

# Deciphering the Lyman- $\alpha$ emission line: towards the understanding of galactic properties extracted from Ly $\alpha$ spectra via radiative transfer modelling

Zhihui Li  <sup>1</sup>★ and Max Gronke  <sup>2,3</sup>†

<sup>1</sup>*Cahill Center for Astrophysics, California Institute of Technology, MC 249-17, 1200 East California Boulevard, Pasadena, CA 91125, USA*

<sup>2</sup>*Max-Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany*

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

Accepted 2022 April 24. Received 2022 April 7; in original form 2021 November 7

## ABSTRACT

Existing ubiquitously in the Universe with the highest luminosity, the Lyman- $\alpha$  (Ly $\alpha$ ) emission line encodes abundant physical information about the gaseous medium it interacts with. Nevertheless, the resonant nature of the Ly $\alpha$  line complicates the radiative transfer (RT) modelling of the line profile. We revisit the problem of deciphering the Ly $\alpha$  emission line with RT modelling. We reveal intrinsic parameter degeneracies in the widely used shell model in the optically thick regime for both static and outflowing cases, which suggest the limitations of the model. We also explore the connection between the more physically realistic multiphase, clumpy model, and the shell model. We find that the parameters of a ‘very clumpy’ slab model and the shell model have the following correspondences: (1) the *total* column density, the effective temperature, and the *average* radial clump outflow velocity of the clumpy slab model are equal to the H I column density, effective temperature, and expansion velocity of the shell model, respectively; (2) large intrinsic linewidths are required in the shell model to reproduce the wings of the clumpy slab models; (3) adding another phase of hot interclump medium increases peak separation, and the fitted shell expansion velocity lies between the outflow velocities of two phases of gas. Our results provide a viable solution to the major discrepancies associated with Ly $\alpha$  fitting reported in previous literature, and emphasize the importance of utilizing information from additional observations to break the intrinsic degeneracies and interpreting the model parameters in a more physically realistic context.

**Key words:** line: formation – radiative transfer – scattering – galaxies: high-redshift – galaxies: ISM.

## 1 INTRODUCTION

Owing to its luminous nature, Lyman- $\alpha$  (Ly $\alpha$ ) is one of the best emission lines to explore the high-redshift Universe, including identifying and studying the formation of distant galaxies as well as probing the reionization era (see a recent review by Ouchi, Ono & Shibuya 2020). Despite all its advantages, the Ly $\alpha$  line is a resonant transition with a large cross-section, making its radiative transfer (RT) process notoriously difficult to model. Initially, the Ly $\alpha$  RT problem was studied analytically for several simple cases, e.g. static plane-parallel slabs (Harrington 1973; Neufeld 1990), a two-phase ISM (Neufeld 1991), static uniform spherical shells (Dijkstra, Haiman & Spaans 2006a), and a uniform neutral IGM with pure Hubble expansion (Loeb & Rybicki 1999). Later on, more and more studies started to employ numerical (mostly Monte Carlo) methods in more sophisticated configurations, e.g. flattened, axially symmetric, rotating clouds (Zheng & Miralda-Escudé 2002), expanding/contracting spherical shells (Zheng & Miralda-Escudé 2002; Ahn, Lee & Lee 2003; Ahn 2004; Verhamme, Schaerer &

Maselli 2006; Dijkstra et al. 2006a; Dijkstra, Haiman & Spaans 2006b; Gronke, Bull & Dijkstra 2015; Song, Seon & Hwang 2020), (moving) multiphase, clumpy medium (Richling 2003; Hansen & Oh 2006; Dijkstra & Kramer 2012; Laursen, Duval & Östlin 2013; Duval et al. 2014; Gronke & Dijkstra 2016), anisotropic gas distributions (Behrens, Dijkstra & Niemeyer 2014; Zheng & Wallace 2014), and gas with power-law density profiles (Lao & Smith 2020), as well as in the context of cosmological simulations (Cantalupo et al. 2005; Tasitsiomi 2006; Laursen & Sommer-Larsen 2007; Verhamme et al. 2012; Smith et al. 2019, 2021).

With the significant advancements on the theoretical side, many attempts have been made to bridge the gap between the simulations and observations, one of which is to match the Ly $\alpha$  spectra derived from the RT models with the observed Ly $\alpha$  profiles. The most widely used RT model for this endeavor is the ‘shell model’, i.e. a spherical, expanding/contracting H I shell. Thus far, the shell model has managed to reproduce a wide variety of Ly $\alpha$  profiles, including typical single and double-peaked profiles from Lyman break galaxies (LBGs), Ly $\alpha$  emitters (LAEs), damped Ly $\alpha$  systems (DLAs) and Green Pea galaxies (e.g. Verhamme et al. 2008; Dessauges-Zavadsky et al. 2010; Vanzella et al. 2010; Krogager et al. 2013; Hashimoto et al. 2015; Yang et al. 2016, 2017), along with the P-Cygni profiles and damped absorption features in nearby starburst galaxies (e.g.

\* E-mail: [zhihui@caltech.edu](mailto:zhihui@caltech.edu)

† Hubble fellow

Atek, Schaerer & Kunth 2009; Martin et al. 2015). Nevertheless, a number of discrepancies between the fitted parameters of the shell model and observational constraints have been observed (e.g. Kulas et al. 2012; Hashimoto et al. 2015; Yang et al. 2016, 2017). Most recently, Orlitová et al. (2018) reported three major discrepancies emerged from shell modelling of the observed Ly $\alpha$  profiles of twelve Green Pea galaxies, namely (1) the required intrinsic Ly $\alpha$  linewidths are on average three times broader than the observed Balmer lines; (2) the inferred outflow velocities of the shell ( $\lesssim 150 \text{ km s}^{-1}$ ) are significantly lower than the characteristic outflow velocities ( $\sim 300 \text{ km s}^{-1}$ ) indicated by the observed ultraviolet (UV) absorption lines of low-ionization-state elements; (3) the best-fitting systemic redshifts are larger (by  $10\text{--}250 \text{ km s}^{-1}$ ) than those derived from optical emission lines. Such inconsistencies suggest the limitations of the shell model and necessitate the development of more realistic RT models.

In addition, it is unclear whether the derived values of the shell model can be directly used to infer other physical properties of the Ly $\alpha$ -emitting object. For example, Verhamme et al. (2015) proposed that low H I column densities ( $\lesssim 10^{18} \text{ cm}^{-2}$ ) inferred from observed Ly $\alpha$  profiles should indicate Lyman-continuum (LyC) leakage. However, it has not been verified quantitatively that a tight correlation does exist between the H I column density inferred from Ly $\alpha$  and the LyC escape fraction as expected theoretically (see equation 4 in Verhamme et al. 2017). The situation is even more complicated when more physics (e.g. turbulence) is considered, e.g. Kakiichi & Gronke (2021) find that a high average H I column density still allows high LyC leakage, as LyC photons can escape through narrow photoionized channels with a large fraction of hydrogen remaining neutral (see their section 4.2; see also Kimm et al. 2019).

The shell model is known as being unrealistically monolithic as it consists of only one phase of H I at  $\sim 10^4 \text{ K}$  (the ‘cool’ phase). Alternatively, Ly $\alpha$  RT has been studied in multiphase, clumpy models (e.g. Neufeld 1991; Hansen & Oh 2006; Dijkstra & Kramer 2012; Laursen et al. 2013; Duval et al. 2014; Gronke & Dijkstra 2016), as numerous observations have revealed the multiphase nature of the interstellar/circumgalactic/intergalactic medium (ISM/CGM/IGM, respectively; see reviews by Cox 2005; McQuinn 2016; Tumlinson, Peebles & Werk 2017). This multiphase, clumpy model consists of two different phases of gas: cool clumps of H I ( $\sim 10^4 \text{ K}$ ) embedded in a hot, highly ionized medium ( $\sim 10^6 \text{ K}$ ). Using the framework in Gronke & Dijkstra (2016), Li et al. (2021a, 2021b) successfully reproduced the spatially resolved Ly $\alpha$  profiles in Ly $\alpha$  blobs 1 and 2 with the multiphase, clumpy model. These results have not only demonstrated the feasibility of the multiphase, clumpy model, but also motivated us to gain a deeper understanding of the physical meaning of the derived model parameters.

The primary goal of this work is to figure out the links between the parameters of the relatively newly developed, more physically realistic multiphase, clumpy model and the commonly adopted shell model, as well as what physical information can be extracted from observed Ly $\alpha$  spectra. A schematic representation of the configuration of these two models is shown in Fig. 1. The shell model only has four most important parameters<sup>1</sup>: the shell expansion velocity ( $v_{\text{exp}}$ ), the shell H I column density ( $N_{\text{H I, shell}}$ ), the shell effective temperature ( $T_{\text{shell}}$ ) or the Doppler parameter ( $b$ ), and the intrinsic Ly $\alpha$  linewidth ( $\sigma_i$ ) (Verhamme et al. 2006; Gronke et al. 2015). The multiphase,

clumpy model has six most crucial parameters: (1) the cloud covering factor ( $f_{\text{cl}}$ ), which is the mean number of clumps per line of sight; (2) the H I column density of the clumps ( $N_{\text{H I, cl}}$ ); (3) the velocity dispersion of the clumps ( $\sigma_{\text{cl}}$ ); (4) the radial outflow velocity of the clumps ( $v_{\text{cl}}$ ); (5) the residual H I number density in the interclump medium (ICM,  $n_{\text{H I, ICM}}$ ); (6) the radial outflow velocity of the ICM ( $v_{\text{ICM}}$ ). For both models, an additional post-processed parameter,  $\Delta v$ , is used to determine the systemic redshift of the Ly $\alpha$  emitting source. The shell model parameters capture different properties of the Ly $\alpha$  spectra:  $v_{\text{exp}}$  determines the blue-to-red peak flux ratio, and sets the position of the absorption trough between two peaks (as  $-v_{\text{exp}}$  corresponds to the largest optical depth);  $N_{\text{H I, shell}}$  dictates the amount of peak separation and the depth of the absorption trough;  $T_{\text{shell}}$  or  $b$  describes the internal kinematics of the shell (including thermal and turbulent velocities) and controls the width of the Ly $\alpha$  profile, but is usually poorly constrained by the data (Gronke et al. 2015);  $\sigma_i$  (if large enough) sets the extent of the wings of the spectrum. The multiphase, clumpy model parameters capture similar spectral properties but in different ways:  $v_{\text{cl}}$  and  $v_{\text{ICM}}$  determines the blue-to-red peak flux ratio<sup>2</sup>;  $n_{\text{H I, ICM}}$  and  $f_{\text{cl}}N_{\text{H I, cl}}$  together dictate the amount of peak separation and the depth of the absorption trough, as both of them contribute to the total H I column density;  $\sigma_{\text{cl}}$  sets the width of the spectrum.

Li et al. (2021a) have observed significant correlations between certain pair of model parameters (namely  $v_{\text{exp}} - v_{\text{cl}}$ ) derived by fitting fifteen observed Ly $\alpha$  spectra. These results are enlightening yet not rigorous and may suffer from parameter degeneracy due to their empirical nature. Motivated by the fact that the multiphase, clumpy model may converge to the shell model in the limit of very high  $f_{\text{cl}}$  (Gronke et al. 2017), in this work, we attempt to find quantitative correlations between the parameters of two models, with the aim of better understanding the physical meaning of model parameters and their relation to Ly $\alpha$  spectral properties.

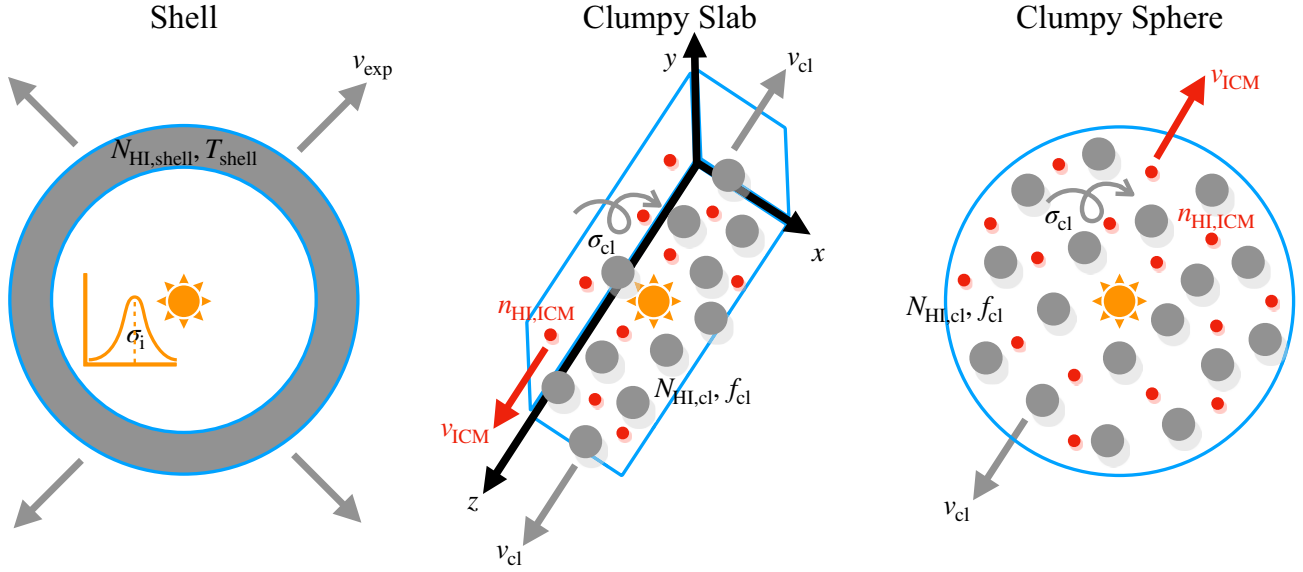
The structure of this paper is as follows. In Section 2, we describe the methodology of this work. In Section 3, we present the intrinsic parameter degeneracies of the shell model. In Section 4, we explore the connection between the shell model and the multiphase, clumpy model. In Section 5, we discuss on how to interpret the physical parameters extracted from Ly $\alpha$  spectra. In Section 6, we summarize and conclude. The physical units used throughout this paper are  $\text{km s}^{-1}$  for velocity,  $\text{cm}^{-2}$  for column density, and K for temperature, unless otherwise specified.

## 2 METHODOLOGY

In this work, we extract physical parameters from Ly $\alpha$  spectra by fitting them with a grid of shell models. The fitted Ly $\alpha$  spectra can be one of the following: (1) a shell model spectrum; (2) a (multiphase) clumpy model spectrum; (3) an observed Ly $\alpha$  spectrum. The grid of shell models that we use was previously described in Gronke et al. (2015). This shell model grid consists of 12960 discrete RT models, with  $[v_{\text{exp}}, \log N_{\text{H I, shell}}, \log T_{\text{shell}}]$  varying between  $[0, 490] \text{ km s}^{-1}$ ,  $[16.0, 21.8] \text{ cm}^{-2}$  and  $[3, 5.8] \text{ K}$ , respectively. Each shell model is calculated via Monte Carlo RT using 20000 Ly $\alpha$  photon packages generated from an a priori Gaussian intrinsic spectrum

<sup>1</sup>Here, we assume that the shell model is dust-free, as the dust content is usually poorly constrained by the observed Ly $\alpha$  spectra (Gronke et al. 2015).

<sup>2</sup>In the multiphase, clumpy model, the absorption trough is not necessarily set by  $-v_{\text{cl}}$  unless the total column density of the clumps is high enough to be optically thick (i.e. the flux density at line centre is close to 0) and the clumps and ICM are co-outflowing at the same velocity (see Section 4.4).



**Figure 1.** A schematic representation of the configuration of the shell model and the multiphase, clumpy model. The four most important parameters in the shell model are the shell expansion velocity ( $v_{\text{exp}}$ ), the shell H I column density ( $N_{\text{HI, shell}}$ ), the shell effective temperature ( $T_{\text{shell}}$ ) or the Doppler parameter ( $b$ ), and the intrinsic Ly $\alpha$  linewidth ( $\sigma_i$ ). The multiphase, clumpy model has six most crucial parameters: (1) the cloud covering factor ( $f_{\text{cl}}$ ), which is the mean number of clumps per line of sight; (2) the H I column density of the clumps ( $N_{\text{HI, cl}}$ ); (3) the velocity dispersion of the clumps ( $\sigma_{\text{cl}}$ ); (4) the radial outflow velocity of the clumps ( $v_{\text{cl}}$ ); (5) the residual H I number density in the interclump medium (ICM,  $n_{\text{HI, ICM}}$ ); (6) the radial outflow velocity of the ICM ( $v_{\text{ICM}}$ ). The orange solar sign represents the central Ly $\alpha$  emitting source in each model. Two different geometries for the multiphase, clumpy model, slab and sphere, are explored in this work.

$N(0, \sigma^2)$ , where  $\sigma = 800 \text{ km s}^{-1}$ . The intrinsic Ly $\alpha$  linewidth,  $\sigma_i \in [1, 800] \text{ km s}^{-1}$ , is accounted for in the form of a weighting function in post-processing. We do not consider the effect of dust in this work as it is usually a poorly constrained parameter<sup>3</sup> (Gronke et al. 2015).

