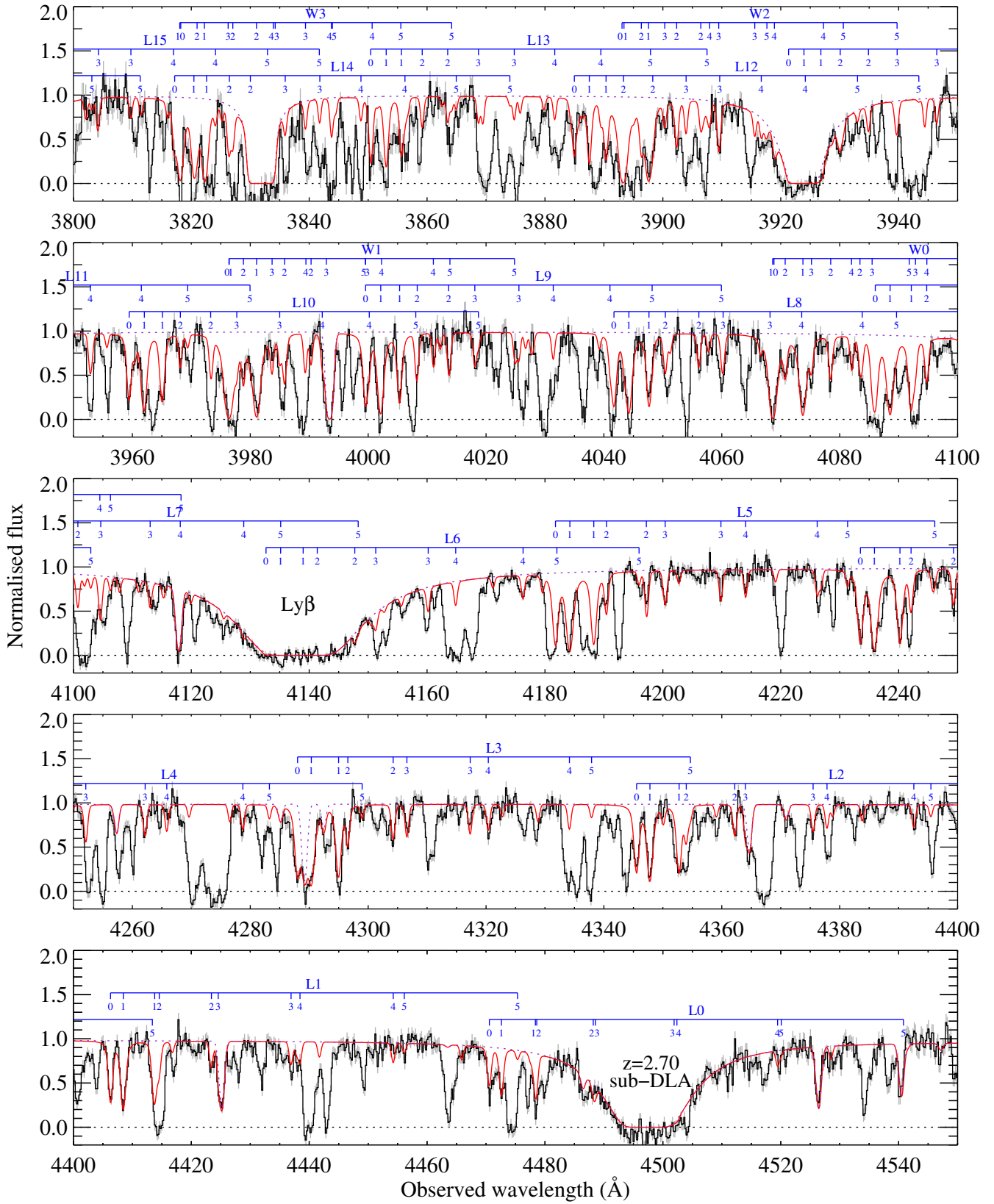


Fig. A.3. Same as Fig. A.2 for the quasar J0059+1124.

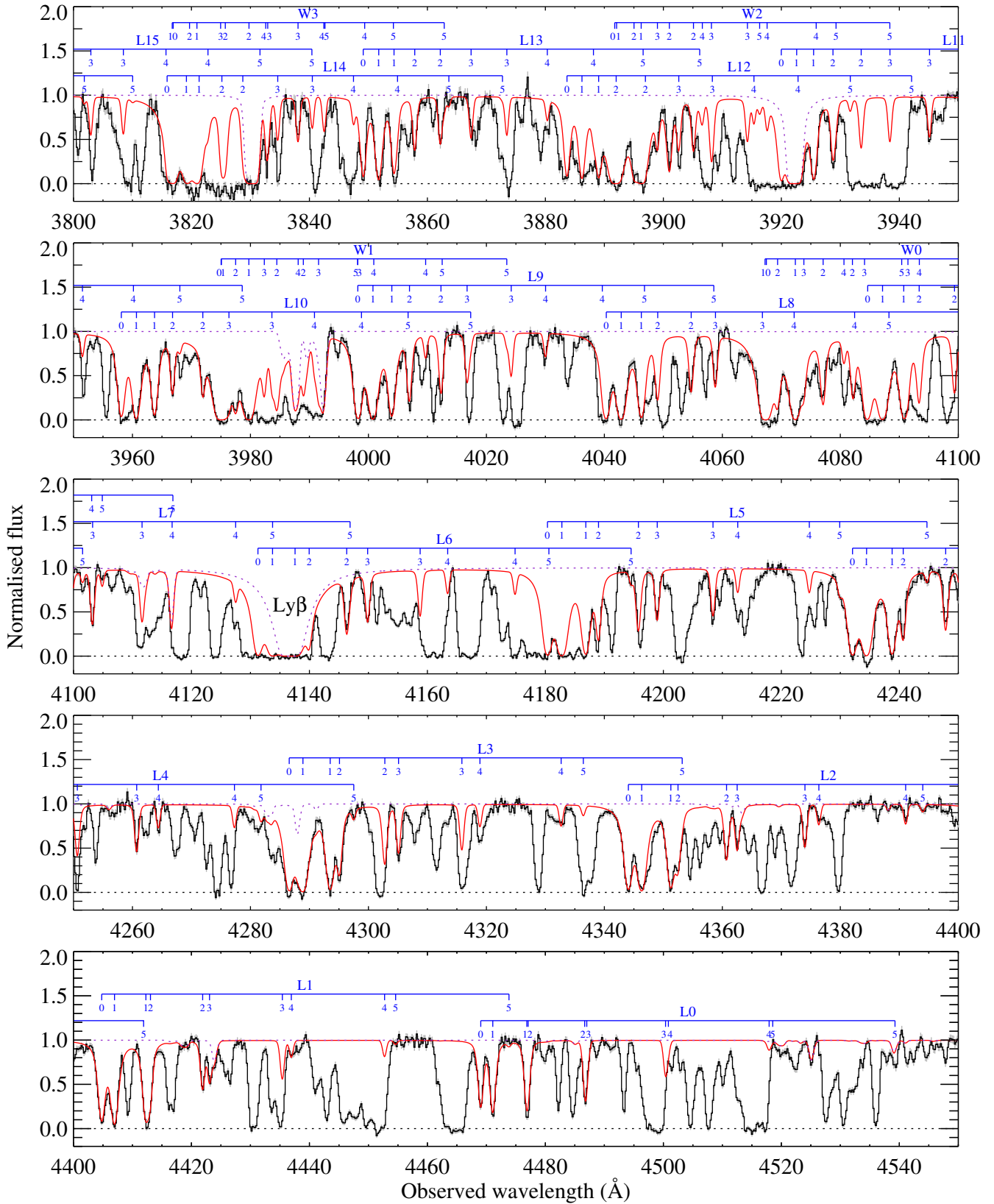


Fig. A.4. Same as Fig. A.2 for the quasar J1236+0010.

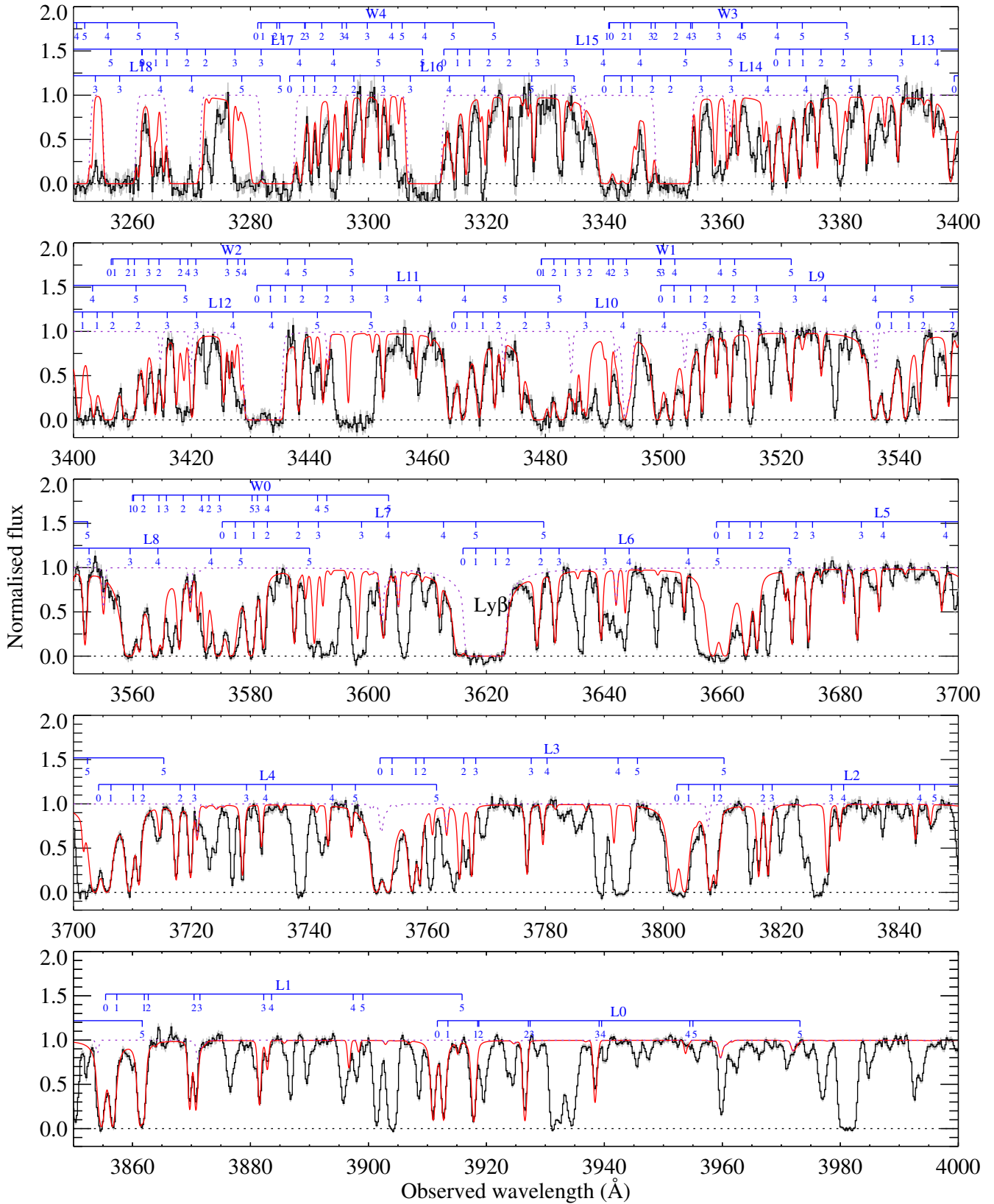
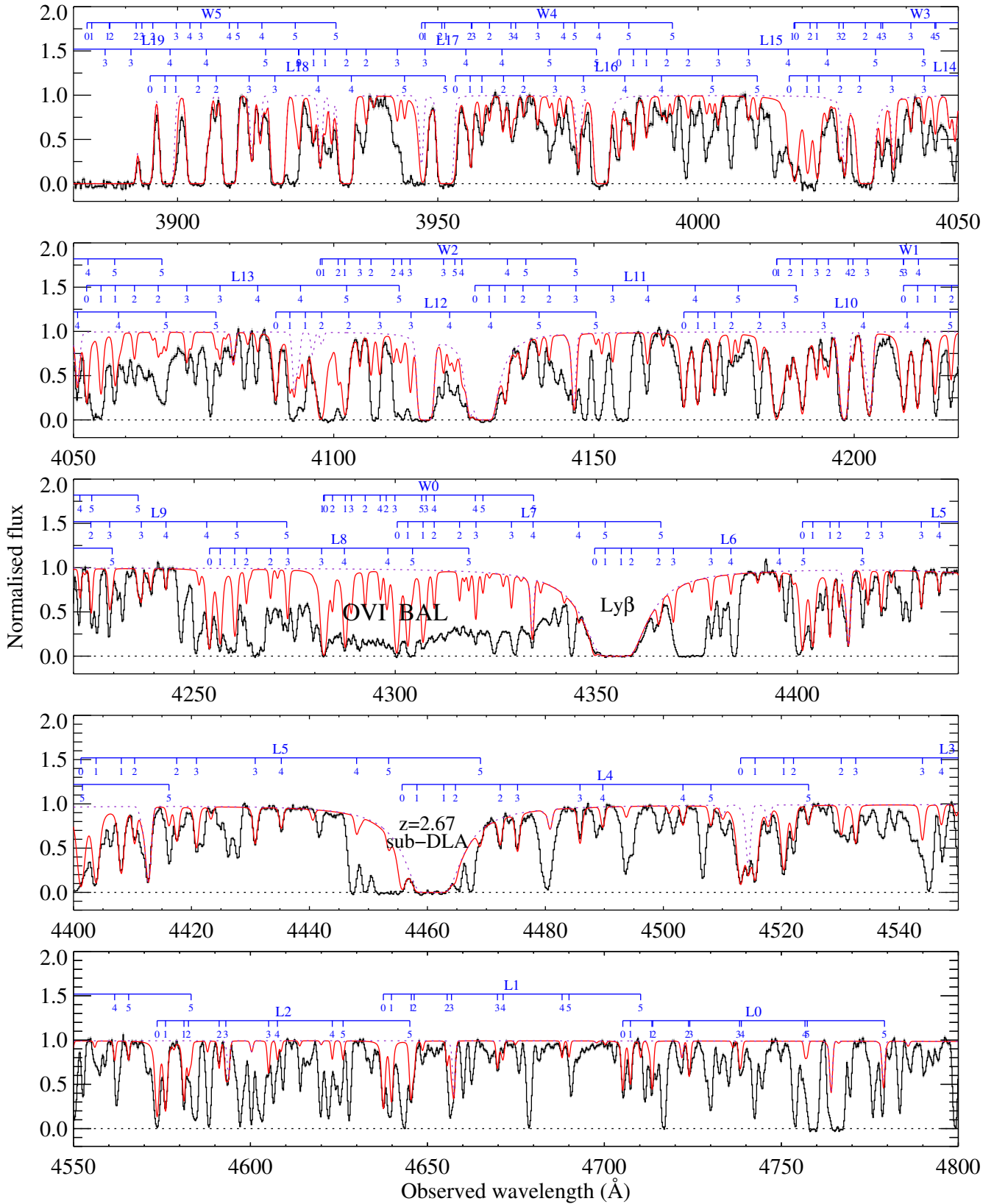


