Confirmation of double peaked Lyα emission at z = 6.593

Witnessing a galaxy directly contributing to the reionisation of the Universe*,**

Jorryt Matthee1, David Sobral2, Max Gronke3, Ana Paulino-Afonso2,4,5, Mauro Stefanon1, and Huub Röttgering1

1 Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands
2 Department of Physics, Lancaster University, LA1 4YB Lancaster, UK
3 Department of Physics, University of California, 9326 Santa Barbara, CA, USA
4 Instituto de Astrofísica e Ciências do Espaco, Universidade de Lisboa, OAL, Tapada da Ajuda, 1349-018 Lisboa, Portugal
5 Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, 1749-016 Lisboa, Portugal

Received 29 May 2018 / Accepted 31 August 2018

ABSTRACT

Distant luminous Lyman-α emitters (LAEs) are excellent targets for spectroscopic observations of galaxies in the epoch of reionisation (EoR). We present deep high-resolution (R ≈ 5000) VLT/X-shooter observations, along with an extensive collection of photometric data of COLA1, a proposed double peaked LAE at z ≈ 6.6. We rule out the possibility that COLA1’s emission line is an [Oii] doublet at z ≈ 1.475 on the basis of i) the asymmetric red line-profile and flux ratio of the peaks (blue/red=0.31±0.03) and ii) an unphysical [Oii]/Hα ratio ([Oii]/Hα > 22). We show that COLA1’s observed B-band flux is explained by a faint extended foreground LAE, for which we detect Lyα and [Oii] at z = 2.142. We thus conclude that COLA1 is a real double-peaked LAE at z = 6.593, the first discovered at z > 6. COLA1 is UV luminous (M1500 = −21.6 ± 0.3), has a high equivalent width (EW_{Lyα} = 120 ± 50 Å) and very compact Lyα emission (r_{FWHM} = 0.33 ± 0.07 kpc). Relatively weak inferred Hβ+[Oii] line-emission from Spitzer/IRAC indicates an extremely low metallicity of Z ≲ 1/20 Z⊙, or reduced strength of nebular lines due to high escape of ionising photons. The small Lyα peak separation of 220 ± 20 km s^{-1} implies a low H\alpha column density and an ionising photon escape fraction of ≈15–30%, providing the first direct evidence that such galaxies contribute actively to the reionisation of the Universe at z > 6. Based on simple estimates, we find that COLA1 could have provided just enough photons to reionise its own ≈0.3 Mpc (2.3 cMpc) bubble, allowing the blue Lyα line to be observed. However, we also discuss alternative scenarios explaining the detected double peaked nature of COLA1. Our results show that future high-resolution observations of statistical samples of double peaked LAEs at z > 5 are a promising probe of the occurrence of ionised regions around galaxies in the EoR.

Key words. galaxies: high-redshift – galaxies: formation – dark ages, reionization, first stars – techniques: spectroscopic – intergalactic medium

1. Introduction

The epoch of reionisation (EoR) is the last phase transition of the Universe. It occurred when the neutral hydrogen of the intergalactic medium (IGM) became reionised (e.g. Madau 2017). In spite of the increasingly precise measurements of the global progress of reionisation and its patchiness (e.g. Becker et al. 2015; Davies et al. 2018; Bañados et al. 2018; Bosman et al. 2018), its evolution and drivers are still largely unknown. One of the main probes of reionisation is the Lyman-α (Lyα) line (see Dijkstra et al. 2007 for a review). Due to the low Lyα transmission in a partially neutral IGM, the progress of reionisation can be mapped out using the strength of detected Lyα emission. This is usually done through the equivalent width distribution among high-redshift galaxies (e.g. Pentericci et al. 2014; Mason et al. 2018), through the evolution in the luminosity function (e.g. Zheng et al. 2017) and/or its clustering signal (e.g. Jensen et al. 2014; Kakiichi et al. 2018).

Recently, wide-field narrow-band surveys have been highly efficient in identifying and confirming luminous Lyman-α emitters (LAEs) into the EoR (z ≲ 7; e.g. Matthee et al. 2015; Hu et al. 2016; Santos et al. 2016; Bagley et al. 2017; Zheng et al. 2017; Shibuya et al. 2018). The number densities of extremely luminous LAEs at z ≈ 7 (L_{Lyα} > 2 × 10^{43} erg s^{-1}) are higher (Matthee et al. 2015; Bagley et al. 2017) than expected based on older, smaller surveys (e.g. Ouchi et al. 2010). While the number densities of faint LAEs decrease by 0.5 dex at z > 6 (Ouchi et al. 2010), the number densities of the most luminous sources are relatively constant between z ≈ 5–7 (e.g. Santos et al. 2016; Zheng et al. 2017), or decrease only slightly by ≈0.2–0.3 dex (e.g. Konno et al. 2018). These luminous sources likely reside in early ionised bubbles (Hu et al. 2016; Stark et al. 2017; Mason et al. 2018), facilitating their Lyα observability during the EoR (e.g. Matthee et al. 2015; Weinberger et al. 2018; Songaila et al. 2018). Due to their luminosity, luminous LAEs are ideal for spectroscopic follow-up (e.g. Sobral et al. 2015; Hu et al. 2016; Matthee et al. 2017a; Shibuya et al. 2018). The most Lyα-luminous example among these sources at z ≈ 6.6 is CR7...
from 0.3 to 2.5 μm and the wealth of public available data from the COSMOS survey which were not explored in Hu et al. (2016).

In this work, we first summarise the available photometric information and present the new X-shooter observations in Sect. 2. We analyse the spectrum in detail and conclude on the redshift of COLA1 in Sect. 3. Section 4 presents the properties of COLA1 that we measure from the spectroscopic and photometric data available. We discuss the implications in Sect. 5 and summarise the results in Sect. 6. Throughout this work we adopt a flat ΛCDM cosmology with Ω_Λ,0 = 0.7, Ω_M,0 = 0.3 and H_0 = 70 km s\(^{-1}\) Mpc\(^{-1}\), a Salpeter (1955) initial mass function and magnitudes in the AB system (Oke 1974).

2. Data

2.1. Photometric measurements

COLA1 (Hu et al. 2016) is located in the COSMOS field (Capak et al. 2007; Scoville et al. 2007) and hence public imaging data are available in ±30 filters from the UV to IR (e.g. Ilbert et al. 2009). This includes high resolution data in the F814W filter from the Hubble Space Telescope (HST)/Advanced Camera for Surveys (Koekemoer et al. 2007). COLA1 is located in a shallow region from the UltraVISTA survey (McCracken et al. 2012) and is covered by Spitzer/IRAC imaging from the SPLASH programme (e.g. Steinhardt et al. 2014)\(^1\). We use optical data from CFHT/Megacam, Subaru/Suprime-cam (Taniguchi et al. 2007)\(^2\) and near-infrared data from UltraVISTA DR3. COLA1 is furthermore covered in Subaru/Suprime-cam NB921 images from Sobral et al. (2012, 2013), similar to the NB921 imaging from Hyper Suprime-Cam from Hu et al. (2016). We show thumbnail images in the u, B, z', NB921, F814W, Y, J, H, [3.6] and [4.5] bands in Fig. 1, where red contours illustrate the location of line-emission (measured from the NB921-z image). We note that we have confirmed the astrometric alignment of the different images using the position of ≃30 stars and galaxies around COLA1. Our photometric measurements are summarised in Table 1. The measurements on ground-based images in bands blue-wards of 900 nm are conducted with 1.2′′ diameter apertures optimised to estimate the S/N of potential detections. Measurements in NB921, z', Y, J and H are performed with 2′′ diameter apertures.

2.1.1. NB921 Narrow-band data

Combining the Suprime-cam NB921 narrow-band data (Sobral et al. 2013) with public z' band data, we measure a total line-flux of \(5.8_{-1.3}^{+1.2} \times 10^{-17}\) erg s\(^{-1}\) cm\(^{-2}\). As the mean transmission of the NB921 filter is 69% at the wavelength where the emission-line is detected, we correct this line-flux to \(8.2_{-1.7}^{+1.4} \times 10^{-17}\) erg s\(^{-1}\) cm\(^{-2}\). Spreading this line-flux over the width of the z' band filter results in \(z = 25.7\), meaning that the full z' band flux can be explained by line-emission. We use the \(Y\) band \((Y = 25.2_{-0.3}^{+0.4}\), \(S/N \approx 3\), measured in a 2′′ diameter aperture and applying an aperture correction of −0.3 mag) to obtain a weak constraint on the observed equivalent width of \(\text{EW}_{\text{line}} = 900_{-340}^{+340}\) Å assuming a UV slope \(\beta = -2.0\). The Y band magnitude is in good agreement with the measurement in Hu et al. (2016) based on Subaru data. The line-flux we measure

\(^1\) Available through http://splash.caltech.edu
\(^2\) Available through http://irsa.ipac.caltech.edu/data/COSMOS/
is a factor of approximately two lower than Hu et al. (2016). We have checked that this is not due to the use of a smaller aperture (2") used herecentrede versus 3" diameter in Hu et al. 2016), as using a 3" aperture would only increase the flux by 10%. We note that our photometry in 3" apertures is consistent with Hu et al. (2016), meaning that the difference is due to the method to calculate the line-flux.