To properly explore the possibly multimodal posterior of the shell model parameters, we use a python package *dynesty* (Skilling 2004, 2006; Speagle 2020) that implements the nested sampling algorithm for our fitting pipeline. The model spectrum of each sampled point in the parameter space is calculated via linear flux interpolation on the model grid rather than running the computationally expensive RT ‘on the fly’. When fitting the Ly $\alpha$  spectra, we manually add a constant  $1\sigma$  uncertainty of about 10 per cent of the maximum flux density to the normalized (mock) data to reflect the typical observational uncertainties. The uncertainties in the fitted parameters are determined as certain quantiles (e.g. 16–84 per cent, or  $1\sigma$  confidence intervals) of the samples in the marginalized posterior probability distributions.

### 3 RESULTS I: INTRINSIC PARAMETER DEGENERACIES OF THE SHELL MODEL

In this section, we show the existence of intrinsic parameter degeneracies in the shell model revealed by fitting, in preparation for our subsequent discussion. We consider the two following cases: static shell and outflowing shell, respectively.

<sup>3</sup>Note that dust plays a similar role in both the shell model and the multiphase, clumpy model (especially in the limit of many clumps per sightline, see e.g. Neufeld 1991; Gronke 2017), which are both assumed to be homogeneous in this work. The effect of dust could be potentially important in an inhomogeneous medium (e.g. where the dust-to-gas ratio is inhomogeneous), but is beyond the scope of this work.

#### 3.1 Static shells: degeneracy between $N_{\text{HI, shell}}$ and $T_{\text{shell}}$

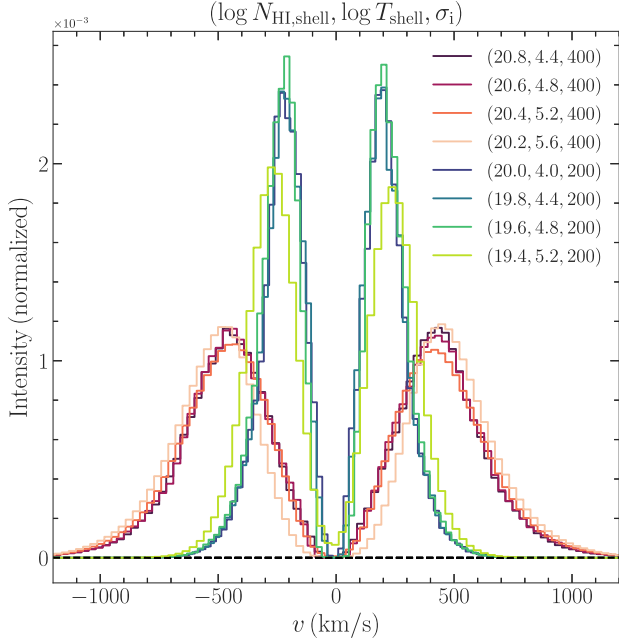
Here, we show that for static shells in the optically thick regime, models with the same  $N_{\text{HI, shell}} T_{\text{shell}}^{0.5}$  exhibit identical Ly $\alpha$  spectra. Theoretically, the angular-averaged Ly $\alpha$  spectral intensity  $J(x)$  emerging from a static, uniform H I sphere is given analytically as (Adams 1972; Harrington 1973; Neufeld 1990; Dijkstra et al. 2006a):

$$J(x) = \frac{\sqrt{\pi}}{\sqrt{24}a\tau_0} \left( \frac{x^2}{1 + \cosh \left[ \sqrt{\frac{2\pi^3}{27}} \frac{|x^3|}{a\tau_0} \right]} \right), \quad (1)$$

where  $x \equiv (\nu - \nu_0)/\Delta\nu_D$  is the unitless frequency, and the Doppler parameter  $\Delta\nu_D = v_{\text{th}}v_0/c = \sqrt{2k_B T/m_{\text{H}}}v_0/c$ , with  $T$  being the H I gas temperature. Here,  $\nu$  is the Ly $\alpha$  photon frequency and  $\nu_0 = 2.47 \times 10^{15} \text{ Hz}$  is the Ly $\alpha$  central frequency. Moreover,  $a = \Delta\nu_L/2\Delta\nu_D \propto T^{-0.5}$  is the Voigt parameter, where  $\Delta\nu_L$  is the natural line broadening;  $\tau_0$  is the H I optical depth at the line centre and  $\tau_0 \propto N_{\text{HI}} T^{-0.5}$ . The complete expressions for  $a$  and  $\tau_0$  can be found, e.g. in Dijkstra et al. (2006a).

One can then switch from the frequency space to the velocity space by converting  $J(x)$  into  $J(v)$  via  $(\nu - \nu_0)/\nu_0 = xv_{\text{th}}/c$ . Then, it is evident that with proper normalization,  $J(v)$  would be identical for different combinations of  $(N_{\text{HI}}, T)$  that give the same  $a\tau_0 v_{\text{th}}^3$ , which is  $\propto N_{\text{HI}} T^{0.5}$ . Alternatively, one can derive this  $N_{\text{HI}} T^{0.5}$  degeneracy by estimating the most likely escape frequency of Ly $\alpha$  photons (see e.g. equation 5 in Gronke et al. 2017, which originally comes from Adams 1972).

We show this degeneracy in Fig. 2 for two sets of static shell models ( $\sigma_i = 200$  and  $400 \text{ km s}^{-1}$ , respectively) with the same  $N_{\text{HI}} T^{0.5}$ . It can be seen that the normalized intensity distributions of each set of models are nearly identical (modulo numerical noise). Note that this degeneracy only exists in the optically thick regime (i.e.  $a\tau_0 \gtrsim 10^3$ ) where equation (1) holds (Harrington 1973; Neufeld 1990). As shown in Fig. 2, models with the same  $N_{\text{HI}} T^{0.5}$  but



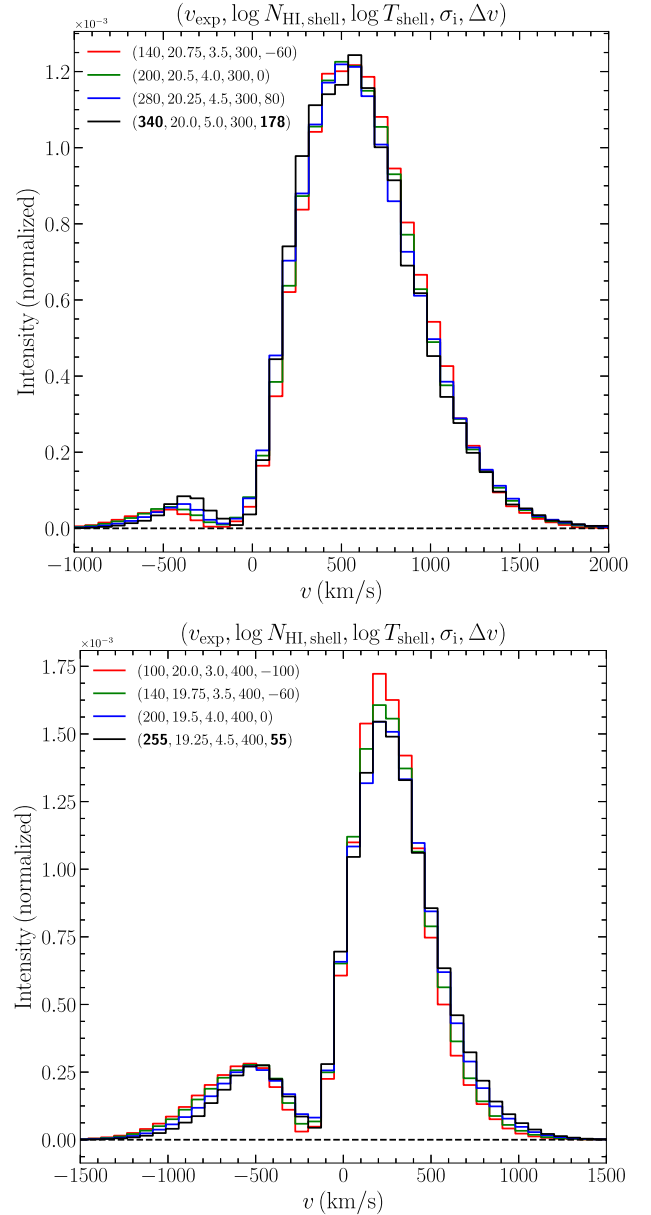
**Figure 2.** Examples of degenerate static shell models with the same  $N_{\text{HI}} T^{0.5}$ . Different coloured curves represent two sets of shell models ( $\sigma_i = 200$  and  $400 \text{ km s}^{-1}$ , respectively) with the same  $N_{\text{HI}} T^{0.5}$ . It can be seen that the normalized intensity distributions of each set of models are nearly identical. Note that this degeneracy only exists in the optically thick regime ( $\tau_0 \gtrsim 10^3$ ); at  $\tau_0 = 2.8 \times 10^{-13} (N_{\text{HI}}/T) \simeq 100$  the models start to deviate from the other degenerate models (the light lime and light pink curves).

$\tau_0 = 2.8 \times 10^{-13} (N_{\text{HI}}/T) \simeq 100$  start to deviate from the other degenerate models, as equation (1) is no longer applicable.

### 3.2 Outflowing shells: degeneracy among $(v_{\text{exp}}, N_{\text{HI, shell}}, T_{\text{shell}}, \Delta v)$

If the shell is outflowing, the  $N_{\text{HI}} T^{0.5}$  degeneracy starts to be broken – in fact, the models with higher  $N_{\text{HI}}$  will have fewer flux in the blue peak, as it is more difficult for the blue photons to escape from the shell. However, such a larger level of asymmetry can be compensated by a lower shell expansion velocity. Heuristically, we find that if we allow the Ly $\alpha$  spectra to shift along the velocity axis (i.e. the systemic velocity of Ly $\alpha$  source is not necessarily at 0; this is often the case for fitting real observed Ly $\alpha$  spectra, where the systemic redshift of the Ly $\alpha$  source has considerable uncertainties), two shell models with  $(v_{\text{exp}}, \log N_{\text{HI, shell}}, \log T_{\text{shell}})$  and  $\sim (2v_{\text{exp}}, \log N_{\text{HI, shell}} - 0.5 \text{ dex}, \log T_{\text{shell}} + 1 \text{ dex}, \Delta v)$  are degenerate with each other, where  $\Delta v$  is the difference in systemic velocity of the two Ly $\alpha$  sources. We have not been able to analytically derive such a quadruple parameter degeneracy rigorously, but we verify its existence numerically in this section.

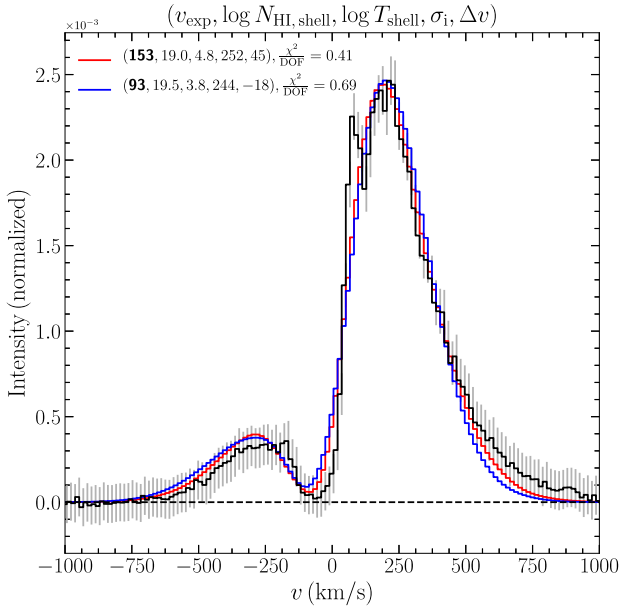
We show this degeneracy with two sets of examples in Fig. 3. Each set contains a series of models with increasing  $T_{\text{shell}}$ , with a step size of 0.5 dex. Accordingly,  $N_{\text{HI, shell}}$  decreases by 0.25 dex and  $v_{\text{exp}}$  increases by a factor of  $\sqrt{2}$ . As can be seen in Fig. 3, each set of spectra appear essentially identical to each other, except for the one with the largest  $T_{\text{shell}}$  (the black curve), which is the only model in the series that does not satisfy  $\tau_0 \gtrsim 10^3$ . We fit this model with our shell model grid by fixing  $N_{\text{HI, shell}}$ ,  $T_{\text{shell}}$ , and  $\sigma_i$  at the expected values and leaving  $v_{\text{exp}}$  and  $\Delta v$  free. It turns out that a decent fit can be achieved, with the best-fitting  $v_{\text{exp}}$  values (shown in bold) slightly



**Figure 3.** Examples of degenerate outflowing shell models. Different coloured curves represent two sets of outflowing shell models ( $\sigma_i = 300$  and  $400 \text{ km s}^{-1}$ , respectively), each consisting of a series of models with increasing  $T_{\text{shell}}$ , with a step size of 0.5 dex. Accordingly,  $N_{\text{HI, shell}}$  decreases by 0.25 dex and  $v_{\text{exp}}$  increases by a factor of  $\sqrt{2}$ . Each set of spectra appear essentially identical to each other, except for the one with the largest  $T_{\text{shell}}$  (the black curve), which is the only model in the series that does not satisfy  $\tau_0 \gtrsim 10^3$ . We fit this model with our shell model grid by fixing  $N_{\text{HI, shell}}$ ,  $T_{\text{shell}}$  and  $\sigma_i$  at the expected values and leaving  $v_{\text{exp}}$  and  $\Delta v$  free. A decent fit is achieved, albeit with the best-fitting  $v_{\text{exp}}$  values (shown in bold) slightly lower than expected.

lower than expected. In other words, this quadruple degeneracy is broken quantitatively but still holds qualitatively.

Such a quadruple degeneracy reminds us of the limitation of shell models in fitting observed Ly $\alpha$  spectra, as  $v_{\text{exp}}$ ,  $N_{\text{HI, shell}}$ ,  $T_{\text{shell}}$ , and  $\Delta v$  cannot be determined independently by merely fitting. Additional constraints (e.g. a very accurate measurement of the systemic redshift of the Ly $\alpha$  emitting source) have to be introduced break the parameter degeneracy.



**Figure 4.** Degeneracy shown in fitting the Ly $\alpha$  spectrum of a Green Pea galaxy, GP 0926+4427. The observed spectrum is shown in black and two degenerate best-fitting shell models are shown in red and blue, respectively. The  $\chi^2$  (per degree of freedom, DOF) values of these two best-fitting models are very close to each other, but the shell expansion velocity of the high-temperature model is about a factor of 2 higher than the low-temperature model (as highlighted in bold), which consequently affects the fitted systemic redshift of the Ly $\alpha$  source. This result may explain the two major discrepancies reported in the literature (see Section 3.2.1 for details).

### 3.2.1 A Real-world example: fitting the Ly $\alpha$ spectrum of a Green Pea galaxy, GP 0926+4427

Here, we further show the quadruple degeneracy with a practical example. We fit an observed Ly $\alpha$  spectrum of a Green Pea galaxy, GP 0926+4427 ( $z = 0.1807$ ; Henry et al. 2015) with our shell model grid. The spectrum is obtained from the Ly $\alpha$  Spectral Database (LASD<sup>4</sup>; Runnholm, Gronke & Hayes 2021). Following Orlitová et al. (2018), we account for the spectral resolution of the *HST* Cosmic Origins Spectrograph (COS) by convolving the shell model spectra with an FWHM = 100 km s<sup>-1</sup> Gaussian before comparing them to the observed Ly $\alpha$  spectrum.<sup>5</sup>

We present two degenerate best-fitting shell models in Fig. 4. These two best-fitting models, whose  $\chi^2$  per degree of freedom are very close to each other, have the parameter degeneracy as described in Section 3.2 – the shell expansion velocity of the high-temperature model is about a factor of 2 higher than the low-temperature model, which consequently affects the fitted systemic redshift of the Ly $\alpha$  source. This result may explain the two major discrepancies reported in Orlitová et al. (2018): (1) the inferred shell outflow velocities are significantly lower than the characteristic outflow velocities indicated by the observed UV absorption lines; (2) the best-fitting systemic redshifts are larger than those derived from optical emission lines. When fitting observed Ly $\alpha$  spectra, the best-fitting model with a low  $v_{\text{exp}}$  may happen to provide the best match for the data, but another degenerate model (or a series of degenerate models) with

much higher  $v_{\text{exp}}$  values can actually fit the data similarly well and hence should also be considered as reasonable solutions.