Fig. A.5. Same as Fig. A.2 for the quasar J1248+0639.



**Fig. A.6.** Same as Fig. A.2 for the quasar J1259+0309. A high-velocity broad absorption line (BAL) system lies in this quasar.

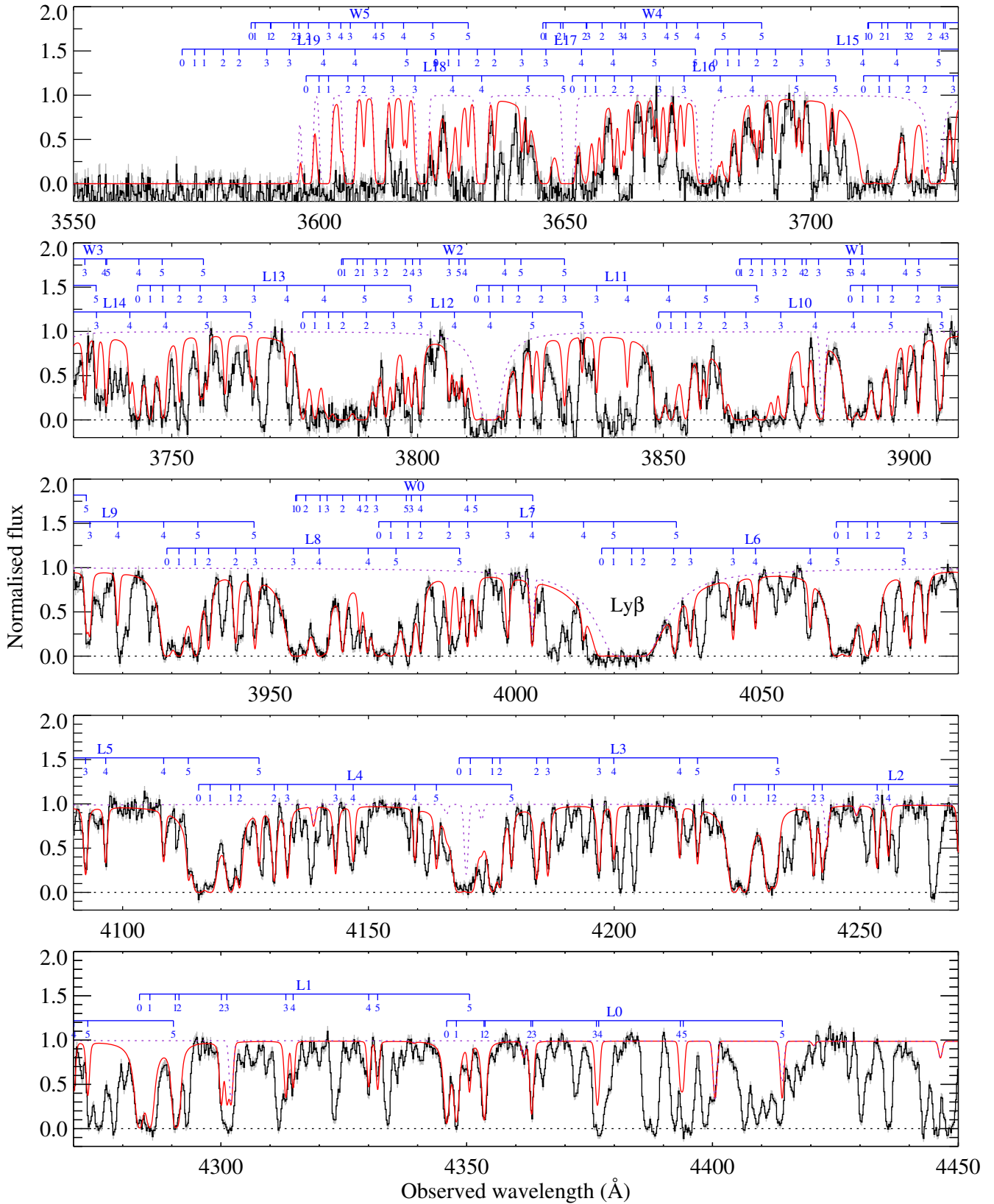


Fig. A.7. Same as Fig. A.2 for the quasar J1331+0206.

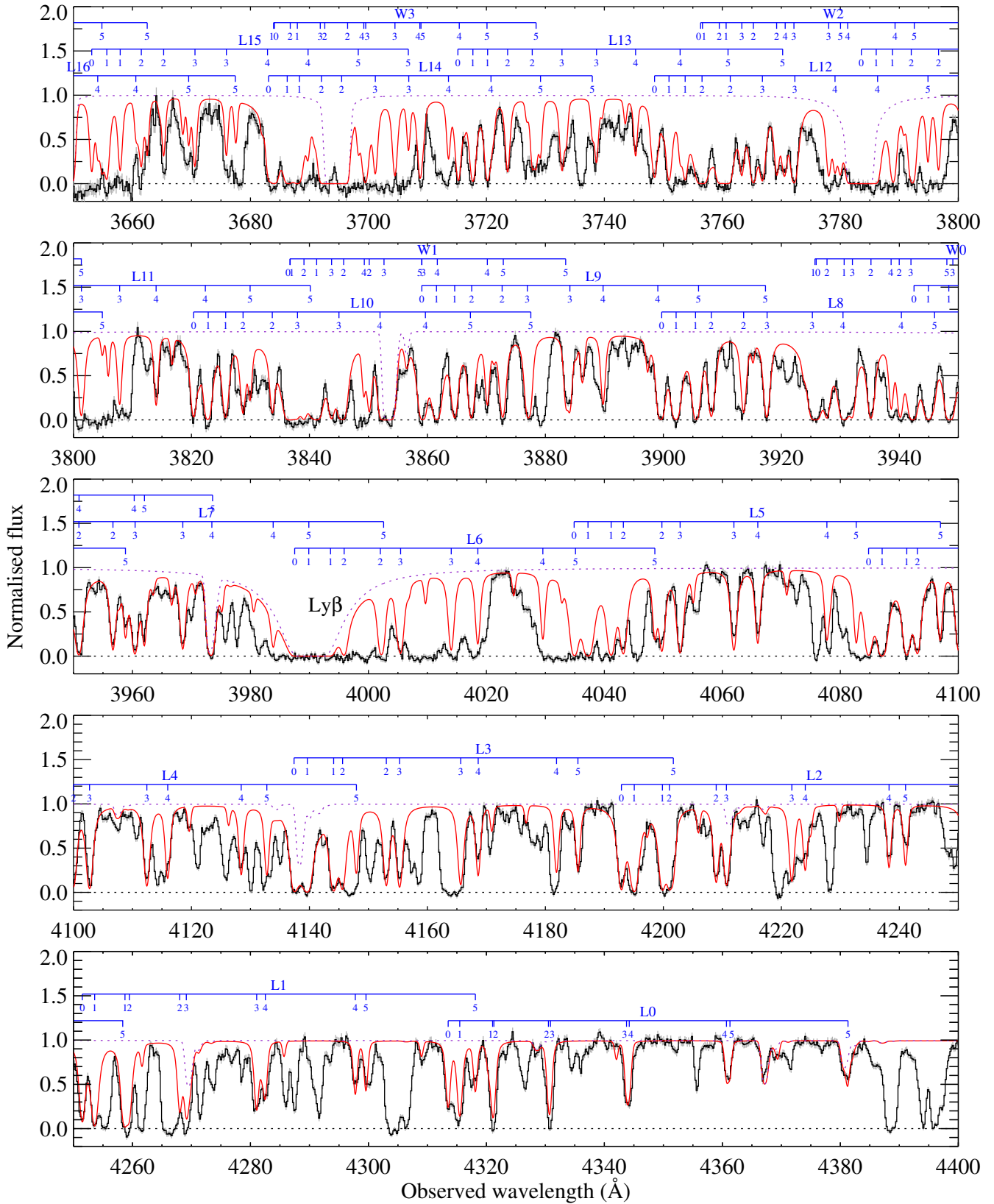


Fig. A.8. Same as Fig. A.2 for the quasar J1358+1410.

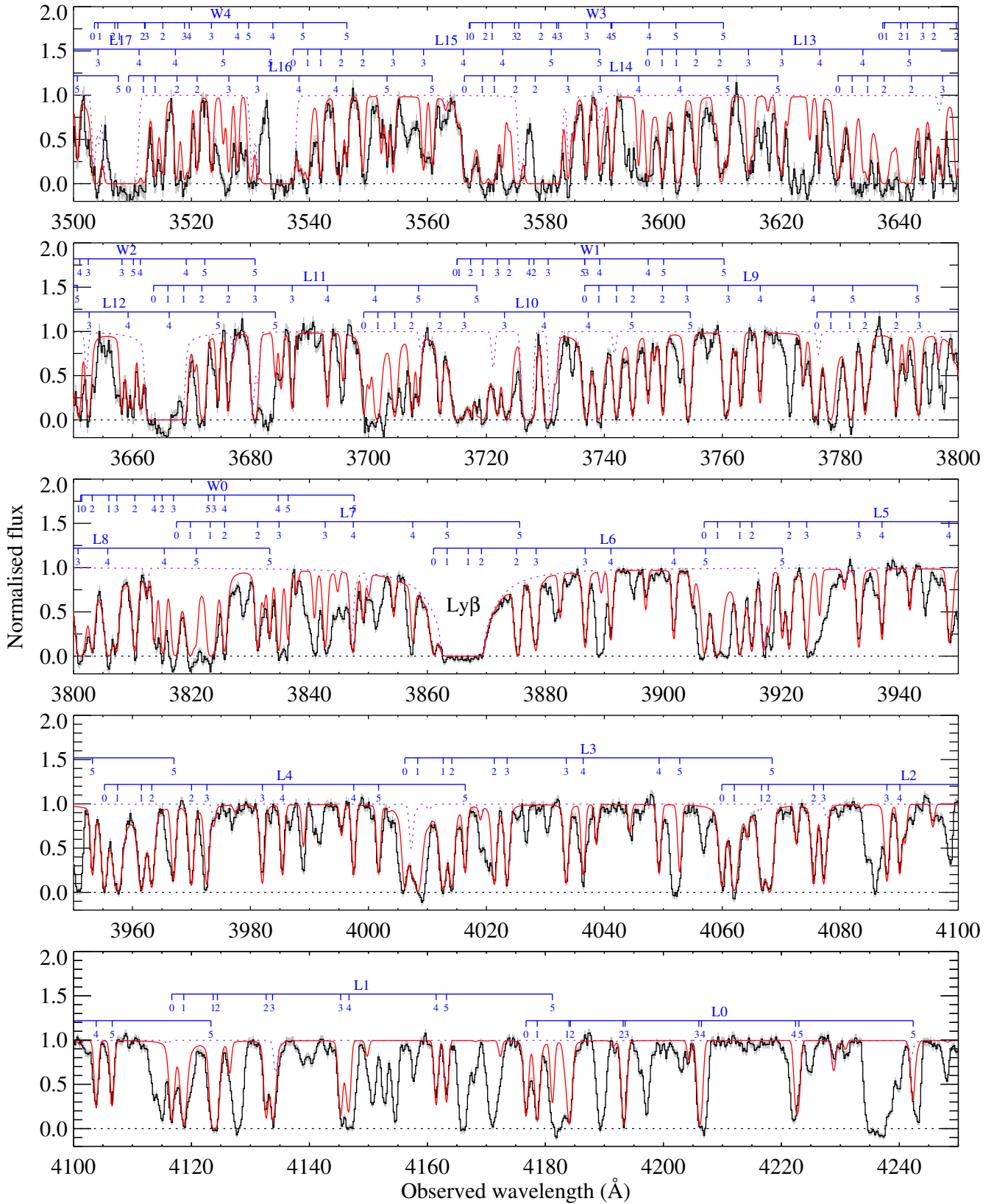


Fig. A.9. Same as Fig. A.2 for the quasar J2228-0221.

