

# Measurement of the primordial helium abundance from the intergalactic medium

Ryan J. Cooke<sup>1,2</sup> & Michele Fumagalli<sup>1,3</sup>

<sup>1</sup>Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

<sup>2</sup>Royal Society University Research Fellow

<sup>3</sup>Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK

**Almost every helium atom in the Universe was created just a few minutes after the Big Bang through a process commonly referred to as Big Bang Nucleosynthesis<sup>1,2</sup>. The amount of helium that was made during Big Bang Nucleosynthesis is determined by the combination of particle physics and cosmology<sup>3</sup>. The current leading measures of the primordial helium abundance ( $Y_P$ ) are based on the relative strengths of H I and He I emission lines emanating from star-forming regions in local metal-poor galaxies<sup>4–7</sup>. As the statistical errors on these measurements improve, it is essential to test for systematics by developing independent techniques. Here we report the first determination of the primordial helium abundance based on a near-pristine intergalactic gas cloud that is seen in absorption against the light of a background quasar. This gas cloud, observed when the Universe was just one-third of its present age ( $z_{\text{abs}} = 1.724$ ), has a metal content  $\sim 100$  times less than the Sun, and  $\gtrsim 30\%$  less metals than the most metal-poor H II region currently known where a determination of the primordial helium abundance is afforded. We conclude that the helium abundance of this intergalactic gas cloud is  $Y = 0.250^{+0.033}_{-0.025}$ , which agrees with the Standard Model primordial value<sup>8–10</sup>,  $Y_P = 0.24672 \pm 0.00017$ . Our determination**

**of the primordial helium abundance is not yet as precise as that derived using metal-poor galaxies, but our method has the potential to offer a competitive test of physics beyond the Standard Model during Big Bang Nucleosynthesis.**

The near-pristine composition of some quasar absorption line systems<sup>11–13</sup>, including systems with apparently pristine composition<sup>14</sup>, makes intergalactic gas clouds well-suited for primordial abundance studies. Intergalactic gas clouds currently offer some of the most constraining observations of Big Bang Nucleosynthesis (BBN)<sup>15</sup>, however, they have hitherto not been exploited to measure the primordial helium abundance. Here we present a novel measure of the helium abundance based on a gas cloud that is seen in absorption along the line-of-sight to the bright  $z_{\text{em}} = 2.7348$  quasar HS 1700+6416. This quasar intersects ten partial Lyman limit systems<sup>16</sup> (pLLSs; defined to have H I column densities in the range  $16.0 \leq \log N(\text{H I})/\text{cm}^{-2} \leq 17.2$ ) and two of these systems are reported to have associated He I absorption based on low signal-to-noise and low resolution data acquired with the *Hubble Space Telescope* (*HST*) Faint Object Spectrograph (FOS)<sup>17,18</sup>.

We retrieved a High Resolution Echelle Spectrometer (HIRES) optical spectrum of HS 1700+6416 from the Keck Observatory Database of Ionized Absorption toward Quasars (KODIAQ)<sup>19,20</sup>. These data cover the H I Ly $\alpha$  absorption line and a selection of metal absorption lines. In order to obtain coverage across the H I Lyman limit of the pLLS ( $\sim 912 \text{ \AA}$  rest frame), we also retrieved the calibrated *HST* FOS data of HS 1700+6416 from the *HST* high level science products archive<sup>21</sup>. Finally, we retrieved the reduced data products of the *HST* Cosmic Origins Spectrograph (COS) G130M observations (program 13491; PI, T. Tripp) of HS 1700+6416 from the Mikulski Archive for Space Telescopes, covering the He I resonant absorption lines of the pLLS at  $z_{\text{abs}} \simeq 1.724$ .

We have adopted an identical analysis procedure to that recently used to determine the primordial deuterium abundance with high precision<sup>15,22</sup>. We simultaneously fit

the quasar emission and the absorption of the pLLS, while keeping blind the total column density of each ion (see Methods for further details). We uniquely determine the total H I column density by fitting the flux decrement at the Lyman limit of the pLLS, modelling the quasar continuum as a power-law over the fitted range. The best-fitting model to the H I Lyman limit absorption is shown in Figure 1a. We simultaneously fit several optically-thin high-order He I absorption lines, whose equivalent widths are directly proportional to the total He I opacity; we therefore uniquely determine the total He I column density, independent of the cloud model. The resulting profile fits are shown in Figure 1b. Finally, we detect a selection of weak metal lines associated with the pLLS, and we model these concurrently with the H I and He I absorption lines (see Figure 2).

We stress that the details of the cloud model and component structure are not important since, for all of the observed ions, there is at least one weak absorption line available to uniquely determine the total column density of H I, He I, and all of the observed metal species. Finally, we note that the entire fitting procedure is conducted in a blind fashion, such that the final parameter values are hidden from view until the analysis is complete. Once we had converged on the best-fitting profile model, the fitting procedure was not changed. The resulting values of the total column density are collected in Table 1. The final chi-squared of our best-fitting model is  $\chi^2/\text{dof} = 2659.0/2354 \simeq 1.130$ .

Based on this absorption line profile analysis, we measure H I and He I column densities of  $\log_{10}(N(\text{H I})/\text{cm}^{-2}) = 16.926 \pm 0.036$  and  $\log_{10}(N(\text{He I})/\text{cm}^{-2}) = 15.851 \pm 0.055$ , respectively, corresponding to a neutral column density ratio  $N(\text{He I})/N(\text{H I}) = 0.085 \pm 0.013$ . This measured ratio may be slightly different from the intrinsic He/H abundance ratio in a pLLS, depending on the incident radiation field. To convert the observed ratio into a determination of the helium number abundance, we construct a grid of *Cloudy*<sup>23</sup> photoionisation simulations, where a gas slab is irradiated by the widely assumed ul-

traviolet background (UVB) of quasars and galaxies<sup>24</sup>. This assumption is supported by the observation that there are no known bright galaxies in close proximity to the pLLS<sup>25</sup>. Our model grid covers a range of carbon metallicity ( $[C/H]$ ; on a logarithmic scale relative to solar<sup>26</sup>), relative C/Si abundance ( $[C/Si]$ ), total H volume density ( $n_H$ ),  $N(H\text{ I})$ , and He/H abundance. Furthermore, we allow for some flexibility in the shape of the UVB by using a radiation field at an ‘effective redshift’,  $z_{\text{eff}}$ , in place of the observed redshift,  $z_{\text{abs}} \simeq 1.724$ .

We linearly interpolate our grid of simulations, and sample the parameter space using a Markov chain Monte Carlo procedure<sup>27</sup> to identify the most likely set of parameters that reproduce the measured column densities. Probability distribution functions of the six model parameters are presented in Figure 3. Based on these calculations, we find that the  $z_{\text{abs}} \simeq 1.724$  pLLS toward HS 1700+6416 has a volume density,  $[C/H]$  abundance, and  $[C/Si]$  relative abundance that are all typical of other pLLSs at this redshift<sup>12</sup>. We note that the metallicity of this absorption line system ( $[C/H] = -1.688 \pm 0.059$  and  $[Si/H] = -1.969 \pm 0.088$ , corresponding to  $\sim 1/50$  and  $\sim 1/100$  solar, respectively) is comparable to the most metal-poor H II region<sup>28</sup> currently used to estimate  $Y_P$ . The abundances we quote here represent the total gas phase abundance of these elements. Since pLLSs are believed to be dust-poor,<sup>29</sup> we do not need to apply a correction for dust.