2.1.2. Ground-based optical data

Visual inspection of the thumbnail images reveals potential detections in the $u$ and $B$ filters. While the potential flux in the $u$ band is offset, this is not as clear in the $B$ band. We measure the significance of the flux in the $B$ band by performing photometry centred on the NB921 detection in an optimised 1.2" diameter aperture. We choose such small aperture to optimise the $S/N$ (the PSF-FWHM of this data is 0.7–0.8") and minimise contamination from nearby objects.\(^3\) The noise level is measured from the standard deviation of 1000 empty aperture measurements located on random sky positions around COLA1. We measure flux at the 2.4 $\sigma$ level with $B = 28.4^{+0.6}_{-0.4}$ (see Fig. 1, where we use a high contrast to highlight potential detections). No flux above the 2$\sigma$ level is detected in the $u$, $V$, $R$ and $I$ filters using the same aperture centred on COLA1. We measure flux at the 2$\sigma$ level in the stacked BVRI image. While the significance levels of these measurements is low, they may be troublesome for a galaxy at $z > 6.5$, motivating the need for careful spectroscopic analysis.

2.1.3. High resolution optical data

The HST imaging in the F814W filter (with >20% transmission between $\lambda = 6988$–9577 Å, $\lambda_{\text{eff}} = 7985$ Å, PSF-FWHM = 0.095", Koekemoer et al. 2007) reveals a faint point-like detection at the position of COLA1 (Fig. 1). Similarly to the $u$ and $B$ imaging, F814W imaging also clearly shows another source 1.2" south-west of COLA1 (identified in the HST thumbnail in Fig. 1). This object has ID 593625 in the Laigle et al. (2016) catalogue (with photometric redshift $z = 1.9^{+0.2}_{-0.1}$) and we will refer to it with that ID from now on. COLA1 is detected at

\[^3\] We note that we do not use measurements with different apertures to measure colours.

---

**Table 1. Coordinates and photometric measurements of COLA1.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Photometry</td>
<td>$B = 28.4^{+0.6}_{-0.4}$</td>
<td>$V &gt; 27.8$</td>
</tr>
<tr>
<td></td>
<td>$R &gt; 27.7$</td>
<td>$I &gt; 27.6$</td>
</tr>
<tr>
<td></td>
<td>$F814W = 26.6^{+0.4}_{-0.3}$</td>
<td>$NB921 = 23.6^{+0.3}_{-0.3}$</td>
</tr>
<tr>
<td></td>
<td>$z' = 25.5^{+0.3}_{-0.3}$</td>
<td>$Y = 25.7^{+0.3}_{-0.3}$</td>
</tr>
<tr>
<td></td>
<td>$J = 24.8^{+0.3}_{-0.3}$</td>
<td>$H &gt; 24.8$</td>
</tr>
<tr>
<td></td>
<td>$K_s &gt; 24.4$</td>
<td>$[3.6] &gt; 24.4$</td>
</tr>
<tr>
<td></td>
<td>$[4.5] &gt; 24.6$</td>
<td></td>
</tr>
</tbody>
</table>

**Notes.** Magnitudes are in the AB system. Magnitude limits are at the 2$\sigma$ level. As the $H$ band photometry is contaminated by ID 593625 we only provide it as a lower limit.
3$r$, with F814W = $26.6^{+0.4}_{-0.3}$ measured in a 0.6′′ diameter aperture and corrected for aperture losses using tabulated encircled fluxes from Bohlin (2016). ID 593625 has a F814W magnitude of F814W = $26.5^{+0.3}_{-0.3}$. If continuum flux at this level contributes to the NB921 photometry it could contribute to the 10% increase in the line-flux of COLA1 measured with 3′′ apertures.

### 2.1.4. Photometric redshift and the Spitzer/IRAC view

COLA1 is present in the public COSMOS2015 catalogue (Laigle et al. 2016; ID 593751), where it has a photometric redshift of $z_p = 0.99^{+0.02}_{-0.02}$. However, the photometry that is used to estimate this redshift is measured with 2′′ diameter apertures and may suffer contamination from ID 593625. We re-measure the photometry in the Spitzer/IRAC [3.6] and [4.5] filters using SPLASH data. We follow the procedure as last described in Stefanon et al. (2017), where the IRAC images are de-confused based on the HST/F814W images using the mophngo software (Labbé et al. 2006, 2015). We measure [3.6] = $24.4^{+0.2}_{-0.1}$ and [4.5] = $24.6^{+0.2}_{-0.2}$ in 1.8′′ diameter apertures and including an aperture correction as described in Labbé et al. (2015), see Fig. 1. This results in a moderately blue colour [3.6] − [4.5] = −0.2 ± 0.3, although with significant uncertainties (see also Harikane et al. 2018, who measure [3.6] − [4.5] = −0.2 ± 0.1).

### 2.2. X-shooter spectroscopic observations and data reduction

We observed COLA1 with the X-shooter echelle spectrograph on UT2 of the Very Large Telescope as part of ESO programme 0100.A-0213 (PI Matthee). X-shooter observes in three arms simultaneously: UVB (λ = 0.30−0.55 μm), VIS (λ = 0.55−1.02 μm) and NIR (λ = 1.0−2.5 μm). Observations were performed on 19 January and 18−19 February 2018 with clear conditions and 0.6−0.8′′ seeing. GD71 was observed as telluric standard star. We used a 1.0′′ slit in the UVB arm and 0.9′′ slits in the VIS and NIR arms.

We first acquired on a $J = 13.5$ magnitude star and then applied a blind offset of 95.86′′ to the position of COLA1 with a position angle of 0 degrees (see Fig. 1). We nodded between two positions along the slit (A and B, offset by 3′′) using the AutoNod0nSlit1 procedure. Integration times were 700 and 780 s for UVB and VIS arms respectively. Four shorter jitters with separations of 1′′ along the direction of the slit were used for NIR exposures of 210 s. This observing sequence was repeated ten times (split over two observing blocks with ABBA sequence and one with an AB sequence), resulting in total exposure times of 7.0, 7.8 and 8.4 ks for the three arms respectively.

Data are reduced using standard procedures (bias and dark subtraction, flat-fielding, sky subtraction, wavelength calibration and flux calibration) incorporated in the X-shooter pipeline (Modigliani et al. 2010). Individual exposures per observing block are combined using the pipeline. We then combine exposures over multiple observing blocks by computing the noise-weighted average as described in detail in Sobral et al. (2018b).

Before extracting 1D spectra, we first smooth the spectrum with 2D gaussian kernels that corresponds to half the resolution ($σ_r$, $σ_y$)UVB,VIS = (0.4 Å, 0.32′′) and ($σ_r$, $σ_y$)NIR = (0.6 Å, 0.21′′), and bin it in the wavelength direction by 3 pixels (0.6 Å, half the resolution). The spatial extraction window in the VIS arm is 2.2′′ which we find to optimise the S/N. Slit losses of the line at 922 nm are estimated using the NB921 image that is convolved to the PSF-FWHM of our spectroscopic observations. We measure the fraction of the total flux that is retrieved within the slit and extraction window. As COLA1’s line-emission is compact in the NB921 data, the estimated slit losses are only 19%. The spatial extraction windows in the UVB and NIR arms are 1.2′′ and we do not apply slit loss corrections. Wavelengths are converted to vacuum wavelengths. We measure the effective spectral resolution using unresolved skylines on our extracted 1D spectrum. We find $R \approx 4000$ at 0.5 μm, $R \approx 5000$ at 0.9 μm and $R \approx 3800$ at 1.6 μm, corresponding to 75, 60 and 80 km s$^{-1}$ respectively.

### 2.3. Spectroscopic measurements

In our X-shooter spectrum, we confirm COLA1’s double peaked emission line at λobs,vac = 9224, 9231 Å ($S/N ≈ 24$), but we also detect faint emission-lines at λobs,vac = 3821 Å ($S/N = 10$) and λobs,vac = 15735 Å ($S/N ≈ 5$). The centroid of these faint emission-lines is however shifted spatially by 1′′ to the west and are therefore not co-located with COLA1 (Fig. 1). We do not detect continuum emission. No other lines are detected in the 0.3−2.5 μm coverage above $S/N > 2$.