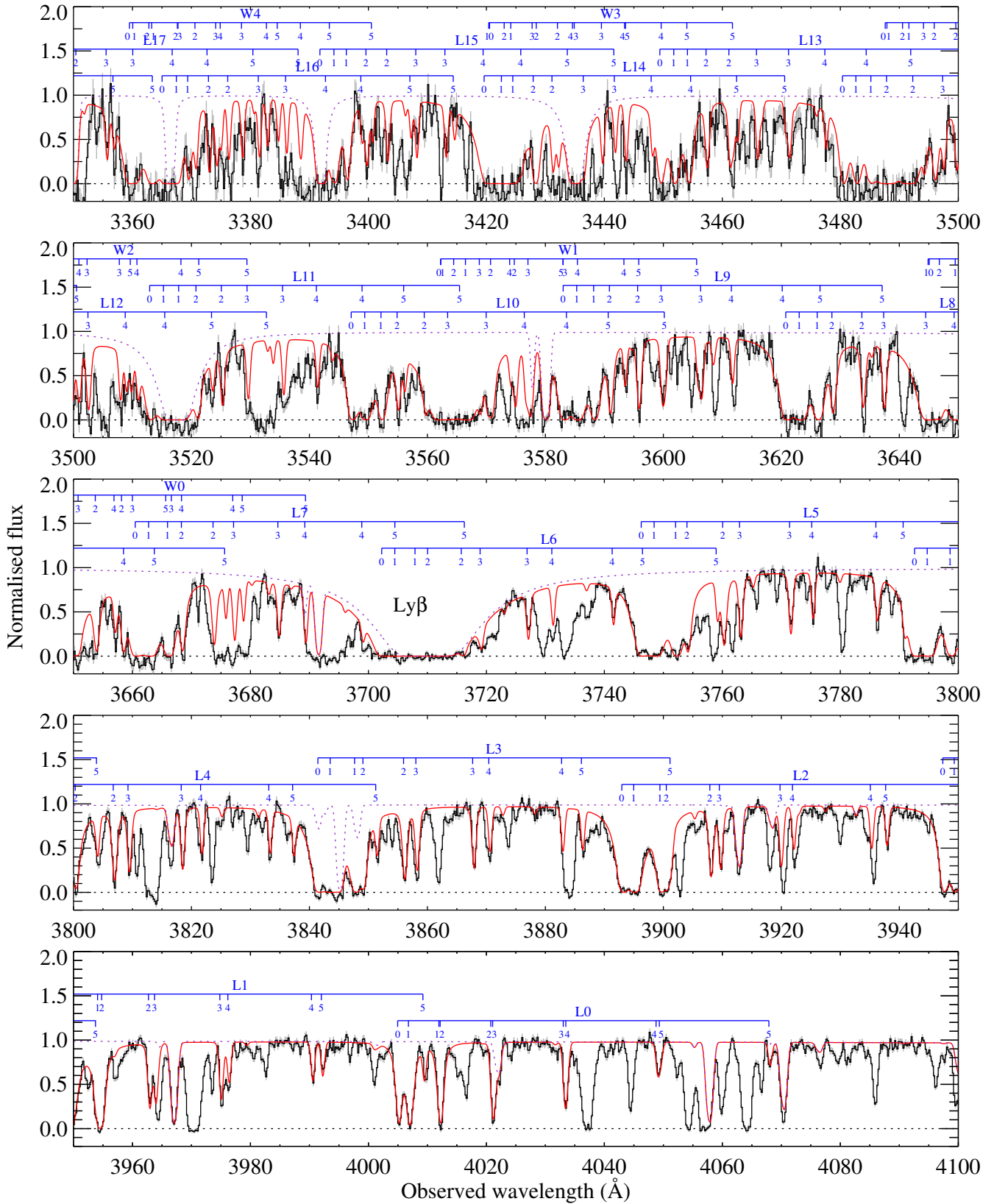
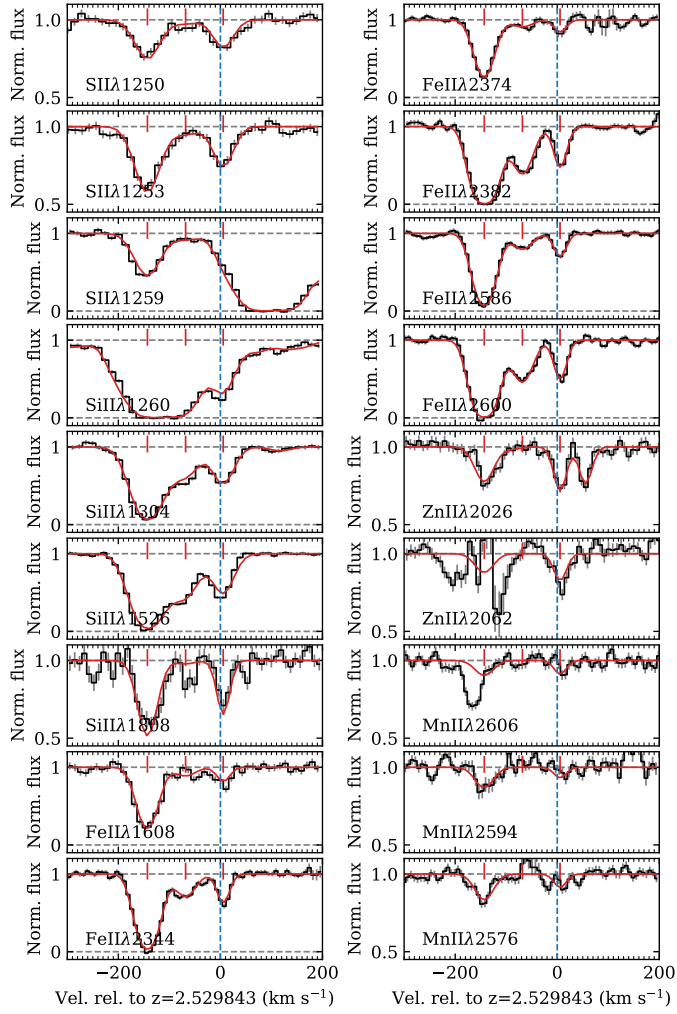
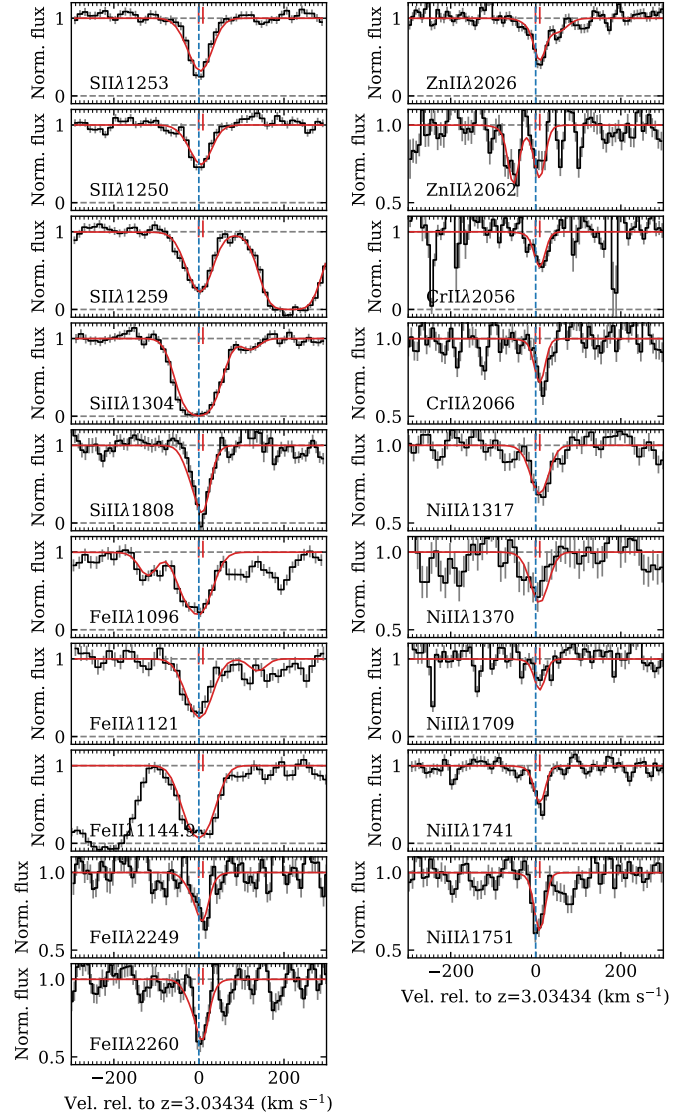


Fig. A.10. Same as Fig. A.2 for the quasar J2325+1539.

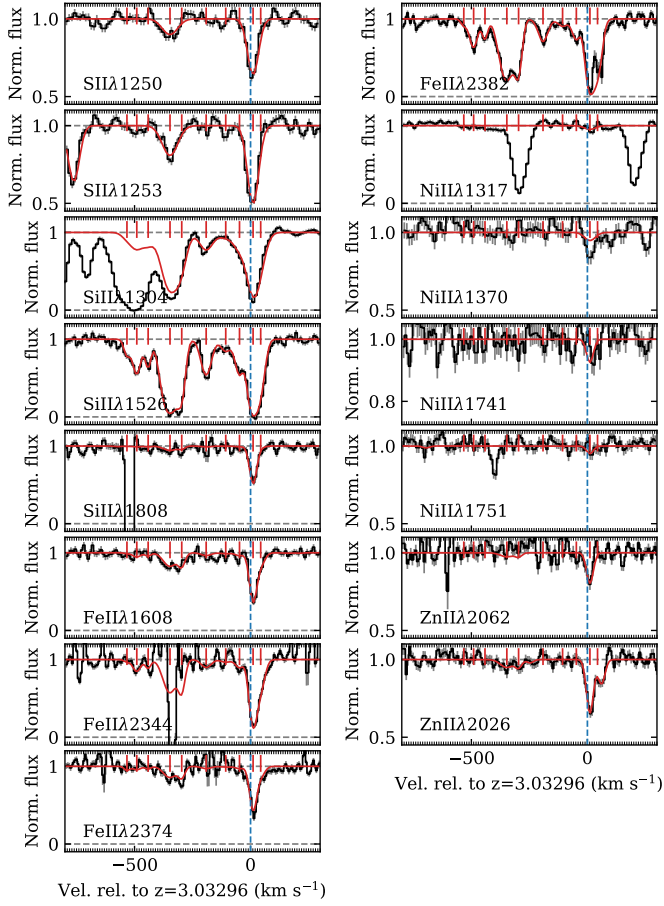
**Appendix B: Low-ionisation metal absorption lines**



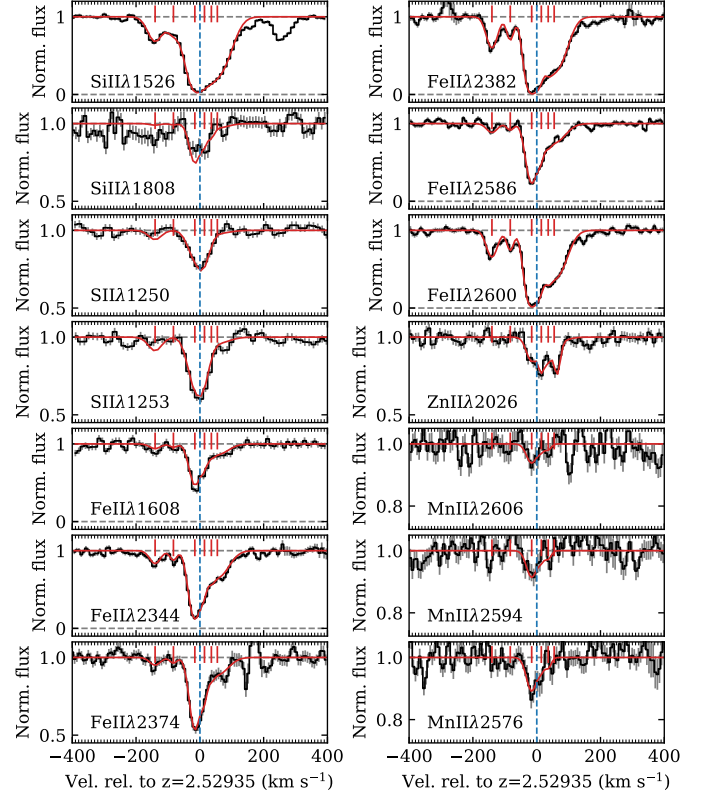
**Fig. B.1.** Main singly ionised metal lines in the proximate system towards J0019–0137 in their ground state. The zero of the velocity scale is set to the redshift of the  $H_2$  component (dashed blue line). Short red ticks mark the position of the different components used in the model (overplotted in red).