## 4 RESULTS II: CONNECTING THE SHELL MODEL TO THE MULTIPHASE, CLUMPY MODEL

In this section, we attempt to connect the shell model parameters to the multiphase, clumpy model parameters. We generate a series of clumpy models as our ‘mock data’ for fitting. We first consider a three-dimensional semi-infinite slab geometry (Section 4.1–4.4) and later we will consider a finite spherical geometry (Section 4.5). This is because for a semi-infinite clumpy slab, it is numerically easier to achieve a very high clump covering factor ( $f_{\text{cl}} \gtrsim 1000$ , i.e. the average number of clumps per line of sight is large enough to be in the ‘very clumpy’ regime, where the clumpy medium is expected to behave like a homogeneous medium in terms of the emergent Ly $\alpha$  spectrum, Gronke 2017), which is prohibitively computationally expensive for a finite clumpy sphere. The clumpy slab models are periodic in the  $x$ - and  $y$ -directions with a half-height  $B$  of 50 pc<sup>6</sup> in the  $z$ -direction. The clumps within the slab are spherical with radius of  $r_{\text{cl}} = 10^{-3}$  pc filled with H I of a column density  $N_{\text{H I, cl}}$ . The clump covering factor is directly proportional to the volume-filling factor of the clumps  $F_V$  via  $f_{\text{cl}} = 3F_V B/4r_{\text{cl}}$  (Dijkstra & Kramer 2012; Gronke et al. 2017).

Each clumpy model is calculated via Monte Carlo RT using 10 000 Ly $\alpha$  photon packages assuming a Gaussian intrinsic spectrum  $N(0, \sigma_i^2)$ , where  $\sigma_i = 12.85$  km s<sup>-1</sup> is the canonical thermal velocity dispersion of  $T = 10^4$  K H I gas in the clumps.<sup>7</sup> Each model spectrum is normalized to a total flux of one before being fitted with the shell model grid.

### 4.1 Clumpy slab with static clumps

We start by correlating the H I column density of the shell model with the equivalent average total column density of the single-phase, clumpy slab model, which is given by  $N_{\text{H I, total}} = \frac{4}{3} f_{\text{cl}} N_{\text{H I, cl}}$ , where the factor  $4/3$  comes from the spherical geometry of the clumps (Gronke et al. 2017). We first generate a series of static, single-phase, clumpy slab models by varying  $N_{\text{H I, cl}}$  of the clumps with a very high covering factor  $f_{\text{cl}}$  (i.e. in the ‘very clumpy’ regime) as the mock data. The clumps are fixed to a temperature of  $T_{\text{cl}} = 10^4$  K and do not have any motions (neither random nor outflow velocities). The parameter values that we use are given in the first row of Table 1.

We first attempt to fit the clumpy slab model spectra with the large grid of shell models that we have described in Section 2. We find that the best-fitting shell models are usually noisy and unsatisfactory due to the low number of effective photon packages – i.e. in order to match the relatively narrow widths of the clumpy slab model spectra (especially the ones with low  $N_{\text{H I, total}}$ ), a weighting function with a small  $\sigma_i$  ( $\lesssim 100$  km s<sup>-1</sup>, the actual intrinsic Ly $\alpha$  linewidth needed) is required, which effectively only includes only a small fraction of modelled photons.<sup>8</sup> Therefore, we build a customized grid of shell

<sup>4</sup><http://lasd.lyman-alpha.com>

<sup>5</sup>Note that different from Orlitová et al. (2018), we do not consider the effect of dust, as the dust optical depth is usually a poorly constrained parameter and may introduce additional degeneracy (Gronke et al. 2015).

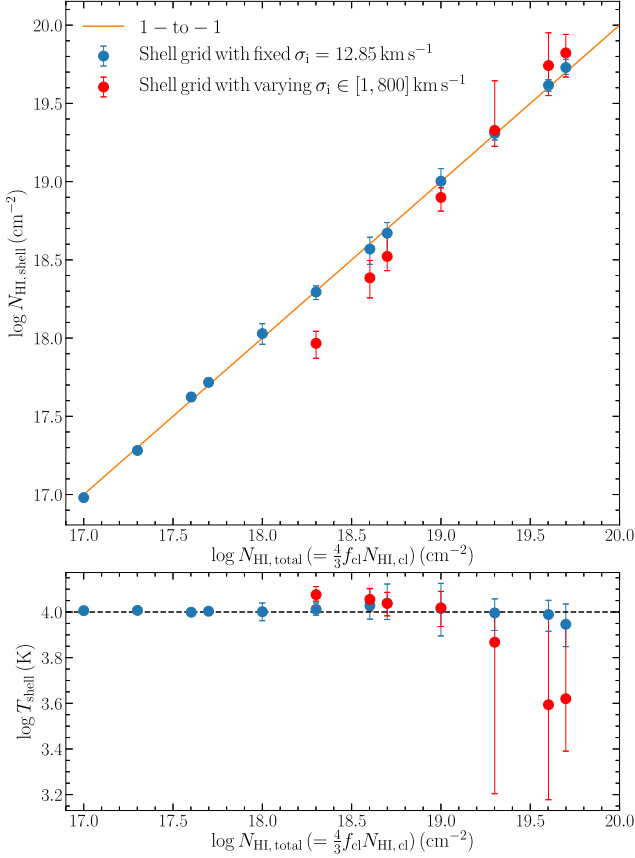
<sup>6</sup>We emphasize that it is the H I column density that actually matters in the RT instead of the physical scales of the models.

<sup>7</sup>In the clumpy model,  $\sigma_i$  is fixed to be small and the clump velocity dispersion is responsible for the broadening of the spectrum.  $\sigma_i$  will not affect Ly $\alpha$  model spectra as long as it is smaller than the clump velocity dispersion (which is almost always the case, see Section 4.2).

<sup>8</sup>This problem is mitigated when the clumps have a considerable random velocity dispersion, which broadens the spectrum significantly (see Section 4.2 and the subsequent sections).

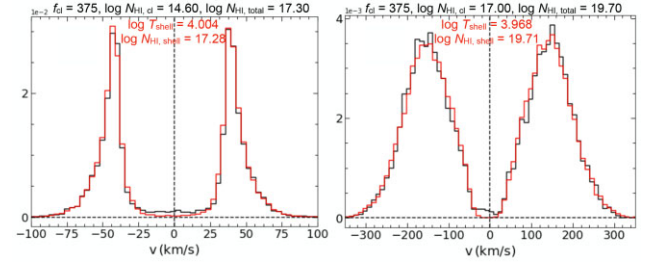
**Table 1.** Parameter values of the clumpy slab models (the mock data).

Model parameter	$F_V$	$f_{cl}$	$\log N_{H\text{I}, cl}$	$\sigma_{cl}$	$v_{cl}$
Definition	Volume-filling factor	Clump covering factor	Clump H I column density	Clump random velocity	Clump outflow velocity
(1)	(2)	(3)	(4)	(5)	(6)
Static clumps	0.1	375	14.3–17.0 $\text{cm}^{-2}$	0	0
Randomly moving clumps	0.02–0.12	750–4500	15.7–17.6 $\text{cm}^{-2}$	(50, 100) $\text{km s}^{-1}$	0
Outflowing clumps	0.08–0.12	3000–4500	15.7–17.0 $\text{cm}^{-2}$	(0, 50, 100) $\text{km s}^{-1}$	50–400 $\text{km s}^{-1}$

**Figure 5.** Results of fitting static clumpy slab models with static shell models. The blue and red points represent the parameter values derived from fitting with the customized, small shell model grid (with fixed  $\sigma_i = 12.85 \text{ km s}^{-1}$ ) and the large shell model grid (with varying  $\sigma_i \in [1, 800] \text{ km s}^{-1}$ ), respectively. Upper panel: The correlation between  $N_{H\text{I}, \text{total}}$  and  $N_{H\text{I}, \text{shell}}$ . A very tight one-to-one correlation is present over three orders of magnitude. Lower panel: The distribution of the best-fitting shell model temperatures  $T_{\text{shell}}$ . In all cases,  $T_{\text{shell}}$  values of  $10^4 \text{ K}$  (the clump temperature) are obtained within  $1\sigma$  uncertainties.

models to fit the clumpy slab model spectra. Such a grid is smaller but similar to the large shell model grid, with two major differences: (1) the shell expansion velocity is fixed to 0; (2) the photon packages are generated from a Gaussian intrinsic spectrum  $N(0, \sigma_i^2)$  with  $\sigma_i = 12.85 \text{ km s}^{-1}$ , i.e. the same as the fitted clumpy slab models. In other words, the intrinsic Ly $\alpha$  spectrum has a fixed small linewidth that is also used to generate the mock data. We find that such a customized grid with only two varying parameters [ $\log N_{H\text{I}, \text{shell}}$ ,  $\log T_{\text{shell}}$ ] can yield better fits (as all the modelled photons contribute to the model spectra) and is significantly faster at fitting the mock data.

As shown in Fig. 5, there is a tight, one-to-one correlation between  $N_{H\text{I}, \text{total}}$  and  $N_{H\text{I}, \text{shell}}$  over three orders of magnitude. Moreover, all

**Figure 6.** Examples of static shell model best fits (obtained by using the shell model grid with fixed  $\sigma_i$ ) to static clumpy slab models. Two panels represent two different [ $f_{cl}$ ,  $N_{H\text{I}, cl}$ ] cases. The black curves represent the static clumpy slab model spectra and the red curves represent the shell model best fits. Both  $T_{\text{shell}}$  and  $N_{H\text{I}, \text{shell}}$  have been obtained at the expected values.

of the  $T_{\text{shell}}$  values are consistent with  $10^4 \text{ K}$  (the clump temperature) within  $1\sigma$  uncertainties. Therefore, we conclude that the equivalent H I column density of a static, very clumpy slab can be exactly reproduced by a shell model with the same H I column density and the same temperature of the clumps. We show two examples of static shell model best fits to static clumpy slab models in Fig. 6.

Despite the shortcomings mentioned above, the large shell model grid is used to fit several static clumpy slab models to verify our results. Several examples<sup>9</sup> are shown in Fig. 5 with red points. We find that  $N_{H\text{I}, \text{shell}}$  and  $T_{\text{shell}}$  can still be roughly obtained at their expected values, albeit with small deviations and larger uncertainties. The required intrinsic linewidths range from  $\sim 50$  to  $\sim 100 \text{ km s}^{-1}$ , depending on the width of the mock data. These intrinsic linewidth values should not have any physical meaning but just ensure that the extent of the wings is proper to yield a good fit.

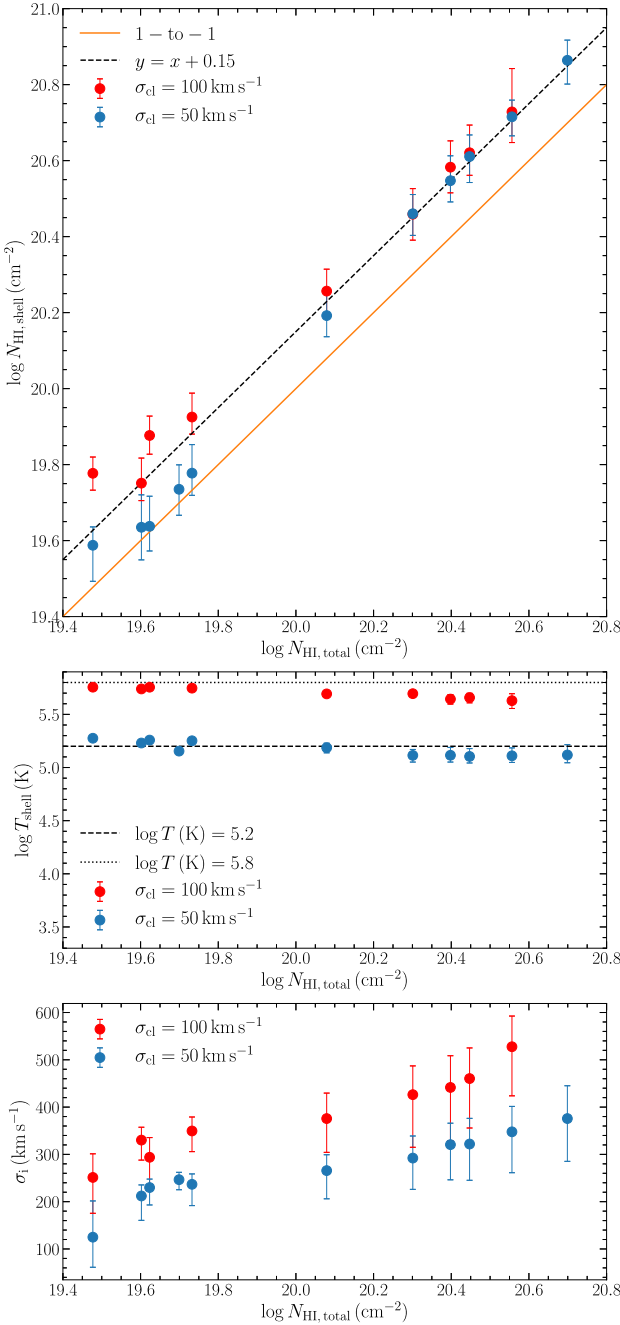
#### 4.2 Clumpy slab with randomly moving clumps

We further add a random velocity dispersion (a Gaussian with a standard deviation of  $\sigma_{cl}$  for all three dimensions) to the clumps and attempt to correlate it with certain shell model parameters, such as the internal random motion (or the effective temperature of the shell,  $T_{\text{shell}}$ ) of the shell model and the linewidth of the intrinsic Ly $\alpha$  emission. We fit the clumpy slab model spectra with the large shell model grid, and the parameter values of the mock data are given in Table 1.

As shown in Fig. 7, we find that:

- (i) The derived  $N_{H\text{I}, \text{shell}}$  values are around the  $N_{H\text{I}, \text{total}}$  values, but a noticeable deviation has emerged. On average,  $N_{H\text{I}, \text{shell}}$  tends to be systemically higher than  $N_{H\text{I}, \text{total}}$  by  $\sim 0.15 \text{ dex}$  (a factor of 1.5), especially at  $N_{H\text{I}, \text{total}} > 10^{20} \text{ cm}^{-2}$ .

<sup>9</sup>These examples have  $N_{H\text{I}, \text{total}}$  high enough to yield  $\sigma_i \gtrsim 50 \text{ km s}^{-1}$ , below which the fraction of photons included is too low to yield a decent fit.

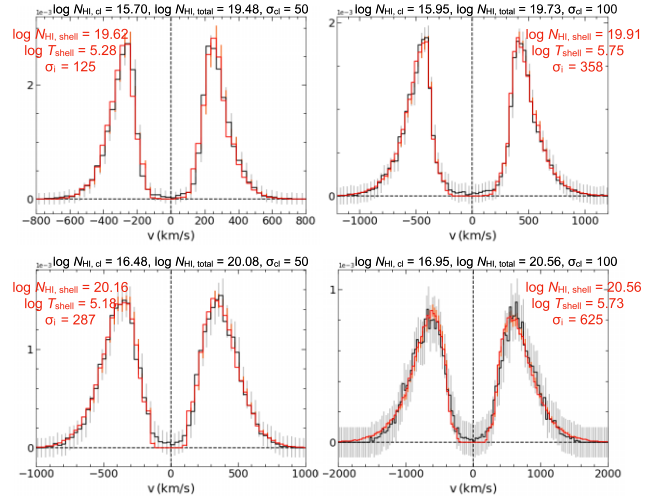


**Figure 7.** Results of fitting clumpy slab models with randomly moving clumps with shell models. Upper panel: The yielded  $N_{\text{HI, shell}}$  values are around the  $N_{\text{HI, total}}$  values, but a noticeable deviation has emerged. Middle panel: The yielded shell temperatures ( $T_{\text{shell}}$ ) are mostly at the effective temperatures of the clumpy slab model. Lower panel: The distribution of the derived linewidths of the intrinsic Ly $\alpha$  emission ( $\sigma_i$ ) of the best-fitting shell models. The blue and red points represent the  $\sigma_{\text{cl}} = 50$  and  $100 \text{ km s}^{-1}$  models, respectively.

(ii) The shell effective temperatures ( $T_{\text{shell}}$ ) are obtained at the effective temperatures of the clumpy slab model, defined as

$$T_{\text{eff, slab}} = T_{\text{cl}} + \frac{\sigma_{\text{cl}}^2 m_{\text{H}}}{2k_{\text{B}}}, \quad (2)$$

where  $T_{\text{cl}}$  is the kinematic temperature of one clump (fixed to  $10^4 \text{ K}$ ),  $m_{\text{H}}$  is the hydrogen atom mass and  $k_{\text{B}}$  is the Boltzmann constant. As



**Figure 8.** Examples of shell model best fits to randomly moving clumpy slab models. Four panels represent four different  $(N_{\text{HI, total}}, \sigma_{\text{cl}})$  cases. The black curves represent the outflowing clumpy slab model spectra and the red curves represent the shell model best fits.  $T_{\text{shell}}$  have been obtained at the expected values from equation (2) within uncertainties.