The helium ionisation correction appears to be minimal in this system; the neutral column density ratio,  $N(\text{He I})/N(\text{H I}) = 0.085 \pm 0.013$ , is consistent with the inferred abundance  $y_P \approx N(\text{He})/N(\text{H}) = 0.085^{+0.015}_{-0.011}$ . Furthermore, the quoted uncertainty is entirely dominated by the measurement uncertainty, and not by the errors associated with the ionisation modelling. We further explore the sensitivity of the ionisation correction to the assumed form of the incident radiation field in the Methods.

The helium number abundance,  $y$ , can also be expressed as a helium mass fraction, which is defined as  $Y \equiv 4y/(1 + 4y)$ . Substitution of our MCMC samples for  $y$

into this equation yields a mass fraction  $Y = 0.250^{+0.033}_{-0.025}$  (68% confidence), constituting an  $\sim 11\%$  determination. Our derived value of  $Y_P$  is consistent with the Standard Model value<sup>9</sup> ( $Y_P = 0.24672 \pm 0.00017$ ), assuming the *Planck* baryon density<sup>8</sup> and the latest determination of the neutron lifetime<sup>10</sup>. Our measure is also in good agreement with the *Planck* determination of the primordial helium abundance<sup>8</sup>, inferred from the damping tail of the CMB anisotropies ( $Y_P = 0.253 \pm 0.021$ ; 68% confidence). However, at present, our measurement of  $Y$  is not as precise as the value of  $Y_P$  derived from the CMB damping tail.

Most previous works have estimated  $Y_P$  using observations of the most metal-poor H II regions in the local universe<sup>4–7</sup>, and usually determine the primordial level by extrapolation to zero metallicity. Our novel measurement technique offers a new approach to pin down the primordial helium abundance in a near-pristine environment. Moreover, our determination is not impacted by the same systematics that currently affect the determination of  $Y_P$  from metal-poor H II regions<sup>30</sup>. The technique that we put forward in this paper offers an independent approach that will be critical to uncover potential systematics with both techniques.

Our analysis raises the exciting prospect of obtaining an independent high precision measure of the primordial helium abundance. Firstly, by collecting high-quality near-ultraviolet data around the H I Lyman limit of the pLLS at  $z_{\text{abs}} \simeq 1.724$  toward HS 1700+6416, it will be possible to pin down the total H I column density to sub-percent level precision by fitting additional weak high-order H I absorption lines jointly with the flux decrement at the Lyman limit; the current  $N(\text{H I})$  uncertainty is  $\sim 8\%$ . Furthermore additional near ultraviolet data of this one sightline towards HS 1700+6416 would yield in excess of 10 independent He/H abundances.<sup>16</sup>

A more precise determination of  $N(\text{He I})$  will require an absorption line system to be identified with more quiescent kinematics than the system analysed here. This will facilitate the detection of additional weak high order He I absorption lines. The

identification of new He I absorbers should be considered a key goal for the future UV space-based observatories<sup>31</sup> that are on the horizon. He I absorbers will be easier to identify along sightlines<sup>32</sup> to quasars in the redshift interval  $1.5 \lesssim z_{\text{qso}} \lesssim 2.0$ , where the contamination due to unrelated absorption near the high order H I and He I absorption lines is considerably less compared with high redshift quasars.

The technique pioneered in this study could deliver a few per cent determination of  $Y_{\text{P}}$  from near-pristine environments in the near-future.  $Y_{\text{P}}$  offers a sensitive probe of the universal expansion rate at the time of BBN,<sup>3</sup> and its measurement is therefore of acute interest in the quest to uncover deviations from the Standard Model of particle physics and cosmology.

1. Alpher, R. A., Bethe, H. & Gamow, G. The Origin of Chemical Elements. *Phys. Rev.* **73**, 803–804 (1948).
2. Hoyle, F. & Tayler, R. J. The Mystery of the Cosmic Helium Abundance. *Nature* **203**, 1108–1110 (1964).
3. Steigman, G., Schramm, D. N. & Gunn, J. E. Cosmological limits to the number of massive leptons. *Phys. Lett. B* **66**, 202–204 (1977).
4. Izotov, Y. I., Thuan, T. X. & Guseva, N. G. A new determination of the primordial He abundance using the He I  $\lambda 10830$  Å emission line: cosmological implications. *Mon. Not. R. Astron. Soc.* **445**, 778–793 (2014).
5. Aver, E., Olive, K. A. & Skillman, E. D. The effects of He I  $\lambda 10830$  on helium abundance determinations. *J. Cosmol. Astropart. P.* **7**, 011 (2015).
6. Peimbert, A., Peimbert, M. & Luridiana, V. The primordial helium abundance and the number of neutrino families. *Rev. Mex. Astron. Astr.* **52**, 419–424 (2016).

7. Fernández, V., Terlevich, E., Díaz, A. I., Terlevich, R. & Rosales-Ortega, F. F. Primordial helium abundance determination using sulphur as metallicity tracer. *Mon. Not. R. Astron. Soc.* **478**, 5301–5319 (2018).
8. Planck Collaboration *et al.* Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.* **594**, A13 (2016).
9. Pitrou, C., Coc, A., Uzan, J.-P. & Vangioni, E. Precision big bang nucleosynthesis with improved Helium-4 predictions. *Phys. Rep.* **10.1016/j.physrep.2018.04.005** (2018).
10. Pattie, R. W. *et al.* Measurement of the neutron lifetime using a magneto-gravitational trap and in situ detection. *Science* **360**, 627–632 (2018).
11. Crighton, N. H. M., O’Meara, J. M. & Murphy, M. T. Possible Population III remnants at redshift 3.5. *Mon. Not. R. Astron. Soc.* **457**, L44–L48 (2016).
12. Lehner, N., O’Meara, J. M., Howk, J. C., Prochaska, J. X. & Fumagalli, M. The Cosmic Evolution of the Metallicity Distribution of Ionized Gas Traced by Lyman Limit Systems. *Astrophys. J.* **833**, 283–316 (2016).
13. Cooke, R. J., Pettini, M. & Steidel, C. C. Discovery of the most metal-poor damped Lyman- $\alpha$  system. *Mon. Not. R. Astron. Soc.* **467**, 802–811 (2017).
14. Fumagalli, M., O’Meara, J. M. & Prochaska, J. X. Detection of Pristine Gas Two Billion Years After the Big Bang. *Science* **334**, 1245–1249 (2011).
15. Cooke, R. J., Pettini, M. & Steidel, C. C. One Percent Determination of the Primordial Deuterium Abundance. *Astrophys. J.* **855**, 102–117 (2018).
16. Fechner, C. *et al.* The UV spectrum of HS 1700+6416. I. Predicting the metal line content of the far UV spectrum. *Astron. Astrophys.* **455**, 73–90 (2006).