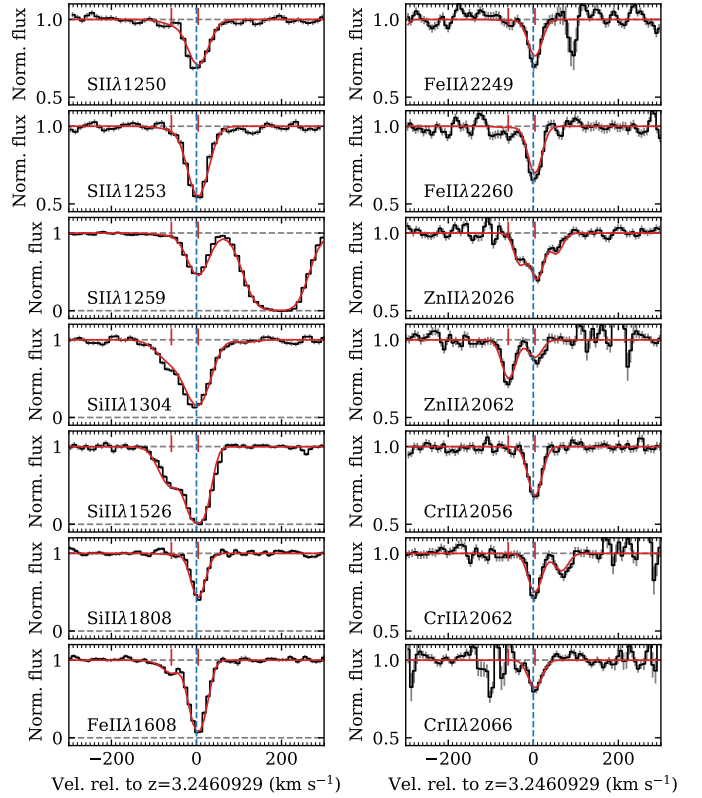
**Fig. B.2.** Singly ionised metal lines in the proximate system towards J0059+1124.



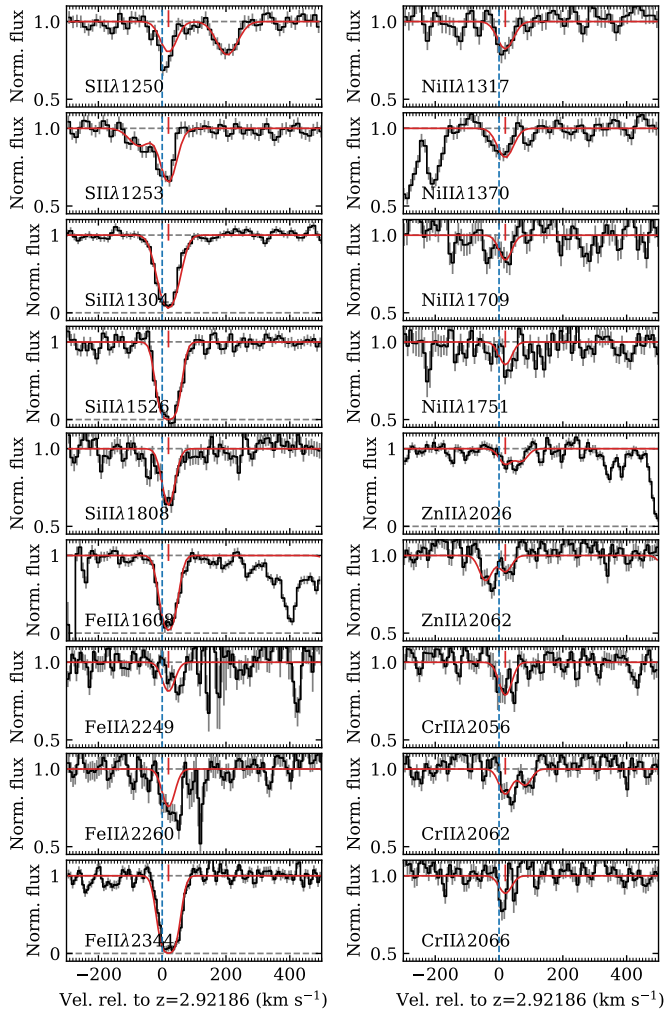
**Fig. B.3.** Singly ionised metal lines in the proximate system towards J1236+0010. The spike at  $v=0$  in Fe II $\lambda$ 2382 is a sky-line residual.



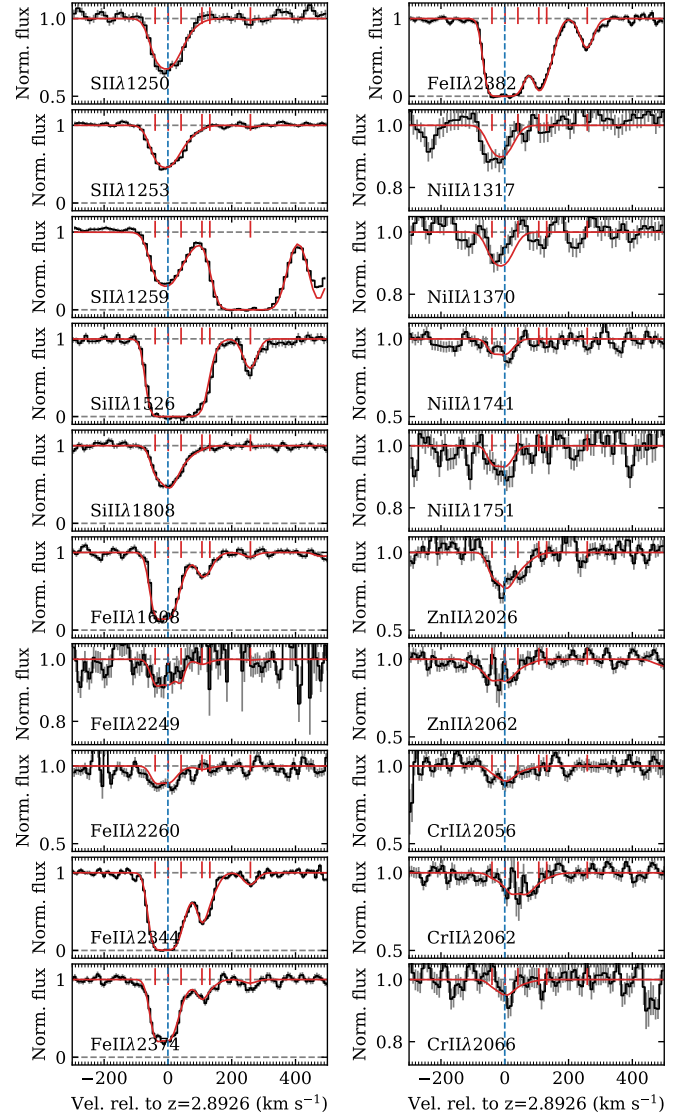
**Fig. B.4.** Singly ionised metal lines in the proximate system towards J1248+0639.



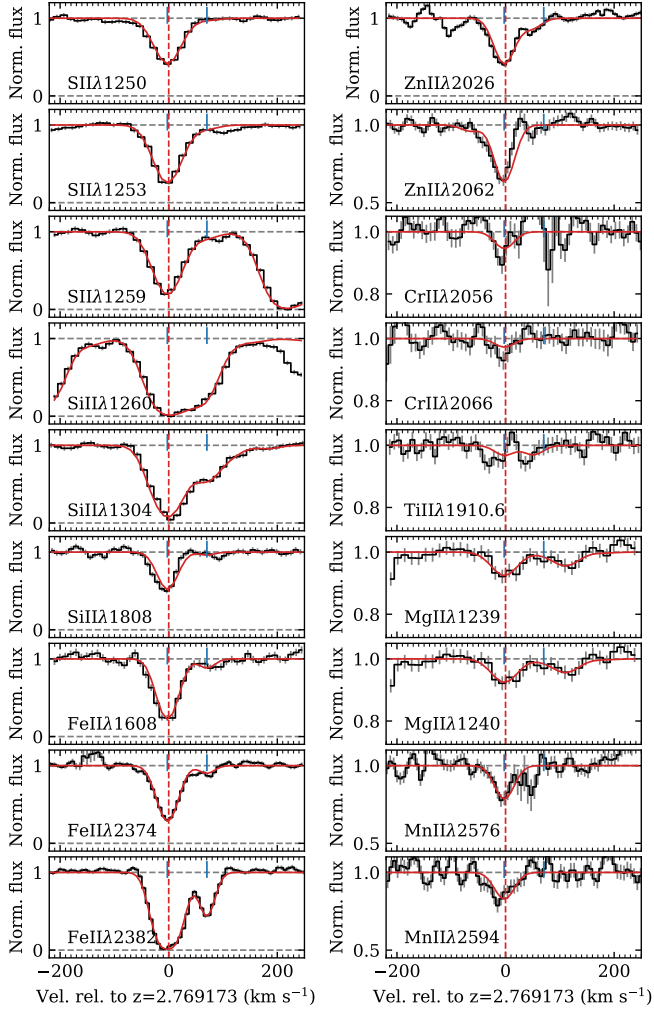
**Fig. B.5.** Singly ionised metal lines in the proximate system towards J1259+0309.



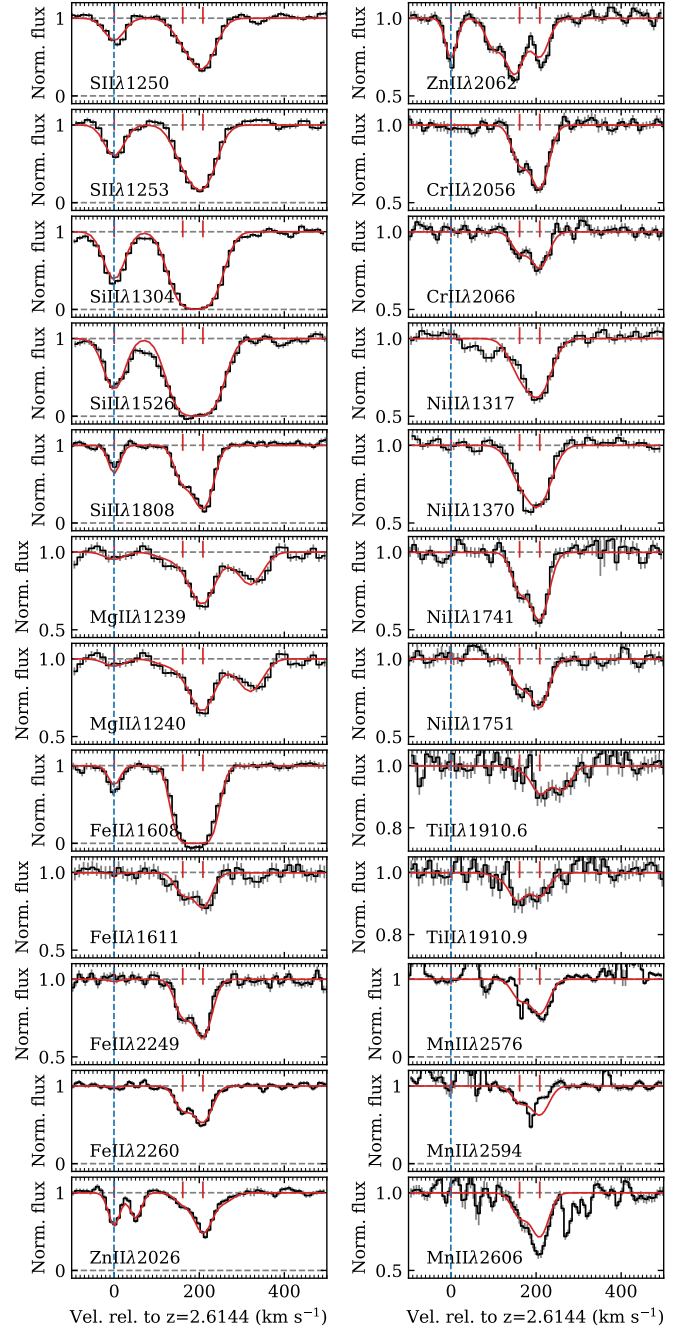
**Fig. B.6.** Singly ionised metal lines in the proximate system towards J1331+0206.



