

# Measuring the properties of reionised bubbles with resolved Lyman alpha spectra

Charlotte A. Mason<sup>1</sup>★† and Max Gronke<sup>2,3</sup>†

<sup>1</sup>*Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA, 02138, USA*

<sup>2</sup>*Department of Physics and Astronomy, University of California, Santa Barbara, 93106, USA*

<sup>3</sup>*Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

Accepted 2020 September 18. Received 2020 September 08; in original form 2020 April 24

## ABSTRACT

Identifying and characterising reionised bubbles enables us to track both their size distribution, which depends on the primary ionising sources, and the relationship between reionisation and galaxy evolution. We demonstrate that spectrally resolved  $z \gtrsim 6$  Lyman-alpha ( $\text{Ly}\alpha$ ) emission can constrain properties of reionised regions. Specifically, the distant from a source to a neutral region sets the minimum observable  $\text{Ly}\alpha$  velocity offset from systemic. Detection of flux on the blue side of the  $\text{Ly}\alpha$  resonance implies the source resides in a large, sufficiently ionised region that photons can escape without significant resonant absorption, and thus constrains both the sizes of and the residual neutral fractions within ionised bubbles. We estimate the extent of the region around galaxies which is optically thin to blue  $\text{Ly}\alpha$  photons, analogous to quasar proximity zones, as a function of the source’s ionising photon output and surrounding gas density. This optically thin region is typically  $\lesssim 0.3$  pMpc in radius (allowing transmission of flux  $\gtrsim -250 \text{ km s}^{-1}$ ),  $\lesssim 20\%$  of the distance to the neutral region. In a proof-of-concept, we demonstrate the  $z \approx 6.6$  galaxy COLA1 – with a blue  $\text{Ly}\alpha$  peak – likely resides in an ionised region  $> 0.7$  pMpc, with residual neutral fraction  $< 10^{-5.5}$ . To ionise its own proximity zone we infer COLA1 has a high ionising photon escape fraction ( $f_{\text{esc}} > 0.50$ ), relatively steep UV slope ( $\beta < -1.79$ ), and low line-of-sight gas density ( $\sim 0.5\times$  the cosmic mean), suggesting it is a rare, underdense line-of-sight.

**Key words:** dark ages, reionisation, first stars – galaxies: high-redshift

## 1 INTRODUCTION

Understanding the process of hydrogen reionisation is one of the frontiers of astronomy. It occurred neither homogeneously nor instantaneously, as ionising photons propagating from nascent galaxies reionised the most overdense regions first, carving out ionised ‘bubbles’ within the then neutral Universe, gradually reionising the entire intergalactic medium (IGM). Measuring the timeline and morphology of reionisation, i.e., studying the redshift evolution and spatial distribution of these ionised regions, is key to understanding how reionisation occurred (e.g., Furlanetto et al. 2004; McQuinn et al. 2007; Mesinger 2016).

A key question is what drove reionisation, that is, where did the ionising photons originate from? Identifying reionised or neutral regions of the IGM not only characterises the morphology of reionisation but enables us to address this question by comparing the properties of observed galaxies in those regions to the local ionisation state (e.g., Beardsley et al. 2015). Regions which reionise early are likely the first overdensities where galaxy formation is ac-

celerated, thus identifying those regions helps to identify the first generations of galaxies.

Mapping reionised bubbles and measuring their size distribution is a goal of future 21 cm intensity experiments. This requires spatial resolution capable of discerning ionised hydrogen gas on scales of  $< 1$  proper Mpc (e.g., Geil et al. 2017). However, sensitivity to these smallest scales is still an observational challenge: it requires large baseline radio telescopes to resolve H I regions (e.g., SKA-low, Koopmans et al. 2015), and detailed spectroscopic follow-up of the galaxies within H II regions to determine their ionising properties. However, estimates of bubble sizes on small scales are currently feasible with Lyman-alpha ( $\text{Ly}\alpha$ , rest wavelength 1216 Å) spectroscopy of high redshift sources.

Due to its high cross-section for absorption by neutral hydrogen,  $\text{Ly}\alpha$  is a sensitive probe of neutral gas. Neutral hydrogen affects both the strength and lineshape of  $\text{Ly}\alpha$  (see, e.g., Dijkstra 2014, for a review). With the advent of sensitive near-IR spectroscopy,  $\text{Ly}\alpha$  emission from galaxies and quasars at  $z > 6$  has been a particularly powerful probe of reionisation (e.g., Malhotra & Rhoads 2006; Fan et al. 2006; Dijkstra et al. 2011; Treu et al. 2013; Mesinger et al. 2015; Davies et al. 2018; Mason et al. 2018a; Greig et al. 2019). In recent years, the declining flux distribution of  $\text{Ly}\alpha$  emission from

★ E-mail: charlotte.mason@cfa.harvard.edu

† Hubble Fellow

galaxies at  $z \gtrsim 6$  has been used to measure the average fraction of the IGM which is neutral at a given redshift (Schenker et al. 2014; Mason et al. 2018a, 2019; Hoag et al. 2019; Whitler et al. 2020).

The lineshape of Ly $\alpha$  also encodes information about neutral hydrogen structures the photons encountered along their path. Within or in close proximity to the emitting galaxy the Ly $\alpha$  spectrum is shaped by strong scattering of photons close to the Ly $\alpha$  resonant wavelength (resonant scattering) with H I that is not necessarily along the line-of-sight (Eide et al. 2018), and typically produces double-peaked emission line profiles due to the high optical depth at line centre (e.g., Neufeld 1990). However, at larger distances, when the probability of scattering back into the line-of-sight becomes negligible, the impact of intervening H I can be treated more simply as absorption. In the following we use ‘resonant absorption’ to refer to the effective absorption of Ly $\alpha$  photons which emerge from galaxies with a blueshift, but encounter significant neutral gas as they redshift into the resonant wavelength. The smooth damping wing, due to  $n_{\text{HI}} \gtrsim 10^{-6} \text{ cm}^3$  gas that can be at large distances, is commonly interpreted as a signature of reionisation (e.g., Miralda-Escude 1998). In this case, the optical depth due to damping wing absorption is a function of the distance to the nearest neutral patch and thus could be used to recover the size of ionised bubbles (Malhotra & Rhoads 2006).

Due to the decreasing recombination time at  $z \gtrsim 6$ , even within ‘ionised’ bubbles there can be significant residual neutral gas. The amount of neutral hydrogen depends on the local ionisation field and can lead to resonant absorption on the blue side of the Ly $\alpha$  resonance (e.g., Gunn & Peterson 1965; Zheng et al. 2010; Laursen et al. 2011; Byrohl & Gronke 2020) – which makes the detection of blue Ly $\alpha$  peaks towards higher redshift increasingly unlikely. In a number of rare sightlines, however, blue Ly $\alpha$  flux *has* been observed at  $z \gtrsim 6$  (Matthee et al. 2018; Songaila et al. 2018; Bosman et al. 2020), implying a low resonant optical depth and thus low residual neutral fraction in reionised regions. In these cases it may be possible to directly measure the properties of individual ionised bubbles.

Here, we demonstrate that  $z > 6$  Ly $\alpha$  emission lineshapes encode information about the ionised bubbles their host galaxies reside in. While previous works have shown that Ly $\alpha$  can be visible early in the epoch of reionisation if galaxies reside in ionised bubbles (Haiman 2002; Mason et al. 2018b) we show here that spectroscopic measurements of such Ly $\alpha$  emitters enable us to calculate the minimum size of the ionised bubble such that Ly $\alpha$  at a given frequency offset is visible to us. Furthermore, we demonstrate that blue-peaked Ly $\alpha$  lines observed at  $z \gtrsim 6$  can be used to constrain the residual neutral gas remaining in reionised bubbles. As a proof-of-concept, we investigate the necessary physical conditions for observing the double-peaked Ly $\alpha$  emitted COLA1 (Hu et al. 2016; Matthee et al. 2018).

This paper is organised as follows: we describe our model for the Ly $\alpha$  optical depth in Section 2 and present our results in Section 3. We discuss our results in Section 4 and present conclusions in Section 5. We use the Planck Collaboration et al. (2015) cosmology:  $(\Omega_{\Lambda}, \Omega_{\text{m}}, \Omega_{\text{b}}, n, \sigma_8, H_0) = (0.69, 0.31, 0.048, 0.97, 0.81, 68 \text{ km s}^{-1} \text{ Mpc}^{-1})$ . Magnitudes are in the AB system. Distances, volumes, and densities are proper unless otherwise stated.

## 2 MODEL

In this section, we describe the two components of our model: the Ly $\alpha$  optical depth as a function of the distance from a galaxy (§ 2.1) and the properties of ionised bubbles (§ 2.2).

The model provides a way to interpret the necessary conditions to observe a blue-shifted Ly $\alpha$  peak emerging from a galaxy at  $z \gtrsim 6$ . Of course, there are numerous scattering processes in the ISM, CGM and IGM which can absorb a blue peak at any redshift (e.g., Gunn & Peterson 1965; Zheng et al. 2010; Laursen et al. 2011), meaning that non-detection of a blue peak does not provide much information about any one of those media. Recently, Hayes et al. (2020) showed that for stacks in the redshift range  $z \sim 3 - 5$  the evolution of the blue peak can be explained entirely by the evolution of the IGM. However, in the rare cases where blue Ly $\alpha$  flux *has* made it through a relatively neutral IGM, this model demonstrates that constraints can be placed on the line-of-sight gas properties in front of the source, in particular, that for blue flux to have been detected at  $z \gtrsim 6$ , the source must reside in a highly ionised region.