For the double peaked emission-line at 923 nm, we measure a line-flux of $5.90 \pm 0.24 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$ and a relative flux ratio between the blue and red component of $0.31 \pm 0.03$. The red line has a full width half maximum (FWHM) of $198 \pm 14$ km s$^{-1}$, while the blue line is narrower with FWHM = $150 \pm 18$ km s$^{-1}$. While the red line is clearly asymmetric (with a weighted skewness of $S_w = 18.0 \pm 0.9$; following the definition from Kashikawa et al. 2006), the blue line is not ($S_w = −0.2 \pm 0.3$ Å). Our measurements are summarised in Table 2. The

![Table 2. Details of the X-shooter observations and spectroscopic measurements for COLA1.](https://example.com/table2)

<table>
<thead>
<tr>
<th>Observations</th>
<th>COLA1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dates</td>
<td>19 Jan 2018, 18–19 Feb 2018</td>
</tr>
<tr>
<td>$R_{0.9}$ μm</td>
<td>5000</td>
</tr>
<tr>
<td>$R_{1.6}$ μm</td>
<td>3800</td>
</tr>
<tr>
<td>$t_{exp, UVB}$</td>
<td>7.0 ks</td>
</tr>
<tr>
<td>$t_{exp, VIS}$</td>
<td>7.8 ks</td>
</tr>
<tr>
<td>$t_{exp, NIR}$</td>
<td>8.4 ks</td>
</tr>
<tr>
<td>Flux limits of interest</td>
<td></td>
</tr>
<tr>
<td>$\text{Hα}_{R}=1.475$</td>
<td>$&lt;5.2 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$</td>
</tr>
<tr>
<td>$\text{[OIII]}_{3899, 3798}=1.475$</td>
<td>$&lt;9.8 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$</td>
</tr>
<tr>
<td>$\text{[OIII]}_{3907, 3795}=1.475$</td>
<td>$&lt;9.0 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$</td>
</tr>
<tr>
<td>$CIV_{1549}=6.591$</td>
<td>$&lt;1.6 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$</td>
</tr>
<tr>
<td>$HeII_{1640}=6.591$</td>
<td>$&lt;0.7 \times 10^{-17}$ erg s$^{-1}$ cm$^{-2}$</td>
</tr>
</tbody>
</table>

Notes. Flux measurements of the 923 nm line are corrected for 19% slit-l losses. Upper and lower limits are at the 2σ level. Flux limits assume a line-width FWHM = 200 km s$^{-1}$.
spectroscopically measured flux differs from the narrow-band measurement at the 1.3σ level, potentially due to systematic uncertainty in the flux calibration or the aperture correction. For the remainder of the paper we use the narrow-band flux corrected for the filter transmission.

The line at 3821 Å has a line-flux of 3.0 ± 0.3 × 10^{-17} erg s^{-1} cm^{-2} and FWHM = 350 km s^{-1}. The line that is observed at 15735 Å has a line-flux of 1.5 ± 0.3 × 10^{-17} erg s^{-1} cm^{-2} and a width FWHM = 260 km s^{-1}. The spatial offset of ≈1″ of these lines coincides with ID 593625.

3. Is COLA1 a LAE at z = 6.59 or an [Oii] emitter at z = 1.47?

The tentative detection in the B band (although at the 2.4σ level) and the unexpected observation of Lyα flux blue-wards of the red asymmetric line might suggest that COLA1 may not be a LAE at z = 6.6, but an [Oii] emitter at z = 1.47 instead. Here, we assess the implications of each observed feature to test this hypothesis.

3.1. 923 nm line strength and consistency with F814W

If the 923 nm line from COLA1 is [Oii], it would yield z = 1.475 and it would have a high luminosity (L_{[Oii]} = 2 × L^⋆; Khostovan et al. 2015). The [Oii] equivalent width would be extreme, EW_{[Oii]} = 360^{+140}_{-100} Å, well above the typical EW for [Oii] emitters (EW_{[Oii]} = 50 ± 20 Å) at z = 1.47 (Khostovan et al. 2016).

We test whether the F814W photometry can be explained by pure line-emission. Assuming negligible contribution from continuum emission and by spreading the line-flux homogeneously over the full transmission region of F814W results in F814W = 26.5 ± 0.2. Hence, the F814W photometry (we measure F814W = 26.6^{+0.6}_{-0.4}) can be perfectly explained by pure line-emission and does not indicate flux blue-wards of the emission-line. Therefore, the F814W detection does not rule out Lyα at z = 6.6.

3.2. Line-detections at 382 nm and 1573.5 nm

As described above, we detect two emission-lines that are offset by ≈1″ to the west of COLA1. This corresponds roughly to the foreground source ID 593625 as can be seen in the B band thumbnail image (Fig. 1), which shows that outskirts of this galaxy fall in the X-shooter slit. We identify the two lines as Lyα and [Oiii] at z = 2.142, as illustrated in Fig. 2, where we have optimised the centroid of spatial extraction aperture for these lines. Lyα is redshifted by 130 km s^{-1} with respect to the systemic redshift, similar to other LAEs at z ∼ 2 (Trainor et al. 2015; Sobral et al. 2018b). Without correcting for slit losses, we measure a Lyα luminosity of 1.0 ± 0.1 × 10^{42} erg s^{-1}, well below the typical Lyα luminosity at z ≈ 2.2 (≈0.4 × L^⋆; Sobral et al. 2017). The B band magnitude corresponding to pure line-emission at this flux-level is B = 28.2 ± 0.1. Lyα emission extends to close to the peak position of COLA1 and could thus contribute significantly to the faint B band detection (B = 28.4^{+0.6}_{-0.4}) at the COLA1 position. The [Oiii] luminosity is 0.5 ± 0.1 × 10^{42} erg s^{-1}. We note that the centroid of the (low S/N) flux in the B- near COLA1 is shifted slightly towards a faint H band detection that could be explained by this [Oii] flux. Therefore, several detections in the images (in particular B and H) can be attributed to LAE 593625 at z = 2.142.

3.3. 923 nm line-profile

We show a detailed figure of the line-profile of the 923 nm line in Fig. 3. In this figure, we shift the spectrum to the rest-frame assuming COLA1 is at z = 1.475. We compare it to a median [Oii] spectrum of galaxies with asymmetric red lines and high red/blue ratios from the VISCOS survey (Paulino-Afonso et al. 2018a) and also show the best fitted [Oii] doublet. While the peak separation (220 ± 20 km s^{-1}) is fully consistent with the peak separation of the [Oii] doublet, the asymmetry of the red line can not be fitted as an [Oii] doublet. In particular, the blue line would have been expected to be asymmetric as well, similar to the red line. Moreover, the absence of flux between the two lines can also not easily be explained in the case of an [Oii] doublet, unless lines are very narrow.

The skewness of COLA1’s red line is high (S_{w,red} = 18.0 ± 0.9 Å; see Table 2), much higher than the typical maximum skewness of a low-redshift galaxy (S_{w} = 3 Å; Kashikawa et al. 2006), which is similar to the skewness of the full doublet. Finally, the line-ratio between [Oiii]_{5007} and [Oii]_{3726} would be significantly lower than the line-ratio in our extreme [Oii] emitter reference sample (with a blue-to-red fraction of >0.65). In fact,
the line-ratio of the blue and red lines of our best-fit (0.44) is significantly lower than the theoretical minimum line-ratio for electron densities as low as 1 cm$^{-3}$ ($\approx$0.65 for an electron temperature of $10^5$ K, Sanders et al. 2016, but also for temperatures between $10^4$–$5^5$ K) and thus unphysical. Therefore, the line-profile strongly disagrees with COLA1 being an [OII] emitter at $z = 1.475$ even though the peak separation is in perfect agreement.

3.4. No line in the NIR associated with $z = 1.475$

As we show in Fig. 4, we do not detect Hα or [OIII]$_{5007}$ if COLA1 would be at $z = 1.475$. We illustrate how these lines would look in our data for the lowest empirical Hα/ [OII] ratio ($\approx 0.2$, Hayashi et al. 2013). These extreme line-ratios would imply both high metallicities and high ionisation states (properties that are typically anti-correlated; Nakajima & Ouchi 2014). If we conservatively assume the spectroscopic flux of the 923 nm line and a FWHM of 200 km s$^{-1}$ (similar to the red line at 923 nm), we measure a 1σ limit on the Hα flux of $2.6 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$ and a lower limit of $9.4 \times 10^{18}$ erg s$^{-1}$ cm$^{-2}$ on the combined [OII]$_{4959,5007}$ line. While the lowest [OIII]/[OII] ratio is within the 1σ noise level, these data rule out Hα/[OII] = 0.2 at 4.7σ, additional clear evidence against the interpretation that COLA1 is an [OII] line at $z = 1.475$.

3.5. Concluding remarks on the redshift of COLA1

As the line-profiles of the red and blue lines differ, the double peaked emission line around 923 nm can not be fitted by an [OII] doublet, even though the peak separation is similar.\footnote{As we show in Sect. 4, the Ly$\alpha$ line profile is well-fitted by a Ly$\alpha$ shell model and shows properties similar to normal LAEs at $z < 6$.}

Photometric indications for COLA1 being at low-redshift (in particular B and F814W detections) are explained due to line-emission from COLA1 itself (F814W) and a foreground LAE at close separation (ID 593625 at $z = 2.142$). The relatively blue Spitzer/IRAC colours, combined with a optical to near-infrared break of $BVRI - Y > 3$ are also not indicative of a red dusty or old interloper at a lower redshift. Finally, if COLA1 would have been an [OII] emitter at $z = 1.475$ the flux ratio of the lines in the [OII] doublet and the limits on [OIII]/Hα indicate unphysical conditions. Combining all the observations from above, we conclude that COLA1 is best explained as a double-peaked LAE at $z_{LY,	ext{red}} = 6.593$, as initially proposed by Hu et al. (2016) and corroborated by Songaila et al. (2018).