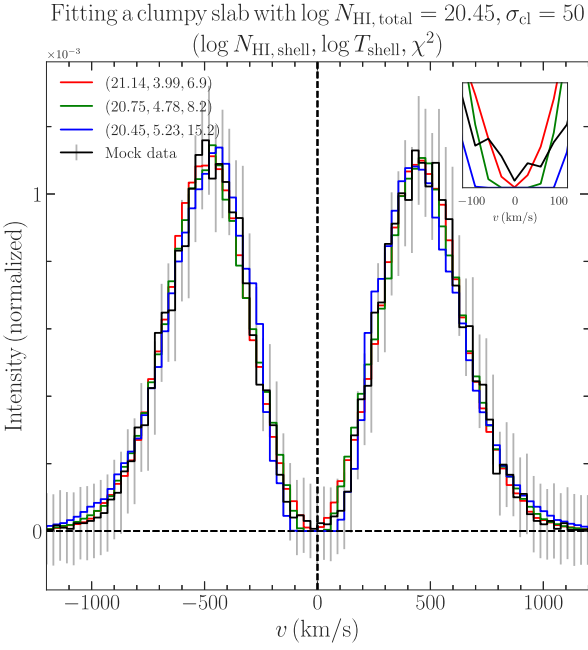
the maximum  $T_{\text{shell}}$  of our large shell model grid is set to be  $10^{5.8} \text{ K}$ , we only explore  $\sigma_{\text{cl}}$  up to  $\sim 100 \text{ km s}^{-1}$ , but we have verified that a larger  $\sigma_{\text{cl}}$  would still correspond to a  $T_{\text{shell}}$  value given by equation (2).

(iii) Large  $\sigma_i$  values (several times of  $\sigma_{\text{cl}}$ ) are required to reproduce the wings of the clumpy slab models. These  $\sigma_i$  values are also positively correlated with  $\sigma_{\text{cl}}$  and  $N_{\text{HI, total}}$ , as shown in the bottom panel of Fig. 7.

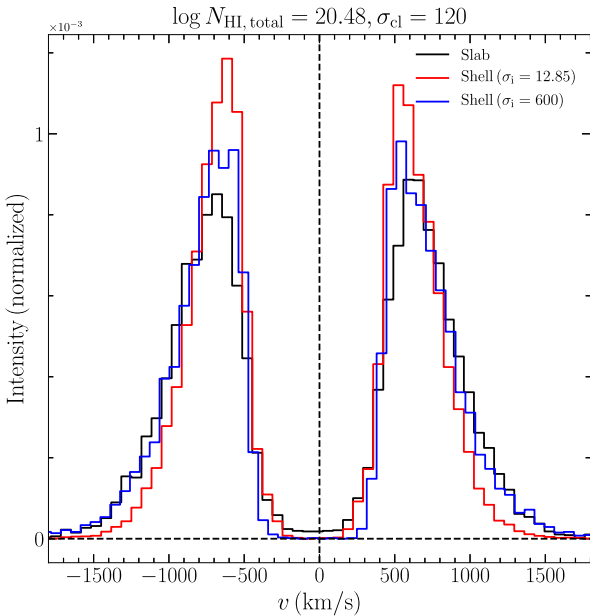
We show four examples of shell model best fits to the clumpy slab models in Fig. 8. We find that (i) is due to the  $N_{\text{HI, shell}} \propto T_{\text{shell}}^{-0.5}$  degeneracy. As we have detailed in Section 3, in the optically thick regime where  $\alpha\tau_0 = 2.8 \times 10^{-13} (N_{\text{HI, shell}}/T_{\text{shell}})^{0.5} \gtrsim 10^3$ , shell models with the same  $N_{\text{HI, shell}} T_{\text{shell}}^{0.5}$  have almost identical line profiles, except that the ones with higher  $T_{\text{shell}}$  have slightly more extended troughs at the line centre.<sup>10</sup> This explains the deviation of  $N_{\text{HI, shell}}$  towards higher values at high H I column densities, where the trough of the clumpy slab model becomes ‘sharper’ as the flux density at line centre approaches 0, and is better fitted at a slightly lower  $T_{\text{shell}}$  (and hence higher  $N_{\text{HI, shell}}$ ). We illustrate this effect in Fig. 9. In other words,  $N_{\text{HI, shell}}$  is still consistent with  $N_{\text{HI, total}}$  if we account for this  $N_{\text{HI, shell}} \propto T_{\text{shell}}^{-0.5}$  degeneracy.

Moreover, (iii) is due to the intrinsic differences between the clumpy slab model and the shell model. As shown in Fig. 10 (cf. fig. 5 in Gronke et al. 2017), for  $\sigma_{\text{cl}} > 0$ , the clumpy slab model (black curve) tends to have lower peaks and larger fluxes near the line centre, as compared to the corresponding homogeneous shell model (red curve). Therefore, in order to obtain a good fit, a large  $\sigma_i$  is required to flatten the peaks and spread the fluxes out into the wings (blue curve). As  $\sigma_{\text{cl}}$  or  $N_{\text{HI, total}}$  increases, the difference between the peak fluxes of two different models becomes larger, which requires a larger  $\sigma_i$ . If we force  $\sigma_i$  to be small, the shell models would fail to fit the clumpy slab model, as a much higher H I column density than  $N_{\text{HI, total}}$  is required to fit the broad wings and it will inevitably yield a significant mismatch in the peaks.

<sup>10</sup>This is because the cross-section function of a higher  $T_{\text{shell}}$  is more extended near the line centre (see equation 54 and 55 in Dijkstra 2017).



**Figure 9.** Model examples showing the degeneracy between  $N_{\text{HI, shell}}$  and  $T_{\text{shell}}$ . The curves with different colours represent the clumpy slab model fitted (black) and the degenerate shell models with different  $(\log N_{\text{HI, shell}}, \log T_{\text{shell}}, \chi^2)$ , obtained by fitting within a certain parameter subspace. As shown in the inset, the models with higher  $T_{\text{shell}}$  have more extended troughs at line centre and thus are less favoured in the fitting to the clumpy slab models (the mock data), which have sharper troughs at high  $N_{\text{HI, total}}$ .



**Figure 10.** Comparison between a clumpy slab model and two shell models with corresponding parameters and  $\sigma_i = 12.85/600 \text{ km s}^{-1}$ . The black curve is an example clumpy slab model with  $N_{\text{HI, total}}$  and  $\sigma_{\text{cl}}$  values labelled on the top of the plot; the red and blue curves are the shell models at the expected  $N_{\text{HI, total}}$  and  $T_{\text{shell}}$  values with  $\sigma_i = 12.85$  and  $600 \text{ km s}^{-1}$ , respectively. It is clear that the clumpy slab model tends to have lower peaks and larger fluxes near the line centre than the corresponding homogeneous shell model with a small  $\sigma_i$ ; such a mismatch is mitigated by the broadening effect of a large  $\sigma_i$ .

As large  $\sigma_i$  values have been shown to be inconsistent with the observed nebular emission linewidths (e.g. H $\alpha$  or H $\beta$ , Orlitová et al. 2018), it is reasonable to postulate that the clumpy model is a more realistic description of the actual gas distribution in ISM/CGM, as it naturally alleviates such discrepancies with moderate velocity dispersions of the clumps (see also Li et al. 2021a). We will further discuss this point in Section 5.1.

#### 4.3 Clumpy slab with outflowing clumps

We further attempt to add a uniform outflow velocity ( $v_{\text{cl}}$ ) to the clumps and correlate it with the shell expansion velocity ( $v_{\text{exp}}$ ). We consider two different cases: (1)  $v_{\text{cl}} > 0$ ,  $\sigma_{\text{cl}} = 0$ ; (2)  $v_{\text{cl}} > 0$ ,  $\sigma_{\text{cl}} > 0$ , i.e. outflowing clumps without and with clump random motion, respectively. We find that in both cases, a considerably large  $\sigma_i$  is still required to achieve decent fits, otherwise the shell model best fit would have a dip between two peaks on the red side, whereas the clumpy slab model has only one smooth red peak.

As shown in Fig. 11, we find that:

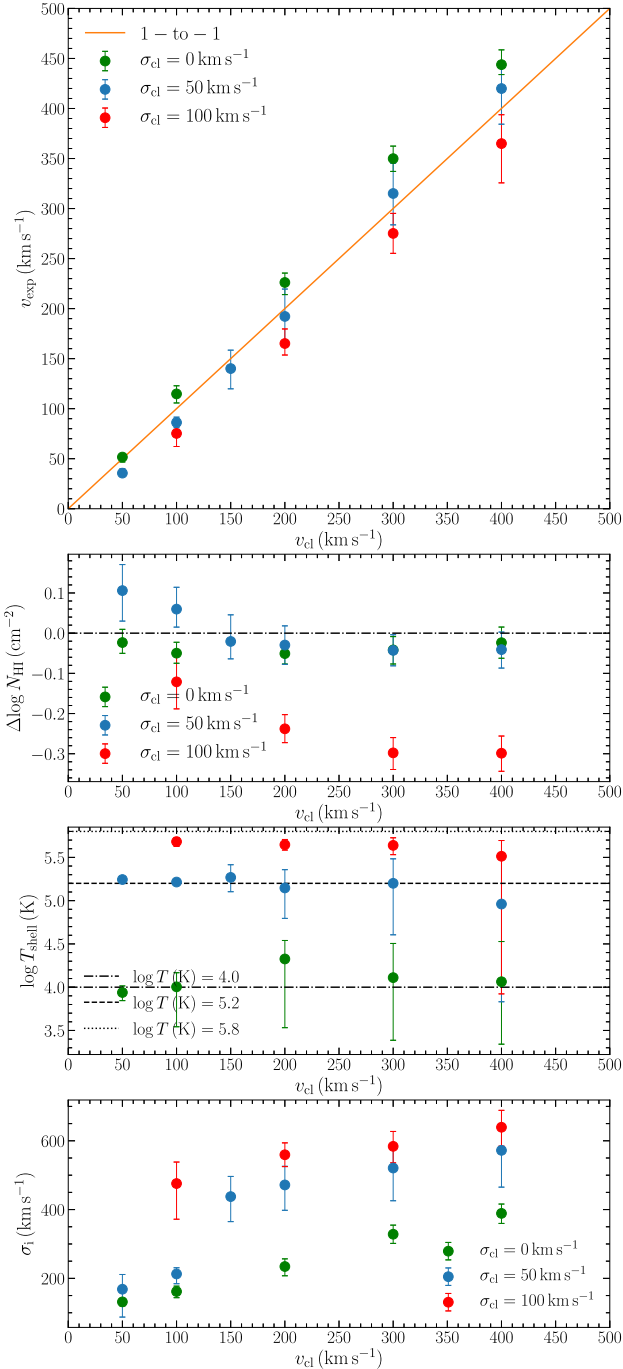
- (i) The fitted  $v_{\text{exp}}$  values are mostly consistent with  $v_{\text{cl}}$  within uncertainties.
- (ii)  $N_{\text{HI, shell}}$  are mostly reproduced at  $N_{\text{HI, total}}$ , albeit with several outliers with  $\Delta \log N_{\text{HI}} \gtrsim 0.3$  dex in the large  $\sigma_{\text{cl}}$  cases.
- (iii)  $T_{\text{shell}}$  are mostly reproduced at  $T_{\text{eff, slab}}$  within uncertainties.
- (iv) Large  $\sigma_i$  values are still required and they increase as  $\sigma_{\text{cl}}$  or  $v_{\text{cl}}$  increases.

We show four examples of outflowing shell model best fits to outflowing clumpy slab models in Fig. 12 to illustrate the quality of the fits.

In addition to adding a uniform clump outflow velocity, we have also experimented with two more sophisticated velocity profiles. First, we add a Hubble flow-like outflow velocity increasing linearly from 0 (at the centre of the simulation region) to  $v_{\text{max}}$  (at the boundary of the simulation region) to the clumps and fit the Ly $\alpha$  model spectra. The results are shown in Fig. 13. The four panels in the first two rows represent four models with linearly increasing outflow velocities, as compared to the lower four panels with constant outflow velocities. It can be seen from the line profiles that for the models with linear increasing outflow velocities: (1) the blue-to-red peak flux ratio is much higher than that of the corresponding uniform outflow model (either with  $v_{\text{cl}} = \frac{1}{2}v_{\text{max}}$  or  $v_{\text{max}}$ ), which is not well captured by the best-fitting shell model, especially at high outflow velocities; (2) the peak separation is also larger than that of the corresponding uniform outflow model. As a result, the fitted  $T_{\text{shell}}$  and  $\sigma_i$  values have been boosted due to (2), but  $N_{\text{HI, shell}} \simeq N_{\text{HI, total}}$  remains true.  $v_{\text{exp}}$  is roughly obtained at  $\frac{1}{2}v_{\text{max}}$ , but should be considered as an upper limit as it actually yields a best-fitting spectrum with a blue-to-red peak flux ratio lower than that of the fitted clumpy model.

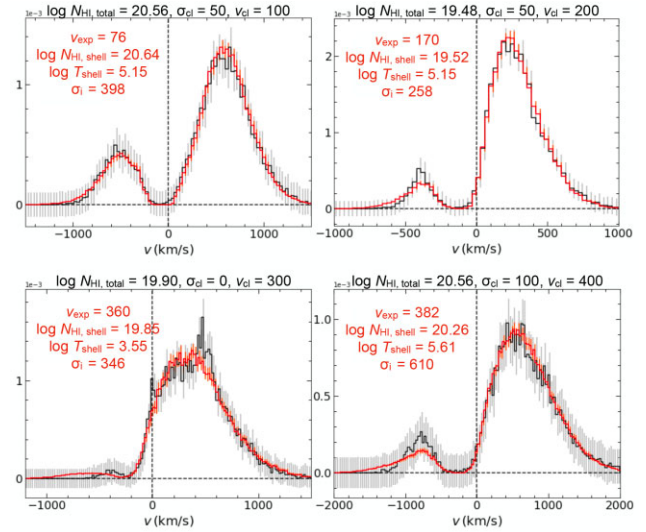
Secondly, we consider a scenario where the clumps are accelerated in a momentum-driven manner and in the meantime, decelerated by a gravitational force. This is motivated by the fact that in real galactic environments, the cool clouds can be accelerated by radiation pressure or ram pressure of the hot wind as they break out of the ISM, as they are decelerating within the gravitational well of the dark matter halo. The momentum equation of a clump can be then written as (Murray, Quataert & Thompson 2005; Dijkstra & Kramer 2012):

$$\frac{dv(r)}{dt} = -\frac{GM(r)}{r^2} + Ar^{-\alpha}, \quad (3)$$



**Figure 11.** Results of fitting clumpy slab models with outflowing moving clumps with shell models. Upper panel: The derived shell expansion velocities are mostly at the clump outflow velocities; Upper middle panel:  $N_{\text{HI, shell}}$  are reproduced mostly at  $N_{\text{HI, total}}$ , albeit with several outliers with  $\Delta \log N_{\text{HI}} \gtrsim 0.3$  dex. Lower middle panel: The derived shell temperatures are mostly at the effective temperatures of the clumpy slab model; Lower panel: The distribution of the derived intrinsic  $\text{Ly}\alpha$  linewidths ( $\sigma_i$ ) of the best-fitting shell models, which increase as  $\sigma_{\text{cl}}$  or  $v_{\text{cl}}$  increases. The green, blue, and red points represent the  $\sigma_{\text{cl}} = 0, 50$ , and  $100 \text{ km s}^{-1}$  models, respectively.

where  $r$  is the clump radial position,  $v(r)$  is the clump radial outflow velocity, and  $M(r)$  is the total mass within  $r$ . Here, the clump acceleration is determined by two competing terms on the right-hand side, the first of which is due to gravitational deceleration and the second of which is an empirical power-law acceleration term



**Figure 12.** Examples of outflowing shell model best fits to outflowing clumpy slab models. Four panels represent four different  $(N_{\text{HI, total}}, \sigma_{\text{cl}}, v_{\text{cl}})$  cases. The black curves represent the outflowing clumpy slab model spectra and the red curves represent the shell model best fits.  $v_{\text{exp}}$  have been obtained at the expected values within uncertainties.