### 2.1 Lyman-alpha optical depth

The Ly $\alpha$  optical depth through hydrogen gas for photon observed at  $\lambda_{\text{obs}} = \lambda_{\text{em}}(1 + z_s)$  to a source at redshift  $z_s$ , observed at  $z_{\text{obs}}$ , is given by:

$$\tau_{\alpha}(\lambda_{\text{obs}}) = \int_{z_{\text{obs}}}^{z_s} dz \, c \frac{dt}{dz} x_{\text{HI}}(z) n_{\text{H}}(z) \sigma_{\alpha} \left( \frac{\lambda_{\text{obs}}}{1+z}, T \right) \quad (1)$$

where  $n_{\text{H}}$  is the total number density of hydrogen and  $x_{\text{HI}}$  is the fraction of hydrogen which is neutral.  $\sigma_{\alpha}(\lambda, T)$  is the Ly $\alpha$  scattering cross-section through an ensemble of hydrogen atoms with a Maxwell-Boltzmann velocity distribution, usually expressed as function of the dimensionless frequency  $x = (\nu - \nu_{\alpha})/\Delta\nu_{\text{D}}$ :

$$\sigma_{\alpha}(x, T) = \sigma_0 \times \phi(x), \quad (2)$$

$$\sigma_0 = \frac{1}{\Delta\nu_{\text{D}} \sqrt{\pi}} \frac{f_{\alpha} \pi e^2}{m_e c} \approx 5.9 \times 10^{-14} \left( \frac{T}{10^4 \text{ K}} \right)^{-1/2} \text{ cm}^2$$

where  $f_{\alpha} = 0.416$  is the Ly $\alpha$  oscillator strength,  $m_e$  and  $e$  are the mass and charge of an electron, and  $\phi(x)$  is the Voigt function:

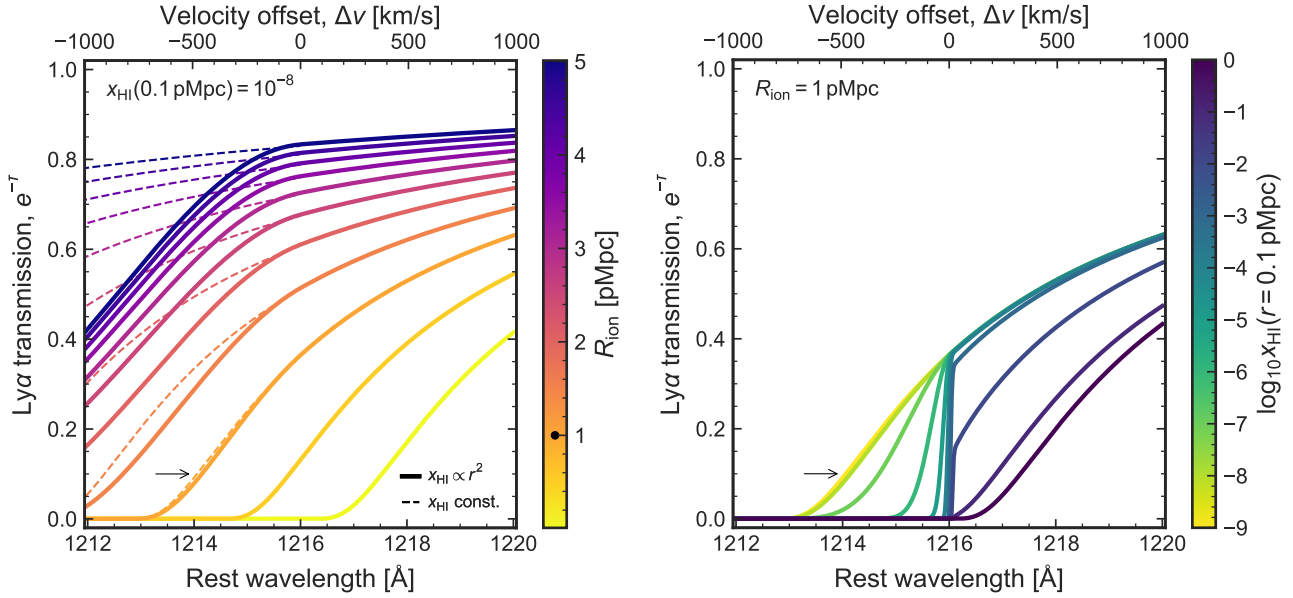
$$\phi(x) = \frac{a_{\text{V}}}{\pi} \int_{-\infty}^{\infty} dy \frac{e^{-y^2}}{(y-x)^2 + a_{\text{V}}^2}. \quad (3)$$

Here,  $\nu_{\alpha} \approx 2.46 \times 10^{15} \text{ Hz}$  is the resonant frequency of Ly $\alpha$ , at wavelength  $\lambda_{\alpha} \approx 1216 \text{ \AA}$ ,  $\Delta\nu_{\text{D}} = \nu_{\alpha} \sqrt{2k_{\text{B}}T/m_{\text{p}}c^2} \equiv \nu_{\alpha} v_{\text{th}}/c$  is the thermally broadened frequency, and the Voigt parameter  $a_{\text{V}} \approx 4.7 \times 10^{-4} (T/10^4 \text{ K})^{-1/2}$ . Equation (3) is normalised such that  $\int dx \phi(x) = 1$ . The cross-section is tightly peaked around the core of the line, but has damping wings which extend out to  $> 1000 \text{ km s}^{-1}$  from the line centre (e.g., Dijkstra 2014). We use the approximation for  $\phi(x)$  given by Tasitsiomi (2006).

Approximating Equation (3) as a Dirac delta function, and assuming constant  $x_{\text{HI}}$ , we obtain the Gunn & Peterson (1965) optical depth for blue photons emitted from a source at  $z_s$ :

$$\tau_{\text{GP}}(z_s) = \frac{f_{\alpha} \pi e^2}{m_e \nu_{\alpha}} \frac{x_{\text{HI}} n_{\text{H}}(z_s)}{H(z_s)} \approx 4.7 \times 10^5 x_{\text{HI}} \Delta \left( \frac{1+z_s}{8} \right)^{3/2}. \quad (4)$$

where  $\Delta = n_{\text{H}}(z)/\bar{n}_{\text{H}}(z)$  is the overdensity of hydrogen gas relative to the cosmic mean. We assume that the source galaxy resides inside an ionised region embedded in a neutral homogeneous intergalactic medium at a distance  $R_{\text{ion}}$ . This is representative of reionisation’s



**Figure 1.** Ly $\alpha$  transmission for a source at  $z_s = 7$  using a  $x_{\text{HI}} \propto r^2$  profile (solid lines). **Left:** Varying the distance to first neutral patch while fixing  $x_{\text{HI}}(r = 0.1 \text{ pMpc}) = 10^{-8}$  ( $n_{\text{HI}} \sim 10^{-12} \text{ cm}^{-3}$ ). For decreasing bubble size, the transmission on the red side of the Ly $\alpha$  line centre decreases, due to the increasing damping wing absorption. However, even in a fully neutral IGM, Ly $\alpha$  can be visible providing it is emitted at  $\gtrsim 300 \text{ km s}^{-1}$ . Thin dashed lines show the transmission if  $x_{\text{HI}} = 10^{-8}$  is constant inside the ionised region. The small black arrow shows  $R_{\text{ion}} = 1 \text{ pMpc}$  to compare with the right panel. **Right:** Changing the residual neutral fraction inside the ionised region while holding the distance to the fully neutral patch,  $R_{\text{ion}} = 1 \text{ pMpc}$ , fixed. For increasing residual neutral fraction, the transmission on the blue side decreases as the gas becomes optically thick to Ly $\alpha$  photons redshifting to the resonant frequency by  $x_{\text{HI}} \gtrsim 10^{-5}$ . For higher neutral fractions, the damping wing absorption due to residual neutral gas inside the ionised region also become significant, reducing transmission on the red side of the line.

early pre-overlap phases when ionised bubbles grow around sources of ionising photons. However, the assumption of an isolated bubble breaks down as reionisation progresses, meaning our method provides only a lower limit on the bubble size.

We construct the optical depth to a source galaxy by modelling the two media separately, i.e. breaking the integral into two components: from  $z_s$  to  $z_{\text{ion}}$  and  $z_{\text{ion}}$  to  $z_{\text{obs}}$  (following, [Haiman 2002](#); [Cen & Haiman 2000](#); [Mesinger et al. 2004](#)). In the ionised bubble we set  $T = 10^4 \text{ K}$  (appropriate for photoionised gas at the mean density, e.g., [Hui & Gnedin 1997](#)) and  $n_{\text{H}}(z) = \Delta \bar{n}_{\text{H}}(z)$ .  $\bar{n}_{\text{H}}(z)$  is the comoving cosmic mean hydrogen number density:  $\bar{n}_{\text{H}}(z) \approx 1.88 \times 10^{-7} (1+z)^3 \text{ cm}^{-3}$ . In the neutral IGM, we set  $T = 1 \text{ K}$ , assuming gas decouples from the CMB at  $z \sim 150$  and cools adiabatically thereafter ([Peebles 1993](#)), and  $n_{\text{H}}(z) = \bar{n}_{\text{H}}(z)$ . This approximation of the density profile as a step function is simplistic, but as we show in Section 3.3.1 observing blue peaks likely requires underdense gas along the line of sight. More realistic model gas density profiles impacting Ly $\alpha$  transmission are discussed by [Santos \(2004\)](#). Our results are not strongly sensitive to these temperatures choices within a physically motivated range ( $T < 10^5 \text{ K}$ ).

The left panel of Figure 1 shows the Ly $\alpha$  transmission,  $e^{-\tau(\lambda)}$ , as a function of wavelength – commonly expressed as velocity offset,  $\Delta v \equiv c(\lambda_{\text{em}}/\lambda_{\alpha} - 1)$  – and ionised bubble radius, assuming the residual neutral fraction inside the ionised bubble is very low ( $x_{\text{HI}} = 10^{-8}$  at 0.1 pMpc from the source, assuming  $x_{\text{HI}} \propto r^2$  – see § 2.2.2), and  $\Delta = 1$ . As the bubble size increases more flux is transmitted on both the blue and red side of the line. Blue photons which redshift into resonance at the edge of bubble or at further distances from the source all encounter fully neutral gas when they reach resonance and are thus absorbed with a high optical depth. Photons which have already redshifted past resonance by the time

they reach the neutral gas experience the damping wing absorption, which smoothly suppresses flux to red wavelengths. For very small bubble sizes, the transmission on the blue side is therefore negligible and the transmission within  $200 \text{ km s}^{-1}$  of the red side can be very low. Ly $\alpha$  lines observed with low velocity offsets from systemic must reside in large ionised regions.

Note that even in a fully neutral IGM ( $R_{\text{ion}} = 0$ ) Ly $\alpha$  flux can still be transmitted on the red side: it is possible to observe Ly $\alpha$  lines at very high redshifts, providing they emit Ly $\alpha \gtrsim 300 \text{ km s}^{-1}$  from systemic ([Dijkstra et al. 2011](#)). Therefore, even at very high redshifts, merely detection of Ly $\alpha$  is not sufficient to identify a reionised bubble: there must be flux  $< 300 \text{ km s}^{-1}$ .

The right panel of Figure 1 shows the transmission through a bubble of fixed size (1 pMpc), but changing the residual neutral fraction in the ionised region around the source. The damping wing set by  $R_{\text{ion}}$  acts as an envelope for the maximum possible transmission: for  $x_{\text{HI}}(r = 0.1 \text{ pMpc}) < 10^{-8}$ , the bubble is fully optically thin and the maximum blue flux allowed given the damping wing shape can be transmitted. As the residual neutral fraction increases, more flux on the blue side of the line is absorbed, with the ionised region becoming optically thick for  $x_{\text{HI}}(r = 0.1 \text{ pMpc}) \gtrsim 10^{-6}$  (corresponding to a neutral hydrogen number density of  $n_{\text{HI}} \gtrsim 5 \times 10^{-10} \text{ cm}^{-3}$ ). For  $x_{\text{HI}} > 10^{-1}$  the transmission displays a strong damping wing on the red side and converges to the  $R_{\text{ion}} = 0$  case in the left plot.

## 2.2 Size and residual neutral fraction of ionised bubbles

The optical depth can be calculated for any values of ionised bubble size,  $R_{\text{ion}}$ , and residual neutral fraction inside the bubble,  $x_{\text{HI}}$ , to estimate those parameters in a model-independent way. In the limiting

case of a single ionising source (plus uniform ionising background) we can also estimate those quantities for a physical model.