4. Properties of COLA1 – a unique LAE at $z = 6.6$

Now we have established that COLA1 is a real LAE at $z = 6.593$, we can have a better look at its properties based on the X-shooter spectrum and available imaging data. The properties are summarised in Table 3.

4.1. Ly$\alpha$ luminosity and spectral energy distribution

The transmission-corrected narrow-band magnitude implies a Ly$\alpha$ luminosity of $L_{\text{Ly}\alpha} = 4.1 \times 10^{43}$ erg s$^{-1}$, which is among the most luminous LAEs know at $z \approx 5$–7 (see a compilation in Matthee et al. 2017a). The Ly$\alpha$ EW is high ($\text{EW}_{\alpha} = 120^{+50}_{-40}$ A, based on the continuum estimated from the Y band flux), but this is a rather common property of LAEs at high-redshift (e.g. Hashimoto et al. 2017; Sobral et al. 2018a) and we note that it is poorly constrained due to uncertainties on the continuum magnitude. The Ly$\alpha$ luminosity implies a comoving number density $\approx 1 \times 10^{-5}$ Mpc$^{-3}$ (Matthee et al. 2015).

COLA1 is weakly detected in the Y and J bands ($S/N \approx 3$ and 2.5, respectively), see Fig. 1. An off-centred H band detection is likely contaminated by strong [OIII]$_{4959,5007}$ emission from the foreground galaxy at $z = 2.14$, and we can only provide a 2σ limit on the magnitude of $H > 24.8$. COLA1 is undetected in the K$_s$ band ($K_s > 24.4$). Due to the large uncertainties in the photometry no meaningful constraints can be obtained on the UV slope, and we assume a flat UV slope $\beta = -2$ in the rest of the paper (this is similar to other LAEs at $z = 6.6$; Ono et al. 2010; Matthee et al. 2017a).

The Y band magnitude implies an absolute magnitude $M_{\text{Y,}5000} = -21.6 \pm 0.3$, slightly above $M_{\text{Y,}5000} = 7$...
Table 3. Derived properties of COLA1 as described in Sect. 4.

<table>
<thead>
<tr>
<th>Property</th>
<th>Best estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral analysis</td>
<td></td>
</tr>
<tr>
<td>$L_{\text{Ly}\alpha}$</td>
<td>$4.1 \pm 0.2 \times 10^{43}$ erg s$^{-1}$</td>
</tr>
<tr>
<td>EW($\text{Ly}\alpha$)</td>
<td>$120^{+30}_{-40}$ Å</td>
</tr>
<tr>
<td>$M_{1500}$</td>
<td>$-21.6 \pm 0.3$</td>
</tr>
<tr>
<td>$M_{\text{star}}$</td>
<td>$\approx 10^{10}$ $M_\odot$</td>
</tr>
<tr>
<td>$r_{50}$</td>
<td>$0.33^{+0.07}_{-0.04}$ kpc</td>
</tr>
<tr>
<td>$f_{\text{esc,LyC}}$</td>
<td>$\approx 15-30%$</td>
</tr>
<tr>
<td>SFR$_{\text{Ly}\alpha}$</td>
<td>$75^{+35}<em>{-25}$ $M</em>\odot$ yr$^{-1}$</td>
</tr>
<tr>
<td>SFR$_{\text{UV,nodust}}$</td>
<td>$27^{+55}<em>{-10}$ $M</em>\odot$ yr$^{-1}$</td>
</tr>
<tr>
<td>$\Sigma_{\text{SFR,UV,nodust}}$</td>
<td>$95^{+30}<em>{-15}$ $M</em>\odot$ yr$^{-1}$ kpc$^{-2}$</td>
</tr>
<tr>
<td>Gas-phase metallicity $Z$</td>
<td>$\lesssim 10^{-2}$ $(1/20 Z_\odot)$</td>
</tr>
<tr>
<td>Shell model fitting</td>
<td></td>
</tr>
<tr>
<td>$v_{\text{exp}}$</td>
<td>$78^{+5}_{-3}$ km s$^{-1}$</td>
</tr>
<tr>
<td>$\log_{10}(N_{\text{HI}}/\text{cm}^{-2})$</td>
<td>17.0$^{+0.3}_{-0.2}$</td>
</tr>
<tr>
<td>$\log_{10}(T/K)$</td>
<td>4.2$^{+0.1}_{-0.2}$</td>
</tr>
<tr>
<td>$\sigma_{\text{intrinsic}}$</td>
<td>159$^{+4}_{-1}$ km s$^{-1}$</td>
</tr>
<tr>
<td>$\tau_d$</td>
<td>4.2$^{+0.5}_{-0.8}$</td>
</tr>
<tr>
<td>$z_{\text{sys}}$</td>
<td>6.5930$^{+0.0001}_{-0.0002}$</td>
</tr>
</tbody>
</table>

Notes. Limits are at the 2σ level.

COLA1’s colours in the shortest Spitzer/IRAC channels are blue, [3.6] − [4.5] = −0.2 ± 0.3. This is likely a consequence of Hβ and/or [OIII] slightly boosting the [3.6] flux more than Hα is boosting the [4.5] flux (which is in a low transmission wavelength of the [4.5] band). However, COLA1 is not nearly as blue as other confirmed galaxies at $z \sim 6.6$. For example, CR7 and Himiko have colours [3.6] − [4.5] = −1.3 ± 0.3 and [3.6] − [4.5] = −0.7 ± 0.4, respectively (Harikane et al. 2018). This could indicate that COLA1’s [OIII] line is relatively weaker due to a lower metallicity. When we compare this colour to the predicted colours at $z = 6.6$ using photoionisation analysis and BPASS SED models presented in Bowler et al. (2016), we find that the implied metallicity lies between $Z \approx 10^{-2} - 10^{-3}$, but always consistent with $Z < 10^{-3} (<12/20 Z_\odot)$ within the uncertainties. Alternatively, the IRAC colour could also indicate that an older stellar population is present in COLA1, or that the strength of nebular lines is lower due to a high LyC escape fraction (e.g. Zackrisson et al. 2017).

4.2. Lyα size and star formation rate density

As discussed in Sect. 3, the detection of COLA1 in the F814W band can be explained by pure Lyα emission. As the HST/ACS imaging has a small PSF, we can therefore use these data to constrain the size of the Lyα emitting region in COLA1 accurately (e.g. Paulino-Afonso et al. 2018b). We use a non-parametric method to measure the median half-light radius ($r_{50,\text{obs}}$) for 5000 random realisations of the image, where each pixel count is drawn from a Poissonian distribution. Then, we correct for PSF-broadening using $r_{50} = \sqrt{r_{50,\text{obs}}^2 - r_{\text{PSF}}^2}$, where $r_{\text{PSF}} = 0.38$ kpc.

This results in $r_{50} = 0.33^{+0.07}_{-0.04}$ kpc. This size is smaller than typical UV-selected galaxies with $M_{1500} \approx -21$, similar to galaxies with $M_{1500} = -18.4$ in the UV luminosity - size relation from Shibuya et al. (2015) and similar to the largest star cluster complexes (Bouwens et al. 2017; Vanzella et al. 2017). As the UV sizes of galaxies are typically much smaller than the Lyα sizes (e.g. Wisotzki et al. 2016), it could be possible that the galaxy is even more compact in the UV. On the other hand, it is likely that more diffuse Lyα emission is resolved out in the broad, high resolution F814W image (similar to for example in CR7; more luminous, extended Lyα emission that is undetected in similar F814W data), and that the core Lyα profile resembles the UV profile. Lyα emission is expected to be compact in galaxies with a low HI column density and high $f_{\text{esc,LyC}}$ (Mas-Ribas et al. 2017) – conditions that are likely present in COLA1 as discussed next.

We can estimate the SFR of COLA1 from the UV continuum or from the Lyα luminosity. Assuming a UV slope $\beta = -2.0$ and a Meurer et al. (1999) attenuation law, we find a dust-corrected SFR$_{\text{UV}} = 40^{+10}_{-7}$ $M_\odot$ yr$^{-1}$ (27$^{+9}_{-6}$ $M_\odot$ yr$^{-1}$ without correcting for dust). Alternatively, the Lyα luminosity indicates SFR$_{\text{Ly}\alpha} = 75^{+35}_{-25}$ $M_\odot$ yr$^{-1}$ (assuming a 60% Lyα escape fraction and 15% LyC escape fraction; following Sobral & Matthee 2018).