(Steidel et al. 2010). Assuming the gravitational potential is of an isothermal sphere, then  $M(r) = 2\sigma_{\text{cl}}^2 r/G$ , where  $\sigma_{\text{cl}}$  is the velocity dispersion of the clumps. equation (3) can then be analytically solved as

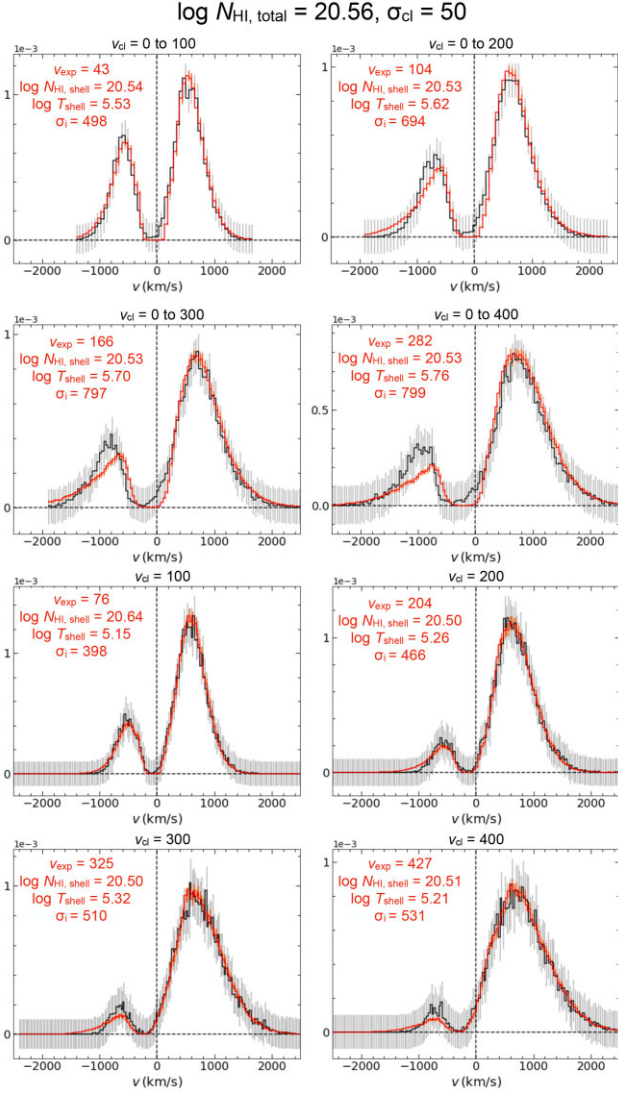
$$v(r) = \sqrt{4\sigma_{\text{cl}}^2 \ln\left(\frac{r_{\text{min}}}{r}\right) + v_{\text{cl}, \infty}^2 \left(1 - \left(\frac{r}{r_{\text{min}}}\right)^{1-\alpha}\right)}, \quad (4)$$

where  $r_{\text{min}}$  is the inner cutoff radius that satisfies  $v(r_{\text{min}}) = 0$ , and  $v_{\text{cl}, \infty} = \sqrt{2Ar_{\text{min}}^{1-\alpha}/(\alpha-1)}$  is the asymptotic maximum outflow velocity if there were no gravitational deceleration. Following Dijkstra & Kramer (2012), we have fixed  $r_{\text{min}} = 1 \text{ pc}$  (note the clumpy slab model has a half height of 50 pc and the model is re-scalable by design) and  $\alpha = 1.4^{11}$  and left  $\sigma_{\text{cl}}$  and  $v_{\text{cl}, \infty}$  as free parameters.

We show five examples of  $v(r)$  with different  $(\sigma_{\text{cl}}, v_{\text{cl}, \infty})$  in Fig. 14. It can be seen that the acceleration decreases with radius, and the velocity either flattens or drops at large  $r$ , depending on the actual values of  $\sigma_{\text{cl}}$  and  $v_{\text{cl}, \infty}$ . We fix  $\sigma_{\text{cl}} = 50 \text{ km s}^{-1}$  and adjust  $v_{\text{cl}, \infty}$  to achieve average radial velocities of  $\bar{v}_{\text{cl}}(r) = 60, 100, 200, 300$ , and  $400 \text{ km s}^{-1}$ . We then assign these radial velocity profiles to the clumps and fit their model spectra with shell models. We find that  $N_{\text{HI, shell}}$  and  $T_{\text{shell}}$  are reproduced at the expected values, and  $v_{\text{exp}}$  is roughly obtained at the average radial velocity,  $\bar{v}_{\text{cl}}(r)$ .

We show that the derived  $v_{\text{exp}}$  is consistent with the average radial velocity  $\bar{v}_{\text{cl}}(r)$  for these two scenarios, as shown in Fig. 15. In reality, if the velocity distribution of the  $\text{Ly}\alpha$  scattering clumps is semilinear or similar to the ‘momentum-driven + gravitational deceleration’ scenario, the shell model fitting will probe the average outflow velocity of the clumps. Assuming the same clumps are responsible for producing metal absorption (and ignoring effects due to an anisotropic gas distribution), the clump velocity distribution can be constrained by observations on UV absorption lines. One should then expect the outflow velocity  $v_{\text{exp}}$  output from shell model fitting

<sup>11</sup>These choices come from the clump radial velocity models that provide good fits to the observed  $\text{Ly}\alpha$  surface brightness profiles (Dijkstra & Kramer 2012).

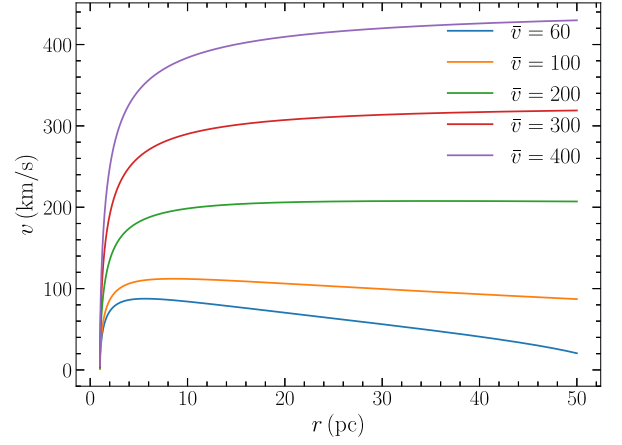


**Figure 13.** Examples of clumpy slab models with linearly increasing outflow velocities, compared to those with constant outflow velocities. The upper two rows are four models with linearly increasing outflow velocities, whereas the lower two rows are four models with constant outflow velocities. For the models with linearly increasing outflow velocities, the blue-to-red peak flux ratio is much higher. The peak separation is also larger, which boosts the fitted  $T_{\text{shell}}$  and  $\sigma_i$  values, but  $N_{\text{HI, shell}} \simeq N_{\text{HI, total}}$  remains true.  $v_{\text{exp}}$  is roughly obtained at  $\frac{1}{2} v_{\text{max}}$ .

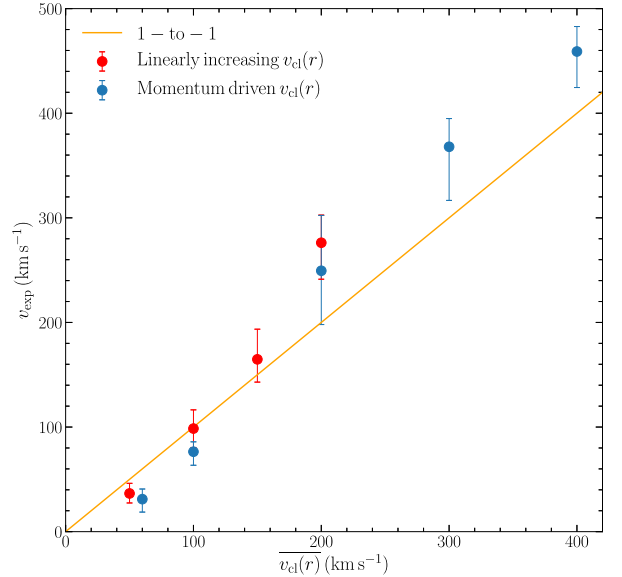
to be consistent with the average of the absorption velocities. If the clump velocity distribution is non-linear,  $v_{\text{exp}}$  may no longer be an average outflow velocity, but should still lie between the minimum and maximum absorption velocities.

#### 4.4 Clumpy slab with an ICM

Motivated by the fact that in real astrophysical environments (e.g. in the CGM; Tumlinson et al. 2017), a hot, highly ionized gas phase with residual H I exists and affects Ly $\alpha$  RT (Laursen et al. 2013), we further add another hot phase of gas between the clumps to the clumpy slab model as the ICM. Although the total column density of the low-density ICM ( $\sim n_{\text{HI, ICM}} r_{\text{gal}} \lesssim 10^{-4} \text{ cm}^{-3} \times \text{kpc} \sim 10^{17} \text{ cm}^{-2}$ ) is supposed to be several orders of magnitude lower than the typical values of  $N_{\text{HI, total}}$  from the cool clumps, it has several non-negligible



**Figure 14.** Examples of momentum-driven radial velocity profiles given by equation (4). Different coloured curves represent five  $v(r)$  profiles with different  $(\sigma_{\text{cl}}, v_{\text{cl}, \infty})$ . Here,  $\sigma_{\text{cl}}$  is fixed to  $50 \text{ km s}^{-1}$  and  $v_{\text{cl}, \infty}$  are adjusted to make the average radial velocity,  $\overline{v(r)} = 60, 100, 200, 300$ , and  $400 \text{ km s}^{-1}$ . The acceleration decreases with radius, and the velocity either flattens or drops at large  $r$ , depending on the actual values of  $\sigma_{\text{cl}}$  and  $v_{\text{cl}, \infty}$ .



**Figure 15.** Relation between the average radial velocity  $\overline{v_{\text{cl}}(r)}$  of clumpy slab models and the derived  $v_{\text{exp}}$  from shell model fitting. The red points represent the linearly increasing scenario, and the blue points represent the momentum-driven scenario. For both scenarios,  $v_{\text{cl}}(r)$  and  $v_{\text{exp}}$  are basically consistent, suggesting that shell model fitting probes the average radial velocity of the clumps.

effects on the Ly $\alpha$  spectra. We find that adding another hot phase of gas at  $10^6 \text{ K}$  (the typical temperature of diffuse gas in a dark matter halo) will (1) deepen the trough at line centre; (2) increase the peak separation; and (3) modify the blue-to-red peak flux ratio of the model Ly $\alpha$  spectrum.<sup>12</sup>

Here, we consider two different scenarios: static ICM and outflowing ICM. For both scenarios, we generate two sets of multi-phase, clumpy slab models with  $n_{\text{HI, ICM}} = 10^{-4}$  and  $10^{-3} \text{ cm}^{-3}$  (or

<sup>12</sup>Regarding the effects of ICM with different temperatures and column densities, we refer the readers to Appendices A and B.

equivalently,  $N_{\text{HI, ICM}} = 10^{16.2}$  and  $10^{17.2} \text{ cm}^{-2}$ ) and fit them using the large shell model grid. The values of the input and output parameters are shown in Fig. 16. We hereby discuss two scenarios, respectively.

#### 4.4.1 Static hot ICM: $T_{\text{ICM}} = 10^6 \text{ K}$ , $v_{\text{ICM}} = 0$

We find that adding a static hot ICM increases the peak separation and increases the blue-to-red peak flux ratio of the model Ly $\alpha$  spectrum. These two effects can be seen by comparing the first and second rows of Fig. 17. In terms of the shell model best-fitting parameters, the former effect increases  $T_{\text{shell}}$  but does not boost  $N_{\text{HI, total}}$  significantly, and the latter effect decreases  $v_{\text{exp}} \ll v_{\text{cl}}$ . These effects are shown in Fig. 16 by green circles and open squares (which correspond to two different ICM H I column densities).

#### 4.4.2 Outflowing hot ICM: $T_{\text{ICM}} = 10^6 \text{ K}$ , $v_{\text{ICM}} > 0$

As shown in the third row of Fig. 17, adding an outflow velocity to the ICM will decrease the blue-to-red peak flux ratio. The first two panels have  $v_{\text{cl}} = 0$  and  $v_{\text{ICM}} > 0$ , whereas the third panel has  $v_{\text{cl}} = v_{\text{ICM}} > 0$ . Notably, the quality of the best fits has become worse, suggesting the non-linear effect of a hot ICM on the Ly $\alpha$  model spectra. The  $T_{\text{shell}}$  and  $N_{\text{HI, total}}$  values are similar to the static ICM case. Interestingly, in the first two panels, where  $v_{\text{cl}} = 0$  and  $v_{\text{ICM}} > 0$ , we have  $v_{\text{exp}} \simeq \frac{1}{2} v_{\text{ICM}}$ ; whereas in the third panel where  $v_{\text{cl}} = v_{\text{ICM}} > 0$ , the blue-to-red peak flux ratio becomes similar to the no-ICM case (the third panel in the first row), and  $v_{\text{exp}} \simeq v_{\text{cl}} = v_{\text{ICM}}$  is obtained.<sup>13</sup>

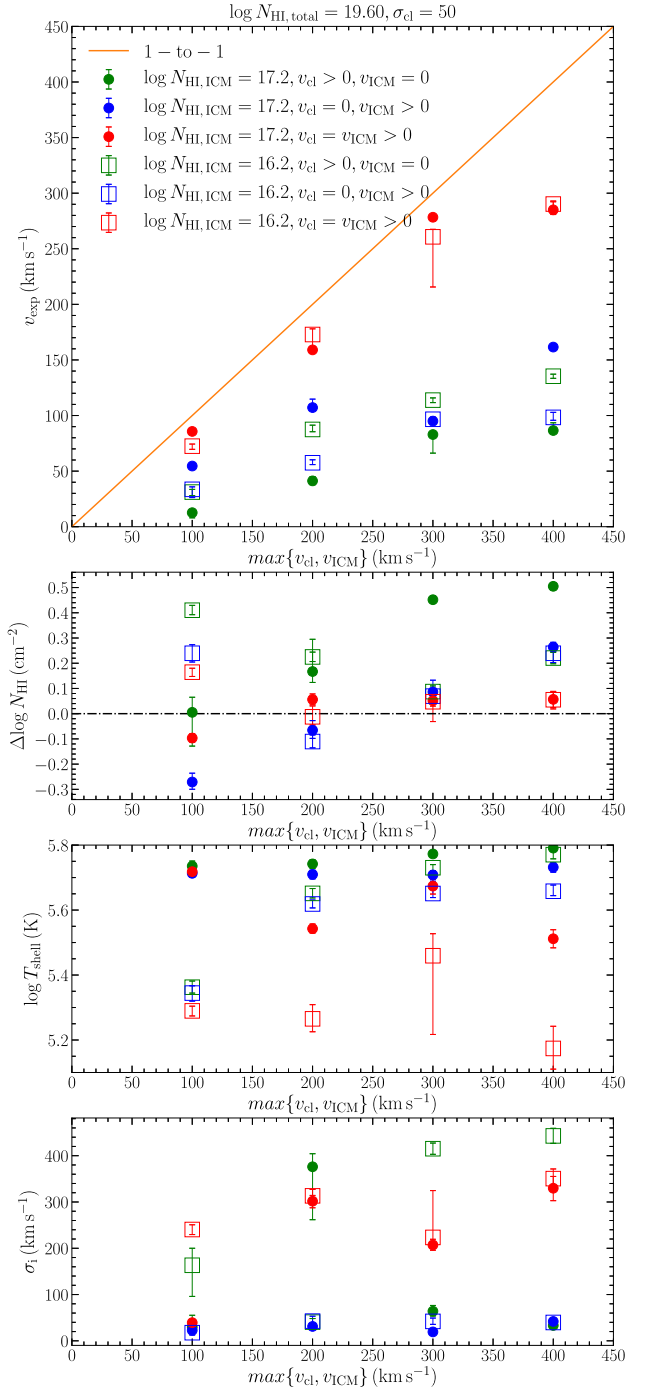
It is therefore evident that if a multiphase, clumpy slab model with  $(v_{\text{cl}}, v_{\text{ICM}})$  and a shell model with  $v_{\text{exp}}$  give the same Ly $\alpha$  spectrum (especially the same blue-to-red peak flux ratio), then  $v_{\text{exp}}$  should lie between  $v_{\text{cl}}$  and  $v_{\text{ICM}}$ . In particular, if  $v_{\text{cl}} = v_{\text{ICM}}$ , i.e. the cool clumps and the hot ICM are co-outflowing at the same speed, we would expect  $v_{\text{exp}} = v_{\text{cl}} = v_{\text{ICM}}$ . In reality, we expect the cool clumps to be entrained by the local flow of hot gas (i.e.  $v_{\text{cl}} \simeq v_{\text{ICM}}$ ), as a large velocity difference between two phases of gas may destroy the cool clumps quickly via hydrodynamic instabilities<sup>14</sup> (see e.g. Klein, McKee & Colella 1994). Therefore, we conclude that the gas outflow velocities extracted from fitting Ly $\alpha$  spectra should be consistent between the shell model and the multiphase, clumpy model.