### 2.2.1 Size of ionised region

Assuming ionisation by a single source at redshift  $z_s$  at the centre of the ionised region, the proper radius of the region can be obtained by solving for the evolution of a ionisation front (e.g. [Shapiro & Giroux 1987](#); [Cen & Haiman 2000](#); [Yajima et al. 2018](#)):

$$\frac{dR_{\text{ion}}^3}{dt} = \frac{3f_{\text{esc}}\dot{N}_{\text{ion}}}{4\pi\bar{n}_{\text{H}}(z_s)} - C_{\text{HII}}\Delta\bar{n}_{\text{H}}(z_s)\alpha_{\text{B}}(T)R_{\text{ion}}^3 + 3H(z)R_{\text{ion}}^3 \quad (5)$$

where the first term is due to ionisations from a source with ionising photon output  $\dot{N}_{\text{ion}}$  (in units of  $\text{s}^{-1}$ ) and ionising escape fraction  $f_{\text{esc}}$ , the second term is due to recombinations – assuming Case B recombination in a clumpy medium.  $C_{\text{HII}}$  is the clumping factor of ionised hydrogen, which describes the enhanced rate of recombinations relative to a uniform medium,  $C_{\text{HII}} \equiv \langle n_{\text{HII}}^2 \rangle / \bar{n}_{\text{HII}}^2$ . For  $\alpha_{\text{B}}$ , we use the approximation from [Hui & Gnedin \(1997\) for the hydrogen recombination coefficient as a function of temperature. The third term of Equation \(5\) is the expansion of the region due to the Hubble flow. As described in Section 2.1 we assume the IGM outside the ionized region to be fully neutral with density  \$n\_{\text{H}}^{\text{IGM}}\(z\) = \bar{n}\_{\text{H}}\(z\)\$ , and assume the gas inside the bubble to be fully ionized \(except for calculating the optical depth, see Section 2.2.2\) and possibly overdense:  \$n\_{\text{H}}^{\text{ion}}\(z\) = \Delta\bar{n}\_{\text{H}}\(z\)\$ .](#)

For constant  $\dot{N}_{\text{ion}}$  and  $f_{\text{esc}}$ , and simplified cosmology, Equation (5) can be solved analytically ([Shapiro & Giroux 1987](#)). For instance, for a luminous source at  $z \lesssim 8$  (when the recombination rate is relatively low), the first term in Equation (5) dominates (Equation 2 of [Cen & Haiman 2000](#)), so that

$$R_{\text{ion}} \approx \left( \frac{3f_{\text{esc}}\dot{N}_{\text{ion}}t_{\text{age}}}{4\pi\bar{n}_{\text{H}}(z_s)} \right)^{1/3} \quad (6)$$

where  $t_{\text{age}}$  is the time since the ionising source has switched on. In reality, due to the Hubble expansion the ionised radius grows more rapidly after  $\sim 10^7$  years than Equation (6). Here, we solve Equation (5) numerically.

The source emissivity  $\dot{N}_{\text{ion}}$ , in  $\text{s}^{-1}$  can be written as

$$\dot{N}_{\text{ion}}(t) = \int_{\nu_{\text{H}}}^{\infty} d\nu \frac{L_{\nu}(t)}{h\nu} \quad (7)$$

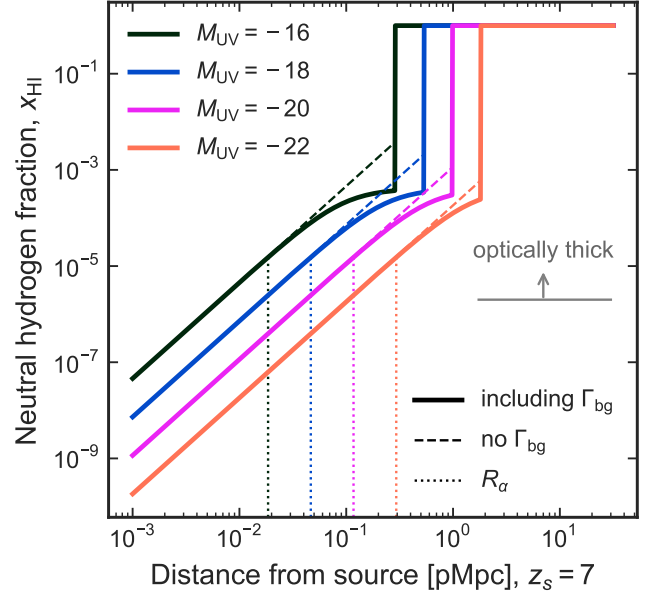
where  $L_{\nu}$  is the ionising spectrum of the source in  $\text{erg s}^{-1} \text{Hz}^{-1}$ . We approximate  $L_{\nu}$  as a double power law:

$$L_{\nu} \propto \begin{cases} \nu^{-\alpha} & \nu > \nu_{\text{H}} \\ \nu^{-\beta} & \nu \leq \nu_{\text{H}} \end{cases} \quad (8)$$

where  $\nu_{\text{H}} \approx 3.3 \times 10^{15} \text{ Hz}$  is the frequency of hydrogen photoionisation,  $\alpha$  is the spectral slope of the ionising continuum, and  $\beta$  is the spectral slope of the non-ionising UV continuum. Typically  $1 \lesssim \alpha \lesssim 2$  for quasars (e.g., [Scott et al. 2004](#); [Stevans et al. 2014](#); [Lusso et al. 2015](#)) and galaxies with massive stars ([Steidel et al. 2014](#); [Feltre et al. 2016](#)), where stripped stars in binaries can cause the spectral slope to reach  $\sim 1$  ([Götberg et al. 2020](#)). For galaxies  $\beta \approx -2$  (e.g., [Dunlop et al. 2013](#); [Bouwens et al. 2014](#)).

Thus we can estimate  $\dot{N}_{\text{ion}}$  for galaxies from a UV magnitude (measured at  $1500 \text{ \AA}$ ) as:

$$\dot{N}_{\text{ion}} \approx \frac{3.3 \times 10^{54}}{\alpha} 10^{-0.4(M_{\text{UV}}+20)} \left( \frac{912}{1500} \right)^{\beta+2} \text{ s}^{-1} \quad (9)$$



**Figure 2.** Residual neutral hydrogen fraction inside an ionised region, with (solid lines) and without (dashed lines) including ionising background flux from [Khaire & Srianand \(2019\)](#), as a function of the distance from the source and source UV magnitude, for a source constantly emitting ionizing photons for  $10^8$  years.  $R_{\text{ion}}$  is determined by solving Equation (5). Calculated assuming  $f_{\text{esc}} = 1$ ,  $C_{\text{HII}} = 1$ ,  $\Delta = 1$ ,  $\beta = -2$ ,  $\alpha = 2$ . The dotted vertical lines mark the radius of the proximity zone,  $R_{\alpha}$ , defined by  $x_{\text{HII}} \lesssim 2 \times 10^{-6}$  (§ 2.2.3).

Galaxy spectra typically have a steeper drop-off beyond the HeII ionising limit (54.4 eV), which isn't captured in our simple power-law approximation. However, we note that this has only a small impact on our estimation of  $\dot{N}_{\text{ion}}$ : assuming the ionizing spectrum is zero for  $\nu > \nu_{\text{HeII}}$  we find  $\dot{N}_{\text{ion}}$  is  $> 0.75\times$  that obtained using Equation (9).

### 2.2.2 Residual neutral fraction of ionised region

Due to the recombination of ionised hydrogen, inside the ionised region there will be some residual neutral fraction. This can be computed by equating the recombination rate to the ionisation rate, assuming ionisation equilibrium. Assuming ionisations due to the central source and some diffuse ionising background, the residual neutral fraction at a proper radius  $r$  from the source is (e.g., [Mesinger et al. 2004](#)):

$$x_{\text{HII}}(r) = C_{\text{HII}}\Delta\bar{n}_{\text{H}}(z)\alpha_{\text{B}}(T) \left( \Gamma_{\text{bg}}(z) + \frac{J_s}{4\pi r^2} \right)^{-1}. \quad (10)$$

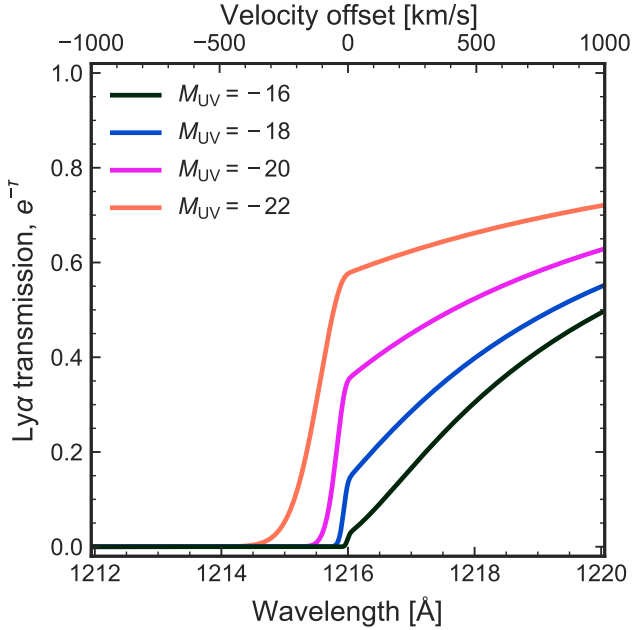
Here,  $\Gamma_{\text{bg}}$  is the hydrogen ionising rate due to the background within the ionised region in  $\text{s}^{-1}$ , and  $J_s$  is the hydrogen ionising emissivity of the central source in  $\text{cm}^2 \text{s}^{-1}$ :

$$J_s = f_{\text{esc}} \int_{\nu_{\text{H}}}^{\infty} d\nu \frac{L_{\nu}}{h\nu} \sigma_{\text{ion}}(\nu) \quad (11)$$

with the hydrogen photoionisation rate  $\sigma_{\text{ion}} = \sigma_{\text{ion},0}(\nu/\nu_{\text{H}})^{-3}$ , where  $\sigma_{\text{ion},0} \approx 6.3 \times 10^{-18} \text{ cm}^2$  (e.g., [Draine 2011](#)). Assuming as above an ionising spectrum  $L_{\nu} \propto \nu^{-\alpha}$  yields

$$J_s = f_{\text{esc}} \dot{N}_{\text{ion}} \frac{\alpha}{\alpha + 3} \sigma_{\text{ion},0} \quad (12)$$





**Figure 3.** Ly $\alpha$  transmission curves for the models shown in Figure 2, calculating the optical depth as described in § 2.1.

where  $\dot{N}_{\text{ion}}$  is given by Equation (9).

The gas reaches ionisation equilibrium with a characteristic timescale  $t_{\text{eq}}^{-1} = \Gamma_{\text{bg}} + J_s/4\pi r^2$ , so the bubble will be in ionisation equilibrium within  $\sim 10^5$  years assuming constant emissivity (e.g., Davies et al. 2020). Thus Equation (10) holds for sources with ionising populations  $> 10^5$  years, which is reasonable for massive galaxies at  $z \sim 6-8$ , though may break down for galaxies with short bursts of star formation.