If we assume that the UV size is similar to the Lyα size, these measurements imply a minimum average SFR surface density $\Sigma_{\text{SFR,UV,nodust}} = 95^{+30}_{-15}$ $M_\odot$ kpc$^{-2}$. This number would be higher if the UV size is more compact or if significant dust attenuation is present. This $\Sigma_{\text{SFR}}$ is significantly higher than other galaxies known at $z > 6$ (e.g. Carniani et al. 2018). The SFR surface density is well above the threshold required to drive galactic outflows (e.g. Heckman et al. 2001), hence such outflows may well be present.
foreground contamination is a major issue \cite{siana2015}, measuring the Lyα peak separation may be a promising alternative.

In order to get a more quantitative view of the ISM properties implied by the observed Lyα line, we fit the Lyα profile of COLA1 using a five-parameter shell model as in \cite{dijkstra2016}. This model (e.g. \cite{ahn2001, verhamme2006}) consists of a shell of neutral gas and dust around a central ionising source and parameters include the HI column density, expansion velocity, temperature and dust optical depth of the shell and the intrinsic width of the Lyα line \cite{gronke2015}. This fit does not include transmission through the IGM.

The best fitted parameters are listed in Table 3. Compared to $z \approx 3$–5 LAEs \cite{gronke2017}, these parameters indicate a relatively low HI column density $N_{\text{HI}} = 10^{17.3 \pm 0.3} \text{cm}^{-2}$ and a high dust optical depth $\tau_d = 4_{-2}^{+6}$, while the other parameters are quite common for LAEs. The low HI column density is inferred from the low peak separation and suggests the possible escape of LyC photons. The high dust optical depth could alternatively be interpreted due to IGM opacity that could lower and narrow the blue peak. The HI column density and expansion velocity are also significantly lower than the inferred column density around CR7 \cite[e.g.][]{dijkstra2016}. We note that the shell-model prediction of the systemic velocity of CR7 presented in \cite{dijkstra2016} agrees perfectly with recent [CII] measurements in \cite{matthee2017}, suggesting that shell-model fitting is a good tool to recover the systemic redshift of a LAE at high redshift. The column density is also significantly lower than the column density inferred from absorption line measurements in local LyC leakers \cite{gazagnes2018}, which lead these authors to conclude that the HI covering fraction in these galaxies is non-uniform. A relatively high column density in these local LyC leakers is consistent with their larger Lyα peak separation compared to COLA1, and their typical escape fraction of $\approx 5\%$.

As detailed in \cite{verhamme2015}, the Lyα line profile can be used as a tracer of $f_{\text{esc, LyC}}$ as both are sensitive to HI column density as $N_{\text{HI}} = -\ln(f_{\text{esc, LyC}}) / \sigma_0$ where $\sigma_0 = 6.3 \times 10^{-18} \text{cm}^2$ is the ionisation cross section. According to this equation, the column density derived from the shell-model fit (see Table 3) implies $f_{\text{esc, LyC}} \approx 50\%$. This is consistent with the result one would infer by using the correlation between Lyα EW_0 and $f_{\text{esc, LyC}}$ at $z \approx 3$ \cite{steidel2018}, although we note this correlation has not been tested beyond EW_0 $\geq 50$ Å and likely breaks down due to a reduced Lyα strength in the high $f_{\text{esc, LyC}}$ regime. The Lyα EW also points towards a $30\%$ escape fraction as implied by comparison with the most recent measurements in local LyC leakers \cite{izotov2018}. Such high escape fractions would also affect the strength of nebular emission lines such as Hα and Hβ and could therefore be tested with future spectroscopic observations with the \textit{James Webb} Space Telescope \cite[e.g.][]{jensen2016, zackrisson2017}. Indeed, the fact that COLA1 shows a [3.6] – [4.5] colour that is much closer to zero than other luminous LAEs at the same redshift could potentially indicate a reduced strength of nebular emission lines, see Sect. 4.1.

4.4. High ionisation rest-UV emission-lines

Thanks to the wavelength coverage of the X-shooter spectrum, we can inspect the spectrum for the presence of strong rest-UV emission line features (e.g. \cite{stark2015}; see Fig. 8). We do not detect a significant feature besides Lyα. We measure 2σ limiting line-fluxes of 1.6 and $0.7 \times 10^{-18} \text{erg} \text{s}^{-1} \text{cm}^{-2}$...
for the CIV$_{1546,1556}$ and HeII$_{1640}$ emission lines at $z = 6.591$, respectively (over 250 km s$^{-1}$ extraction boxes), corresponding to EW$_{O,CIV} < 25$ Å and EW$_{O,HeII} < 12$ Å. These non-detections are not surprising given the fact that the limits are shallower than other unsuccessful spectroscopic follow-up of LAEs at $z = 6.6$ (e.g. Shibuya et al. 2018), and EW limits are higher than detections in for example Stark et al. (2017).

5. Discussion: witnessing a galaxy reionising its surroundings

5.1. The unlikelihood of observing a blue Lyα line at $z > 6$

Lyα photons resonantly scatter in the presence of neutral hydrogen. Once Lyα photons are absorbed, they are re-emitted in a random direction in the rest-frame of the absorbing hydrogen atom, resulting in a diffusion process in real and frequency space. Analytical models show that this process results in a double peaked Lyα spectrum in a static medium (e.g. Neufeld 1990). In the presence of outflows, Lyα photons see a larger optical depth towards the blue, resulting in a redshifted asymmetric spectrum. Hence, the Lyα profile is sensitive to the neutral hydrogen content and the velocity field of the gas in a galaxy (e.g. Loeb & Rybicki 1999; Santos 2004; Dijkstra et al. 2007).

After escaping from the ISM, Lyα photons are also affected by HI in the circum-galactic medium (CGM). The CGM predominantly transmits red Lyα photons, enhancing the asymmetry between the red and the blue peak (e.g. Laurensen et al. 2011). This is consistent with observations at high-redshift ($z > 2$), where high HI column densities in the CGM result in a majority of Lyα profiles that consist of a single, red asymmetric line (e.g. Erb et al. 2014). Due to the increasing neutral fraction of the IGM, the asymmetry between the transmission of the red and blue line increases and the chances of observing double peaked emission decrease.

This is also found in semi-analytical and hydrodynamical models of the EoR, which typically predict negligible transmission bluewards of the systemic velocity at $z > 6$ (Dijkstra et al. 2007; Laurensen et al. 2011; Weinberger et al. 2018). In the radiative transfer simulations from Laurensen et al. (2011) the median IGM transmission at the blue peak is $>30\%$ only at $z < 4.5$, while it is $<0.1\%$ at $z > 5$ (similar to more recent simulations by Weinberger et al. 2018). Given that these models predict such low transmission bluewards of the systemic redshift, how is it possible that we observe a strong blue peak in the Lyα profile of COLA1 at $z = 6.59$, a redshift where this is highly unexpected? Here, we propose three scenarios that facilitate the transmission of blue Lyα photons and that may explain COLA1’s Lyα profile.

5.2. Three scenarios to explain double peaked Lyα emission at $z = 6.6$

In Fig. 9, we sketch three different physical scenarios that may explain the Lyα profile of COLA1. We first describe these models and then we propose observations that may test and differentiate between them. We note that the transmission, line-widths and offsets in Fig. 9 are chosen for illustrative purposes.

In Model 1, COLA1 is surrounded by a highly ionised bubble that is large enough for Lyα photons to redshift out of resonance due to the expansion of the Universe before encountering a relatively neutral IGM (e.g. Haiman 2002; Hu et al. 2016). The Lyα profile is therefore only determined by the ISM conditions of the galaxy. In Model 2 the bubble itself is smaller, but is embedded in a larger relatively ionised region. The Lyα profile here is mostly affected by the gas in the CGM and IGM. Model 3 attempts to explain COLA1’s Lyα profile without invoking IGM attenuation, but rather by an infalling self-shielded cloud of neutral hydrogen. This neutral hydrogen cloud absorbs part of the emitted Lyα line, resulting in a double peaked spectrum.

5.2.1. Model 1: Large, highly ionised bubble in a neutral IGM

If COLA1 resides in a large, highly ionised region, it is possible that blue Lyα photons redshift out of the resonance wavelength due to the Hubble expansion prior to encountering a neutral IGM (e.g. Malhotra & Rhoads 2006). Depending on the size of this highly ionised region, it is possible that only (part of) the blue peak is attenuated (e.g. Haiman 2002). In the left panel of Fig. 9, we illustrate this scenario where a double peaked Lyα line escapes from the ISM of the galaxy (resulting from resonant scattering effects, e.g. Neufeld 1990; Gronke et al. 2017), before the blue part of the line is attenuated by a sigmoid transmission function of the IGM.