17. Reimers, D. & Vogel, S. He I Absorption Lines in High Redshift Lyman Limit Systems of the QSO HS 1700+6416. *Astron. Astrophys.* **276**, L13–L16 (1993).
18. Syphers, D. & Shull, J. M. The He II Post-reionization Epoch: HST/COS Observations of the Quasar HS1700+6416. *Astrophys. J.* **765**, 119–128 (2013).
19. O'Meara, J. M. *et al.* The First Data Release of the KODIAQ Survey. *Astron. J.* **150**, 111–121 (2015).
20. O'Meara, J. M. *et al.* The Second Data Release of the KODIAQ Survey. *Astron. J.* **154**, 114–118 (2017).
21. Evans, I. N. & Koratkar, A. P. A Complete Atlas of Recalibrated Hubble Space Telescope Faint Object Spectrograph Spectra of Active Galactic Nuclei and Quasars. I. Pre-COSTAR Spectra. *Astrophys. J. Suppl.* **150**, 73–164 (2004).
22. Cooke, R. J., Pettini, M., Jorgenson, R. A., Murphy, M. T. & Steidel, C. C. Precision Measures of the Primordial Abundance of Deuterium. *Astrophys. J.* **781**, 31–46 (2014).
23. Ferland, G. J. *et al.* The 2017 Release Cloudy. *Rev. Mex. Astron. Astr.* **53**, 385–438 (2017).
24. Haardt, F. & Madau, P. Radiative Transfer in a Clumpy Universe. IV. New Synthesis Models of the Cosmic UV/X-Ray Background. *Astrophys. J.* **746**, 125–144 (2012).
25. Shapley, A. E. *et al.* Ultraviolet to Mid-Infrared Observations of Star-forming Galaxies at  $z \sim 2$ : Stellar Masses and Stellar Populations. *Astrophys. J.* **626**, 698–722 (2005).



26. Asplund, M., Grevesse, N., Sauval, A. J. & Scott, P. The Chemical Composition of the Sun. *Annu. Rev. Astron. Astr.* **47**, 481–522 (2009).
27. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: The MCMC Hammer. *Publ. Astron. Soc. Pac.* **125**, 306–312 (2013).
28. Izotov, Y. I. & Thuan, T. X. Heavy-Element Abundances in Blue Compact Galaxies. *Astrophys. J.* **511**, 639–659 (1999).
29. Fumagalli, M., O’Meara, J. M. & Prochaska, J. X. The physical properties of  $z > 2$  Lyman limit systems: new constraints for feedback and accretion models. *Mon. Not. R. Astron. Soc.* **455**, 4100–4121 (2016).
30. Izotov, Y. I., Thuan, T. X. & Stasińska, G. The Primordial Abundance of  $^4\text{He}$ : A Self-consistent Empirical Analysis of Systematic Effects in a Large Sample of Low-Metallicity H II Regions. *Astrophys. J.* **662**, 15–38 (2007).
31. Muslimov, E. *et al.* POLLUX: a UV spectropolarimeter for the LUVUOIR space telescope project. *ArXiv e-prints* (2018).
32. O’Meara, J. M., Prochaska, J. X., Worseck, G., Chen, H.-W. & Madau, P. The HST/ACS+WFC3 Survey for Lyman Limit Systems. II. Science. *Astrophys. J.* **765**, 137–165 (2013).

Correspondence and requests for materials should be addressed to R.J.C.  
(ryan.j.cooke@durham.ac.uk).

**Acknowledgements** We thank Gabor Worseck for a helpful discussion about the HST+COS data of this sightline. During this work, R. J. C. was supported by a Royal Society University Research Fellowship. We acknowledge support from STFC (ST/L00075X/1, ST/P000541/1). This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research

and innovation programme (grant agreement No 757535). This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, and STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. This research has made use of NASA's Astrophysics Data System.

**Author Contributions** Both authors participated in the interpretation and commented on the manuscript. R.J.C. led the project and analysis, and was responsible for the text of the paper. M.F. was responsible for the ionisation calculations.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper.

**Figure 1 | Hydrogen and helium absorption from an intergalactic gas cloud toward the quasar HS 1700+6416. a**, High order H I Lyman series absorption lines (indicated by the red tick marks above the spectrum) and H I Lyman continuum absorption (blueward of an observed wavelength  $\lambda \lesssim 2483\text{\AA}$ ). The best-fit model is shown by the continuous red curve and the power-law fit to the quasar continuum is shown by the long dashed blue line. **b**, Four He I absorption lines associated with the pLLS at redshift  $z_{\text{abs}} \simeq 1.724$ . The red tick marks and associated numbers above the spectrum indicate the locations of the components comprising the absorption model. The blue points below all spectra are the normalised fit residuals,  $(\text{data} - \text{model})/\text{error}$ , of all pixels that were used in the analysis, and the grey bands represent a confidence interval of  $\pm 2\sigma$ . Labels at the top of all panels indicate the transition shown.

**Figure 2 | Heavy element absorption due to an intergalactic gas cloud toward**

**the quasar HS 1700+6416.** The red tick marks and associated numbers above each spectrum indicate the locations of the components comprising the absorption model. Fitted blends that are unrelated to the pLLS at  $z_{\text{abs}} \simeq 1.724$  are marked with blue tick marks above the spectrum. For example, the absorption feature at  $v \simeq +100 \text{ km s}^{-1}$  relative to C IV  $\lambda 1548$  is an unrelated Ly $\alpha$  absorption line at  $z_{\text{abs}} = 2.4703$ , which is fully accounted for by simultaneously fitting the higher order H I Lyman lines of this system. Similarly, the absorption feature at  $v \simeq +50 \text{ km s}^{-1}$  relative to Si II  $\lambda 1260$  is an unrelated N II  $\lambda 1083$  absorption line at  $z_{\text{abs}} = 2.1679$ . The blue points below all spectra are the normalised fit residuals,  $(\text{data} - \text{model})/\text{error}$ , of all pixels that were used in our analysis, and the grey bands represent a confidence interval of  $\pm 2\sigma$ . Note the different y-axis scale used on each panel.

**Figure 3 | Confidence regions of the model parameters.** The 1D posterior distribution of He/H is shown in the bottom right panel. The remaining panels show the 68% and 95% confidence contours (dark and light shades, respectively) illustrating how each model parameter depends on the helium abundance.

**Table 1 | Ion column densities of the pLLS at  $z_{\text{abs}} \simeq 1.724$  toward HS 1700+6416.**

Ion	Transitions used	$\log N(X)/\text{cm}^{-2} \text{ }^\dagger$
H I	Ly $\alpha$ , Ly7–Lyman Limit	$16.926 \pm 0.036$
He I	537.0, 515.6, 507.7, 506.2	$15.851 \pm 0.055$
C II	1334	$13.189 \pm 0.043$
C IV	1548, 1550	$13.446 \pm 0.047$
Si II	1193, 1260	$12.586 \pm 0.059$
Si IV	1393, 1402	$13.002 \pm 0.028$

<sup>†</sup> Quoted uncertainties represent the 68 per cent confidence interval.

## Methods

**Observational Data.** The  $z_{\text{em}} = 2.7348$  quasar HS 1700+6416 was first identified as part of the Hamburg quasar survey<sup>33</sup>. Since its discovery, HS 1700+6416 has been observed multiple times, and across many wavebands. We utilise archival data taken with the W. M. Keck Observatory High Resolution Echelle Spectrometer (HIRES) in addition to archival data acquired with *HST* FOS and *HST* COS.

**HIRES data.** Using custom built routines, we combined all HIRES data from the KODIAQ<sup>19,20</sup> database that were acquired through decker C1 ( $0.861'' \times 7.0''$ ) over the lifetime of HIRES, amounting to a total exposure time of 30,417 s. The C1 decker has a nominal spectral resolution of  $R = 48,000$ , corresponding to a velocity full-width at half-maximum (FWHM) resolution of  $v_{\text{FWHM}} = 6.25 \text{ km s}^{-1}$ . The combined data cover a wavelength range 3109–6076 Å, which includes H I Ly $\alpha$  and a selection of metal absorption lines. The final combined spectrum has a signal-to-noise ratio per pixel of  $S/N \sim 30$  near Ly $\alpha$  and  $S/N \sim 50$  near the metal absorption lines.