We use the ionising background model by Khaire & Srianand (2019) but note that it does not significantly impact the residual neutral fraction for bubbles at  $z > 6$  as the background is low compared to the local ionisation field, leading to  $x_{\text{HI}} \propto r^2$ . Strictly,  $\Gamma_{\text{bg}}$  accounts for other ionising sources nearby (e.g., satellite galaxies in the vicinity of the central source) and will therefore vary depending on the density of the environment. We expect  $\Gamma_{\text{bg}} \lesssim 6\langle\Gamma_{\text{bg}}\rangle$  based on fluctuations of density and mean free path (Mesinger & Dijkstra 2008; Davies & Furlanetto 2016). For reference  $\langle\Gamma_{\text{bg}}(z \sim 7)\rangle \approx 0.2 \times 10^{-12} \text{ s}^{-1}$  in the Khaire & Srianand (2019) model, and has been measured to be  $\lesssim 0.3 \times 10^{-12} \text{ s}^{-1}$  at  $z \sim 6$  (Wyithe et al. 2010; Calverley et al. 2011).

Figure 2 shows some typical neutral fraction profiles inside a HII region. Here, and below, we assume that the neutral fraction is unity outside the HII region at the radius determined by Equation 5. We see that more luminous galaxies produce bubbles which are both larger (Equation 6) and more highly ionised at a fixed distance from the source. Figure 3 shows the Ly $\alpha$  for the same set of models. Only UV bright galaxies are capable of producing a sufficiently large ionised region to allow blue flux to be observed.

### 2.2.3 Optically thin region within ionised region

Importantly, due to the high cross-section of Ly $\alpha$  for scattering around the resonant wavelength (Equation (2)) an ionised bubble can still be optically thick to Ly $\alpha$ . Thus blue Ly $\alpha$  flux can be suppressed

by residual neutral gas *within* an ionised bubble. The proper radius at which the bubble becomes optically thick to Ly $\alpha$  is the radius where the Gunn & Peterson (1965) optical depth (Equation (4), using  $x_{\text{HI}}(r)$  given by Equation 10) exceeds an optical depth threshold  $\tau_{\text{lim}} \sim 2.3$  (i.e. transmission  $\sim 10\%$ ):

$$R_\alpha = \left( \frac{J_s}{4\pi} \right)^{1/2} \left[ \frac{C_{\text{HII}} \Delta^2 \bar{n}_{\text{H}}^2(z) \alpha_{\text{B}}(T) f_\alpha \pi e^2}{H(z) \tau_{\text{lim}} m_e v_\alpha} - \Gamma_{\text{bg}}(z) \right]^{-1/2} \quad (13)$$

$$\approx 0.1 \left( \frac{\tau_{\text{lim}}}{2.3} \right)^{1/2} \left( \frac{f_{\text{esc}} \dot{N}_{\text{ion}}}{3.3 \times 10^{54} \text{ s}^{-1}} \right)^{1/2} \left( \frac{2.5\alpha}{\alpha + 3} \right)^{1/2}$$

$$\times \frac{1}{C_{\text{HII}}^{1/2} \Delta} \left( \frac{T}{10^4 \text{ K}} \right)^{0.4} \left( \frac{1+z}{8} \right)^{-9/4} \text{ Mpc}$$

where  $J_s$  is given by Equation (12). For the latter equality we assumed  $\Gamma_{\text{bg}} = 0$ ,  $\alpha_{\text{B}}(T) \approx 2.6 \times 10^{-13} (T/10^4 \text{ K})^{-0.8} \text{ cm}^3 \text{ s}^{-1}$  and used  $\dot{N}_{\text{ion}}$  for a  $M_{\text{UV}} = -20$  galaxy (Equation 9).

This radius corresponds to reaching a neutral hydrogen number density of  $n_{\text{HI}} \gtrsim 2 \times 10^{-10} (\tau_{\text{lim}}/2.3) [(1+z)/8]^{3/2} \text{ cm}^{-3}$  – or  $x_{\text{HI}} \gtrsim 2 \times 10^{-6} \Delta^{-1} (\tau_{\text{lim}}/2.3) [(1+z)/8]^{-3/2}$  – in the ionised region. At higher densities/neutral fractions the gas is optically thick to Ly $\alpha$  photons.

This is analogous to the quasar near/proximity zones described by Bolton & Haehnelt (2007), except here we include the contribution of other, diffuse sources of ionising photons. As discussed in Section 2.2.2 we assume the reionised region is in ionisation equilibrium, which is valid for sources with ionising populations  $> 10^5$  years. See Davies et al. (2020) for discussion of the time evolution of such proximity zones around quasars.

A lower limit on  $R_\alpha$  can be estimated from the minimum observable blue Ly $\alpha$  velocity offset  $\Delta v_\alpha^{\text{min}}$ . To be transmitted to us, blue photons must redshift beyond the Ly $\alpha$  resonant wavelength (i.e.  $\Delta v_\alpha = 0$ ) *within* the proximity zone  $R_\alpha$ . Thus, the minimum distance photons travel while they redshift from  $\Delta v_\alpha^{\text{min}}$  into the resonant wavelength is:

$$R_\alpha > \frac{|\Delta v_\alpha^{\text{min}}|}{H(z_s)} \quad (14)$$

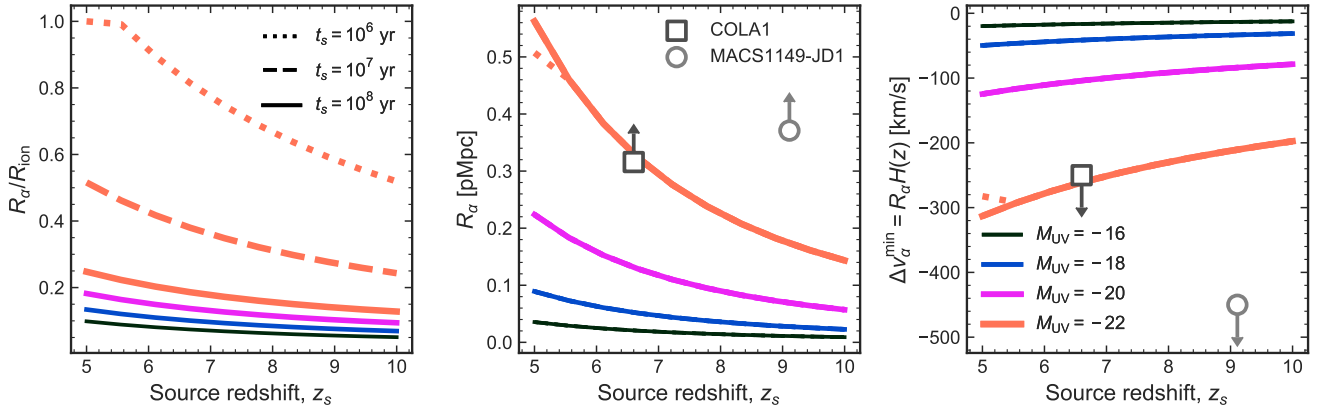
Previous works, which assumed ionised bubbles are optically thin to Ly $\alpha$  (e.g., Matthee et al. 2018; Hashimoto et al. 2018) estimated  $R_{\text{ion}} > |\Delta v_\alpha^{\text{min}}|/H(z_s)$ . From the above, we see this is actually measuring  $R_\alpha$  and is an underestimate of  $R_{\text{ion}}$ . We will show below in Section 3.1 that  $R_\alpha \ll R_{\text{ion}}$ .

## 3 RESULTS

### 3.1 Evolution of optically thin regions around galaxies

The Ly $\alpha$  optical depth (Equation 1) decreases with decreasing redshift, due to the increasing ionising output of sources, and the reducing density of neutral gas due to cosmic expansion. Thus, we expect the proximity zones around galaxies in reionising bubbles to grow with decreasing redshift, increasing the observable blue flux.

Figure 4 shows the total ionised radius ( $R_{\text{ion}}$ ) and the optically thin radius ( $R_\alpha$ ) as a function of source redshift, fixing  $f_{\text{esc}} = 1$ ,  $C_{\text{HII}} = 1$ ,  $\Delta = 1$ ,  $\alpha = 2$ ,  $\beta = -2$ . We compare the sizes of the bubbles and proximity zones for sources with different UV luminosities ( $M_{\text{UV}} = -16, -18, -20, -22$ ) and age ( $10^6, 10^7, 10^8$  yrs). Except for very young sources  $R_\alpha \lesssim 0.1 R_{\text{ion}}$ .  $R_\alpha$  does not change in size with age for constant emissivity once ionisation equilibrium is reached. As noted above, previous works, which assumed ionised bubbles are fully ionised when estimating blue peak transmission,



**Figure 4.** Radii of ionised bubbles and Ly $\alpha$  proximity zones around galaxies as a function of source redshift and age. Note these show optimistic upper limits on the sizes assuming  $f_{\text{esc}} = 1$ ,  $C_{\text{HII}} = 1$ ,  $\Delta = 1$ ,  $\alpha = 2$ ,  $\beta = -2$ . **Left:** Ratio of the Ly $\alpha$  proximity zone ( $R_\alpha$ , Equation 13) to the radius of ionised sphere ( $R_{\text{ion}}$ , Equation 5) for galaxies with  $M_{\text{UV}} = -16, -18, -20, -22$  (black, blue, pink, orange lines respectively). We show the time evolution of  $R_\alpha/R_{\text{ion}}$  for the brightest source. Early on, before the ionisation front can grow significantly,  $R_\alpha \sim 0.3 - 0.5 R_{\text{ion}}$ , however, for sources  $> 100$  Myr old,  $R_\alpha \lesssim 0.1 R_{\text{ion}}$ . **Center:** The proper size of the proximity zone  $R_\alpha$  in Mpc as a function of source redshift. All lines are the same as the left panel. **Right:** Radius of proximity zone expressed as a Ly $\alpha$  velocity offset, i.e. the maximum Ly $\alpha$  blue peak velocity offset that would be observable from a Ly $\alpha$  emitter inside this ionised region. We also show the observed blue-peaked Ly $\alpha$  emitters COLA1 (Hu et al. 2016; Matthee et al. 2018) and MACS1149-JD1 (Hashimoto et al. 2018). NEPLA4 (Songaila et al. 2018) is at the same redshift as COLA1 and has a similar blue peak velocity.

underestimated the total extent of the ionised region when using the observed blue Ly $\alpha$  peak. By including recombinations we see the blue Ly $\alpha$  flux only probes the much smaller proximity zone. In the next section, we show that there are model-independent ways to estimate a lower bound on the size of the full ionised region.