The Lyα photons in the blue peak of COLA1 need to redshift out of the resonance wavelength by at least $\gtrsim 250$ km s$^{-1}$ (the maximum velocity at which we observe blue Lyα flux compared to line-centre, see Fig. 7) before encountering significant amounts of neutral hydrogen. This requires a large ionised region. Ignoring peculiar velocities of inflowing gas, we can calculate the required size as follows:

$$d_{\text{prop}} = \frac{\Delta \nu}{H_{z=6.59}} \text{Mpc}. \quad (1)$$

Here, $d_{\text{prop}}$ is the proper distance photons are required to travel, $\Delta \nu$ the required velocity offset and $H_{z=6.59}$ is the Hubble parameter at $z = 6.59$, which is $H_{z=6.59} = 803.85$ km s$^{-1}$ Mpc$^{-1}$ in our assumed cosmological model. Therefore, in order for Lyα photons to redshift by $\gtrsim 250$ km s$^{-1}$ would require an ionised sightline or region of at least 0.3 mPc; or $\gtrsim 2.3$ cMpc.
Simulations indicate that this bubble-size is similar to the characteristic bubble size for a global ionised fraction of $x_{\text{HI}} \approx 50\%$ (e.g. Furlanetto et al. 2006; Lin et al. 2016). As the ionised region needs to be at least 0.3 pMpc, and could be larger, this means the bubble size of COLA1 would correspond to a mean IGM $x_{\text{HI}} < 0.5$ at $z = 6.6$. On the other hand, the bubble around COLA1 may also be an outlier as COLA1 is likely a relatively massive galaxy. This could imply a higher neutral fraction.

We now assess whether such ionised region can be explained using the observed properties of COLA1. Following Haiman (2002), we estimate the maximum proper radius of the ionised region around a galaxy (in the absence of neighbouring ionising sources and ignoring recombinations):

$$R_s = 2.5 \left( f_{\text{esc, LyC}} \frac{Q_{\text{ion}}}{10^{54} \text{s}^{-1}} \right)^{1/3} \left( b_{\text{burst}} / 10^7 \text{yr} \right)^{1/3} (1 + z)^{-1} \text{Mpc}.$$  \hspace{1cm} (2)

Here, $R_s$ is the radius of the Strömgren sphere, $f_{\text{esc, LyC}}$ the escape fraction of ionising photons, $Q_{\text{ion}}$ the produced number of ionising photons per second and $t$ the age of the burst of star formation. As listed in Table 4, we estimate $f_{\text{esc, LyC}} = 0.15$ (based on the Lyα peak offset), $Q_{\text{ion}} = 5 \times 10^{54} \text{s}^{-1}$ (based on either the Lyα luminosity, or based on the UV luminosity, combined with a ionising photon production efficiency of $\xi_{\text{ion}} = 10^{25.2} \text{Hz erg}^{-1}$; Bouwens et al. 2016) and $b_{\text{burst}} = 10^7 \text{yr}$ (based on the high Lyα EW; e.g. Charlot & Fall 1993) at $z = 6.591$. This results in $R_{s,\text{max, Lyα}} = 0.33^{+0.07}_{-0.06} \text{pMpc}$ and $R_{s,\text{max, UV}} = 0.29^{+0.03}_{-0.03} \text{pMpc}$, corresponding to $\approx 2.5 \text{cMpc}$. The maximum radius would marginally increase by 0.03 pMpc when correcting the UV luminosity for a (high) attenuation $A_{\text{UV}} = 0.45$.

Therefore, under basic assumptions, the star-formation in COLA1 may provide enough photons that can ionise a large enough region for allowing the blue peak to be observed up to $\approx 250 \text{ km s}^{-1}$ from line-centre. However, there are important caveats that require attention. For example, our estimate conservatively assumes that the IGM does not affect Lyα photons red-wards of line-centre (as illustrated in the left panel of Fig. 9) and the real required ionised region may have to be significantly larger.

Moreover, the calculation so far also ignores peculiar velocities that typically blueshift Lyα photons with respect to neutral gas in the IGM (Laursen et al. 2011). Our calculation

\footnote{We note that these uncertainties are the propagated errors corresponding to the UV and Lyα luminosities and ignore uncertainties in the age of the burst of star formation and the escape fraction.} also ignores self-shielded neutral regions in the CGM around galaxies, that could be challenging to ionise by the galaxy itself and therefore may form a major source of opacity (e.g. Mesinger et al. 2015; Sadoun et al. 2017). The recent simulations from Weinberger et al. (2018) suggest that these self-shielded regions may be more common for galaxies that reside in halos with $M_{\text{halo}} \sim 10^{11} M_\odot$, for which the fraction of sight-lines that has a significant transmission on the blue side of the systemic redshift is consequently extremely low. Sight-lines with high blue-transmission do exist in a significant fraction of low mass halos ($M_{\text{halo}} \sim 10^7 M_\odot$). These simulations therefore suggest that COLA1 resides in a low mass halo, unless the ionisation from galaxies themselves have been under-estimated.

Finally, as noted in Haiman (2002), the radius of the ionised sphere is over-estimated in case recombinations are important, for example when the clumping factor in the IGM is high ($C > 10$) or the age of the star formation burst is $> 10^7 \text{yr}$. It is thus unclear whether the ionised bubble can also be sustained, in particular when the star formation rate of COLA1 would decline or if the escape fraction would decrease. A solution would be if a quasar or (faint) neighbouring galaxy contribute to the local ionising budget (e.g. Kakiichi et al. 2018). COLA1 is not located nearby a known quasar (Bañados et al. 2016). However, COLA1 is at a relatively close separation to CR7, the most luminous LAE known at $z = 6.6$. The projected distance on the sky of $34.17^\circ$, which corresponds to a comoving distance of 77 Mpc (proper distance 10.1 Mpc) and the velocity difference is $\approx 350 \text{ km s}^{-1}$. No (faint) neighbouring galaxies are known around COLA1 (Bowler et al. 2014; Matthee et al. 2015), although the sensitivity to neighbouring galaxies is limited (corresponding to SFR $\gtrsim 25 M_\odot \text{yr}^{-1}$) and deeper observations are required.

Our calculations differ somewhat from those in Hu et al. (2016), who argue that a change in the Lyα properties is expected when the number of escaping ionising photons is $\approx 10^{54} \text{s}^{-1}$ (corresponding to $R_s \approx 0.3 \text{pMpc}$ and hence consistent with our assumptions above). While Hu et al. (2016) find that this is achieved with $f_{\text{esc, LyC}} = 1\%$ in case of $f_{\text{esc, Lyα}} = 100\%$ and a Lyα luminosity $\approx 2.5 \times 10^{45} \text{erg s}^{-1}$, we find that this requires a higher escape fraction $f_{\text{esc, LyC}} = 30\%$ at fixed luminosity. This can be mitigated to $f_{\text{esc, LyC}} \approx 20\%$ in case of a slightly higher Lyα luminosity and lower Lyα escape fraction (see Table 4), indicating that a change in the Lyα line properties is still expected around COLA1’s Lyα luminosity.

5.2.2. Model 2: Large mildly ionised region

In our second scenario, COLA1 resides in a relatively large ionised region with a high transmission away from line-centre, while the galaxy itself is not able to ionise gas clouds in its direct vicinity, causing the low transmission at line-centre. This scenario is illustrated in the centre panel of Fig. 9. Here, the double peaked emission does not emerge from the galaxy itself, but is caused by the environment of COLA1 that transmits flux at different velocities differentially.

Here, the gas in the CGM of the galaxy still contains significant amounts of neutral hydrogen (for example due to self-shielding), while the gas at larger distances from the galaxy is mildly ionised and has a transmission of $\approx 50\%$ at $<1000 \text{ km s}^{-1}$ from the systemic redshift (similar to the simulated IGM properties at $z = 4$ in Laursen et al. 2011). The double peaked emission originates from residual hydrogen in the vicinity of COLA1, absorbing flux around the systemic velocity. In this scenario, the intrinsic Lyα profile (that escapes the ISM) is broader than the observed red Lyα line, but the
Fig. 9. Sketches of different scenarios that may explain the observed Lyα line-profile of COLA1. Top row: sketches of the physical scenarios and bottom row: emerging Lyα spectrum. Left panel: Model 1 shows a double peaked emission line emerging from scattering in the ISM that is redshifted due to the expansion of the Universe before encountering the IGM, that preferentially attenuates the blue component. Middle panel: Model 2 shows that a double peak may originate due to a low transmission at line centre (due to neutral hydrogen in the CGM), while there is a relatively high transmission at further distances from the galaxy due to a large relatively ionised region (similar to the IGM transmission curve at \( z \approx 4 \) in Laursen et al. 2011). Right panel: Model 3 shows that a double peaked profile can also arise without IGM attenuation in case there is an HI absorber slightly blue shifted with respect to the systemic Lyα velocity (which in this model is significantly redshifted with respect to the IGM).

Table 4. Parameters used in calculations of the ionised bubble around COLA1 at \( z = 6.591 \) (Sect. 5.2.1).