#### 4.5 Multiphase clumpy sphere

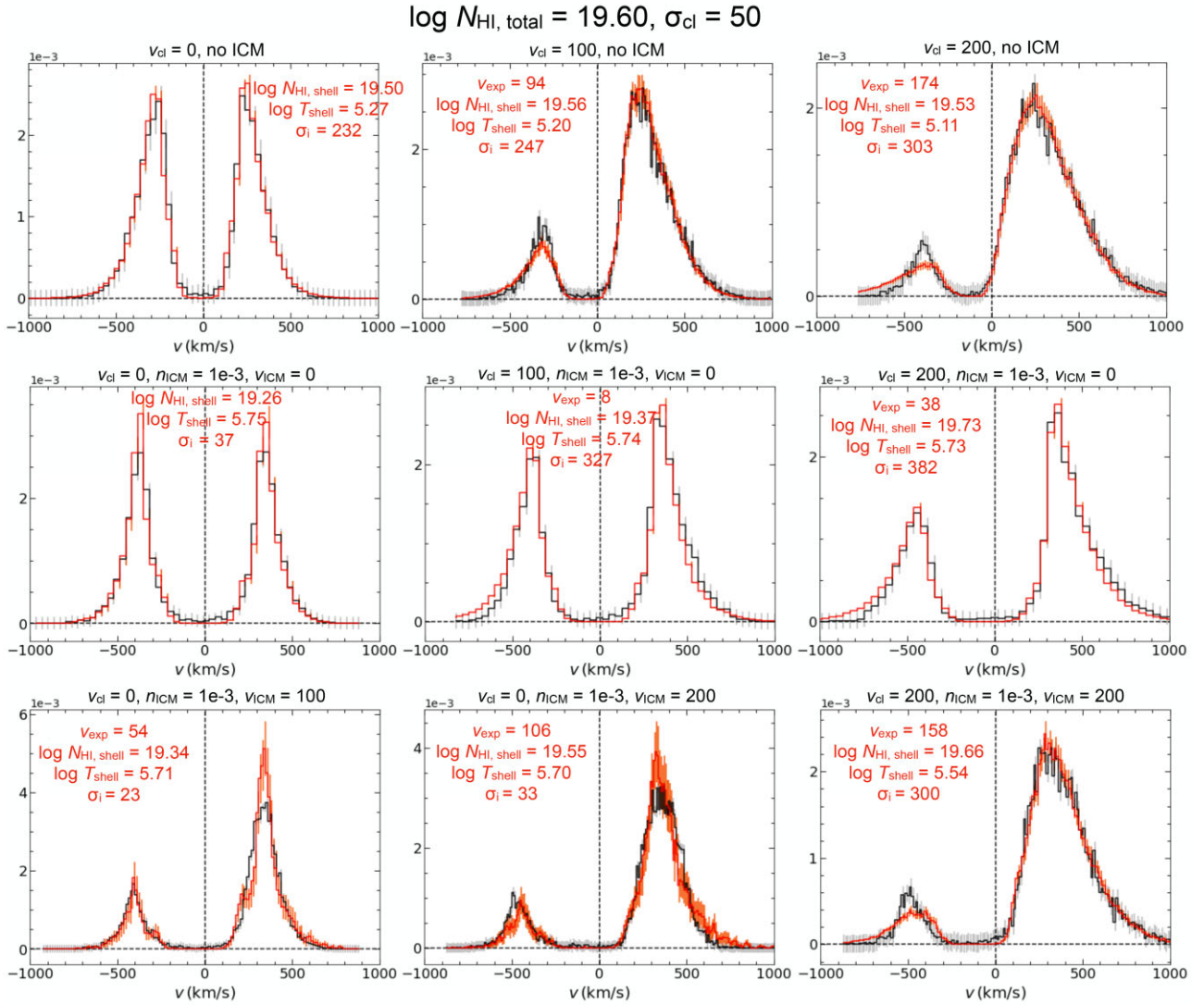
We further consider a more physically realistic gas geometric distribution, i.e. a multiphase clumpy sphere, which we have adopted in fitting observed spatially resolved Ly $\alpha$  spectra in Li et al. (2021b) and Li et al. (2021a). As a multiphase sphere model has an upper limit for the clump volume-filling factor  $F_{\text{V}} (\lesssim 0.7 \text{ for numerical reasons})$  and hence for the covering factor  $f_{\text{cl}} (= 3r_{\text{gal}}/4r_{\text{cl}} F_{\text{V}} = 150 F_{\text{V}})$ , it cannot have an as high covering factor as the multiphase clumpy slab model (i.e. ‘less clumpy’). However, as long as  $f_{\text{cl}}$  is much larger than a critical value  $f_{\text{cl, crit}}$ , the clumpy model would be sufficiently similar to a homogeneous model (Gronke et al. 2017). If the condition  $f_{\text{cl}} \gg f_{\text{cl, crit}}$  is not satisfied, a considerable number of Ly $\alpha$  photons would escape near the line centre and yields residual fluxes at line centre in the emergent spectrum. Here, we explore the connection between

<sup>13</sup>Here, ‘ $\simeq$ ’ refers to a difference within a factor of 4.

<sup>14</sup>Note that in a relatively rare scenario (e.g. very close to the launching radius of a galactic wind), the hot phase may be moving faster than the cool clumps, i.e.  $v_{\text{ICM}} > v_{\text{cl}}$ , and the shell model fitting would obtain  $v_{\text{ICM}} > v_{\text{exp}} > v_{\text{cl}}$ .



**Figure 16.** Results of fitting clumpy slab models with static and outflowing ICM with shell models. Upper panel: The derived shell expansion velocities are much smaller than the clump/ICM outflow velocities unless both components are co-outflowing, which yields  $v_{\text{exp}} \simeq v_{\text{cl}} = v_{\text{ICM}}$ ; Upper middle panel:  $N_{\text{HI, shell}}$  are mostly at  $N_{\text{HI, total}}$ , albeit with several outliers with  $\Delta \log N_{\text{HI}} \gtrsim 0.3$  dex; Lower middle panel: The derived shell temperatures are boosted by the hot ICM to be higher than the effective temperatures of the clumpy slab model; Lower panel: The distribution of the derived intrinsic Ly $\alpha$  linewidths ( $\sigma_i$ ) of the best-fitting shell models, which increase as  $\sigma_{\text{cl}}$  or  $v_{\text{cl}}$  increases. The green, blue, and red points represent (1)  $v_{\text{cl}} > 0$ ,  $v_{\text{ICM}} = 0$  (2)  $v_{\text{cl}} = 0$ ,  $v_{\text{ICM}} > 0$  (3)  $v_{\text{cl}} = v_{\text{ICM}} > 0$  models, respectively.



**Figure 17.** Examples of shell model best fits to outflowing clumpy slab models with and without ICM. The first row represents three  $v_{\text{cl}} > 0$  cases without ICM. The second row represents three  $v_{\text{cl}} > 0$  cases with a static,  $T = 10^6$  K,  $n_{\text{HI}} = 10^{-3} \text{ cm}^{-3}$  ICM. Adding this hot phase of static ICM tends to (1) deepen the trough at line centre; (2) increase the peak separation; and (3) increase the blue-to-red peak flux ratio. The third row represents three  $v_{\text{cl}} > 0$  cases with an outflowing ICM. A  $v_{\text{ICM}} > 0$  ICM will further decrease the blue-to-red peak flux ratio and increase the  $v_{\text{exp}}$  of the shell model best fit. In particular, the model with the same clump and ICM outflow velocity prefers a shell expansion velocity  $v_{\text{exp}} \simeq v_{\text{cl}} \simeq v_{\text{ICM}}$ .

multiphase, clumpy spherical models and shell models in these two physical regimes, respectively:  $f_{\text{cl}} \gg f_{\text{cl, crit}}$  and  $f_{\text{cl}} \simeq f_{\text{cl, crit}}$ .

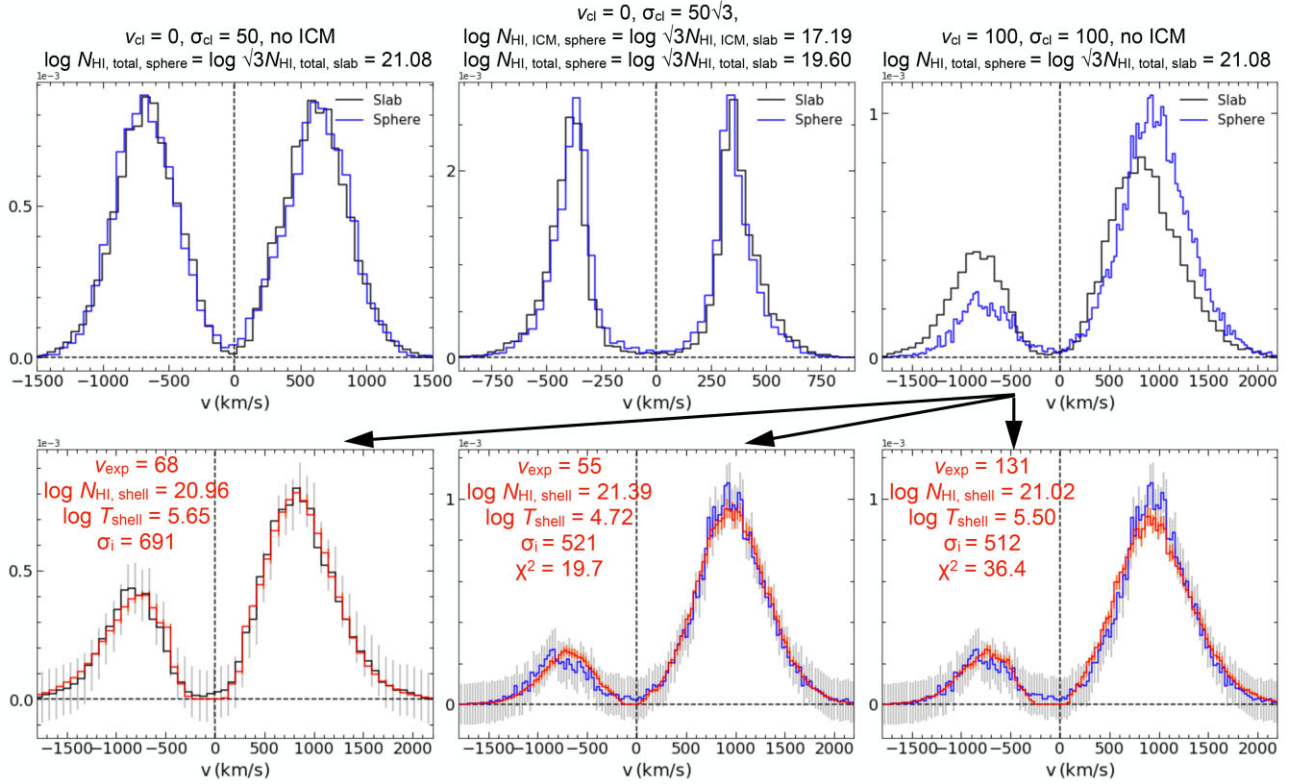
#### 4.5.1 Very clumpy sphere: $f_{\text{cl}} \gg f_{\text{cl, crit}}$

For a very clumpy spherical model, i.e.  $f_{\text{cl}} \gg f_{\text{cl, crit}}$ , we find that a non-outflowing ( $v_{\text{cl}} = v_{\text{ICM}} = 0$ ) spherical model with  $(\sigma_{\text{cl}}, \sqrt{3}N_{\text{HI, total}}, \sqrt{3}N_{\text{HI, ICM}})$  gives an identical spectrum to a slab model with  $(\sigma_{\text{cl}}, N_{\text{HI, total}}, N_{\text{HI, ICM}})$ , and hence yields the same shell model best-fitting parameters. This is shown in the first two panels in the first row of Fig. 18. The factor  $\sqrt{3}$  should arise from the geometrical difference a sphere and a slab. Specifically, in the optically thick and  $f_{\text{cl}} \gg f_{\text{cl, crit}}$  regime, the mean path-length of Ly $\alpha$  photons is  $\sqrt{3}B$  for a slab and  $R$  for a sphere, where  $B$  is the slab half-height and  $R$  is the sphere radius (Adams 1975).

However, adding two different models the same outflow velocity to either the clumps or the ICM yields a mismatch, as shown

in the third panel in the first row of Fig. 18. Such a mismatch should be due to the geometrical difference as well, yet we are unable to relate the two models with a scale factor in their outflow velocities [e.g. a spherical model with  $(\sigma_{\text{cl}}, \sqrt{3}N_{\text{HI, total}}, \sqrt{3}v_{\text{cl}})$  is still different from a slab model with  $(\sigma_{\text{cl}}, N_{\text{HI, total}}, v_{\text{cl}})$ ]. Therefore, we speculate that the geometrical difference has a non-linear effect on the propagation of the Ly $\alpha$  photons through the outflowing H I gas.

Nevertheless, we attempt to fit an outflowing clumpy spherical model with the shell model grid. The results are shown in the second row of Fig. 18. The first panel shows the shell model best fit to a clumpy slab model (with  $v_{\text{exp}} \simeq v_{\text{cl}}$ ,  $N_{\text{HI, shell}} \simeq N_{\text{HI, total}}$  and  $T_{\text{shell}} \simeq T_{\text{eff, slab}}$  as expected), and the second panel shows the best fit to the corresponding clumpy spherical model, where  $v_{\text{exp}}$  and  $T_{\text{shell}}$  are lower than expected and  $N_{\text{HI, shell}}$  is higher than expected. However, we find that such a mismatch is due to the intrinsic parameter degeneracy of the shell model (see Section 3.2). If we restrict  $T_{\text{shell}}$  to be  $\geq 10^{5.5}$  K, the best-fitting  $v_{\text{exp}}$  and  $N_{\text{HI, shell}}$  become consistent with the



**Figure 18.** Comparison between clumpy sphere models and clumpy slab models. Upper row: The first two panels show that a non-outflowing ( $v_{cl} = v_{ICM} = 0$ ) spherical model with  $(\sigma_{cl}, \sqrt{3}N_{HI, total}, \sqrt{3}N_{HI, ICM})$  gives an identical spectrum to a slab model with  $(\sigma_{cl}, N_{HI, total}, N_{HI, ICM})$ , and hence yields the same shell model best-fitting parameters. The factor  $\sqrt{3}$  should arise from the geometrical difference a sphere and a slab (see Section 4.5.1 for details). The third panel shows that adding the same  $v_{cl}$  yields a mismatch between the two models, which should be a non-linear effect due to the geometrical difference. Lower row: The first panel shows the shell model best fit to a clumpy slab model (with  $v_{exp} \simeq v_{cl}$ ,  $N_{HI, shell} \simeq N_{HI, total}$  and  $T_{shell} \simeq T_{eff, slab}$ ), and the second panel shows the best fit to the corresponding clumpy spherical model, where  $v_{exp}$  and  $T_{shell}$  are lower than expected and  $N_{HI, shell}$  is higher than expected. The third panel shows that if we restrict  $T_{shell}$  to be  $\geq 10^{5.5}$  K, the best-fitting  $v_{exp}$  and  $N_{HI, shell}$  become closer to the expected values, although the best fit gives a higher  $\chi^2$  due to the mismatch in the red peak.

expected values, although the best fit gives a higher  $\chi^2$  due to a larger mismatch in the trough and red peak (as shown in the third panel). We therefore speculate that the correspondence between shell models and clumpy slab models still roughly holds for clumpy spherical models, with slightly larger uncertainties due to the inessential geometrical difference.

#### 4.5.2 Moderately clumpy sphere: $f_{cl} \simeq f_{cl, crit}$

If the spherical model is only moderately clumpy, i.e.  $f_{cl} \simeq f_{cl, crit}$ , the Ly $\alpha$  optical depth at line centre will be low enough for photons to escape, which yields a significant non-zero residual flux density at line centre.<sup>15</sup> We find that as  $f_{cl}/f_{cl, crit}$  decreases, in the beginning the shell model is still able to produce a decent fit with reasonable parameters (albeit with the mismatch at line centre), but eventually the fit fails at  $f_{cl}/f_{cl, crit} \simeq 10$ . We illustrate this result in Fig. 19. In general, in the  $f_{cl} \simeq f_{cl, crit}$  regime, no direct correlation has been found between the shell and clumpy model parameters due to the efficient escape of Ly $\alpha$  photons at line centre in the clumpy model (Gronke & Dijkstra 2016).

## 5 DISCUSSION

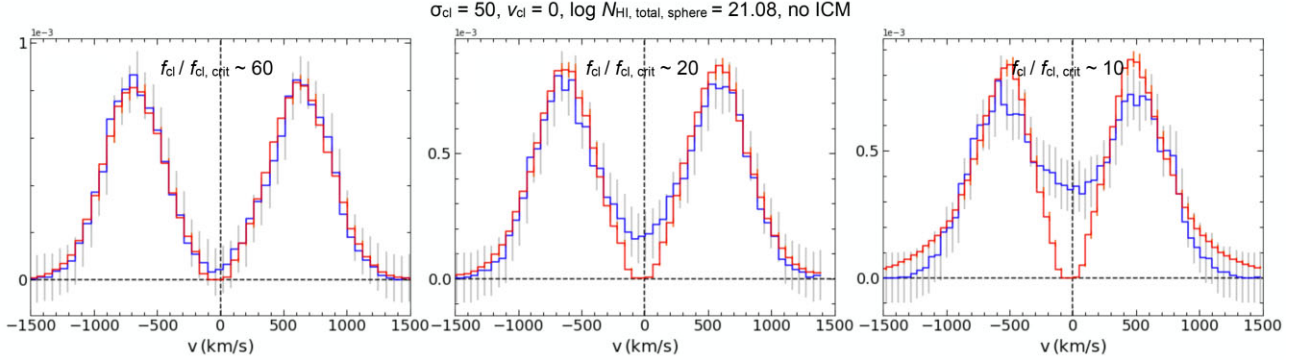
### 5.1 Interpretation of model parameters

Assuming that the degeneracy, we present in Section 3 can be somehow broken (see possible examples in the following Section 5.2), it is of great interest to decipher the crucial physical properties (kinematics, column density, etc.) of the Ly $\alpha$  scattering gaseous medium encoded in observed Ly $\alpha$  spectra, which exist ubiquitously in the Universe and often exhibit a diversity of morphology, e.g. different numbers and shapes of peaks, peak flux ratios, and peak separations. In this section, we summarize our findings on how one should interpret the parameters of the shell or clumpy model derived from fitting observed Ly $\alpha$  profiles. We will focus on the ‘very clumpy’ regime ( $f_{cl} \gg f_{cl, crit}$ ) unless otherwise noted.

(i) H I column density: This parameter can be constrained by the peak separation and the extent of the wings of the Ly $\alpha$  profile. The best-fitting shell model gives the H I column density of the shell  $N_{HI, shell}$ , whereas the best-fitting clumpy model gives the *total* H I column density within the clumps,<sup>16</sup> given by  $N_{HI, total} = \frac{4}{3} f_{cl} N_{HI, cl}$ .