### 3.2 Observable Ly $\alpha$ lines in a mostly neutral medium

Figure 5 shows the minimum Ly $\alpha$  velocity offset,  $\Delta v$ , observable as a function of the distance to the first neutral region ( $R_{\text{ion}}$ ) and residual neutral fraction in the bubble ( $x_{\text{HI}}$ ). This can be interpreted as the necessary conditions in the galaxy’s surroundings for us to observe an emission line with given  $\Delta v$ . Our estimate does not depend on the intrinsic emission line shape, only on the transmission possible given the conditions inside the bubble. This means that our observations always provide lower limits on  $R_{\text{ion}}$  and upper limits on  $x_{\text{HI}}$ —Figure 5 shows that if  $R_{\text{ion}} = 1$  pMpc we could observe blue flux at  $-400 \text{ km s}^{-1}$ , however if a galaxy inside an ionized region of that size only emits flux at  $-100 \text{ km s}^{-1}$  we infer  $R_{\text{ion}} > 0.5$  pMpc.

To calculate the minimum  $\Delta v$  we compute Ly $\alpha$  transmission  $e^{-\tau(\Delta v)}$  on a grid of  $R_{\text{ion}}$  and  $x_{\text{HI}}$  values as described in Section 2.1. Some of the resulting transmission curves are shown in Figure 1. For each pair of  $R_{\text{ion}}$  and  $x_{\text{HI}}$  values we compute the minimum velocity offsets observable if  $> 10\%$  of the flux emitted at that velocity offset is transmitted through the IGM. We choose 10% as assuming an emitted Ly $\alpha$  EW of  $200 \text{ \AA}$ , this transmitted flux should be observable with current facilities. By using a grid of  $R_{\text{ion}}-x_{\text{HI}}$  values, our estimate does not assume any particular ionizing model (such as those described in Section 2.2), and thus provides a model-independent estimate of the properties of an ionised bubble based on Ly $\alpha$  transmission.

For small bubbles with high  $x_{\text{HI}}$ , it is only possible to observe Ly $\alpha$  which is significantly redshifted. Conversely, it is only possible to observe blue Ly $\alpha$  flux if there is a significant distance to the first neutral patch ( $> 0.5$  pMpc) and the HII region is highly ionised ( $x_{\text{HI}} < 10^{-5}$ ).

The two panels in Figure 5 compare the minimum observable

Ly $\alpha$  velocity offsets in the case of a homogeneous residual neutral fraction in the bubble ( $x_{\text{HI}} = \text{constant}$ ), approximating reionisation by a uniform ionising background of ultra-faint sources) and in the limiting case of the Ly $\alpha$  emitter as the sole reionising source ( $x_{\text{HI}} \propto r^2$ , Equation 10), with the value on the y-axis  $x_{\text{HI}}(r = 0.1 \text{ pMpc})$ . The trend of increased red-blue visibility with increasing bubble size is the same in both cases.

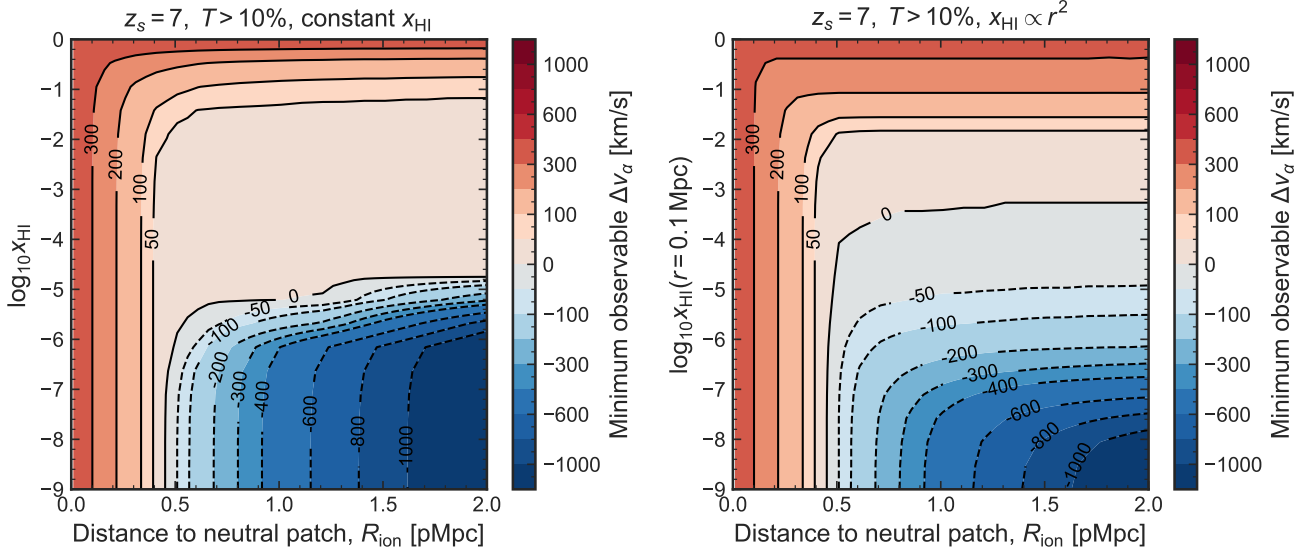
### 3.3 The ionised environment around observed blue-peaked Ly $\alpha$ emitters

#### 3.3.1 COLA1

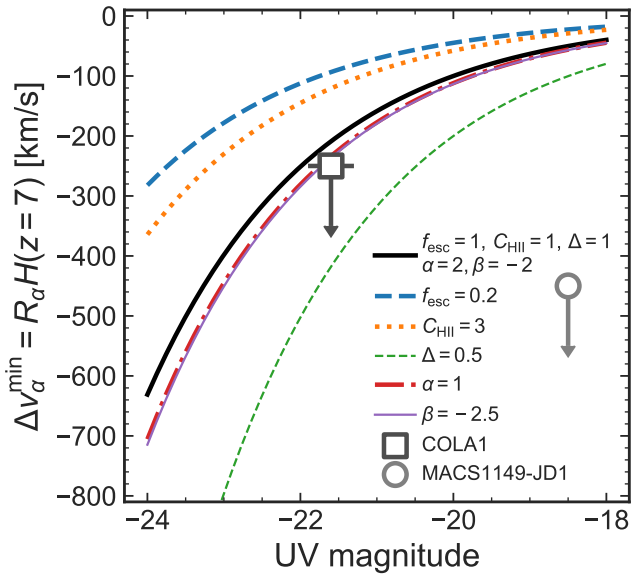
COLA1 is a  $z \approx 6.6$  galaxy with a blue Ly $\alpha$  peak with flux up to  $-250 \text{ km s}^{-1}$  from systemic (Hu et al. 2016; Matthee et al. 2018). From Figure 5 we see that this requires it to reside in an ionised region at least  $0.7$  pMpc to the nearest neutral patch, with a residual neutral fraction  $x_{\text{HI}} < 10^{-6}$ . Our estimate of the extent of the ionised region is roughly double than that of Matthee et al. (2018), who estimated  $0.3$  pMpc ( $2.3 \text{ cMpc}$ ). This is due to their assumption that the entire bubble is optically thin. As discussed in Section 3.1, with just the minimum observed blue flux velocity we can only calculate  $R_\alpha$  (Equation 14), but the total ionised region is much larger.

Figure 6 shows the radius of the proximity zone as a function of source magnitude. We compare  $R_\alpha$  to the blue peak velocity offsets from Matthee et al. (2018) and Hashimoto et al. (2018), and vary  $f_{\text{esc}} = \{1, 0.2\}$ ,  $C_{\text{HII}} = \{1, 3\}$ ,  $\Delta = \{1, 0.5\}$ ,  $\alpha = \{1, 2\}$  and  $\beta = \{-2, -2.5\}$ . For COLA1 to ionise its own proximity zone likely requires a high escape fraction, steep UV spectral slope  $\beta$ , and low gas density, whilst the ionising spectral slope makes a negligible impact on the proximity zone size.

To investigate in more detail the necessary conditions for COLA1’s blue peak to be observable we perform a Bayesian inference to infer the parameters in Equation (13). We define the likelihood to observe blue Ly $\alpha$  flux at  $-250 \pm 10 \text{ km s}^{-1}$  (the minimum velocity of observed blue flux and its uncertainty, J. Matthee private communication), from a galaxy with  $M_{\text{UV}} = -21.6 \pm 0.3$  at  $z = 6.6$ . For us to observe a blue peak at  $\Delta v$ , the photons must have travelled



**Figure 5.** Minimum observable Ly $\alpha$  velocity offset,  $\Delta v_\alpha$  as a function of bubble size and residual neutral fraction, assuming  $> 10\%$  transmission on the red or blue sides. **Left** for a constant residual neutral fraction,  $x_{\text{HI}}$ , inside the ionised bubble. **Right**  $x_{\text{HI}} \propto r^2$ , with the quoted value at 0.1 Mpc.



**Figure 6.** Radius of proximity zone produced by a single source of a given UV magnitude. We show the minimum observable blue-shifted velocity offset from systemic. The black solid line shows a fiducial galaxy model ( $f_{\text{esc}} = 1$ ,  $C_{\text{HII}} = 1$ ,  $\Delta = 1$ ,  $\alpha = 2$ ,  $\beta = -2$ ). The other lines show the impact of changing one of these parameters to a more extreme value:  $f_{\text{esc}} = 0.2$  (blue dash),  $C_{\text{HII}} = 3$  (orange dots),  $\Delta = 0.5$  (thin green dash),  $\alpha = 2$  (red dot-dash)  $\beta = -2.5$  (thin purple solid). We also show the observed blue-peaked Ly $\alpha$  emitters COLA1 (Hu et al. 2016; Matthee et al. 2018) and MACS1149-JD1 (Hashimoto et al. 2018).