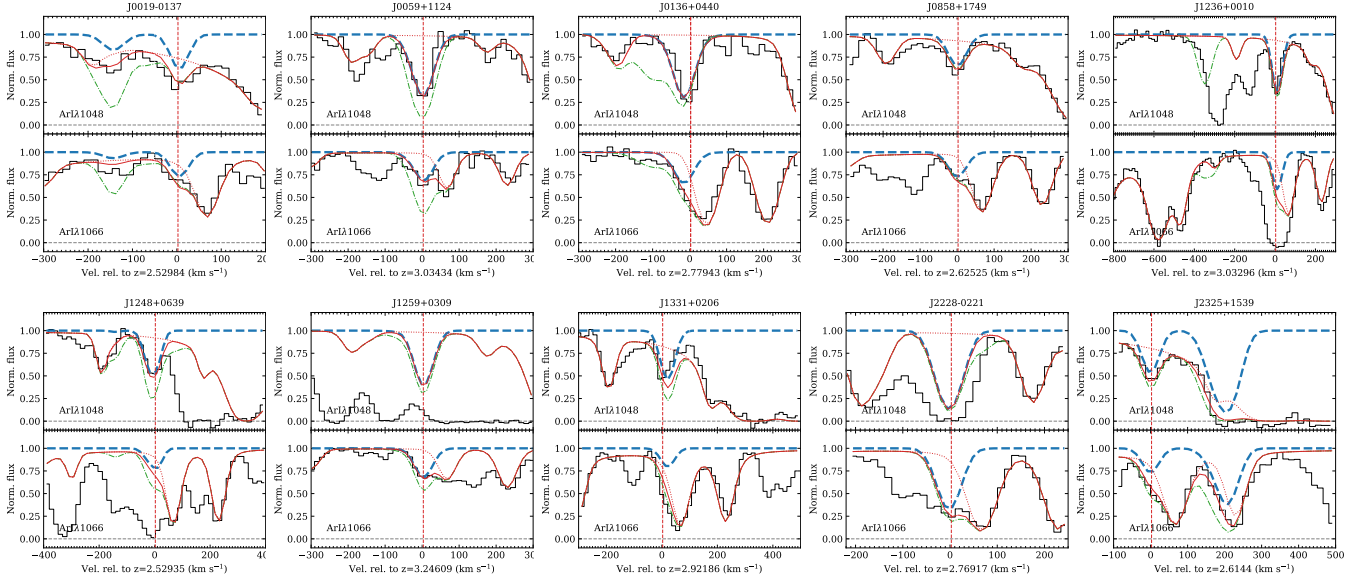
**Fig. B.7.** Singly ionised metal lines in the proximate system towards J1358+1410.



**Fig. B.8.** Singly ionised metal lines in the proximate system towards J2228-0221..



**Fig. B.9.** Singly ionised metal lines in the proximate system towards J2325+1539. Mn II is close to the sky residuals.



**Fig. B.10.** Neutral argon absorption lines. The dashed blue profile and dotted red and solid red represents the Ar I H<sub>2</sub> and total absorption, respectively. The dashed green profile represents the expected total absorption obtained by scaling the S II profile assuming all argon to be in neutral form and a solar [Ar/S] ratio. The vertical dotted line shows the location of the H<sub>2</sub> component.

**Table B.1.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J0019–0137.

$z$	$b$ (km s <sup>-1</sup> )	log $N$ [cm <sup>-2</sup> ]				
		S II	Si II	Fe II	Zn II	Mn II
2.5282	19.4±0.2	15.03±0.02	15.50±0.04	14.53±0.01	12.54±0.05	12.52±0.04
2.5290	20.9±0.8	13.97±0.10	14.01±0.02	13.36±0.01	...	...
2.5299	5.7±0.2	14.89±0.06	15.29±0.09	13.48±0.03	12.60±0.05	12.06±0.24
Total	-	15.29±0.03	15.71±0.04	14.59±0.01	12.87±0.04	12.65±0.07

**Notes.** The adopted spectral resolutions are 44.3 and 27.0 km s<sup>-1</sup> for the UVB and VIS arm, respectively.

**Table B.2.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J0059+1124.

$z$	$b$ (km s <sup>-1</sup> )	log $N$ [cm <sup>-2</sup> ]					
		S II	Si II	Fe II	Zn II	Cr II	Ni II
3.0343	29.0±1.1	15.22±0.05	15.72±0.09	15.02±0.06	12.38±0.24	12.77±0.55	13.13±0.70
3.0345	7.8±2.2	15.26±0.23	16.09±0.45	14.77±0.14	12.86±0.11	13.57±0.10	14.10±0.10
Total	-	15.54±0.12	16.25±0.32	15.21±0.06	12.98±0.10	13.64±0.11	14.14±0.11

**Notes.** The adopted spectral resolutions are 45.0 and 25.0 km s<sup>-1</sup> for the UVB and VIS arm, respectively.

**Table B.3.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J1236+0010.

$z$	$b$ ( $\text{km s}^{-1}$ )	$\log N [\text{cm}^{-2}]$				
		S II	Si II	Fe II	Zn II	Ni II
3.0258	5.8±0.5	...	13.18±0.11	12.12±0.15	...	...
3.0264	17.3±1.8	...	13.70±0.04	12.99±0.03	...	...
3.0270	5.9±1.0	...	13.57±0.10	12.74±0.07	...	...
3.0283	35.0±1.0	14.71±0.04	14.61±0.03	13.73±0.02	12.02±0.29	...
3.0290	11.5±1.0	...	14.12±0.08	13.43±0.05	11.73±0.24	...
3.0304	25.1±1.0	...	13.79±0.02	12.90±0.04	...	...
3.0315	23.3±3.4	...	13.21±0.02	12.40±0.11	...	...
3.0323	18.3±0.3	...	13.71±0.03	12.83±0.05	...	...
3.0331	12.9±0.4	15.30±0.03	15.43±0.04	14.24±0.02	12.69±0.03	13.16±0.10
3.0335	18.0±0.9	...	13.95±0.04	13.48±0.04	...	...
Total	-	15.40±0.03	15.55±0.03	14.51±0.01	12.81±0.06	13.16±0.10

**Notes.** The adopted spectral resolutions are 50.0 and 26.5  $\text{km s}^{-1}$  for the UVB and VIS arm, respectively.

**Table B.4.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J1248+0639.

$z$	$b$ ( $\text{km s}^{-1}$ )	$\log N [\text{cm}^{-2}]$				
		S II	Si II	Fe II	Zn II	Mn II
2.5277	19.6±0.8	14.01±0.08	13.62±0.02	13.03±0.02	...	...
2.5284	8.5±0.6	...	13.35±0.04	12.72±0.03	...	...
2.5292	18.7±0.3	14.79±0.03	15.08±0.04	14.11±0.01	12.30±0.07	12.27±0.06
2.5295	2.7±0.4	15.12±0.26	14.45±0.28	13.52±0.10	12.25±0.14	11.53±0.33
2.5298	3.4±0.9	...	14.25±0.31	12.87±0.09	11.86±0.33	11.53±0.37
2.5300	44.3±1.2	14.02±0.18	14.37±0.01	13.69±0.01	11.96±0.53	...
Total	-	15.33±0.16	15.29±0.05	14.37±0.02	12.73±0.11	12.41±0.08

**Notes.** The adopted spectral resolutions are 40.0 and 21.0  $\text{km s}^{-1}$  for the UVB and VIS arm, respectively.

**Table B.5.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J1259+0309.

$z$	$b$ ( $\text{km s}^{-1}$ )	$\log N [\text{cm}^{-2}]$				
		S II	Si II	Fe II	Zn II	Cr II
3.2453	32.2±0.5	13.59±0.29	13.95±0.02	13.53±0.04	...	...
3.2461	14.5±0.2	15.15±0.02	15.59±0.02	15.01±0.02	12.46±0.05	13.46±0.03
Total	-	15.17±0.02	15.60±0.02	15.03±0.02	12.46±0.05	13.46±0.03

**Notes.** The adopted spectral resolutions are 50.0 and 28.5  $\text{km s}^{-1}$  for the UVB and VIS arm, respectively.

**Table B.6.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J1331+0206.

$z$	$b$ ( $\text{km s}^{-1}$ )	$\log N [\text{cm}^{-2}]$					
		S II	Si II	Fe II	Zn II	Cr II	Ni II
2.9221	21.5±0.4	14.92±0.04	15.35±0.05	14.97±0.04	12.50±0.09	13.32±0.07	13.78±0.05
Total	-	14.92±0.04	15.35±0.05	14.97±0.04	12.50±0.09	13.32±0.07	13.78±0.05

**Notes.** The adopted spectral resolutions are 45.0 and 26.0  $\text{km s}^{-1}$  for the UVB and VIS arm, respectively.

**Table B.7.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J1358+1410.

$z$	$b$ ( $\text{km s}^{-1}$ )	$\log N [\text{cm}^{-2}]$					
		S II	Si II	Fe II	Zn II	Cr II	Ni II
2.8921	15.8±0.6	14.79±0.04	15.10±0.08	14.38±0.03	12.12±0.19	12.26±0.67	13.14±0.16
2.8926	30.6±1.4	15.09±0.03	15.62±0.04	14.66±0.02	12.57±0.11	12.98±0.18	13.51±0.08
2.8931	48.5±2.8	14.80±0.05	15.28±0.07	13.83±0.06	12.42±0.16	12.66±0.44	...
2.8940	15.0±2.3	...	13.75±0.31	13.62±0.04	...	...	...
2.8943	28.4±1.1	...	13.02±0.28	13.46±0.03	...	...	...
2.8959	23.7±1.0	13.54±0.57	13.66±0.02	13.08±0.03	...	...	...
Total	-	15.40±0.02	15.87±0.03	14.93±0.02	12.88±0.08	13.20±0.18	13.66±0.07

**Notes.** The adopted spectral resolutions are 43.0 and 26.0  $\text{km s}^{-1}$  for the UVB and VIS arm, respectively.

**Table B.8.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J2228-0221..

$z$	$b$ ( $\text{km s}^{-1}$ )	$\log N [\text{cm}^{-2}]$							
		S II	Si II	Fe II	Zn II	Cr II	Mn II	Mg II	Ti II
2.7691	18.5±0.2	15.59±0.02	15.49±0.03	14.50±0.01	13.09±0.02	12.64±0.13	12.64±0.05	15.36±0.10	12.44±0.17
2.7701	10.3±0.5	13.92±0.15	14.15±0.06	13.28±0.03	...	...	...	14.10±1.62	...
Total	-	15.60±0.02	15.51±0.02	14.52±0.01	13.09±0.02	12.64±0.13	12.64±0.05	15.38±0.12	12.44±0.17

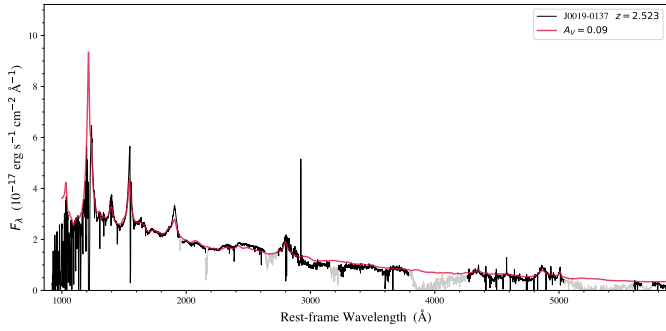
**Notes.** The adopted spectral resolutions are 45.5 and 28.0  $\text{km s}^{-1}$  for the UVB and VIS arm, respectively.