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{\text{esc,LyC}} )</td>
<td>15%</td>
<td>(Conservative) Lyα peak separation and blue-to-red flux ratio in local LyC leakers</td>
</tr>
<tr>
<td>( f_{\text{esc,Lyα}} )</td>
<td>( 60_{-10}^{+20} % )</td>
<td>Lyα EW calibration, tested at ( z = 0 - 2.6 ) (Sobral &amp; Matthee 2018)</td>
</tr>
<tr>
<td>( Q_{\text{ion,Lyα}} )</td>
<td>( 7.5 \times 10^{54} ) s(^{-1} )</td>
<td>Lyα luminosity, ( f_{\text{esc,LyC}} ) and ( f_{\text{esc,Lyα}} ) (Sobral &amp; Matthee 2018)</td>
</tr>
<tr>
<td>( Q_{\text{ion,UV}} )</td>
<td>( 5 \times 10^{54} ) s(^{-1} )</td>
<td>UV luminosity (no dust correction) and ( \xi_{\text{ion}} = 10 ) Hz erg(^{-1} )</td>
</tr>
<tr>
<td>( t_{\text{burst}} )</td>
<td>10(^7) yr</td>
<td>High Lyα EW (Charlot &amp; Fall 1993)</td>
</tr>
<tr>
<td>( R_{\text{s,required}} )</td>
<td>&gt; 0.3 pMpc</td>
<td>Velocity offset ( \geq 250 ) km s(^{-1} ) due to expansion of the Universe</td>
</tr>
<tr>
<td>( R_{\text{s,max,Lyα}} )</td>
<td>( 0.33^{+0.07}_{-0.06} ) pMpc</td>
<td>Eq. (2), assuming no recombinations and ionisation by the Lyα source of COLA1</td>
</tr>
<tr>
<td>( R_{\text{s,max,UV}} )</td>
<td>( 0.29^{+0.03}_{-0.06} ) pMpc</td>
<td>Eq. (2), assuming no recombinations and ionisation by the UV source of COLA1</td>
</tr>
</tbody>
</table>

resulting blue and red lines have similar widths. While this scenario still requires a highly ionised region (a transmission of 50% corresponds to \( \tau = 0.7 \), or a mean neutral fraction of \( \sim 2 \times 10^{-6} \) at \( z = 6.6 \)), less ionising photons originating from COLA1 are required compared to Model 1.

We point out that double peaked Lyα profiles with low flux at line centre are the norm among Green Pea galaxies in the local Universe (\( z \approx 0.3 \); e.g. Henry et al. 2015; Yang et al. 2017), emerging from scattering in the ISM. Therefore, it would be a remarkable coincidence if an emerging spectrum at \( z = 6.6 \) would have a similar profile, but due to the effect of the IGM (and without scattering in the ISM).

5.2.3. Model 3: Large velocity offset and intervening absorber

The last scenario follows an alternative explanation which does not require a strong imprint from the IGM (either highly ionised, neutral, or partly ionised) in the vicinity of the emitting galaxy. Here, the emergent Lyα spectrum consists of a significantly redshifted broad peak (e.g. due to an outflow), such that line is not affected by HI in the CGM or IGM. The right panel of Fig. 9 shows that the double peaked profile originates from an absorbing system on top of this emission profile. This scenario is similar to the “cold flow” scenario that is used to explain the Lyα profiles of high-redshift...
radio galaxies (e.g. van Ojik et al. 1997), but with the difference that the column density is much lower.

The absorbing system could be interpreted as an inflow (e.g. Matsuda et al. 2006), or potentially a wall between two ionised regions (although we note that the IGM in Model 3 can also be fully neutral), such as a filament that has not yet reionised (e.g. Finlator et al. 2009). We explore the H\textsc{i} column density that such an absorber would imply by fitting the line-profile with a voigt profile absorber on top of a gaussian emission line. The upper H\textsc{i} absorber would imply by fitting the line-profile with a voigt profile absorber on top of a gaussian emission line. The upper H\textsc{i} column density limit is mostly determined by the width of the absorber, which implies a marginally self-shielded column density $N_{\text{H}i} \gtrsim 10^{17.7} \text{ cm}^{-2}$, blue shifted by $\approx 80 \text{ km s}^{-1}$ with respect to the peak of the emission. The intrinsic line-width FWHM is $380 \pm 20 \text{ km s}^{-1}$.

How likely is such a scenario? In this model, both observed lines are redshifted with respect to the systemic redshift. This is not seen in $z = 0$–3 galaxies with double peaked Ly\textsc{a} emission, where the systemic redshift typically lies between peaks (e.g. Kulas et al. 2012; Henry et al. 2015; Sobral et al. 2018b). Vanzella et al. (2018) identify a luminous system at $z = 4$ ($M_{1500} = -22.2$, $L_{\text{Ly} \alpha} = 1.4 \times 10^{43} \text{ erg s}^{-1}$) that has four Ly\textsc{a} components, of which only one is redshifted by $\approx 350 \text{ km s}^{-1}$, while the others are either systemic or blueshifted.

Multiple redshifted lines are seen in high-redshift radio galaxies (van Ojik et al. 1997), but these lines are typically much broader ($FWHH \gtrsim 1000 \text{ km s}^{-1}$) and likely a consequence of large amounts of hydrogen present around the massive galaxies that harbour radio AGN. Therefore, unless particular conditions in the EoR are important, this scenario may be quite unlikely.

5.2.4. Summary and predictions

The sketched scenarios in Fig. 9 may be over-simplified and reality may be a combination of different aspects of the models. Even though Models 2 and 3 may be unlikely, some aspects of these models may be combined with the general scenario in Model 1. For example, the absorber from Model 3 could reside in the highly ionised region from Model 1. Other complications would arise if the galaxy is undergoing a major merger, with multiple emission line regions at different relative velocities (although there are currently no indications for this scenario). How can observations discriminate between the different explanations and make progress?

As illustrated in Fig. 9, it is clear that the key quantities that would allow to differentiate between the models are the intrinsic line-width and the systemic redshift (marked as dashed, vertical lines in Fig. 9). These properties can be measured from the H\textsc{ii} or the H\textsc{ii} lines, accessible with the James Webb Space Telescope. The systemic redshift can alternatively also be measured using [C\textsc{i}] FIR emission with ALMA (e.g. Matthee et al. 2017b), which can also test whether a merger of multiple components is ongoing. Furthermore, if COLA1 is surrounded by a larger ionised region, additional sources of ionising photons are likely present. It is therefore expected that COLA1 resides in a larger scale over-density of faint galaxies, potentially observable using deep Ly\textsc{a} observations with VLT/MUSE and/or with ALMA observations that can identify [C\textsc{i}].

5.3. Implications and considerations for surveys

Why have only a few double peaked LAEs been found at $z = 5$–7 (e.g. Songaila et al. 2018)? Why do we find a double peaked LAE at a redshift where the IGM transmission towards the blue is likely very low, instead of around $z \approx 5$–6? Could it be that more double-peaked LAEs have been found, but they have been discarded as [O\textsc{ii}] interlopers?

5.3.1. Observational biases against identifying double peaked LAEs?

As already noted in Dijkstra et al. (2007), LAEs with a strong blue peak show a relatively low asymmetry (quantified by skewness) and could thus be classed as low-redshift interlopers in the case of low resolution spectroscopy. Moreover, the range of observed peak separations in double peaked LAEs ($\Delta v \approx 200$–400 km s$^{-1}$; Verhamme et al. 2017) corresponds to observed wavelength differences $\Delta \lambda_{\text{obs}} = (1+z_{\text{ff}}) \times [0.8$–1.6] \AA, similar to the observed wavelength difference of the [O\textsc{ii}] doublet $\Delta \lambda_{\text{obs,[OII]}} = (1+z_{\text{ff}}) \times 2.79 \text{ \AA at all observed wavelengths in the optical. Therefore, “[single” line-identifications may be biased against confirming double peaked LAEs. This bias may particularly be important when spectral wavelength coverage is limited to the optical, but less important for $z_{\text{ff}} \leq 4.8$, as potential interlopers of LAEs at these wavelengths will likely be detected in multiple lines (such as H\textsc{ii} and [O\textsc{ii}], besides [O\textsc{ii}]).

There are also other explanations of why the number of known double-peaked LAEs at $z = 5$–6 may be low. One explanation is that the fraction of double-peaked LAEs increases strongly with Ly\textsc{a} luminosity (as discussed in Hu et al. 2016) and current samples at $z = 5$–6 do not yet include enough luminous sources (as the majority of very wide-field surveys prioritised $z > 6$; e.g. Songaila et al. 2018). As the typical Ly\textsc{a} luminosity increases with redshift, this explanation would indirectly imply that the LyC escape fraction increases with Ly\textsc{a} luminosity and with redshift (as $L_{\text{Ly} \alpha}^{*}$ increases; Sobral et al. 2018a).

Alternatively, it may be possible that spectroscopic follow-up observations and/or publication efforts may be biased towards $z > 6$ and double-peaked LAEs at $z = 3$–6 simply have not yet been observed spectroscopically or published; or that the typical S/N of spectroscopic observations at $z > 4$ is too low to identify similar blue peaks. The flux in the blue peak is a factor $\approx 0.3$ lower than the brighter red component and can therefore only be detected at $3\sigma$ significance if the red peak is detected at $10\sigma$ significance, which is not always the case.