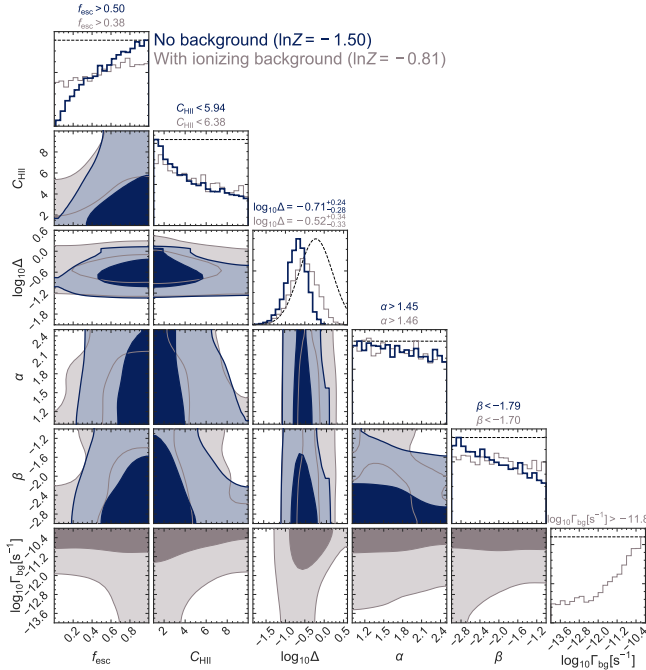
through a optically thin region of at least  $R_\alpha \geq |\Delta v|/H(z)$  before they redshifted into the Ly $\alpha$  resonant frequency. The likelihood of  $R_\alpha \geq |\Delta v|/H(z)$  given the model parameters  $\theta$  is:

$$p(R_\alpha \geq |\Delta v|/H(z) | \theta) = \frac{1}{2} \text{erfc} \left( \frac{|\Delta v|/H(z) - R_{\alpha, \text{mod}}(\theta)}{\sqrt{2}\sigma_R} \right) \quad (15)$$

where we have assumed the probability of a proximity zone having radius  $R_\alpha$ ,  $p(R_\alpha | \theta)$ , is a normal distribution with mean  $R_{\alpha, \text{mod}}(\theta)$  (Equation 13) and variance  $\sigma_R^2 = [\sigma_{\Delta v}/H(z)]^2$ . We use uniform priors on the parameters  $[f_{\text{esc}}, C_{\text{HII}}, \alpha, \beta, \log_{10} \Gamma_{\text{bg}}]$ :  $0 \leq f_{\text{esc}} \leq 1$ ,  $0.2 < C_{\text{HII}} < 10$ ,  $1 \leq \alpha \leq 2.5$ ,  $-3 < \beta < -1$ , and  $-14 < \log_{10}[\Gamma_{\text{bg}}/\text{s}^{-1}] < -10$ . We use a log-normal prior on  $\Delta$ :  $p(\ln \Delta) = \mathcal{N}(\sigma_0/\sqrt{2}, \sigma_0)$ , where  $\sigma_0$  is the variance of matter fluctuations on the filtered on the scale of the Jeans mass of the IGM. We use  $\sigma_0 \approx 1$  appropriate for ionized IGM at  $z = 6.6$  (e.g., Bi & Davidsen 1997; Bi et al. 2003).

We run two versions of the model: one where we fix the ionising background  $\Gamma_{\text{bg}} = 0$  (i.e. assuming COLA1 ionises its proximity zone alone), and one where  $\Gamma_{\text{bg}}$  is a free parameter. To estimate the posteriors and the evidence for each model,  $Z = \int d\theta p(\theta | R_\alpha)$ , we use Dynamic Nested Sampling implemented in dynesty (Speagle 2020).

Figure 7 shows the posteriors for these parameters and their median and 16 – 84% credible intervals or 68% upper/lower limits, and the evidence  $Z$ . In both cases, gas density is inferred to be low compared the mean ( $\log_{10} \Delta = -0.71^{+0.24}_{-0.28}$  with no ionizing background,  $\log_{10} \Delta = -0.52^{+0.34}_{-0.33}$  with an ionizing background, the UV slope  $\beta$  is inferred to be relatively steep ( $\beta < -1.79$   $1\sigma$  for  $\Gamma_{\text{bg}} = 0$ ) and the spectral slope of the ionising continuum,  $\alpha$ , is not particularly well constrained by the proximity zone, due to the smaller range of possible  $\alpha$  having a minimal impact on the size of the proximity zone (see Figure 6). In the model without an ionising background, high escape fractions are inferred ( $> 0.50$ ,  $1\sigma$ ), while when we include an ionising background there is a degeneracy between high single source  $f_{\text{esc}}$  and low ionising background, or low  $f_{\text{esc}}$  and high ionising background. In general, a high ionizing background ( $\log_{10}[\Gamma_{\text{bg}}/\text{s}^{-1}] \gtrsim -11$ ) will produce a large optically thin region, regardless of the other parameters, thus in the extreme case of a very high ionizing background the posteriors for the other parameters are prior-dominated. However, note that the maximum of our  $\Gamma_{\text{bg}}$  prior still constrains  $\Delta \lesssim 1.5$ : to see blue peaks in very overdense regions requires  $\Gamma_{\text{bg}} \gg 10^{-10} \text{s}^{-1}$ . Neither model is strongly preferred, with a Bayes factor  $Z_{\text{bg}}/Z_{\text{nobg}} \approx 2$  (e.g., Trotta 2008).



**Figure 7.** Posterior distributions for  $f_{\text{esc}}$ ,  $C_{\text{HII}}$ ,  $\Delta$ ,  $\alpha$ ,  $\beta$  and  $\Gamma_{\text{bg}}$  inferred from the observed maximum blue Ly $\alpha$  peak of COLA1 ( $-250 \text{ km s}^{-1}$ ). We show  $1\sigma$  and  $2\sigma$  contours of the 2D posteriors, and histograms of marginalised 1D posteriors for the parameters. Blue lines show the model with  $\Gamma_{\text{bg}} = 0$ , grey lines the model with  $\Gamma_{\text{bg}}$  as a free parameter. The likelihood and priors are described in Section 3.3.1.

### 3.3.2 Other $z > 6$ blue peaks

Hashimoto et al. (2018) reported a  $4\sigma$  detection of a Ly $\alpha$  line in the  $z = 9.11$  source MACS1149-JD1. The Ly $\alpha$  line is offset by  $-450 \pm 60 \text{ km s}^{-1}$  from their detection of [OIII]88  $\mu\text{m}$ .

Based on Figure 5, if the Ly $\alpha$  comes from the same source as the [OIII], the environment of MACS1149-JD1 must be extremely highly ionised ( $> 10^{-6}$ ) and in a bubble  $\gtrsim 1 \text{ pMpc}$ . Figure 6 shows the size of the proximity zones produced by galaxies of a given  $M_{\text{UV}}$ . Given the observed faintness of MACS1149-JD1 ( $M_{\text{UV}} = 18.5 \pm 0.1$  based on lens modelling and fits to photometry and grism spectroscopy, Hoag et al. 2018), it is impossible for it to produce such a large proximity zone, even with high  $f_{\text{esc}}$ , steep  $\alpha$  and low gas density. We thus agree with a possible interpretation by Hashimoto et al. (2018) that the Ly $\alpha$  emission comes from a different, slightly lower redshift, source compared to the [OIII] emission in MACS1149-JD1.

NEPLA4 (Songaila et al. 2018) is a narrow-band selected Ly $\alpha$  emitter at  $z = 6.6$  with blue flux up to  $\sim -250 \text{ km s}^{-1}$ , similar to COLA1. As the UV continuum is not known we cannot place it on Figure 6.

## 4 DISCUSSION

### 4.1 Ly $\alpha$ constraints on bubble properties

Blue Ly $\alpha$  peaks can reveal conditions inside individual reionised bubbles, however we expect blue peaks to be rare at  $z > 6$ , due to the high IGM opacity (see § 3.1 and e.g., Laursen et al. 2011). By contrast, Ly $\alpha$  lines which are redshifted with respect to systemic

may arise more often at high redshift due to outflows, which aids the transmission of photons through the IGM (Dijkstra et al. 2011). Figure 5 demonstrates that the velocity offset of an observed red peak can also place lower limits on the size of an ionised region (see also Malhotra & Rhoads 2006).

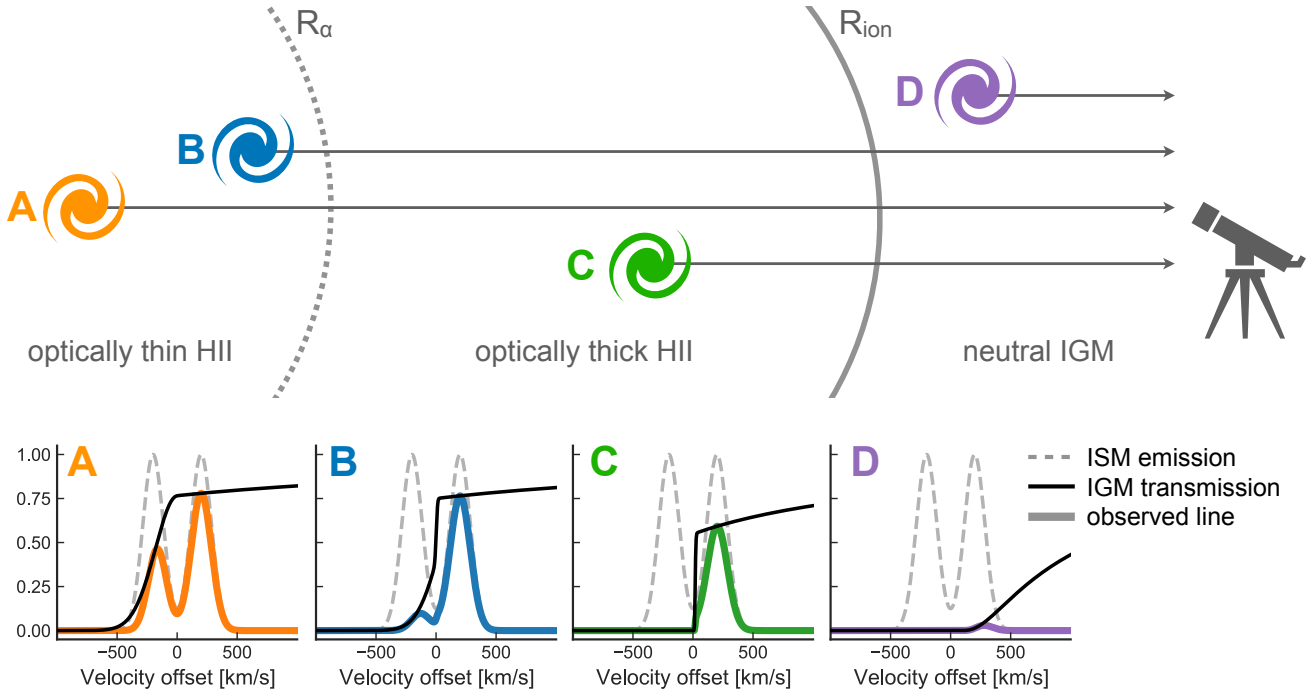
Not only can a single source place constraints on bubbles sizes, with a deep spectroscopic survey in a single field it could be possible to map an ionised bubble directly. Due to the ionisation gradient across bubbles the transmission of Ly $\alpha$  will vary radially across a bubble (see Figure 8). As we demonstrate in Figure 5, the observable minimum velocity offset from systemic varies as a function of the distance from the nearest neutral region. Likewise, the transmitted Ly $\alpha$  flux will decrease for sources further from the center of bubbles. If the faint-end of Ly $\alpha$  luminosity function is steep (e.g.,  $\alpha \sim -2.8$  Drake et al. 2017) at  $z \sim 7$ , using the luminosity function model by Gronke et al. (2015) (setting  $\alpha = -2.8$ ) we expect  $\sim 12$  Ly $\alpha$  emitting galaxies with a luminosity  $L \gtrsim 10^{41} \text{ erg s}^{-1}$  (flux  $\gtrsim 1 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2}$ ) located within  $R_{\alpha} \sim 0.3 \text{ pMpc}$  (the lower limit on COLA1’s proximity zone – see Figure 4). Note, this estimate does not account for galaxy clustering – if Ly $\alpha$  emitters live in overdense regions (Ouchi et al. 2018) we expect higher number counts. Within this proximity zone we would expect the fraction of galaxies with blue Ly $\alpha$  peaks to be comparable to those seen at lower redshifts, when the IGM is highly ionised.