**Table B.9.** Results of Voigt-profile fitting to singly ionised metal absorption lines proximate to J2325+1539.

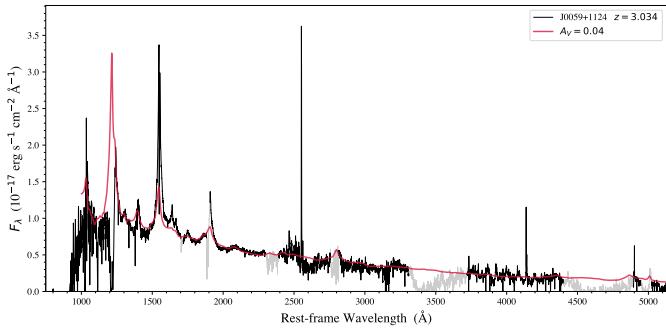
$z$	$b$ ( $\text{km s}^{-1}$ )	$\log N [\text{cm}^{-2}]$								
		S II	Si II	Fe II	Zn II	Cr II	Ni II	Mg II	Ti II	Mn II
2.6144	9.5±0.3	15.19±0.04	15.24±0.06	13.63±0.02	12.80±0.03	...	...	15.14±0.24	...	...
2.6163	19.4±0.5	15.01±0.04	15.46±0.03	15.05±0.02	12.29±0.07	13.35±0.03	13.80±0.03	15.29±0.17	11.97±0.78	12.73±0.03
2.6169	23.3±0.3	15.68±0.02	15.96±0.02	15.34±0.02	12.93±0.04	13.68±0.02	14.17±0.02	16.14±0.03	12.90±0.10	13.09±0.02
Total	-	15.87±0.02	16.14±0.02	15.53±0.01	13.23±0.02	13.84±0.01	14.32±0.01	16.23±0.04	12.95±0.12	13.25±0.01

**Notes.** The adopted spectral resolutions are 50.0 and 26.5  $\text{km s}^{-1}$  for the UVB and VIS arm, respectively.

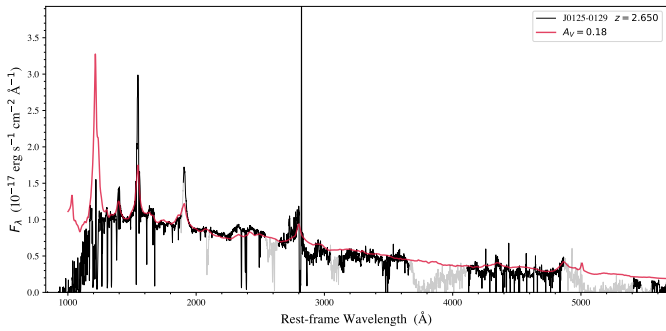
Appendix C: Measurement of dust extinction



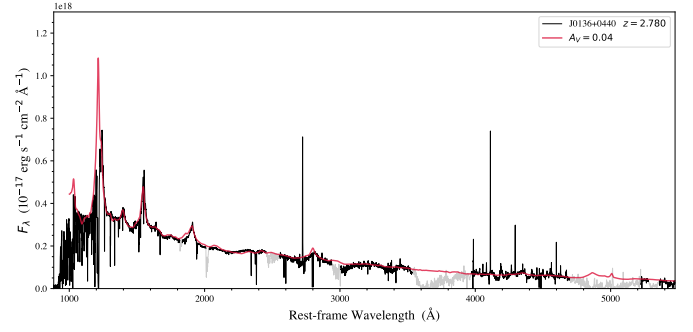
**Fig. C.1.** Measurement of the dust extinction by matching the X-shooter spectrum of J0019–0137 (black) with a reddened quasar composite spectrum from [Selsing et al. \(2016\)](#) (red). Grey regions were masked out during the fit.



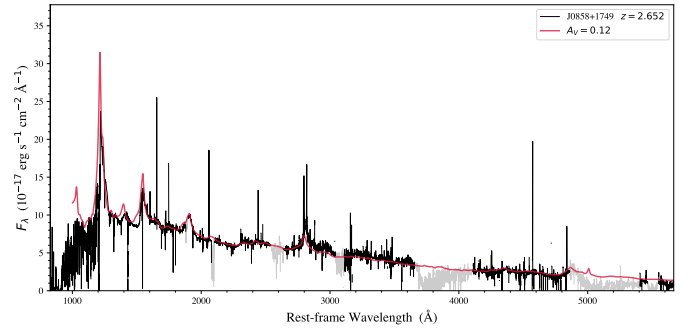
**Fig. C.2.** Same as Fig. C.1 for the quasar J0059+1124.



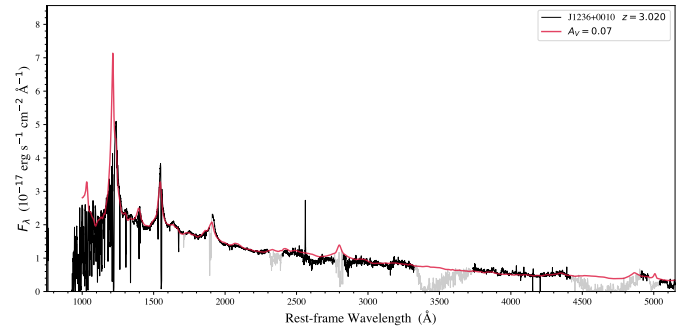
**Fig. C.3.** Same as Fig. C.1 for the quasar J0125–0129.



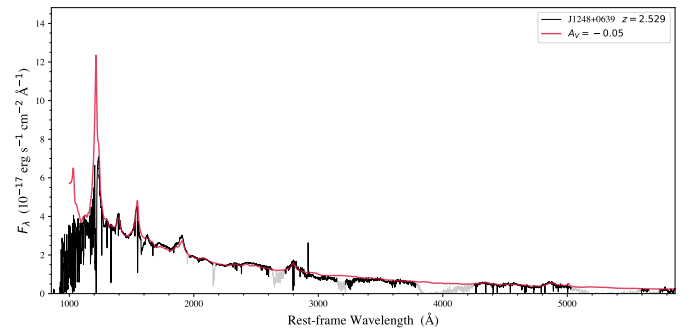
**Fig. C.4.** Same as Fig. C.1 for the quasar J0136+0440.



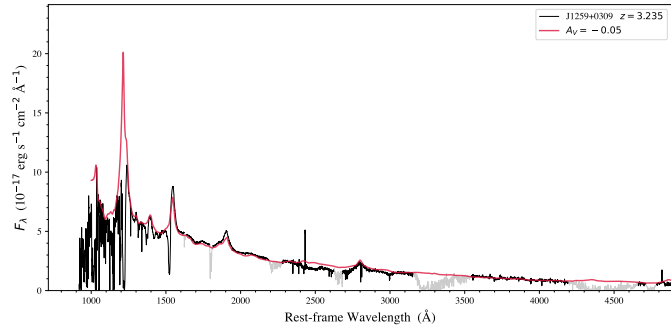
**Fig. C.5.** Same as Fig. C.1 for the quasar J0858+1749.



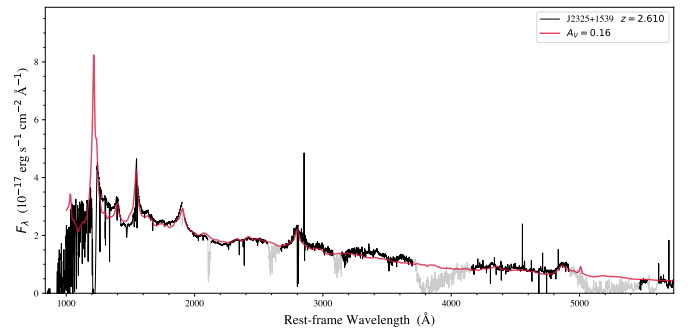
**Fig. C.6.** Same as Fig. C.1 for the quasar J1236+0010.



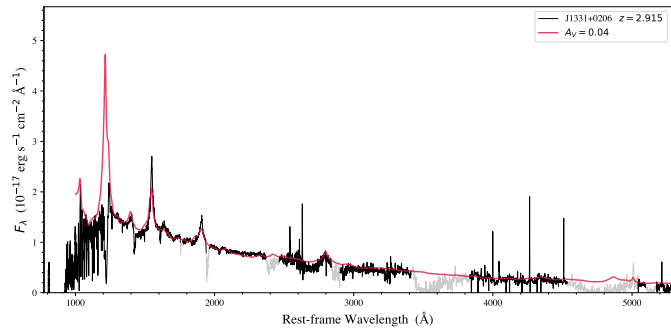
**Fig. C.7.** Same as Fig. C.1 for the quasar J1248+0639.



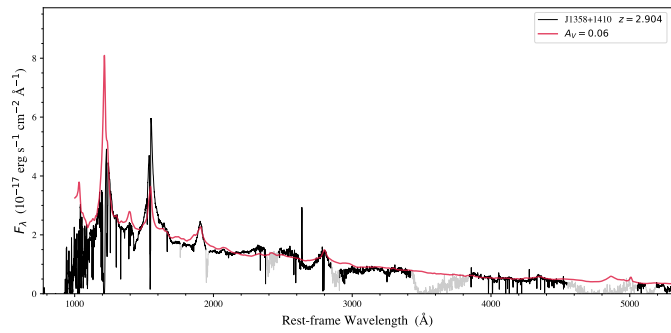
**Fig. C.8.** Same as Fig. C.1 for the quasar J1259+0309.



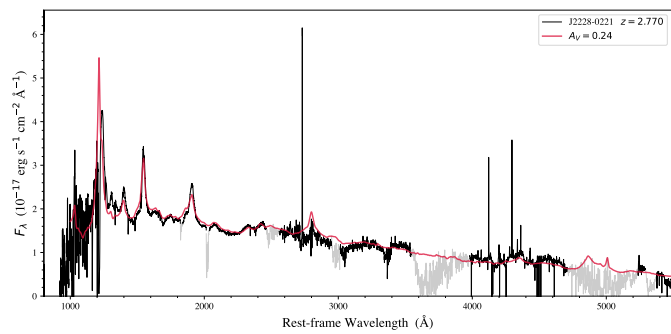
**Fig. C.12.** Same as Fig. C.1 for the quasar J2325+1539.