**HST FOS data.** We retrieved the fully reduced, coadded and calibrated *HST* FOS data of HS 1700+6416 from the *HST* high level science products archive<sup>21</sup>. The full data cover the wavelength range 1140–3277 Å, with a spectral resolution  $R \simeq 1,300$ . Near the H I Lyman limit, the data exhibit a signal-to-noise ratio per pixel of  $S/N \simeq 25$ .

**HST COS data.** We retrieved the final calibrated data products of the *HST* COS G130M observations of HS 1700+6416 from the Mikulski Archive for Space Telescopes. The processed X1DSUM files were combined using version 2.1 of the COADD\_X1D software package. These high quality COS data cover the He I resonant absorption lines of the pLLS at  $z_{\text{abs}} \simeq 1.724$ . The final combined signal-to-noise ratio per  $2.2 \text{ km s}^{-1}$  pixel of the data near He I  $\lambda 507.7$  is  $S/N \simeq 20$ .

**Atomic Data.** The atomic data used in our analysis were retrieved from several sources. For the H I and metal absorption lines, we use the atomic data compiled by Morton<sup>34</sup>. In order to fit the H I Lyman limit absorption, we take the photoionisation

cross-section of H I from the Verner et al. database<sup>35</sup>. Finally, we retrieved all of the He I vacuum wavelengths ( $\lambda_0$ ) and oscillator strengths ( $f$ ) from the National Institutes of Standards and Technology (NIST) Atomic Spectral Database (ASD)<sup>36</sup>. We computed the associated natural damping constant ( $\gamma_{ul}$ ) of all transitions by summing the spontaneous transition probabilities to all lower levels

$$\gamma_{ul} = \sum_l A_{ul} \quad (1)$$

All of the relevant values are collected in Supplementary Table 1.

The pLLS at  $z_{\text{abs}} \simeq 1.724$  toward the quasar HS 1700+6416 is detected in He I absorption from  $\lambda 537.0$  to  $\lambda 505.5$ . However, the  $f$ -values and damping constants of these He I absorption lines are only available for wavelengths longward of He I  $\lambda 507.0$ ; the atomic data (both  $f$  and  $\gamma_{ul}$ ) of the higher order He I lines are not available in the NIST ASD. To overcome this limitation, we fit a low-order polynomial to the available atomic data and extrapolate the polynomial fit to estimate the unknown values. We use a third-order polynomial for the  $f$ -values and a second-order polynomial for the damping constants. To test the accuracy of this extrapolation, we applied an identical procedure to the atomic data of neutral hydrogen, by fitting the atomic data of Ly $\alpha$ -Ly9 and compare the extrapolated values of Ly10-Ly14 to the true values reported by Morton<sup>34</sup>. We found that this simple extrapolation allows us to predict the H I Ly10-Ly14  $f$ -values to be better than 2%, and the damping constants to within 25%.

The only extrapolated value that is used in our analysis is that of He I  $\lambda 506.2$ . Since this line is weak, the accuracy of the damping constant is not important. However, the total He I column density is degenerate with the  $f$ -value. We estimate that the systematic uncertainty of the He I  $\lambda 506.2$   $f$ -value to be  $\lesssim 0.5\%$ , which translates to a systematic limit of 0.5% on the total He I column density. As shown below, this level of precision is well below the measurement precision of the He I column density.

**Fitting Procedure.** We use the Absorption Line Software (ALIS) to derive the column densities of the absorption lines associated with the pLLS at  $z_{\text{abs}} \simeq 1.724$  toward HS 1700+6416. ALIS uses a chi-squared minimisation algorithm to determine the model parameters that best fit the data, weighted by the inverse variance of each pixel that is used in the analysis.

We perform a simultaneous fit to the quasar emission and the absorption of the pLLS. This ensures that any uncertainty associated with the emission profile is folded into the ion column density uncertainties. We model the quasar emission by a low-order Legendre polynomial locally around each absorption line. Typically, we use a polynomial of degree  $\lesssim 4$  to describe the quasar continuum. All absorption lines are modelled as a Voigt profile, which is characterised by a column density, a redshift, and a Doppler parameter. Since the pLLS studied here contains several absorption components, we fit directly for the total column density of each ion. For all fitted ions, we only include pixels in our analysis that we deem are free from unrelated absorption. In some cases, we identify blends and fully account for these by simultaneously fitting a Voigt profile. Finally, we include two global model parameters to define the zero level of the HIRES and COS data, to account for any small residual to the background subtraction.

For the H I Ly $\alpha$  profile, we include D I absorption at the primordial ratio<sup>15</sup> (i.e.  $N(\text{D I}) = N(\text{H I}) - 4.5974$  for each component of H I), and several unrelated blends. Since the equivalent width of the H I Ly $\alpha$  absorption line lies on the flat part of the curve of growth, the column density is degenerate with the cloud model. However, the total H I column density is uniquely determined by fitting the flux decrement at the Lyman limit of the pLLS. The optical depth beyond the H I Lyman limit is modelled as  $\tau_\nu = N(\text{H I}) \sigma_\nu$ , where  $\sigma_\nu$  is the photoionisation cross-section of H I.

In order to fit the observed component structure of the metal absorption lines, we require seven absorption components. The redshift and Doppler parameter of each

component are allowed to vary, but are forced to be the same across all species. We only include pixels in the metal absorption line fitting that we deem to be free of unrelated absorption. Where possible, we account for line blending due to unrelated absorption features provided that the blend can be fully accounted for using another transition. For example, the absorption feature at  $v \simeq +100 \text{ km s}^{-1}$  relative to C IV  $\lambda 1548$  corresponds to Ly $\alpha$  absorption at  $z_{\text{abs}} = 2.4703$ , which we account for by simultaneously fitting the corresponding higher order Lyman series lines.

We use the same component redshifts to fit the He I absorption lines, but allow the Doppler parameter of each line to vary independently of the metal absorption lines. In our analysis, we only include the He I absorption lines that are entirely free of line blending, to reduce the risk of contamination from the rich forest of metal absorption lines<sup>16</sup> associated with absorption systems in the range  $0.2 < z < 2.6$ . The spectra near all of the available He I resonant absorption lines are overlaid with the best-fitting model in Supplementary Figure 1.

**Ionisation Modelling.** Since the gas associated with pLLSs is mostly ionised, both He I and H I are minor ions; thus, He I may not precisely trace the H I gas, and a small ionisation correction may be required to convert the measured  $N(\text{He I})/N(\text{H I})$  ratio into a He/H abundance.

To estimate the magnitude of this correction, we use version 17.0 of the *Cloudy* photoionisation simulation software<sup>23</sup> to model the pLLS as a slab of constant density gas in ionisation and thermal equilibrium, irradiated by the CMB and ultraviolet background (UVB) of galaxies and quasars. The UVB radiation field at a given redshift is defined by a shape and intensity (see Supplementary Figure 2 for five example spectra). We allow for some flexibility in the shape of the UVB by irradiating the pLLS by the widely-assumed Haardt & Madau<sup>24</sup> UVB at an ‘effective redshift’  $z_{\text{eff}}$ ; we do not enforce a prior of the UVB based on the measured redshift,  $z_{\text{abs}}$ . The incident radiation field of our fiducial model is therefore  $J(\nu) = J_{\text{CMB}}(\nu, z_{\text{abs}}) + J_{\text{HM12}}(\nu, z_{\text{eff}})$ .