With a large near infra-red spectroscopic survey to measure Ly $\alpha$  flux and high S/N resolved lineshapes, as well as systemic redshift from other emission lines (e.g. rest-frame optical lines visible with JWST) it could be possible to directly map ionised bubbles and place constraints on the bubble size distribution during reionisation using observations of both red and blue peaks. To accurately measure double-peaked Ly $\alpha$  line shapes requires a spectral resolution  $R \gtrsim 4000$  with  $\text{S/N} \gtrsim 2$  per pixel (e.g., Verhamme et al. 2015). These deep measurements will be feasible with 30 m telescope spectroscopy (e.g., E-ELT/MOSAIC is expected to reach  $1 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2}$  in 40 hrs, Evans et al. 2015).

To fully interpret such data requires a more realistic treatment of the neutral gas distribution than the uniform model used here. For example, (Gronke et al. 2020) explores the prevalence of blue Ly $\alpha$  peaks in the cosmological radiative hydrodynamical simulation CoDaII (Ocvirk et al. 2020), and finds large line-of-sight variation in Ly $\alpha$  transmission due to inhomogeneous gas distributions.

While the visibility of blue peaks (or red peaks with small velocity offset) can put lower limits on  $R_{\alpha}$ , measuring the size more accurately is difficult. However, given sufficient spectral quality, it might be possible to detect a sharp cutoff on the blue side of a blue peak – a signature of absorption, since frequency redistribution yields smoother profiles towards the wings (Neufeld 1990). A sharp cutoff would be expected if there is a sharp transition from an optically thin ionised region to one that is optically thick, for example, from galaxies sustaining their own  $R_{\alpha}$  surrounded by homogeneous  $10^{-5} \lesssim x_{\text{HII}} \lesssim 10^{-3}$  reionised gas (which may be typical of the IGM at the end of reionisation e.g., Fan et al. 2006). This would enable a direct measurement of  $R_{\alpha}$  and thus tighter constraints on  $f_{\text{esc}}$ . In contrast, a sharp cutoff of the red peak towards line centre can be due to either the IGM or radiative transfer effects and is commonly found also at low redshifts (e.g., Rivera-Thorsen et al. 2015; Yang et al. 2017), thus, this is harder to use as a measurement for  $R_{\alpha}$ . However, one could use this signature in a statistical sense: intra-galactic radiative transfer sets a characteristic correlation between the width and velocity offset of red Ly $\alpha$  peaks (e.g., Neufeld 1990; Verhamme et al. 2018), and this trend should be altered at high





**Figure 8.** Illustration of IGM attenuation for galaxies at different radial positions from the center of an ionised bubble. We show relative (not to scale) positions and observed line profiles of galaxies: at the center of the bubble, inside the proximity zone (A), close to the edge of the proximity zone (B), and within the optically thick region (C), in the neutral IGM (D). The plots show the observed Ly $\alpha$  emission lines expected if all the galaxies had the same double-peaked emission line emerging from the ISM. Galaxy A can be observed with significant blue flux, whereas galaxies C and D have no observed blue flux. A small fraction of Ly $\alpha$  is still visible from galaxy D, but only at  $\Delta v \gtrsim 300 \text{ km s}^{-1}$ .

redshift due to the IGM absorption yielding a flatter slope in the width-offset relation.

Naturally, the above discussion depends on the ‘intrinsic’ Ly $\alpha$  line, i.e., the one shaped by radiative transfer in the ISM/CGM, with the most and least constraining intrinsic lines (see Figure 5) being a wide double peak (with significant flux on the blue side), and a single red peak with large velocity offset, respectively. While in principle, the intrinsic spectrum (and its evolution) is unknown, we can assume a similar fraction of  $\sim 20 - 50\%$  of intrinsic spectra with significant blue flux – as seen in on low redshift observations (Yamada et al. 2012; Henry et al. 2015; Yang et al. 2016; Rivera-Thorsen et al. 2015; Erb et al. 2014; Herenz et al. 2017). The assumption of weak evolution in the blue peak fraction with redshift is supported by high redshift studies that find similar spectral properties to low redshift galaxies (Matthee et al. 2017; Songaila et al. 2018), and, in any case, would only affect our estimates, e.g., on the number of detected Ly $\alpha$  emitters with a blue peak by a factor of  $\sim 2$ .

#### 4.2 Impact of resonant absorption on EoR inferences

The increase in the size of the proximity zone means that more blue flux will be observed at lower redshifts, even if the size of the ionised bubble remains the same. The evolving flux distribution of Ly $\alpha$  emission from  $z > 6$  galaxies has been used to measure the average neutral hydrogen fraction of the IGM (Mesinger et al. 2015; Mason et al. 2018a), under the assumption that the declining flux is due to damping wing absorption alone. However, an increase in resonant absorption may also account for some of the decrease observed in the Ly $\alpha$  flux distribution at  $z > 6$  (see e.g., Bolton & Haehnelt 2013; Mesinger et al. 2015).

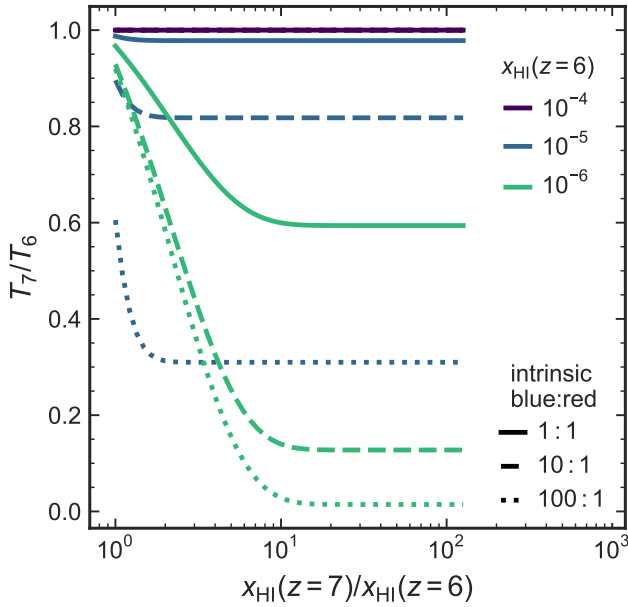
Given current constraints on the neutral fraction at the end of

reionisation ( $z < 6$ , Fan et al. 2006), we can ask under what conditions could the increased optical depth due to an increased residual neutral fraction in ionised bubbles cause the observed decline in Ly $\alpha$  emission.

Assuming a double-peaked Gaussian Ly $\alpha$  lineprofile, with red:blue flux ratio  $R : 1$ , the resulting Ly $\alpha$  transmission fraction through the IGM can be written as:

$$\frac{T(z=7)}{T(z=6)} = \frac{e^{-\tau_{\text{GP}}(z=7)} + R}{e^{-\tau_{\text{GP}}(z=6)} + R} \quad (16)$$

Figure 9 shows this as a function of the relative increase in the average residual neutral fraction between  $z \sim 6$  and  $z \sim 7$ ,  $x_{\text{HI}}(z=7)/x_{\text{HI}}(z=6)$ . In order to produce a drop in the observed Ly $\alpha$  fraction of  $T_7/T_6 \sim 0.5$ , either the neutral fraction at  $z=6$  must be  $< 10^{-6}$ , i.e. not optically thick, or if  $x_{\text{HI}}(z=6) \sim 10^{-5}$  the blue peak flux must be  $\gtrsim 100\times$  stronger than the red peak. Both of these scenarios are unlikely: the  $z \sim 6$  Ly $\alpha$  forest is optically thick on average (e.g., Fan et al. 2006, find  $x_{\text{HI}}(z \sim 6) \gtrsim 10^{-4}$ ) and the observed Ly $\alpha$  line shapes of galaxies at all redshifts show dominant red peaks (e.g., Trainor et al. 2015; Rivera-Thorsen et al. 2015; Yang et al. 2017; Steidel et al. 2018). Therefore, while the residual neutral fraction in ionised regions will increase between  $z \sim 6$  to  $z \sim 7$  due to a lower ionising background and increased recombinations, it is unlikely to significantly impact the integrated transmission of Ly $\alpha$  on average, and thus has a small impact on reionisation inferences.



**Figure 9.**  $\text{Ly}\alpha$  transmission ratio  $z = 7$  to  $z = 6$  given by Equation (16), assuming all of the optical depth to  $\text{Ly}\alpha$  is due to resonant absorption of the blue peak (i.e. no damping wing), as a function of the relative increase in the residual neutral fraction  $x_{\text{HI}}(z = 7)/x_{\text{HI}}(z = 6)$ . The coloured lines show the ratio for different values of  $x_{\text{HI}}(z = 6)$  and the line styles for different intrinsic blue:red peak ratios. For an increase in residual neutral fraction to explain the drop in  $\text{Ly}\alpha$  transmission  $T_7/T_6 \sim 0.5$  requires the  $z \sim 6$  IGM to be highly neutral  $x_{\text{HI}} \lesssim 10^{-5}$  and the blue peak to be significantly stronger than the red peak, contrary to observations at  $z \lesssim 5$  which show majority of dominant red peaks (e.g., Trainor et al. 2015; Rivera-Thorsen et al. 2015; Yang et al. 2017; Steidel et al. 2018).

## 5 CONCLUSIONS

We have used an analytic model to estimate the  $\text{Ly}\alpha$  optical depth within reionised bubbles and investigate the impact of reionisation on  $\text{Ly}\alpha$  lineshapes. Our conclusions are as follows:

- (i) Both the size of, and residual neutral hydrogen fraction within, reionised bubbles affect the observed lineshape of  $\text{Ly}\alpha$  emission during reionisation. As such, measurements of the  $\text{Ly}\alpha$  velocity offset from systemic can provide lower limits on the source’s distance to the first large neutral patch along the line of sight, and upper limits on the residual neutral fraction inside its HII bubble.
- (ii) Galaxies with  $\text{Ly}\alpha$  lines observed with low velocity offsets from systemic must reside in large reionised regions. Detecting blue  $\text{Ly}\alpha$  peaks during reionisation requires the source galaxy live in a  $R \gtrsim 0.5$  pMpc, and highly ionised ( $x_{\text{HI}} > 10^{-5}$ ) bubble. By contrast,  $\text{Ly}\alpha$  can be detected even from fully neutral regions, providing it was emitted at high velocity offset ( $\gtrsim 300 \text{ km s}^{-1}$ ).
- (iii) Around individual galaxies, we predict the regions of low  $\text{Ly}\alpha$  optical depth – proximity zones – to be typically  $< 0.3$  pMpc, allowing blue-shifted  $\text{Ly}\alpha$  to be detected out to  $\gtrsim -250 \text{ km s}^{-1}$ .
- (iv) The observed blue-peaked  $\text{Ly}\alpha$  emitter COLA1, with blue flux up to  $-250 \text{ km s}^{-1}$ , must reside in a highly ionised region ( $x_{\text{HI}} < 10^{-5.5}$ ) at least  $0.7$  pMpc from the nearest large scale neutral patch. For COLA1 to have generated its own proximity zone requires it to have a high escape fraction,  $f_{\text{esc}} > 0.50$ , and steep UV spectral slope,  $\beta < -1.79$ , and for the total gas density along the line of sight

to be low  $\sim 0.2\bar{n}_{\text{H}}$ . Including an ionising background alleviates the need for a high  $f_{\text{esc}}$  and steep  $\beta$ , but still requires a low gas density.