**Fig. C.9.** Same as Fig. C.1 for the quasar J1331+0206.

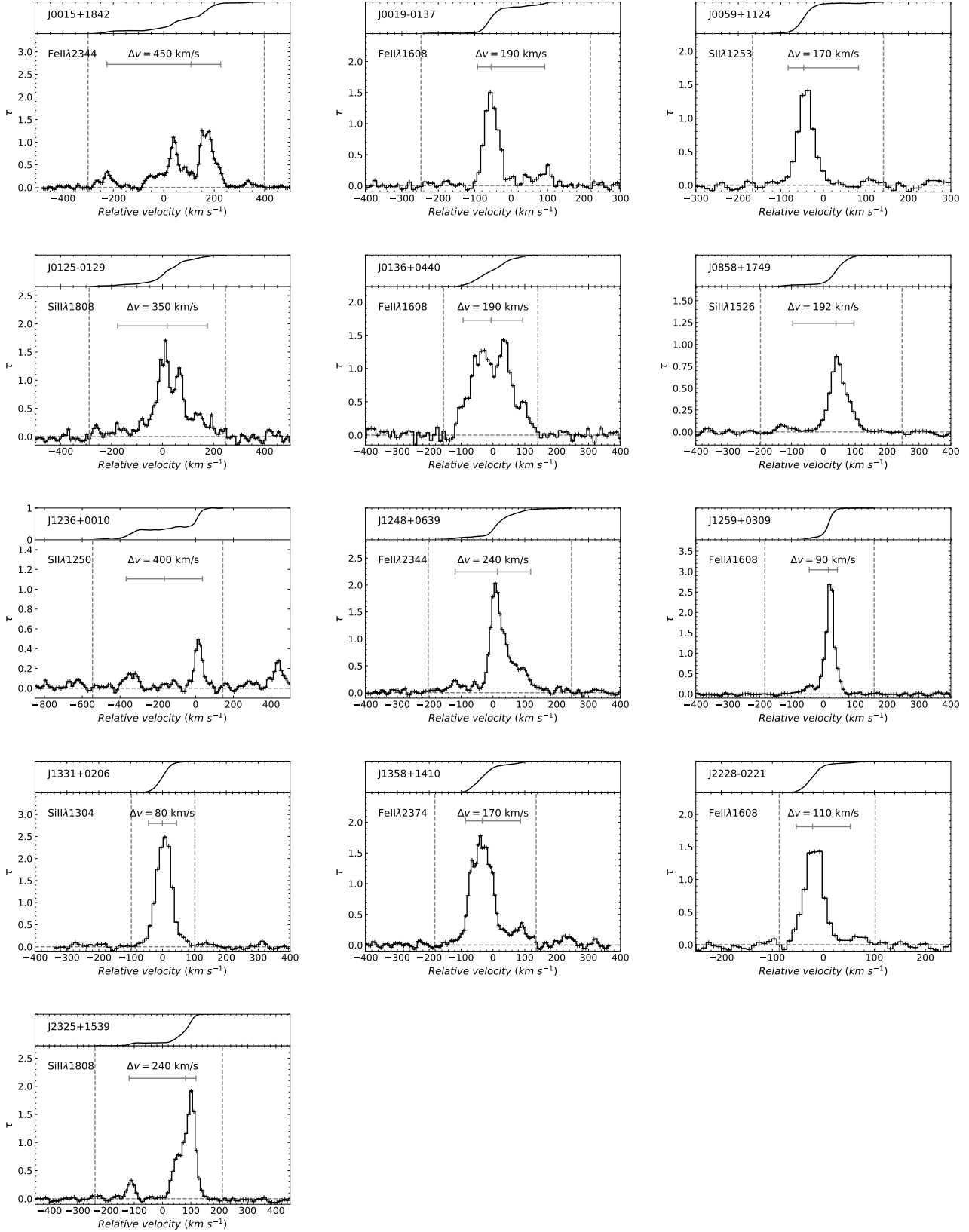


**Fig. C.10.** Same as Fig. C.1 for the quasar J1358+1410.



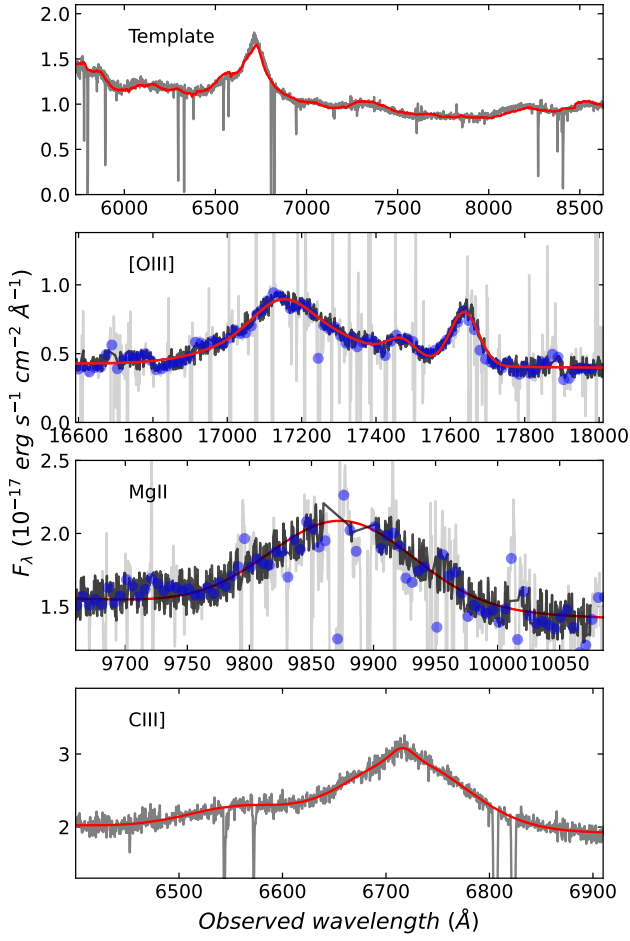
**Fig. C.11.** Same as Fig. C.1 for the quasar J2228-0221..

Appendix D:  $\Delta v_{90}$  measurements

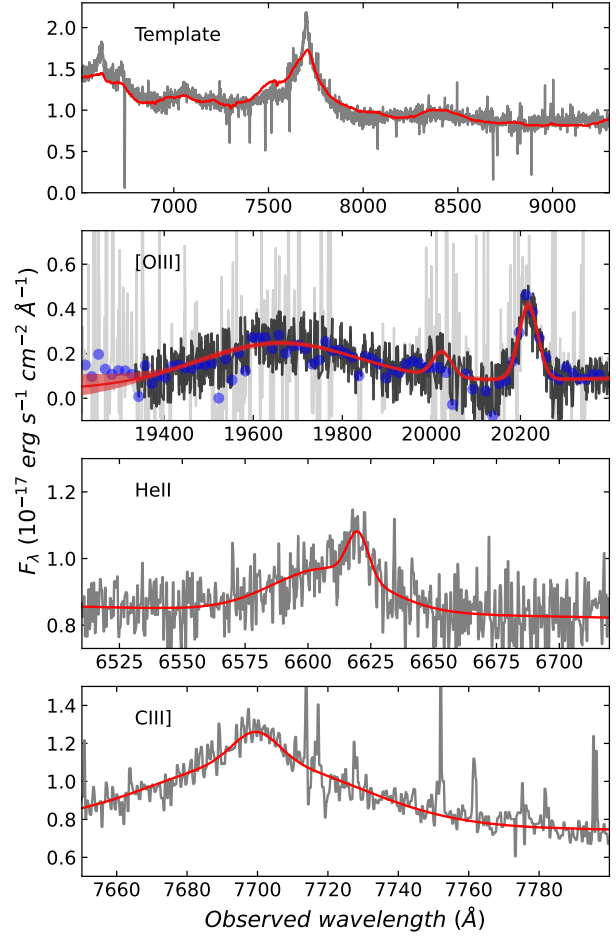


**Fig. D.1.** Measurements of  $\Delta v_{90}$  for low-ionisation line profiles. The different panels show the apparent optical depth ( $\tau$ ) in the selected transition line for a given system, and the cumulative apparent optical depth is shown on top. The segments represent the locations of the 5%, 50%, and 95% percentiles.

Appendix E: Optical and NIR emission lines



**Fig. E.1.** Emission lines from J0019–0137, detected in the visual and near-infrared. In the case of NIR, we show in dark grey the spectrum after iteratively rejecting deviant points. Blue dots show the median in 20-pixel bins. The best-fit model is shown by the solid red line.



**Fig. E.2.** Same as Fig. E.1 for the quasar J0059+1124.

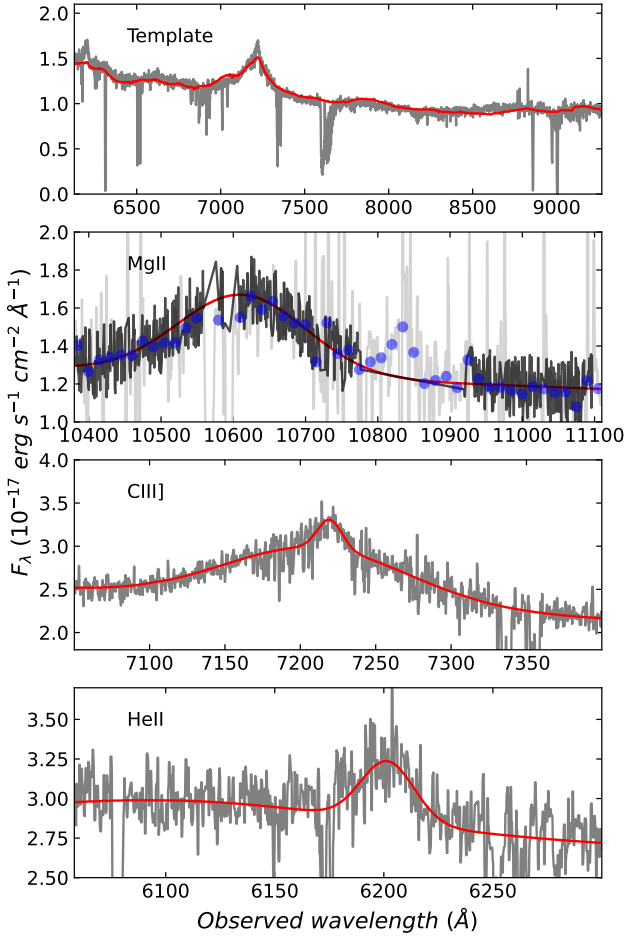


Fig. E.3. Same as Fig. E.1 for the quasar J0136+0440.

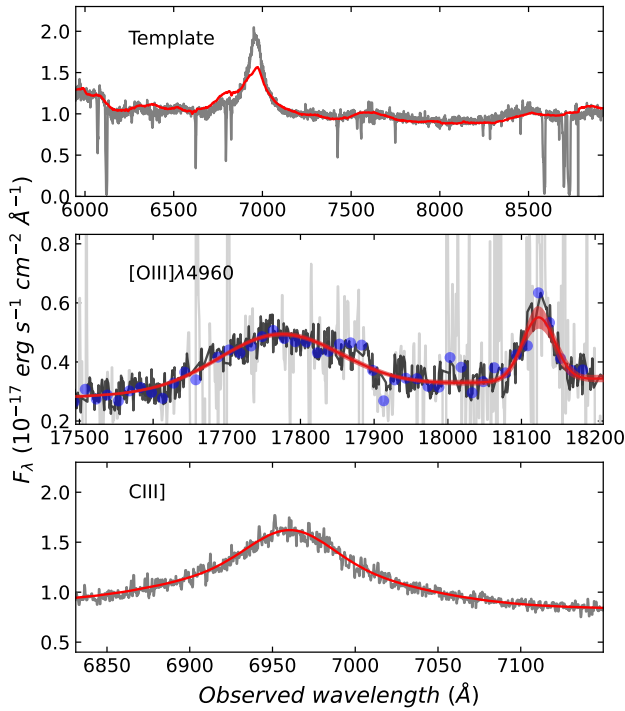


Fig. E.4. Same as Fig. E.1 for the quasar J0125-0129.

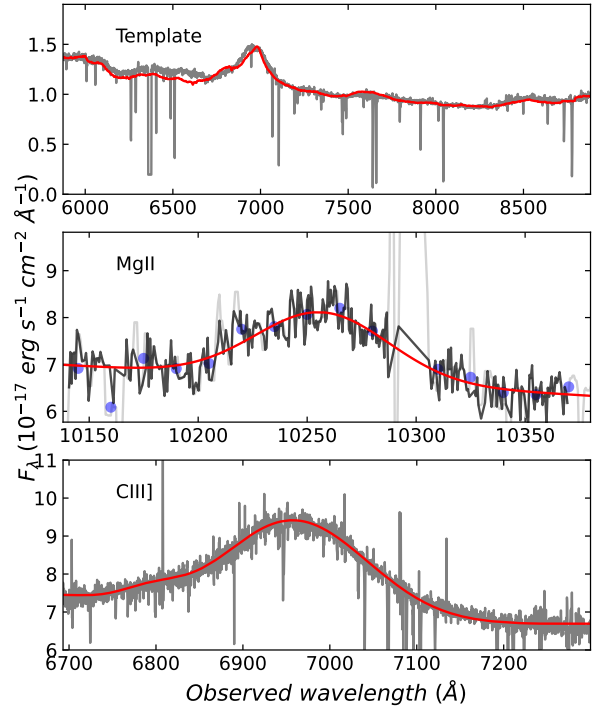


Fig. E.5. Same as Fig. E.5 for the quasar J0858+1749.

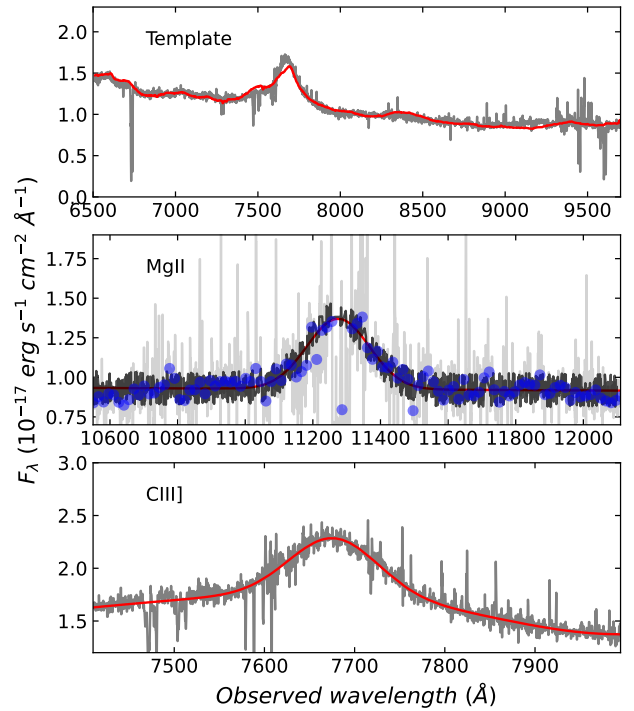


Fig. E.6. Same as Fig. E.1 for the quasar J1236+0010.

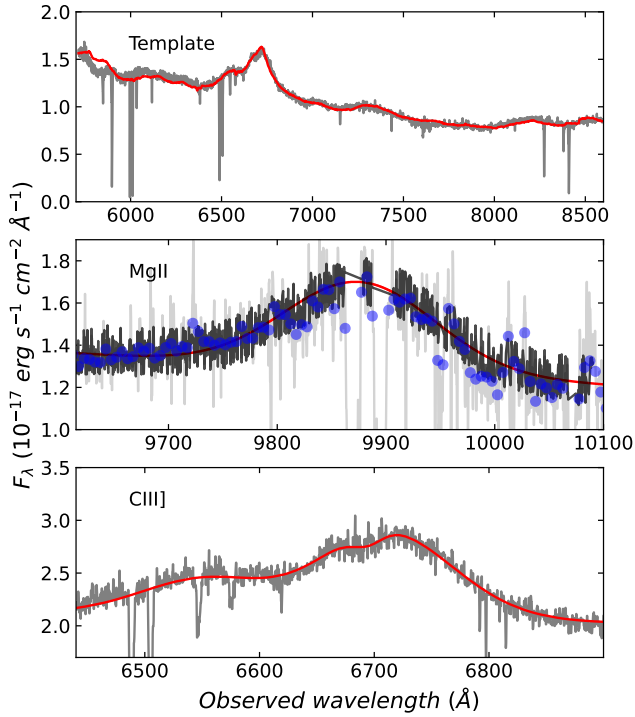


Fig. E.7. Same as Fig. E.1 for the quasar J1248+0639.