We consider a grid of UVB radiation fields between  $0 \leq z_{\text{eff}} \leq 6$ , sampled in steps of  $\Delta z_{\text{eff}} = 0.25$  when  $z_{\text{eff}} < 2$  and in steps of  $\Delta z_{\text{eff}} = 0.5$  when  $z_{\text{eff}} > 2$ . These UVB models set both the intensity and shape of the radiation field; to allow for some flexibility in the ionisation parameter at a given redshift, we construct a grid of models to cover a range of volume density ( $-4.6 < \log_{10}(n_{\text{H}}/\text{cm}^{-3}) < +0.2$ , in steps of 0.2 dex). Note that the ionisation parameter, which is defined as the ratio of gas number density to photon number density, is what determines the ionisation state of the gas.

Our grid of *Cloudy* models also includes a range of metallicity ( $-4.0 < [\text{X}/\text{H}] < -0.8$ , in steps of 0.2 dex), H I column density ( $16.6 < \log_{10}(N(\text{H I})/\text{cm}^{-2}) < 17.2$ , in steps of 0.15 dex), and helium abundance ( $0.05 < (\text{He}/\text{H})/0.084249 < 1.5$ , in linear steps of 0.05). Grains are not included in our calculations.

Our *Cloudy* models are stopped once the input neutral hydrogen column density is reached, at which point we store the column densities of each ion stage of H, He, C, and Si. Our grid of models is linearly interpolated and sampled using a Markov chain Monte Carlo procedure<sup>27</sup> to identify the parameter values that reproduce the measured column densities listed in Table 1 (see Supplementary Figure 3). Our model contains six parameters ( $\text{He}/\text{H}$ ,  $z_{\text{eff}}$ ,  $n_{\text{H}}$ ,  $[\text{C}/\text{H}]$ ,  $[\text{C}/\text{Si}]$ , and  $N(\text{H I})$ ), which are exactly determined by the six reported column density measurements. The resulting parameter values, together with their 68% confidence intervals, are listed in the penultimate column of Table 2.

One assumption of our fiducial model is that the incident radiation field is that of the metagalactic UVB calculated by Haardt & Madau<sup>24</sup>. However, due to the uncertainty of the UVB<sup>37</sup> and the fact that some pLLSs arise in the outskirts of a nearby galaxy<sup>38</sup>, we have tested the sensitivity of the helium ionisation correction to the assumed radiation field. In what follows, we assume that the incident radiation field comprises the CMB, the Haardt & Madau<sup>24</sup> UVB at a redshift  $z_{\text{abs}} \simeq 1.724$ , and a galaxy spectrum,  $J_{\text{G}}(\nu)$ . This is similar to an approach recently used to model Ly-



man Limit Systems<sup>29</sup>. The incident radiation field in this case is of the form  $J(\nu) = J_{\text{CMB}}(\nu, z_{\text{abs}}) + J_{\text{HM12}}(\nu, z_{\text{abs}}) + j_{\text{gal}} J_{\text{G}}(\nu)$ , where  $j_{\text{gal}}$  sets the intensity of the galactic spectrum at the surface of the pLLS, and  $J_{\text{G}}(\nu)$  is defined such that when  $j_{\text{gal}} = 1$ , it has the same intensity of H Lyman continuum photons as  $J_{\text{HM12}}(\nu, z_{\text{abs}})$ . We define the spectral shape of  $J_{\text{G}}(\nu)$  by synthesising a STARBURST99 model<sup>39</sup> (v7.0.1) using the default input parameters and assuming a continuous star formation rate of  $\dot{\psi} = 1 M_{\odot} \text{ yr}^{-1}$ . Our galaxy model employs the Geneva (2012)<sup>40</sup> solar metallicity stellar models, without rotation. The galactic radiation field, when added to the UVB, results in a softer spectrum relative to the UVB alone; we illustrate five example spectra of  $J(\nu)$  in Supplementary Figure 4.

With this alternative definition of  $J(\nu)$ , we generated a grid of *Cloudy* models that are otherwise identical to that described earlier, and sampled the grid with a Markov chain Monte Carlo procedure (see Supplementary Figure 5). The maximum likelihood parameter values, along with their corresponding 68% confidence intervals, are listed in the final column of Table 2. These parameter values are all within  $1\sigma$  of our fiducial model. The known galaxy population<sup>25</sup> along this sightline consists of four galaxies with redshifts that are consistent with  $z_{\text{abs}} \simeq 1.724$ . The closest galaxy (BX 756) is  $\sim 500$  kpc in projection from the pLLS. The remaining three galaxies are  $\gtrsim 1$  Mpc in projection from the pLLS. Thus there does not appear to be a significant bright galaxy population in the immediate vicinity of the pLLS. This is consistent with the maximum likelihood value  $j_{\text{gal}} = 0$ , implying that the radiation field incident on the pLLS at  $z_{\text{abs}} \simeq 1.724$  is likely dominated by the metagalactic UVB. We therefore conclude that our assumed shape for  $J(\nu)$  in the fiducial model is supported by the data in hand.

By changing the form of the incident radiation field, the central He/H values of our two models differ by  $\sim 9\%$  (compare the values listed in the final row of Table 2). Thus, the shape of the incident radiation field must be known before a robust ionisation correction can be applied. In principle, the shape of the radiation field can be inferred by

measuring the column density ratio of successive ions of the same element, particularly the ions that bracket the ionisation potentials of H and He. Thus, in order to pin down the radiation field with the accuracy demanded by precision cosmology, it will be important to reliably determine the column densities of metal absorption lines from a range of ionisation stages.

**Table 2 | Best-fitting parameter values of our ionisation modelling.**

Parameter	UVB only	UVB+galaxy
$\log N(\text{H I})$	$16.908^{+0.039}_{-0.034}$	$16.911^{+0.038}_{-0.035}$
$\log n_{\text{H}}$	$-2.28^{+0.08}_{-0.24}$	$-2.23^{+0.10}_{-0.03}$
$z_{\text{eff}}$	$1.21^{+0.60}_{-0.38}$	...
$\log j_{\text{gal}}$	...	$\leq -0.45^{\dagger}$
[C/H]	$-1.696^{+0.063}_{-0.055}$	$-1.732^{+0.075}_{-0.229}$
[C/Si]	$+0.255^{+0.060}_{-0.037}$	$+0.255^{+0.052}_{-0.055}$
He/H	$0.085^{+0.015}_{-0.011}$	$0.077^{+0.015}_{-0.037}$

$^{\dagger}$  95% Confidence Interval

**Data availability.** The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request. The Keck HIRES data are available as a high level science product, housed at the Keck Observatory Data Archive, which is available from <https://koa.ipac.caltech.edu/cgi-bin/KOA/nph-KOAllogin>.

The fully reduced *HST* FOS data are available from [https://archive.stsci.edu/prepds/fos\\_agn/](https://archive.stsci.edu/prepds/fos_agn/).