<sup>15</sup>Note that even in the  $f_{cl} \gg f_{cl, crit}$  regime, residual flux density at line centre still exists, although to a lesser extent (see e.g. Figs 6 and 8).

<sup>16</sup>The total H I column density in the ICM is usually negligible compared to that within the clumps, as the ICM is usually much hotter ( $\gtrsim 10^6$  K) and has a much lower H I number density.



**Figure 19.** Shell model fits to clumpy spherical models with decreasing  $f_{\text{cl}}/f_{\text{cl, crit}}$ . The three panels correspond to three  $(\sigma_{\text{cl}}, v_{\text{cl}}, \log N_{\text{HI, total}}) = (50, 0, 21.08)$  models with different  $f_{\text{cl}}/f_{\text{cl, crit}}$  values:  $\sim 60$ ,  $20$ , and  $10$ , respectively. The blue curves are the clumpy spherical model spectra and the red curves are the shell model best fits. As  $f_{\text{cl}}/f_{\text{cl, crit}}$  decreases, in the beginning, the shell model is still able to produce a decent fit with reasonable parameters (albeit with the mismatch at line centre), but eventually the fit fails at  $f_{\text{cl}}/f_{\text{cl, crit}} \approx 10$ .

As we have shown in previous sections,  $N_{\text{HI, shell}} \approx N_{\text{HI, total}}$  usually holds for the clumpy model and the best-fitting shell model, suggesting that the H I column density can be robustly determined from fitting.

However, this parameter should be treated with at least two caveats: (1) As we have shown in Section 3 and 4.2, in the optically thick regime, the (total) H I column density is degenerate with the shell effective temperature (or the random velocity of the clumps). Therefore, in order to get a well-constrained H I column density by fitting observed Ly $\alpha$  spectra, additional constraints are needed to break the degeneracy; (2) As both the shell model and the clumpy model assume an isotropic H I gas distribution, whereas in actual astrophysical environments the gas distribution is more likely to be anisotropic, the derived  $N_{\text{HI}}$  value should be regarded only as an average value along the paths of escape of the Ly $\alpha$  photons (which is actually not necessarily the average column density either along the line of sight or of all angles).

(ii) Shell effective temperature (or Doppler parameter)/clump velocity dispersion: This parameter can be constrained by the width of the Ly $\alpha$  profile, i.e. the full width at half-maximum (FWHM) of the peak(s). As we have shown, the shell effective temperature is usually equal to the effective temperature of the clumpy model with velocity dispersion  $\sigma_{\text{cl}}$  (see equation 2). In other words, the turbulent velocity term in the Doppler parameter of the shell model (see equation 3 in Verhamme et al. 2006) is equivalent to  $\sigma_{\text{cl}}$  of the clumpy model for the same Ly $\alpha$  profile.

We hereby highlight a scenario where the fitted  $T_{\text{eff}}$  of the shell model cannot be interpreted literally. If a Ly $\alpha$  spectrum is very broad and has very extended wings, it may require a high  $T_{\text{eff}} \gtrsim 10^6$  K,<sup>17</sup> or equivalently, Doppler parameter  $b \gtrsim 100$  km s<sup>-1</sup>. Such a high Doppler parameter already corresponds to a very high internal turbulent Mach number  $\mathcal{M}_{\text{cl}}^{\text{turb}} \sim 10$ , which is enough to disintegrate the H I shell or shock-ionize the H I gas. Alternatively, we should interpret this as a clump velocity dispersion  $\sigma_{\text{cl}} \gtrsim 100$  km s<sup>-1</sup> of the clumpy model, which is physically reasonable in a strong gravitational field and/or in the presence of feedback.

(iii) Shell expansion/clump outflow velocity: This parameter can be constrained by the blue-to-red peak flux ratio of the Ly $\alpha$  profile.

As we have shown,  $v_{\text{exp}} \approx v_{\text{cl}}$  usually holds for the clumpy model and the best-fitting shell model, suggesting that this parameter can also be robustly determined from fitting. However, the fitted shell expansion velocity ( $v_{\text{exp}}$ ) should not be interpreted literally as the bulk outflowing velocity of the H I gas in at least two cases: (1) The actual velocity field of the H I gas varies spatially. For example, in Section 4.3, we find that the best-fitting shell model to a clumpy slab model with either a linearly increasing outflow or a ‘momentum-driven + gravitational deceleration’ velocity profile has  $v_{\text{exp}} \approx \overline{v_{\text{cl}}(r)}$ . Moreover, if the UV absorption lines suggest a series of outflow velocities,  $v_{\text{exp}}$  is expected to lie between the minimum and maximum absorption velocities; (2) The fitted Ly $\alpha$  spectrum emerges from a multiphase scattering medium. For example, in Section 4.4 we show that the best-fitting shell model to a multiphase clumpy slab model with cool clumps outflowing at  $v_{\text{cl}}$  and a hot ICM outflowing at  $v_{\text{ICM}}$  has  $v_{\text{exp}} < \max\{v_{\text{cl}}, v_{\text{ICM}}\}$  unless  $v_{\text{cl}} = v_{\text{ICM}}$ . Therefore, the fitted  $v_{\text{exp}}$  should be interpreted as an average outflow velocity – both space-wise and phase-wise.

(iv) Intrinsic linewidth: This parameter can be constrained by the extent of the wings of the Ly $\alpha$  profile (for the shell model only; in the clumpy model, the intrinsic linewidth is fixed to be small and  $\sigma_{\text{cl}}$  is responsible for the broadening of the wings). It is well known that the fitted intrinsic linewidth  $\sigma_i$  of the shell model is usually overly large compared to the widths of the observed Balmer lines (Orlitová et al. 2018). In this work, we have shown in Section 4.2 that a large intrinsic linewidth  $\sigma_i$  is always required for a shell model to fit a clumpy slab model with randomly moving clumps, and (1)  $\sigma_i$  increases as  $\sigma_{\text{cl}}$  increases; (2)  $\sigma_i > \sigma_{\text{cl}}$  always holds (see Fig. 7). This is due to the intrinsic difference between a shell model and a clumpy model with corresponding parameters: compared to the shell model, the clumpy model naturally has more extended wings and lower but more extended peaks (see Fig. 10), and is better suited for fitting broad Ly $\alpha$  spectra with extended wings. This  $\sigma_{\text{cl}} < \sigma_i$  trend, together with the quadruple degeneracy that we have discussed in Section 3, provides a viable solution to the three major discrepancies emerged from shell model fitting as reported by Orlitová et al. (2018). It also suggests that the large  $\sigma_i$  values required in shell model fitting may simply imply a clumpy gas distribution (with a considerable velocity dispersion).

(v) Systemic redshift: When fitting an observed Ly $\alpha$  profile, a parameter that dictates the systemic redshift of the modelled Ly $\alpha$  source is usually introduced in post-processing. As Ly $\alpha$  profile fitting is usually done in velocity space, this parameter can be specified as

<sup>17</sup>The FWHM of a Ly $\alpha$  profile is positively correlated with  $N_{\text{HI, shell}}$  and  $T_{\text{eff}}$ ; for a static sphere,  $\text{FWHM} \approx 320 \left( \frac{N_{\text{HI, shell}}}{10^{20} \text{ cm}^{-2}} \right)^{1/3} \left( \frac{T_{\text{eff}}}{10^4 \text{ K}} \right)^{1/6} \text{ km s}^{-1}$  (see equation 87 from Dijkstra 2017).

$\Delta v$ , which is the difference between the systemic velocity of the modelled Ly $\alpha$  source and the zero velocity of the observed Ly $\alpha$  profile. For a typical double-peak Ly $\alpha$  profile with a central trough between two peaks, the  $\Delta v$  of the best-fitting shell model is correlated with  $v_{\text{exp}}$ , as the optical depth is maximum at  $\sim -v_{\text{exp}}$  (i.e. the trough location; see Orlitová et al. 2018). In other words,  $\Delta v$  and  $v_{\text{exp}}$  are intrinsically degenerate with each other.

Now the clumpy model offers us more possibilities to solve this issue with more flexibility. Although a single-phase slab with a very high clump covering factor ( $f_{\text{cl}} \gg f_{\text{cl, crit}}$ ) basically converges to the shell model, a multiphase clumpy medium can produce many different trough shapes: for example, for  $f_{\text{cl}} \gg f_{\text{cl, crit}}$  with a static or outflowing ICM, the trough can extend to both sides of the zero velocity (see Fig. 17); for  $f_{\text{cl}} \simeq f_{\text{cl, crit}}$  with a static ICM, the trough has residual flux and is always located at the line centre. More modelling of observed Ly $\alpha$  profiles is needed to examine whether these possibilities are physically reasonable.

## 5.2 Breaking the degeneracy

The intrinsic parameter degeneracy of the shell model (and the clumpy model as well, at least in the ‘very clumpy’ regime) that we have described in Section 3 concerns us that how much meaningful physical information, if any, can be extracted from Ly $\alpha$  spectra via RT modelling. In this section, we speculate several scenarios where the intrinsic parameter degeneracy can be broken and the physical properties of the Ly $\alpha$  scattering medium can actually be constrained.

(i) An accurate measurement of the systemic redshift of the Ly $\alpha$  source: Assuming that all the Ly $\alpha$  photons are generated from recombination and nebular emission line(s) are clearly detected (e.g. H $\alpha$ , H $\beta$ , or [O III]), the systemic redshift (i.e. the  $\Delta v$  parameter) of the Ly $\alpha$  source can be constrained reasonably well, and hence breaks the degeneracy. However, this requires that (1) the observed Ly $\alpha$  spectrum has a clear trough between the double peaks so that  $v_{\text{exp}}$  can be constrained; (2) the asymmetry of the observed Ly $\alpha$  spectrum is significant enough so that the corresponding  $v_{\text{exp}}$  is much higher than the uncertainty of  $\Delta v$ .

(ii) Additional observational constraints on the gas outflow velocity/velocity dispersion/H I column density: If additional information is available from other observations, it may also help break the parameter degeneracy. Nevertheless, as such quantities are derived rather than directly observed (e.g. the gas outflow velocity can be deduced from UV absorption lines, and the gas velocity dispersion can be inferred from the widths of nebular emission lines), it is more reasonable to treat them as priors that confine the parameter space. Therefore, unless these additional constraints are reasonably stringent, the output parameters will still suffer from the degeneracy (which actually exists continuously across the parameter space).

(iii) The Ly $\alpha$  profile corresponds to an optically thin regime: As we have only found the parameter degeneracy in the optically thick regime, it is anticipated that if the Ly $\alpha$  profile does not belong to this regime, it may not be heavily affected by the parameter degeneracy. Ly $\alpha$  spectra emerged from H I with very low column densities ( $\lesssim 10^{18} \text{ cm}^{-2}$ ) will be naturally in the optically thin regime – they often exhibit narrow peak separations and/or residual flux at line centre. However, objects that produce such Ly $\alpha$  profiles are presumably LyC leakers and are rare in the Universe (Cooke et al. 2014; Verhamme et al. 2015).

In short, one should be cautious when interpreting the extracted parameters from fitting observed Ly $\alpha$  spectra with idealized RT

models. Additional observations on other lines may help break the intrinsic parameter degeneracy and better constrain the properties of the gaseous medium, although sometimes different types of constraints may contradict each other and yield unsuccessful fits (see e.g. section 4.1 in Orlitová et al. 2018). In that case, development of more advanced RT models that are more physically realistic and flexible may help solve this issue in the future.

## 6 CONCLUSIONS

In this work, we have explored what physical properties can be extracted from Ly $\alpha$  spectra via RT modelling. The main conclusions of this work are as follows:

(i) Intrinsic parameter degeneracies exist in the widely used shell model in the optically thick regime. For static shells, models with the same  $N_{\text{H I, shell}} T_{\text{shell}}^{0.5}$  exhibit nearly identical Ly $\alpha$  spectra. For outflowing shells, a quadruple degeneracy exists among ( $v_{\text{exp}}$ ,  $N_{\text{H I, shell}}$ ,  $T_{\text{shell}}$ ,  $\Delta v$ ). This finding reveals the limitations of the shell model and cautions against making any reasonable statements about the physical properties of the Ly $\alpha$  scattering medium with only shell model fitting (cf. Section 3);

(ii) The parameters of a ‘very clumpy’ slab model have a close correspondence to the parameters of the shell model. Specifically, (1) the total column density of the clumpy slab model,  $N_{\text{H I, total}} = \frac{4}{3} f_{\text{cl}} N_{\text{H I, cl}}$  is equal to the H I column density of the shell model,  $N_{\text{H I, shell}}$ ; (2) the effective temperature of the clumpy slab model,  $T_{\text{eff, slab}} = T_{\text{cl}} + \frac{\sigma_{\text{cl}}^2 m_{\text{H}}}{2k_{\text{B}}}$ , where  $\sigma_{\text{cl}}$  is the 1D velocity dispersion of the clumps, is equal to the effective temperature of the shell model,  $T_{\text{shell}}$ ; (3) the average radial clump outflow velocity,  $\overline{v_{\text{cl}}(r)}$ , is equal to the shell expansion velocity,  $v_{\text{exp}}$ . This reminds us that the shell model parameters should be interpreted in a more physically realistic context rather than literally;

(iii) In the shell model, large intrinsic linewidths (several times of  $\sigma_{\text{cl}}$ ) are required to reproduce the wings of the clumpy slab models, reflecting the intrinsic difference between two different models. This  $\sigma_{\text{cl}} < \sigma_{\text{I}}$  trend, together with the quadruple degeneracy, provides a viable solution to the three major discrepancies emerged from shell model fitting as reported by Orlitová et al. (2018);

(iv) Adding another phase of hot ICM to the clumpy slab model will increase peak separation and boost  $T_{\text{shell}}$ , but keeps  $N_{\text{H I, shell}} \simeq N_{\text{H I, total}}$ . The fitted  $v_{\text{exp}}$  lies between  $v_{\text{cl}}$  and  $v_{\text{ICM}}$ . In particular, if  $v_{\text{cl}} = v_{\text{ICM}}$ , i.e. the cool clumps and the hot ICM are co-outflowing at the same speed, we get  $v_{\text{exp}} \simeq v_{\text{cl}} = v_{\text{ICM}}$ ;

(v) For multiphase, clumpy spherical models, if  $f_{\text{cl}}$  is much larger than a critical value  $f_{\text{cl, crit}}$ , the parameter correspondence still holds, albeit with larger uncertainties due to the geometrical difference; whereas if  $f_{\text{cl}} \simeq f_{\text{cl, crit}}$ , no direct correlation has been found between the shell and clumpy model parameters.

In general, in order to obtain meaningful constraints on the physical properties of the Ly $\alpha$  scattering gaseous medium, one should try to break the intrinsic parameter degeneracies revealed in this work with extra information from additional observations, rather than merely rely on fitting observed Ly $\alpha$  spectra with idealized RT models. Moreover, the model parameters derived from Ly $\alpha$  spectra fitting should not be understood literally – instead, they should be interpreted in a more physically realistic context, e.g. in a multiphase, clumpy medium that we have explored in this work. Efforts in building more advanced RT models (e.g. with more realistic geometries) will also be helpful in the future.

## ACKNOWLEDGEMENTS

We acknowledge the anonymous referee for the detailed reading of the manuscript and constructive feedback. We thank Phil Hopkins for providing computational resources. MG was supported by NASA through the NASA Hubble Fellowship grant *HST*-HF2-51409 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 thanks the Max Planck Society for support through the Max Planck Research Group. Numerical calculations were run on the Caltech compute cluster ‘Wheeler,’ allocations from XSEDE TG-AST130039 and PRAC NSF.1713353 supported by the NSF, and NASA HEC SMD-16-7592. We also acknowledge the use of the the following software packages: Astropy (Astropy Collaboration et al. 2018), the SCIPY and NUMPY system (Harris et al. 2020; Virtanen et al. 2020).