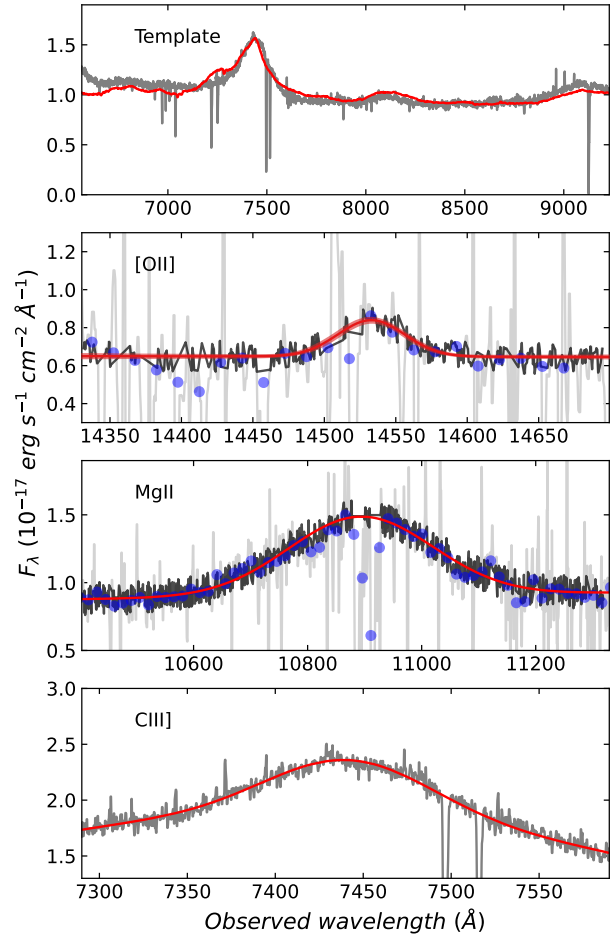


Fig. E.9. Same as Fig. E.1 for the quasar J1358+1410.

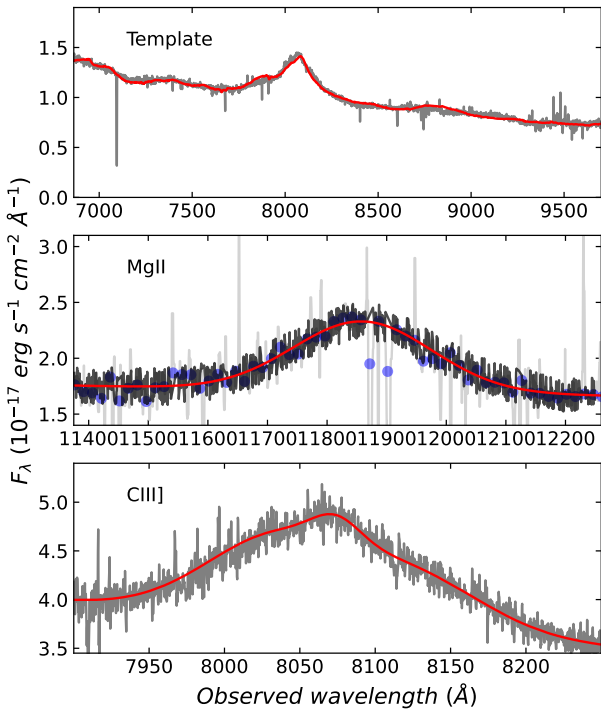


Fig. E.8. Same as Fig. E.1 for the quasar J1259+0309.

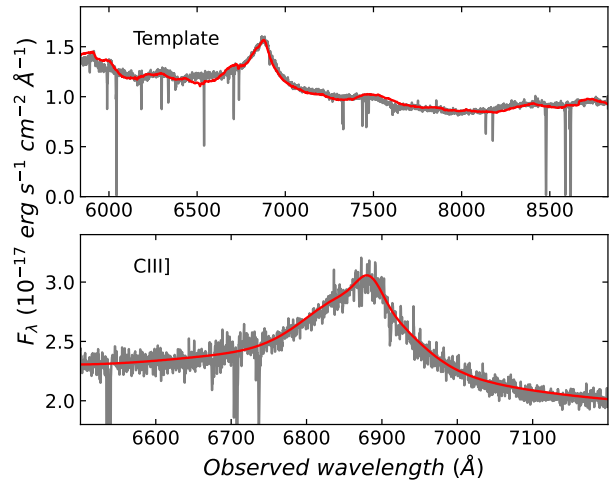


Fig. E.10. Same as Fig. E.1 for the quasar J2325+1539

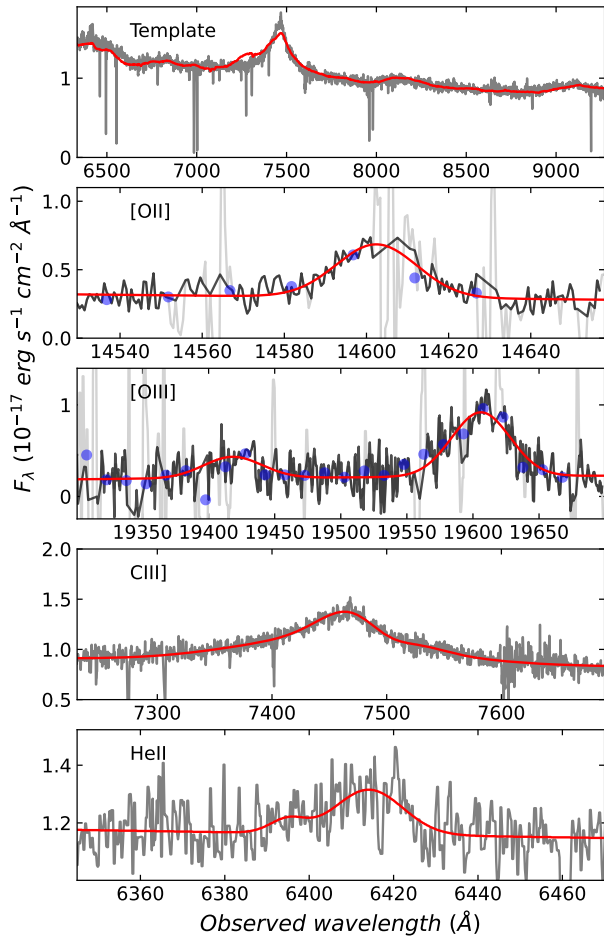


Fig. E.11. Same as Fig. E.1 for the quasar J1331+0206.

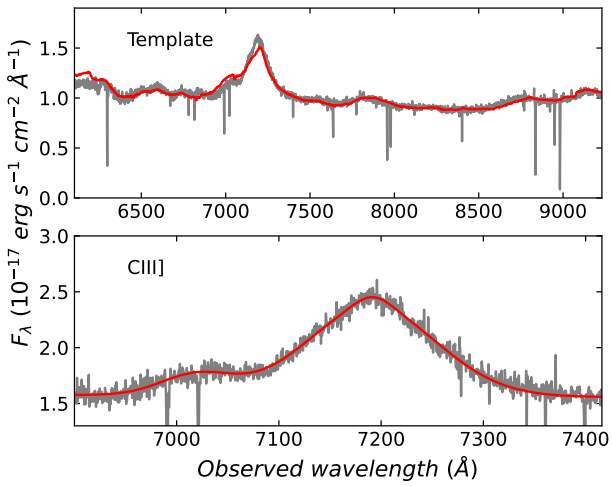
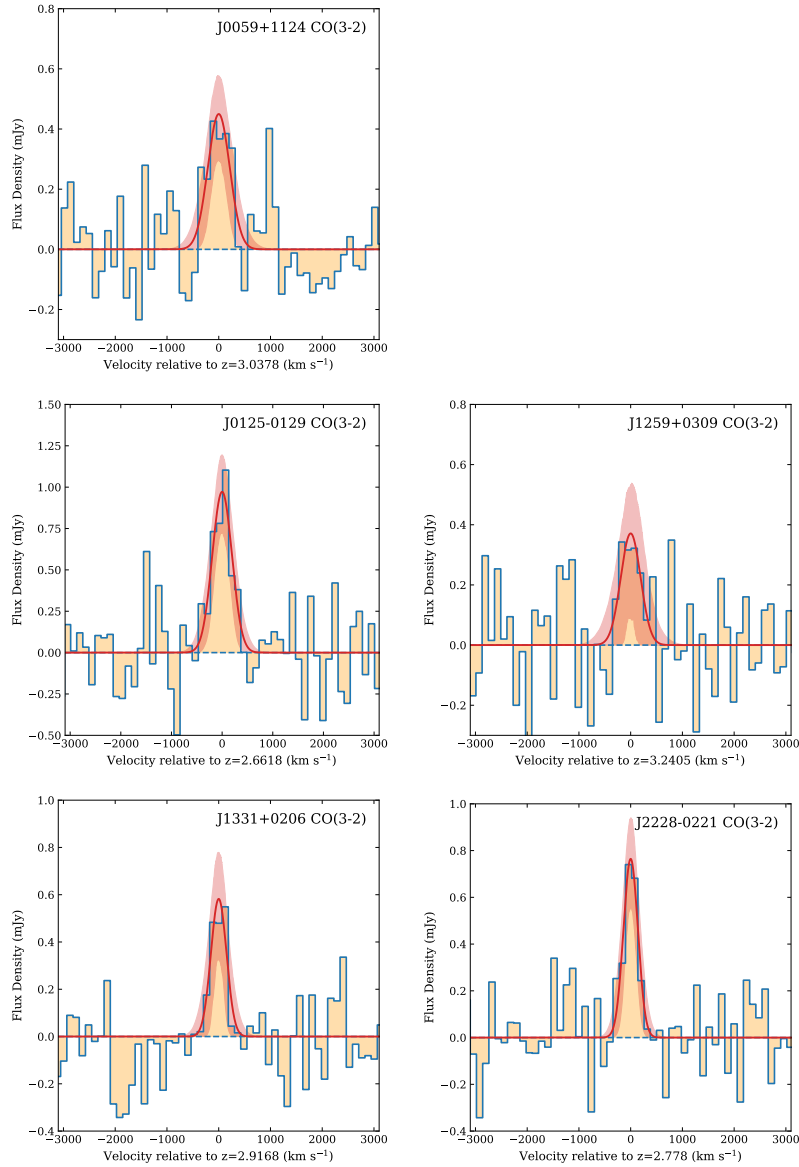


Fig. E.12. Same as Fig. E.1 for the quasar J2228-0221.

## Appendix F: CO(3-2) emission lines



**Fig. F.1.** Detection of CO(3-2) emission from 3 mm NOEMA spectra extracted in  $120 \text{ km s}^{-1}$  bins. A single-component Gaussian fit is superimposed in red together with its 90% confidence interval. A sixth quasar (J0015+1842) in this sample has also been observed with NOEMA. The corresponding CO(3-2) detection was presented in Paper III.

**Table F.1.** Results from Gaussian fits of CO(3-2) emission lines.

Quasar	$z_{CO}$	FWHM ( $\text{km s}^{-1}$ )	$F_{CO(3-2)}$ ( $\text{Jy km s}^{-1}$ )	$M(\text{H}_2)$ ( $10^{10} M_{\odot}$ )
J0015+1842	2.6285	$1010 \pm 120$	$1.1 \pm 0.2$	3.4 - 17
J0059+1124	3.0378	$500 \pm 110$	$0.24 \pm 0.07$	0.9 - 4.6
J0125-0129	2.6618	$460 \pm 80$	$0.5 \pm 0.1$	1.5 - 7.4
J1259+0309	3.2405	$450 \pm 180$	$0.2 \pm 0.1$	0.8 - 3.9
J1331+0206	2.9168	$350 \pm 100$	$0.22 \pm 0.08$	0.8 - 4.0
J2228-0221	2.7780	$330 \pm 60$	$0.27 \pm 0.07$	0.9 - 4.5

**Notes.** Molecular masses are calculated from the CO(3-2) luminosities assuming  $r_{31} = 0.97$  (Carilli & Walter 2013). The range corresponds to assumed CO-to- $\text{H}_2$  conversion factors either typical for quasars  $\alpha_{CO} = 0.8 M_{\odot}(\text{km s}^{-1} \text{pc}^2)^{-1}$  (see e.g. Bolatto et al. 2013) or using the standard value  $\alpha_{CO} = 4 M_{\odot}(\text{km s}^{-1} \text{pc}^2)^{-1}$ .