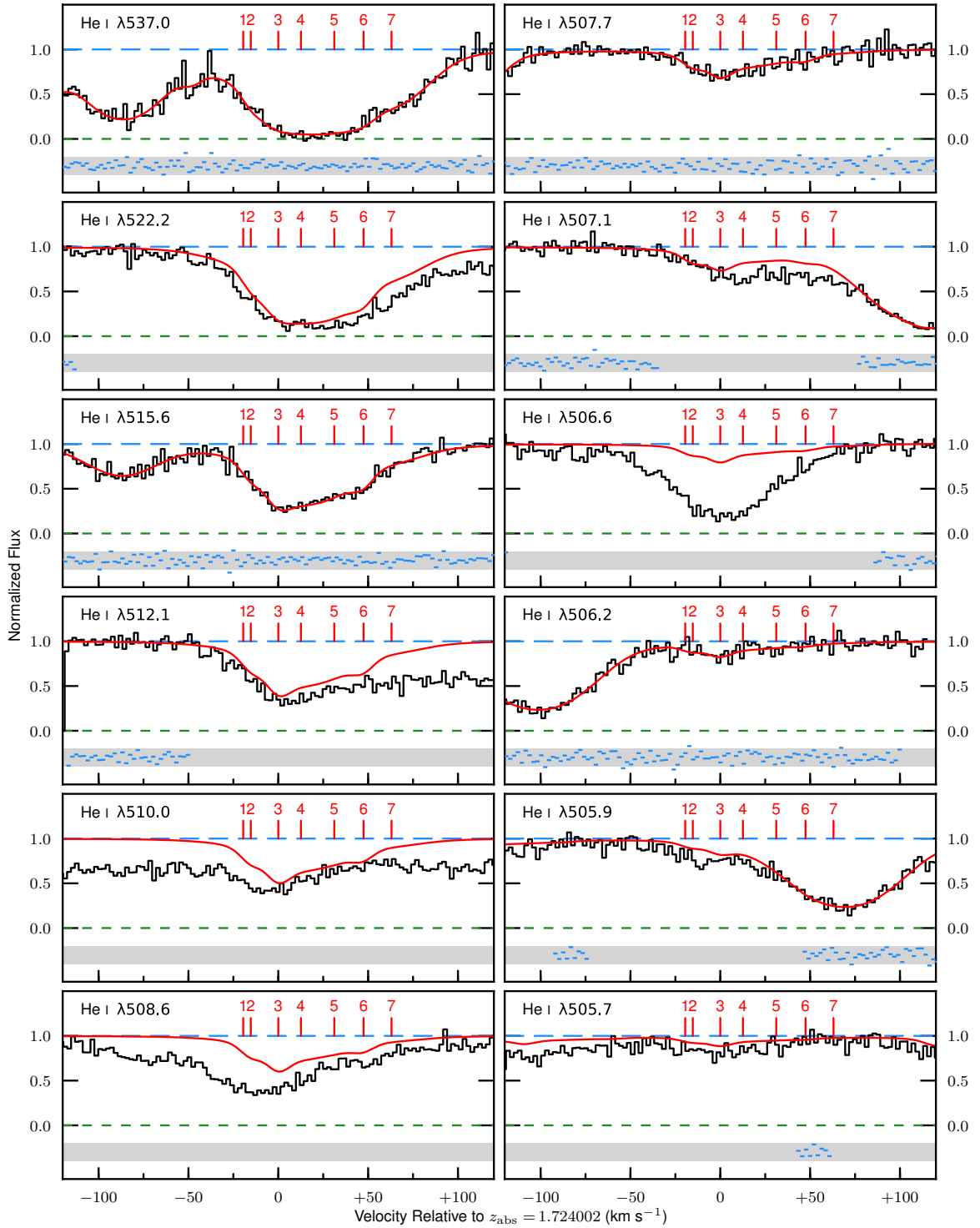
The reduced *HST* COS data are available at [http://archive.stsci.edu/proposal\\_search.php?mission=hst&id=13491](http://archive.stsci.edu/proposal_search.php?mission=hst&id=13491).

33. Reimers, D. *et al.* The luminous quasar HS1700+6416 and the shape of the 'big bump' below 500 Å. *Astron. Astrophys.* **218**, 71–77 (1989).
34. Morton, D. C. Atomic Data for Resonance Absorption Lines. III. Wavelengths Longward of the Lyman Limit for the Elements Hydrogen to Gallium. *Astrophys. J. Suppl.* **149**, 205–238 (2003).
35. Verner, D. A., Ferland, G. J., Korista, K. T. & Yakovlev, D. G. Atomic Data for Astrophysics. II. New Analytic FITS for Photoionization Cross Sections of Atoms and Ions. *Astrophys. J.* **465**, 487–498 (1996).
36. Kramida, A., Yu. Ralchenko, Reader, J. & and NIST ASD Team. NIST Atomic Spectra Database (ver. 5.5.6), [Online]. Available: <https://physics.nist.gov/asd> [2017, April 9]. National Institute of Standards and Technology, Gaithersburg, MD. (2018).
37. Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M. & Hernquist, L. A New Calculation of the Ionizing Background Spectrum and the Effects of He II Reionization. *Astrophys. J.* **703**, 1416–1443 (2009).
38. Fumagalli, M. *et al.* Absorption-line systems in simulated galaxies fed by cold streams. *Mon. Not. R. Astron. Soc.* **418**, 1796–1821 (2011).
39. Leitherer, C. *et al.* Starburst99: Synthesis Models for Galaxies with Active Star Formation. *Astrophys. J. Suppl.* **123**, 3–40 (1999).
40. Ekström, S. *et al.* Grids of stellar models with rotation. I. Models from 0.8 to 120  $M_{\odot}$  at solar metallicity ( $Z = 0.014$ ). *Astron. Astrophys.* **537**, A146 (2012).

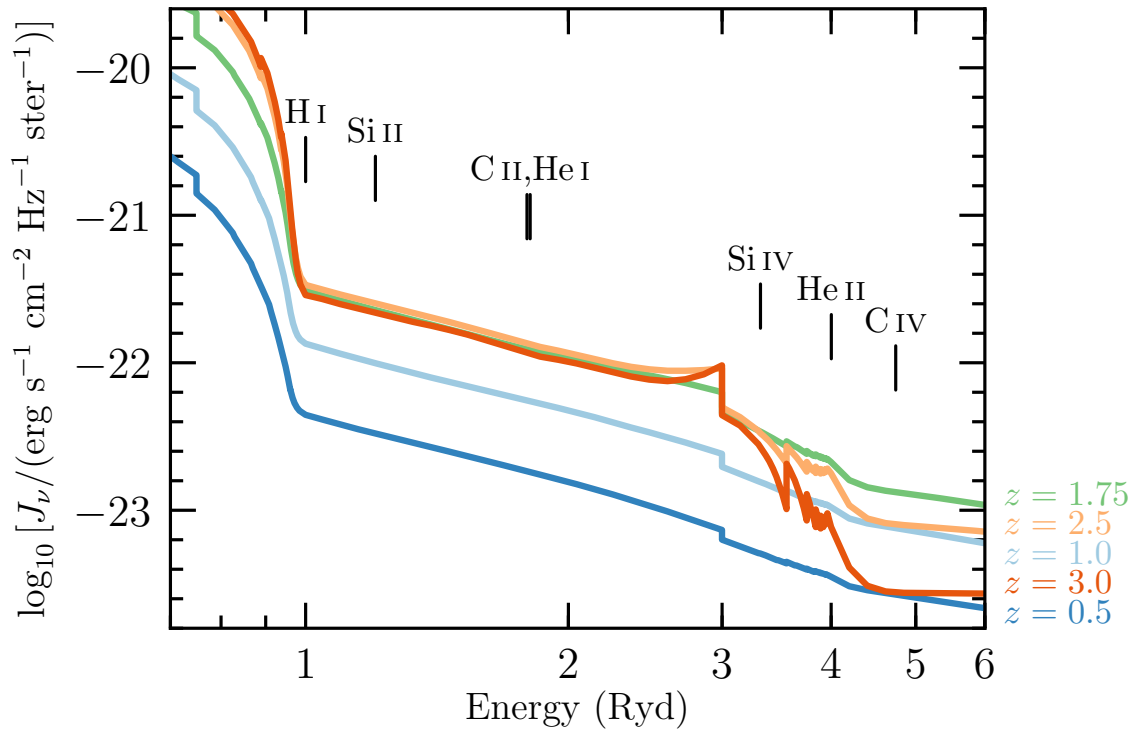
**Supplementary Table 1 | Atomic data of the He I resonant absorption lines.**