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Adams T. F., 1972, *ApJ*, 174, 439  
 Adams T. F., 1975, *ApJ*, 201, 350  
 Ahn S.-H., 2004, *ApJ*, 601, L25  
 Ahn S.-H., Lee H.-W., Lee H. M., 2003, *MNRAS*, 340, 863  
 Astropy Collaboration et al., 2018, *AJ*, 156, 123  
 Atek H., Schaerer D., Kunth D., 2009, *A&A*, 502, 791  
 Behrens C., Dijkstra M., Niemeyer J. C., 2014, *A&A*, 563, A77  
 Cantalupo S., Porciani C., Lilly S. J., Miniati F., 2005, *ApJ*, 628, 61  
 Cooke J., Ryan-Weber E. V., Garel T., Díaz C. G., 2014, *MNRAS*, 441, 837  
 Cox D. P., 2005, *ARA&A*, 43, 337  
 Dessauges-Zavadsky M., D’Odorico S., Schaerer D., Modigliani A., Tapken C., Vernet J., 2010, *A&A*, 510, A26  
 Dijkstra M., 2017, preprint (arXiv:1704.03416)  
 Dijkstra M., Kramer R., 2012, *MNRAS*, 424, 1672  
 Dijkstra M., Haiman Z., Spaans M., 2006a, *ApJ*, 649, 14  
 Dijkstra M., Haiman Z., Spaans M., 2006b, *ApJ*, 649, 37  
 Duval F., Schaerer D., Östlin G., Laursen P., 2014, *A&A*, 562, A52  
 Gronke M., 2017, *A&A*, 608, A139  
 Gronke M., Dijkstra M., 2016, *ApJ*, 826, 14  
 Gronke M., Bull P., Dijkstra M., 2015, *ApJ*, 812, 123  
 Gronke M., Dijkstra M., McCourt M., Peng Oh S., 2017, *A&A*, 607, A71  
 Hansen M., Oh S. P., 2006, *MNRAS*, 367, 979  
 Harrington J. P., 1973, *MNRAS*, 162, 43  
 Harris C. R. et al., 2020, *Nature*, 585, 357  
 Hashimoto T. et al., 2015, *ApJ*, 812, 157  
 Henry A., Scarlata C., Martin C. L., Erb D., 2015, *ApJ*, 809, 19  
 Kakiichi K., Gronke M., 2021, *ApJ*, 908, 30  
 Kimm T., Blaizot J., Garel T., Michel-Dansac L., Katz H., Rosdahl J., Verhamme A., Haehnelt M., 2019, *MNRAS*, 486, 2215  
 Klein R. I., McKee C. F., Colella P., 1994, *ApJ*, 420, 213  
 Krogager J.-K. et al., 2013, *MNRAS*, 433, 3091  
 Kulas K. R., Shapley A. E., Kollmeier J. A., Zheng Z., Steidel C. C., Hainline K. N., 2012, *ApJ*, 745, 33  
 Lao B.-X., Smith A., 2020, *MNRAS*, 497, 3925  
 Laursen P., Sommer-Larsen J., 2007, *ApJ*, 657, L69  
 Laursen P., Duval F., Östlin G., 2013, *ApJ*, 766, 124  
 Li Z., Steidel C. C., Gronke M., Chen Y., Matsuda Y., 2021a, *MNRAS*  
 Li Z., Steidel C. C., Gronke M., Chen Y., 2021b, *MNRAS*, 502, 2389  
 Loeb A., Rybicki G. B., 1999, *ApJ*, 524, 527  
 Martin C. L., Dijkstra M., Henry A., Soto K. T., Danforth C. W., Wong J., 2015, *ApJ*, 803, 6  
 McQuinn M., 2016, *ARA&A*, 54, 313  
 Murray N., Quataert E., Thompson T. A., 2005, *ApJ*, 618, 569

- Neufeld D. A., 1990, *ApJ*, 350, 216  
 Neufeld D. A., 1991, *ApJ*, 370, L85  
 Orlitová I., Verhamme A., Henry A., Scarlata C., Jaskot A., Oey M. S., Schaerer D., 2018, *A&A*, 616, A60  
 Ouchi M., Ono Y., Shibuya T., 2020, *ARA&A*, 58, 617  
 Richling S., 2003, *MNRAS*, 344, 553  
 Runnholm A., Gronke M., Hayes M., 2021, *PASP*, 133, 034507  
 Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, AIP Conf. Ser. Vol. 735. Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Am. Inst. Phys., New York, NY. p. 395  
 Skilling J., 2006, *Bayesian Anal.*, 1, 833  
 Smith A., Ma X., Bromm V., Finkelstein S. L., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2019, *MNRAS*, 484, 39  
 Smith A. et al., 2021, preprint (arXiv:2111.13721)  
 Song H., Seon K.-I., Hwang H. S., 2020, *ApJ*, 901, 41  
 Speagle J. S., 2020, *MNRAS*, 493, 3132  
 Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, 717, 289  
 Tasitsiomi A., 2006, *ApJ*, 645, 792  
 Tumlinson J., Peebles M. S., Werk J. K., 2017, *ARA&A*, 55, 389  
 Vanzella E. et al., 2010, *A&A*, 513, A20  
 Verhamme A., Schaerer D., Maselli A., 2006, *A&A*, 460, 397  
 Verhamme A., Schaerer D., Atek H., Tapken C., 2008, *A&A*, 491, 89  
 Verhamme A., Dubois Y., Blaizot J., Garel T., Bacon R., Devriendt J., Guiderdoni B., Slyz A., 2012, *A&A*, 546, A111  
 Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, *A&A*, 578, A7  
 Verhamme A., Orlitová I., Schaerer D., Izotov Y., Worseck G., Thuan T. X., Guseva N., 2017, *A&A*, 597, A13  
 Virtanen P. et al., 2020, *Nat. Meth.*, 17, 261  
 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., Miralda-Escudé J., 2002, *ApJ*, 578, 33  
 Zheng Z., Wallace J., 2014, *ApJ*, 794, 116

APPENDIX A: EFFECT OF ICM TEMPERATURE ON LY $\alpha$  MODEL SPECTRA

Here, we show that adding a hot phase of ICM does not necessarily affect the Ly $\alpha$  model spectrum unless it satisfies a certain condition. Specifically, the transmission function  $\mathcal{T}(x) = e^{-\tau(x)} = e^{-\sigma_{\text{H I}}(x, T) N_{\text{H I}}}$  needs to be wider for the hot phase than the cool phase.

In the core of the Ly $\alpha$  line, the H I cross-section,  $\sigma_{\text{H I}} = \sigma_0 H(a_v, x) \sim \sigma_0 e^{-x^2}$ , where  $H(a_v, x)$  is the Voigt function, and  $\sigma_0 \approx 5.895 \times 10^{-14} (T/10^4 \text{ K})^{-1/2} \text{ cm}^2$  is the H I cross-section at line centre. Assuming that at a certain optical depth  $\tau$ ,  $\mathcal{T}(x)$  becomes sufficiently small and reaches a threshold  $\mathcal{T}_0$  (e.g. for  $\tau(x) \gtrsim 7$ ,  $\mathcal{T}(x) \lesssim 10^{-3}$ ), and denoting  $C_1 = 5.895 \times 10^{-12} \text{ cm}^2$  so that  $\sigma_0 = C_1/\sqrt{T}$ , we have

$$e^{-\frac{C_1}{\sqrt{T}} e^{-x^2} N_{\text{H I}}} = \mathcal{T}_0 \quad (\text{A1})$$

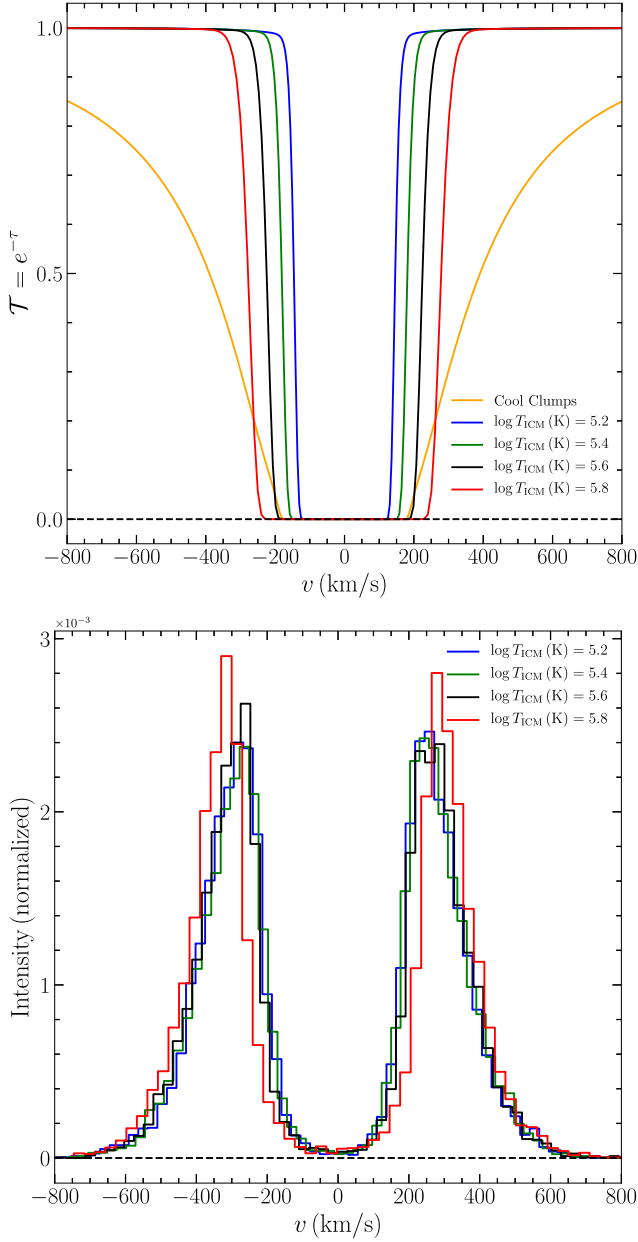
so the threshold frequency can be solved as

$$x_0 = \sqrt{-\ln \left( -\frac{\sqrt{T}}{C_1 N_{\text{H I}}} \ln \mathcal{T}_0 \right)} \quad (\text{A2})$$

which corresponds to a threshold velocity:

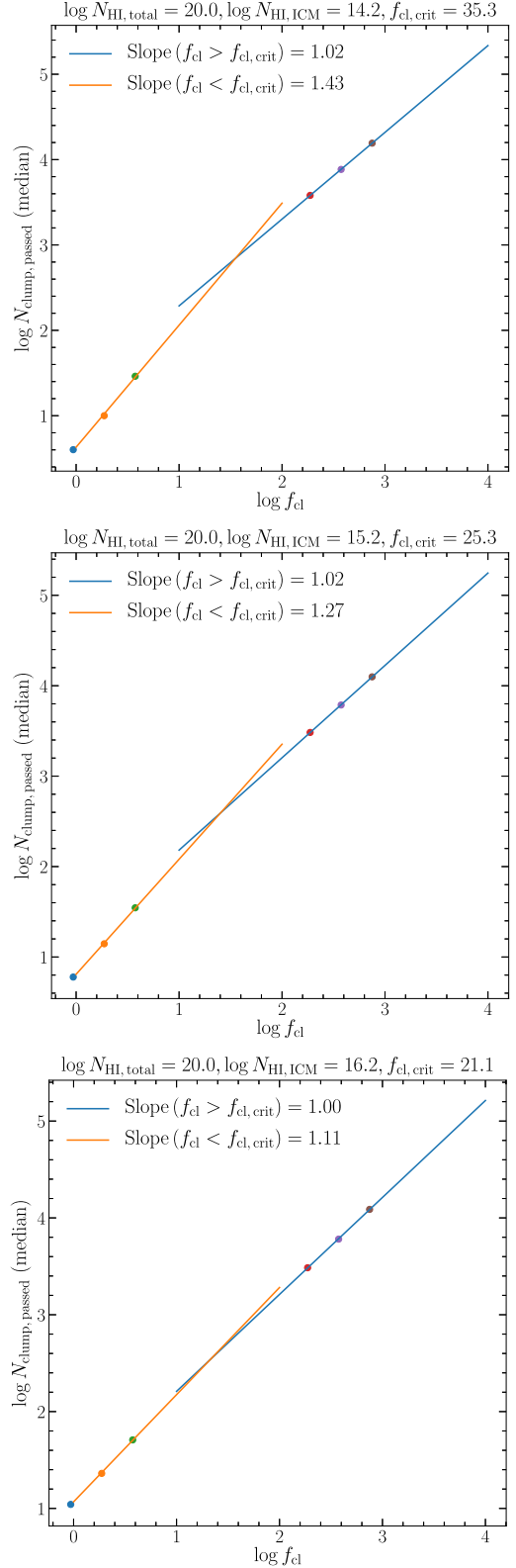
$$v_0 = x_0 v_{\text{th}} = \sqrt{-\ln \left( -\frac{\sqrt{T}}{C_1 N_{\text{H I}}} \ln \mathcal{T}_0 \right)} \sqrt{\frac{2k_B T}{m_{\text{H}}}} \quad (\text{A3})$$

In order to have an impact on the Ly $\alpha$  profile, the hot phase of ICM needs to have a threshold velocity larger than that of the cool phase, i.e.  $v_{0, \text{ICM}} > v_{0, \text{cool}}$ . We hereby consider the following example: a two-phase scattering medium that consists of



**Figure A1.** Effect of ICM temperature on Ly $\alpha$  transmission function and Ly $\alpha$  model spectra. Here, we show one example: a two-phase scattering medium that consists of cool clumps with velocity dispersion  $\sigma_{cl} = 50 \text{ km s}^{-1}$  (hence effective temperature  $\log T_{eff}(K) = 5.2$ ) and total H I column density  $\log N_{HI, total, cool} = 19.6$ , and a hot ICM with total H I column density  $\log N_{HI, total, ICM} = 17.2$ . Top panel: The transmission function of the cool clumps (the orange curve) as compared to those of the ICM at different temperatures. Bottom panel: The model Ly $\alpha$  profile as a function of the ICM temperature. As inferred from equation (A3), at  $\log T_{ICM}(K) \sim 5.6$  (the black curves) the ICM starts to have an impact on both the transmission function and the model Ly $\alpha$  profile.

cool clumps with velocity dispersion  $\sigma_{cl} = 50 \text{ km s}^{-1}$  (hence effective temperature  $\log T_{eff}(K) = 5.2$ ) and total H I column density  $\log N_{HI, total, cool} = 19.6$ , and a hot ICM with total H I column density  $\log N_{HI, total, ICM} = 17.2$ . Using equation (A3) and demanding  $v_0, ICM > v_0, cool$  yields that  $\log T_{ICM}(K) \gtrsim 5.6$  is required for the ICM to have a wider transmission function than the cool clumps, and hence have a visible impact on the Ly $\alpha$  profile. We illustrate this result in Fig. A1 by showing a series of ICM transmission functions and model Ly $\alpha$  profiles with different  $T_{ICM}$  values. It can be clearly



**Figure B1.** Effect of ICM with different column densities on the critical clump covering factor,  $f_{cl, crit}$ . Here, we show one set of examples: a two-phase clumpy slab with  $\log N_{HI, total} = 20.0$  in the static clumps, and a hot ICM with total H I column density  $\log N_{HI, ICM} = 14.2-16.2$  (or equivalently,  $n_{HI, ICM} = 10^{-6}-10^{-4} \text{ cm}^{-3}$ ). With  $\log N_{HI, ICM}$  varying by two orders of magnitudes,  $f_{cl, crit}$  only changes by a factor of  $\sim 1.5$ , suggesting that the hot ICM only has a minor effect on  $f_{cl, crit}$  and the boundaries of different RT regimes.

seen that at  $\log T_{\text{ICM}}(\text{K}) \sim 5.6$  the ICM starts to have an impact on both the transmission function and the model Ly $\alpha$  profile.

## APPENDIX B: EFFECT OF ICM ON THE CRITICAL COVERING FACTOR

In Gronke et al. (2017), an important physical quantity is defined – the critical covering factor of the clumps,  $f_{\text{cl, crit}}$ . It is the critical average number of clumps per line of sight, above which the clumpy scattering medium will behave like a homogeneous medium (i.e. a homogeneously filled shell or slab), and below which a significant number of Ly $\alpha$  photons will escape near the line centre. In this section, we test how much impact the hot ICM component has on  $f_{\text{cl, crit}}$ , and hence on the boundaries of different RT regimes.

The value of  $f_{\text{cl, crit}}$  sets the transition between two physical regimes of Ly $\alpha$  resonant scattering. Assuming that the ensemble of the clumps is optically thick at the Ly $\alpha$  line centre (which is always true throughout this work), if  $f_{\text{cl}} \lesssim f_{\text{cl, crit}}$ , photons scatter off the clumps in a random-walk manner, and the number of clumps a photon

intercepts scales as  $N_{\text{cl}} \propto f_{\text{cl}}^2$ ; whereas if  $f_{\text{cl}} \gtrsim f_{\text{cl, crit}}$ , photons escape

via a frequency excursion (i.e. a series of wing scatterings), and the number of clumps a photon intercepts scales as  $N_{\text{cl}} \propto f_{\text{cl}}$ . Therefore,  $f_{\text{cl, crit}}$  can be estimated by determining the turning point of the scaling relation between  $N_{\text{cl}}$  and  $f_{\text{cl}}$  (see figs 2, 4 and 6 from Gronke et al. 2017).

Here, we show one set of examples in Fig. B1: a two-phase clumpy slab with  $\log N_{\text{H I, total}} = 20.0$  in the clumps (which are static), and a hot ICM with total H I column density  $\log N_{\text{H I, ICM}} = 14.2\text{--}16.2$  (or equivalently,  $n_{\text{H I, ICM}} = 10^{-6}\text{--}10^{-4} \text{ cm}^{-3}$ ). It can be seen that with  $\log N_{\text{H I, ICM}}$  varying by two orders of magnitudes,  $f_{\text{cl, crit}}$  only changes by a factor of  $\sim 1.5$ . This result suggests that although under certain conditions, the hot ICM can have a significant impact on the model Ly $\alpha$  spectrum (see Section 4.4 and 4.5), it only has a minor effect on  $f_{\text{cl, crit}}$  and the boundaries of different RT regimes.

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