5.3.2. A preference for observing narrow peak separations only in luminous LAEs at $z > 5$?

Multiple peaked Ly\textsc{a} profiles have regularly been observed in LAEs at $z = 2$–3 (e.g. Rauch et al. 2008, 2011; Kulas et al. 2012; Yamada et al. 2012; Trainor et al. 2015; Vanzella et al. 2016; Rivera-Thorsen et al. 2017), but also in quasars and in high-redshift radio galaxies (e.g. van Ojik et al. 1997; Miley & De Breuck 2008). As COLA1 does not show evidence for AGN activity, we focus our comparison to LAEs at $z = 2$–3 (combining data from Yamada et al. 2012, Saez et al. 2015 and Sobral et al. 2018b) and Ly\textsc{a} analogues (i.e. Green Pea galaxies; GPs; Yang et al. 2017). The fraction of double peaked LAEs at $z = 2$–3 is $\approx 30$–50% (Kulas et al. 2012; Yamada et al. 2012; Trainor et al. 2015), who find typical peak separations of $\approx 500$–$750 \text{ km s}^{-1}$, although we note that the spectral resolution of these observations ($\approx 150$–$200 \text{ km s}^{-1}$) may not identify peaks with low velocity separations. The peak separation in COLA1 is smaller than the peak separation of the ultra-faint ($M_{1500} = -17$), lensed LAE at $z = 3.12$ identified in Vanzella et al. (2016),
which is an extremely compact, highly ionising star-forming galaxy.

As shown in Fig. 10, there is no clear relation between peak separation and Lyα luminosity for LAEs at $z \approx 3$ (although the dynamic range is limited). On the other hand, a clear anti-correlation is seen between peak separation and luminosity for Lyα emitting GPs. COLA1, which is the most luminous LAE with the lowest known peak separation, seems to follow this relation. However, we note that this anti-correlation may be a selection effect, as GPs are selected on compactness and strong [O\textsc{ii}]/[O\textsc{ii}] ratios that may trace density bounded H\textalpha regions, and not a pure Lyα selection.

If Lyα luminosity is anti-correlated with the peak separation at $z \geq 6$, this may explain why double peaked Lyα profiles may only be observed in the most luminous LAEs. As the IGM transmission strongly decreases towards the blue of the systemic redshift (e.g. Laursen et al. 2011; Weinberger et al. 2018), the blue lines of faint LAEs are more likely to be boosted with respect to the observability of fainter galaxies (e.g. Stark et al. 2017). While recent models with uniform ionising backgrounds reproduce this result (e.g. Mason et al. 2018b), they can not yet explain the highest Lyα equivalent widths. We speculate that this could be mitigated in case the local ionising background is enhanced around bright galaxies (i.e. similar to Model 1 discussed here).

6 NEPLA4 has a larger peak separation than COLA1, inconsistent with the peak separation that an [O\textsc{ii}] interloper would have.

6. Summary
We have presented a detailed analysis of the COLA1 galaxy, a double peaked LAE at $z = 6.59$ (Hu et al. 2016), using deep multi-wavelength photometry and new deep, high resolution optical and near infrared spectroscopy with VLT/X-shooter. Due to the opacity of the CGM and IGM for blue Lyα photons, detecting a blue peak at $z > 6$ is highly unlikely. We therefore test whether the line profile is interpreted correctly as Lyα at $z = 6.6$ and study the properties of COLA1 and their implications. Our results are summarised as follows:

- We rule out that the double peaked line is [O\textsc{ii}] emission at $z = 1.475$ based on (1) the shape of the line-profile: the asymmetric wing of the red line is not observed in the blue line and no flux is detected in the middle of the blue and red lines; (2) the extremely low blue-to-red flux ratio that is inconsistent with [O\textsc{ii}] and (3) the non-detection of H\textbeta at $z = 1.475$ ruling out even the lowest observed ratios with respect to [O\textsc{ii}].
- A tentative detection in the B band is explained by a contribution from Lyα emission in a foreground LAE at $z = 2.142$, confirmed through [O\textsc{iii}]$_{5007}$ emission in addition to Lyα.
- We confirm that COLA1 is a double peaked LAE at $z_{Ly\alpha,red} = 6.593$ and summarise its properties in Table 3. The Lyα luminosity and EW are high ($L_{Ly\alpha} = 4.1 \pm 0.2 \times 10^{41}$ erg s$^{-1}$ and $EW_{Ly\alpha} = 120^{+50}_{-40}$ Å) and the Lyα emission is compact ($r_{90} = 0.33^{+0.07}_{-0.04}$ kpc based on HST imaging).
- The Lyα peak separation is $220 \pm 20$ km s$^{-1}$ and the flux of the blue peak is $0.31 \pm 0.03$ times the flux in the red peak. The Lyα lines are narrow ($FWHM = 198 \pm 14$ km s$^{-1}$) and $150 \pm 18$ km s$^{-1}$ for the red and blue peak, respectively and the full Lyα line-profile resembles the profile of a luminous LAE at $z = 2.2$ with a prominent blue peak (Fig. 6). Due to the narrowness of the Lyα line, COLA1 is unlikely to be powered by an AGN.
- By modelling the Lyα profile with a five-parameter shell model, we find that the line-profile is characterised by a very low H\textalpha column density ($N_{H\alpha} = 10^{17.7 \pm 0.3}$ cm$^{-2}$), indicating a non-zero escape of ionising photons. Based on correlations between the escape fraction and the peak separation in LyC leaking galaxies, we infer an escape fraction of $\approx 15\%$, but potentially up to $\approx 30\%$. Other inferences imply escape fractions up to $f_{esc,LyC} \sim 50\%$. This means that we are witnessing a star-forming galaxy that is actively contributing to the reionisation of the Universe.
- COLA1 has a high UV luminosity, $M_{1500} = -21.6 \pm 0.3$, implying a SFR $\approx 30 M_{\odot}$ yr$^{-1}$. The Spitzer/IRAC colours indicate that the CGM/IGM attenuated a significant amount of the flux at even bluer wavelengths. Hence, it may simply be expected that observed peak separations at $z > 6$ are lower than at $z < 5$. A crucial test would therefore be to map out the evolution of the relation between peak separation and Lyα luminosity from $z \approx 5$–7 with high resolution spectroscopy, and explore how the blue-to-red flux ratio evolves with redshift.

The Lyα observability of luminous galaxies at $z \approx 7$ appears to be boosted with respect to the observability of fainter galaxies (e.g. Stark et al. 2017). While recent models with uniform ionising backgrounds reproduce this result (e.g. Mason et al. 2018b), they can not yet explain the highest Lyα equivalent widths. We speculate that this could be mitigated in case the local ionising background is enhanced around bright galaxies (i.e. similar to Model 1 discussed here).
As we have shown, the confirmation of COLA1 as a double peaked LAE at $z=6.59$ allows us to witness directly that galaxies contribute to the reionisation of the Universe. Several assumptions made here can be overcome with future observations of the intrinsic line-profile and systemic redshift of COLA1. In particular, future high resolution spectroscopic observations of LAEs may reveal the dependence on double peaked Lyα profiles on luminosity and redshift, providing a new and deeper view of the reionisation process.

Acknowledgements. JM acknowledges the award of a Huygens PhD fellowship from Leiden University. MG acknowledges support from NASA grant NNX17AK58G. APA, PhD::SPACE fellow, acknowledges support from the FCT scholarship from Leiden University. MG acknowledges support from NASA grant 80NSSC18K1711. BPH acknowledges support from the FCT. We thank the referee for their comments that improved the paper. We also thank Christoph Behrens, Len Cowie, Koki Kakiichi, Peter Laursen, Charlotte Mason, Eriko Matsuda, and Simon Sobral for useful discussions. This research was supported by the Munich Institute for Astro- and Particle Physics. The detectability of blue Lyα emission from COLA1 implies a low Lyα optical depth out to $\approx 250$ km s$^{-1}$ of the galaxy, potentially due to a surrounding $\approx 0.3$ Mpc (2.3 Mpc) ionised region. Based on the observed properties of COLA1 and simple assumptions (detailed in Table 4), we find that enough ionising photons have escaped COLA1 to ionise a $\approx 0.3$ Mpc (2.3 Mpc) bubble. However, keeping this region ionised likely requires a contribution from faint neighbour- ing sources. No faint neighbouring galaxies are known, but COLA1 is likely in the same large scale over-density as CR7 with a separation of $\approx 30$ ($\approx 10$ Mpc) and $\approx 350$ km s$^{-1}$.

As we have shown, the confirmation of COLA1 as a double peaked LAE at $z=6.59$ allows us to witness directly that galaxies contribute to the reionisation of the Universe. Several assumptions made here can be overcome with future observations of the intrinsic line-profile and systemic redshift of COLA1. In particular, future high resolution spectroscopic observations of LAEs may reveal the dependence on double peaked Lyα profiles on luminosity and redshift, providing a new and deeper view of the reionisation process.