Ion	$\lambda_0$ (Å)	$f$	$\gamma_{ul}$ (s <sup>-1</sup> )
He I $\lambda$ 584.3	584.3344	0.2763	1.799E9
He I $\lambda$ 537.0	537.0299	0.07346	5.799E8
He I $\lambda$ 522.2	522.2131	0.02988	2.520E8
He I $\lambda$ 515.6	515.6168	0.01504	1.308E8
He I $\lambda$ 512.1	512.0986	0.008630	7.634E7
He I $\lambda$ 510.0	509.9983	0.005407	4.939E7
He I $\lambda$ 508.6	508.6434	0.003611	3.248E7
He I $\lambda$ 507.7	507.7181	0.002531	2.287E7
He I $\lambda$ 507.1	507.0580	0.001842	1.670E7
He I $\lambda$ 506.6	506.5705	0.001382 <sup>†</sup>	1.221E7 <sup>†</sup>
He I $\lambda$ 506.2	506.2003	0.001064 <sup>†</sup>	9.247E6 <sup>†</sup>
He I $\lambda$ 505.9	505.9125	0.0008363 <sup>†</sup>	7.143E6 <sup>†</sup>
He I $\lambda$ 505.7	505.6843	0.0006693 <sup>†</sup>	5.609E6 <sup>†</sup>
He I $\lambda$ 505.5	505.5004	0.0005440 <sup>†</sup>	4.469E6 <sup>†</sup>

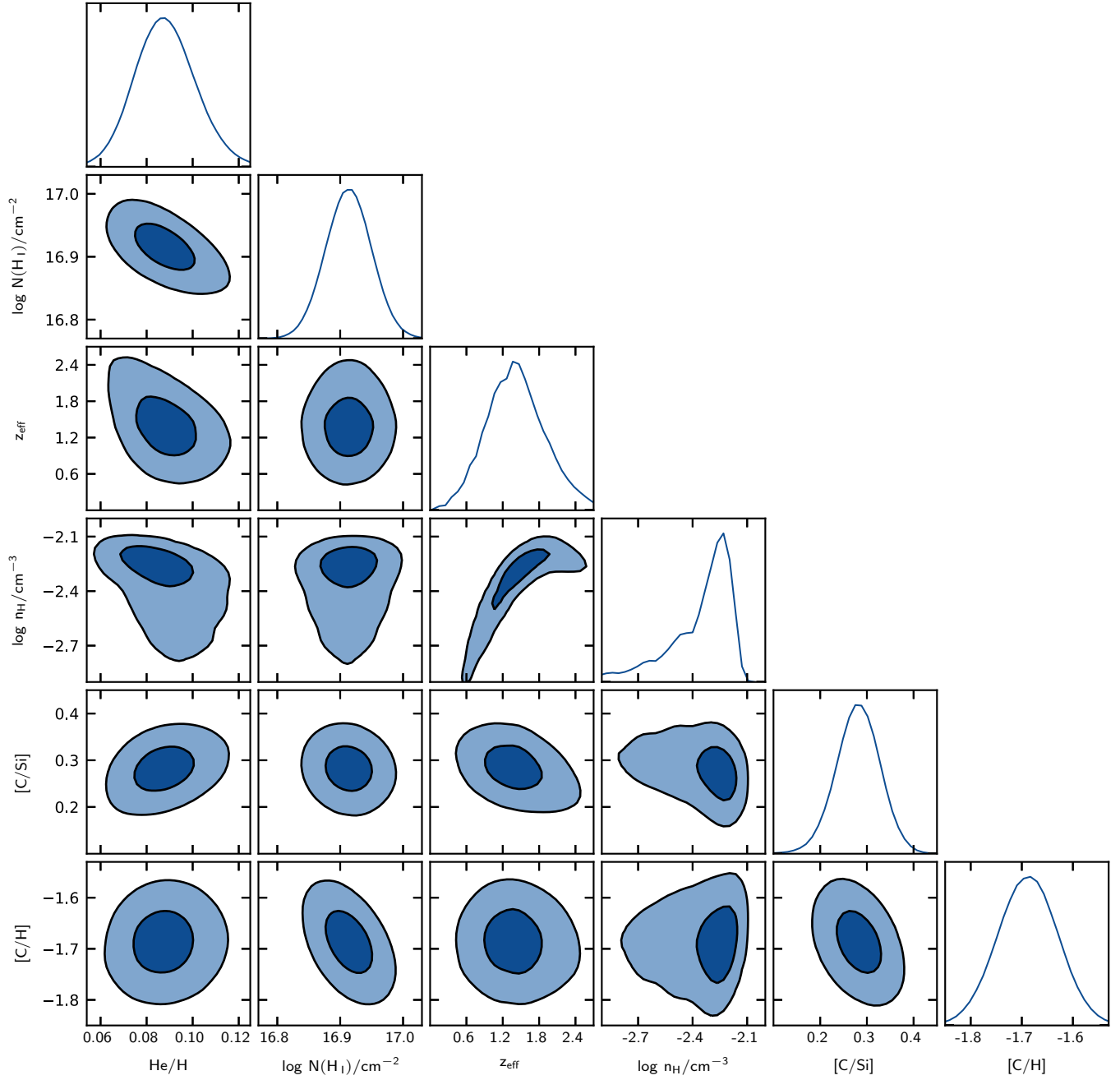
<sup>†</sup>No atomic data exist for these transitions. Instead, these values were extrapolated based on a polynomial fit to the known measures from lower order lines (see Methods).