Detailed measurements of  $\text{Ly}\alpha$  lineshapes and velocity offsets can be used to constrain the properties of reionised regions, including to place lower limits on ionised bubble sizes. With rest-UV – optical spectroscopy with, e.g., the James Webb Space Telescope (JWST) these methods enable direct comparison between galaxy properties and their ionised regions.

## ACKNOWLEDGEMENTS

The authors thank the ENIGMA group at UCSB (especially Vikram Khaire, Fred Davies and Joe Hennawi), Jorryt Matthee, Zoltan Haiman, Steve Finkelstein, Dan Stark and Tommaso Treu for useful discussions. CAM acknowledges support by NASA Headquarters through the NASA Hubble Fellowship grant HST-HF2-51413.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. MG was supported by NASA through the NASA Hubble Fellowship grant HST-HF2-51409 and acknowledges support from HST grants HST-GO-15643.017-A, HST-AR-15039.003-A, and XSEDE grant TG-AST180036.

*Software:* IPython (Pérez & Granger 2007), matplotlib (Hunter 2007), NumPy (Van Der Walt et al. 2011), SciPy (Oliphant 2007), Astropy (Robitaille et al. 2013), dynesty (Speagle 2020).

## DATA AVAILABILITY

The source code underlying this article is available on GitHub at <https://github.com/charlottenosam/LyaLineshapes>.

## REFERENCES

- Beardsley A. P., Morales M. F., Lidz A., Malloy M., Sutter P. M., 2015, *ApJ*, 800, 128
- Bi H., Davidsen A. F., 1997, *ApJ*, 479, 523
- Bi H., Fang L.-Z., Feng L., Jing Y., 2003, *ApJ*, 598, 1
- Bolton J. S., Haehnelt M. G., 2007, *MNRAS*, 374, 493
- Bolton J. S., Haehnelt M. G., 2013, *MNRAS*, 429, 1695
- Bosman S. E. I., Kakiichi K., Meyer R. A., Gronke M., Laporte N., Ellis R. S., 2020, *ApJ*, 896, 49
- Bouwens R. J., et al., 2014, *ApJ*, 793, 115
- Byrohl C., Gronke M., 2020, arXiv e-prints, p. arXiv:2006.10041
- Calverley A. P., Becker G. D., Haehnelt M. G., Bolton J. S., 2011, *MNRAS*, 412, 2543
- Cen R., Haiman Z., 2000, *ApJ*, 542, L75
- Davies F. B., Furlanetto S. R., 2016, *MNRAS*, 460, 1328
- Davies F. B., et al., 2018, *ApJ*, 864, 142
- Davies F. B., Hennawi J. F., Eilers A.-C., 2020, *MNRAS*, 493, 1330
- Dijkstra M., 2014, *Publ. Astron. Soc. Aust.*, 31, e040
- Dijkstra M., Mesinger A., Wyithe J. S. B., 2011, *MNRAS*, 414, 2139
- Draine B. T., 2011, *Physics of the Interstellar and Intergalactic Medium*
- Drake A. B., et al., 2017, *A&A*, 608, A6
- Dunlop J. S., et al., 2013, *MNRAS*, 432, 3520
- Eide M. B., Gronke M., Dijkstra M., Hayes M., 2018, *ApJ*, 856, 156
- Erb D. K., et al., 2014, *ApJ*, 795, 33
- Evans C., et al., 2015, p. arXiv:1501.04726
- Fan X., et al., 2006, *AJ*, 132, 117
- Feltre A., Charlot S., Gutkin J., 2016, *MNRAS*, 456, 3354
- Furlanetto S. R., Zaldarriaga M., Hernquist L., 2004, *ApJ*, 613, 16

- Geil P. M., Mutch S. J., Poole G. B., Duffy A. R., Mesinger A., Wyithe J. S. B., 2017, *MNRAS*, **472**, 1324
- Götberg Y., de Mink S. E., McQuinn M., Zapartas E., Groh J. H., Norman C., 2020, *A&A*, **634**, A134
- Greig B., Mesinger A., Bañados E., 2019, *MNRAS*, **484**, 5094
- Gronke M., Dijkstra M., Trenti M., Wyithe S., 2015, *MNRAS*, **449**, 1284
- Gronke M., et al., 2020, arXiv e-prints, p. [arXiv:2004.14496](https://arxiv.org/abs/2004.14496)
- Gunn J. E., Peterson B. A., 1965, *ApJ*, **142**, 1633
- Haiman Z., 2002, *ApJ*, **576**, L1
- Hashimoto T., et al., 2018, *Nature*, **557**, 392
- Hayes M. J., Runnholm A., Gronke M., Scarlata C., 2020, arXiv e-prints, p. [arXiv:2006.03232](https://arxiv.org/abs/2006.03232)
- Henry A., Scarlata C., Martin C. L., Erb D., 2015, *ApJ*, **809**, 19
- Herenz E. C., et al., 2017, *A&A*, **606**, A12
- Hoag A., et al., 2018, *ApJ*, **854**, 39
- Hoag A., et al., 2019, *ApJ*, **878**, 12
- Hu E. M., Cowie L. L., Songaila A., Barger A. J., Rosenwasser B., Wold I. G. B., 2016, *ApJ*, **825**, L7
- Hui L., Gnedin N. Y., 1997, *MNRAS*, **292**, 27
- Hunter J. D., 2007, *Comput. Sci. Eng.*, **9**, 99
- Khaire V., Srianand R., 2019, *MNRAS*, **484**, 4174
- Koopmans L., et al., 2015, in *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*. p. 1 ([arXiv:1505.07568](https://arxiv.org/abs/1505.07568))
- Laursen P., Sommer-Larsen J., Razoumov A. O., 2011, *ApJ*, **728**, 52
- Lusso E., Worseck G., Hennawi J. F., Prochaska J. X., Vignali C., Stern J., O’Meara J. M., 2015, *MNRAS*, **449**, 4204
- Malhotra S., Rhoads J. E., 2006, *ApJ*, **647**, L95
- Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018a, *ApJ*, **856**, 2
- Mason C. A., et al., 2018b, *ApJ*, **857**, L11
- Mason C. A., et al., 2019, *MNRAS*, **485**, 3947
- Matthee J., Sobral D., Darvish B., Santos S., Mobasher B., Paulino-Afonso A., Röttgering H., Alegre L., 2017, *MNRAS*, **472**, 772
- Matthee J., Sobral D., Gronke M., Paulino-Afonso A., Stefanon M., Röttgering H., 2018, *A&A*, **619**, A136
- McQuinn M., Lidz A., Zahn O., Dutta S., Hernquist L., Zaldarriaga M., 2007, *MNRAS*, **377**, 1043
- Mesinger A., 2016, *Understanding the Epoch of Cosmic Reionization: Challenges and Progress*. Vol. 423, Springer, [doi:10.1007/978-3-319-21957-8](https://doi.org/10.1007/978-3-319-21957-8)
- Mesinger A., Dijkstra M., 2008, *MNRAS*, **390**, 1071
- Mesinger A., Haiman Z., Cen R., 2004, *ApJ*, **613**, 23
- Mesinger A., Aykutalp A., Vanzella E., Pentericci L., Ferrara A., Dijkstra M., 2015, *MNRAS*, **446**, 566
- Miralda-Escude J., 1998, *ApJ*, **501**, 15
- Neufeld D. A., 1990, *ApJ*, **350**, 216
- Ocvirk P., et al., 2020, *MNRAS*, **496**, 4087
- Oliphant T. E., 2007, *Comput. Sci. Eng.*, **9**, 10
- Ouchi M., et al., 2018, *PASJ*, **70**, S13
- Peebles P. J. E., 1993, *Principles of Physical Cosmology*
- Pérez F., Granger B. E., 2007, *Comput. Sci. Eng.*, **9**, 21
- Planck Collaboration et al., 2015, *A&A*, **594**, A8
- Rivera-Thorsen T. E., et al., 2015, *ApJ*, **805**, 14
- Robitaille T. P., et al., 2013, *A&A*, **558**, A33
- Santos M. R., 2004, *MNRAS*, **349**, 1137
- Schenker M. A., Ellis R. S., Konidaris N. P., Stark D. P., 2014, *ApJ*, **795**, 20
- Scott J. E., Kriss G. A., Brotherton M., Green R. F., Hutchings J., Shull J. M., Zheng W., 2004, *ApJ*, **615**, 135
- Shapiro P. R., Giroux M. L., 1987, *ApJ*, **321**, L107
- Songaila A., Hu E. M., Barger A. J., Cowie L. L., Hasinger G., Rosenwasser B., Waters C., 2018, *ApJ*, **859**, 91
- Speagle J. S., 2020, *MNRAS*, **493**, 3132
- Steidel C. C., et al., 2014, *ApJ*, **795**, 165
- Steidel C. C., Bogosavljevic M., Shapley A. E., Reddy N. A., Rudie G. C., Pettini M., Trainor R. F., Strom A. L., 2018, *ApJ*, **869**, 123
- Stevens M. L., Shull J. M., Danforth C. W., Tilton E. M., 2014, *ApJ*, **794**, 75
- Tasitsiomi A., 2006, *ApJ*, **645**, 792
- Trainor R. F., Steidel C. C., Strom A. L., Rudie G. C., 2015, *ApJ*, **809**, 89
- Treu T., Schmidt K. B., Trenti M., Bradley L. D., Stiavelli M., 2013, *ApJ*, **775**, L29
- Trotta R., 2008, *Contemporary Physics*, **49**, 71
- Van Der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, **13**, 22
- Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, *A&A*, **578**, A7
- Verhamme A., et al., 2018, *MNRAS*, **478**, L60
- Whitler L. R., Mason C. A., Ren K., Dijkstra M., Mesinger A., Pentericci L., Trenti M., Treu T., 2020, *MNRAS*, **495**, 3602
- Wyithe J. S. B., Hopkins A. M., Kistler M. D., Yüksel H., Beacom J. F., 2010, *MNRAS*, **401**, 2561
- Yajima H., Sugimura K., Hasegawa K., 2018, *MNRAS*, **477**, 5406
- Yamada T., Matsuda Y., Kousai K., Hayashino T., Morimoto N., Umemura M., 2012, *ApJ*, **751**, 29
- Yang H., Malhotra S., Gronke M., Rhoads J. E., Dijkstra M., Jaskot A., Zheng Z., Wang J., 2016, *ApJ*, **820**, 130
- Yang H., et al., 2017, *ApJ*, **844**, 171
- Zheng Z., Cen R., Trac H., Miralda-Escudé J., 2010, *ApJ*, **716**, 574

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.