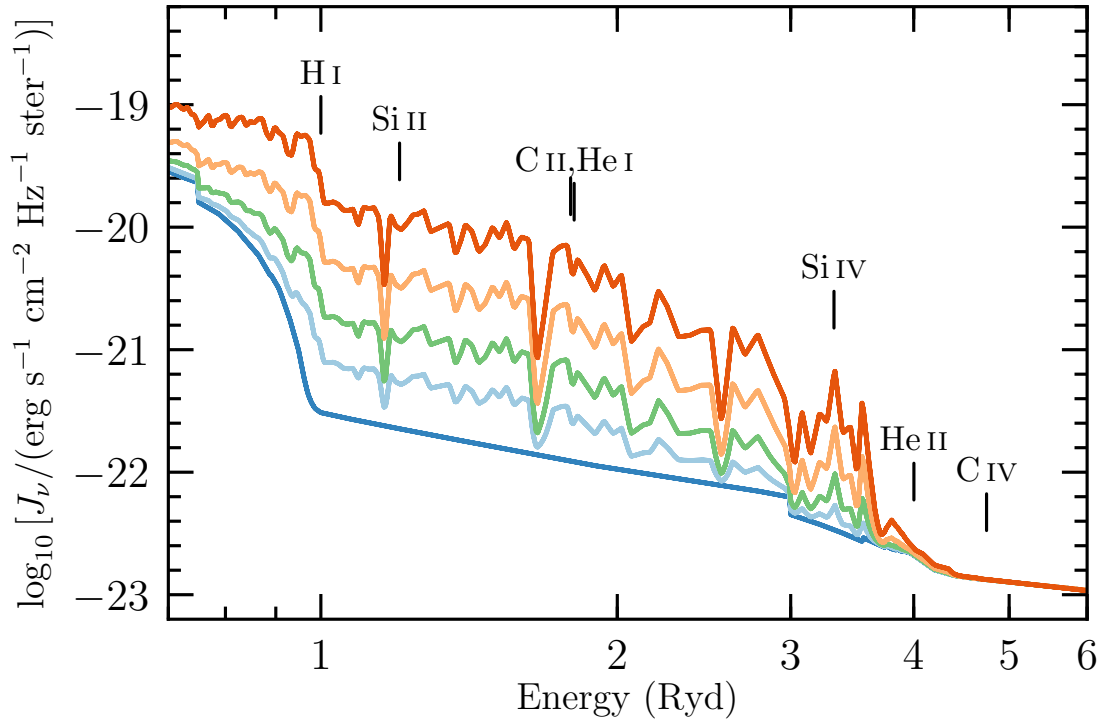
**Supplementary Figure 1 | Resonant absorption lines of neutral helium from an intergalactic gas cloud at  $z_{\text{abs}} \simeq 1.724$  toward the quasar HS 1700+6416.** The red curve shows the best-fitting model profile. The data and model are normalised to the best-fitting continuum model (long dashed blue line) and corrected for the fitted zero-level (short dashed green line). The red tick marks and associated numbers above the spectrum indicate the locations of the components comprising the absorption model. The He I absorption lines that are used to pin down the total He I column density include: He I  $\lambda 537.0$ ,  $\lambda 515.6$ ,  $\lambda 507.7$ , and  $\lambda 506.2$ . The blue points below all spectra are the normalised fit residuals,  $(\text{data} - \text{model})/\text{error}$ , of all pixels that were used in the analysis, and the grey bands represent a confidence interval of  $\pm 2\sigma$ . A label in the top left corner of every panel indicates the He I absorption line shown.



**Supplementary Figure 2 | Five sample spectra illustrate how the Haardt & Madau UVB radiation field varies as a function of redshift.** The redshift corresponding to each spectrum is indicated at the bottom right of the figure. Labelled tick marks above the spectra indicate the ionisation potentials of the relevant ions that are used in this letter.

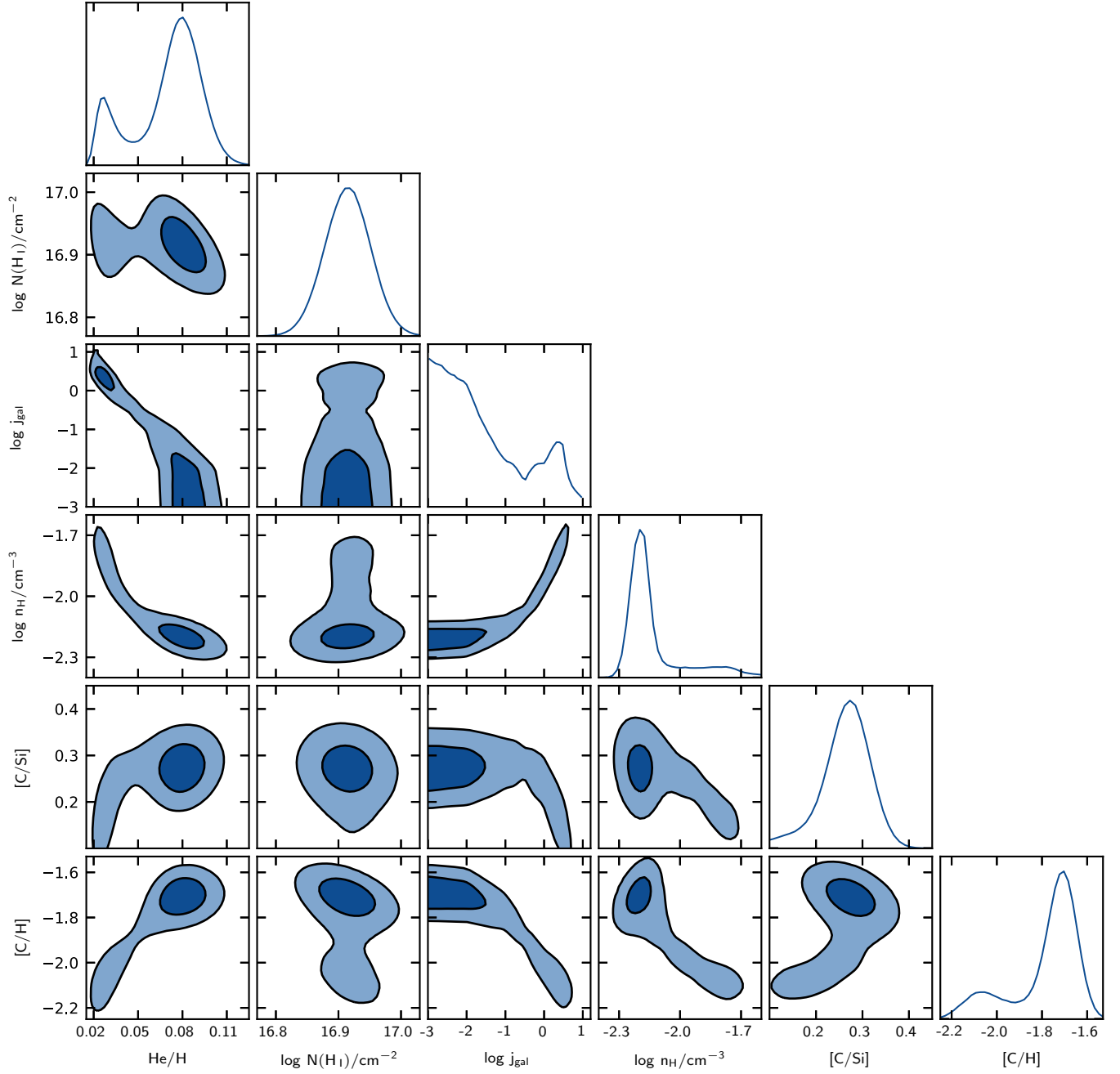


**Supplementary Figure 3 | Confidence regions of the model parameters.** The 1D posterior distributions of each parameter are shown along the diagonal panels, while the off-diagonal panels show the two dimensional projections of each parameter combination, illustrating the covariance between model parameters. Dark and light shades represent the 68% and 95% confidence regions, respectively.



**Supplementary Figure 4 | Five sample radiation fields that include a contribution from the metagalactic UVB radiation field and a galactic radiation field.** The lowest blue curve shows the Haardt & Madau UVB at  $z_{\text{abs}} \simeq 1.724$ , while the remaining curves (bottom to top) include an increasing contribution from a galactic radiation field (corresponding to  $\log j_{\text{gal}} = -\infty, -0.5, 0.0, +0.5, +1.0$ ). Labelled tick marks above the spectra indicate the ionisation potentials of the relevant ions that are used in this letter.





**Supplementary Figure 5 | Confidence regions of the model parameters when using an incident radiation field that consists of a galaxy and the UVB.** The curves and contours have the same meaning as in Supplementary Figure 3.