

# The less and the more IGM transmitted galaxies from $z \sim 2.7$ to $z \sim 6$ from VANDELS\* and VUDS†

R. Thomas<sup>1</sup>, L. Pentericci<sup>2</sup>, O. Le Fèvre<sup>3, \*\*</sup>, A. M. Koekemoer<sup>16</sup>, M. Castellano<sup>2</sup>, A. Cimatti<sup>4, 5</sup>, F. Fontanot<sup>6, 7</sup>, A. Gargiulo<sup>8</sup>, B. Garilli<sup>8</sup>, M. Talia<sup>4, 9</sup>, R. Amorín<sup>10, 11</sup>, S. Bardelli<sup>4</sup>, S. Cristiani<sup>6</sup>, G. Cresci<sup>5</sup>, M. Franco<sup>12, 13</sup>, J. P. U. Fynbo<sup>14, 15</sup>, N. P. Hathi<sup>16</sup>, P. Hibon<sup>1</sup>, Y. Khusanova<sup>3, 17</sup>, V. Le Brun<sup>3</sup>, B. C. Lemaux<sup>18</sup>, F. Mannucci<sup>5</sup>, D. Schaerer<sup>19</sup>, G. Zamorani<sup>4</sup>, and E. Zucca<sup>4</sup>

<sup>1</sup> European Southern Observatory, Av. Alonso de Córdova 3107, Vitacura, Santiago, Chile ,  
e-mail: rthomas@eso.org

<sup>2</sup> INAF - Osservatorio Astronomico di Roma, via Frascati 33, I-00078 Monteporzio Catone, Italy

<sup>3</sup> Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France

<sup>4</sup> INAF - Osservatorio di Astrofisica e Scienza dello Spazio, via Gobetti 93/3, 40129 Bologna Italy

<sup>5</sup> INAF - Osservatorio Astrofisico di Arcetri, largo E. Fermi 5, 50125, Firenze, Italy

<sup>6</sup> INAF-Astronomical Observatory, via G.B. Tiepolo 11, I-34143 Trieste, Italy

<sup>7</sup> IFPU - Institute for Fundamental Physics of the Universe, via Beirut 2, 34151, Trieste, Italy

<sup>8</sup> INAF IASF-Milano, via A.Corti 12, 20133, Milano, Italy

<sup>9</sup> University of Bologna - Department of Physics and Astronomy, Via Gobetti 93/2, I-40129, Bologna, Italy

<sup>10</sup> Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raúl Bitrán 1305, La Serena, Chile

<sup>11</sup> Departamento de Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, La Serena, Chile

<sup>12</sup> AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France

<sup>13</sup> Centre for Astrophysics Research, University of Hertfordshire, Hatfield, AL10 9AB, UK

<sup>14</sup> Cosmic DAWN Center, Vibenshuset, Lyngbyvej 2, DK-2100 Copenhagen, Denmark

<sup>15</sup> Niels Bohr Institute, Copenhagen University, Lyngbyvej 2, DK-2100 Copenhagen O, Denmark

<sup>16</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<sup>17</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>18</sup> Department of Physics, University of California, Davis, One Shields Ave., Davis, CA 95616, USA

<sup>19</sup> Observatoire de Genève, Université de Genève, 51 Ch. des Maillettes, 1290 Versoix, Switzerland

Received ; accepted

## ABSTRACT

**Aims.** Our aim is to analyse the variance of the Inter-Galactic Medium transmission (IGM) by studying this parameter in the rest-frame UV spectra of a large sample of high redshift galaxies.

**Methods.** We make use of the VIMOS Ultra Deep Survey and the VANDELS public survey to have an insight into the far UV spectrum of  $2.7 < z < 6$  galaxies. Using the SPARTAN fitting software, we estimate the IGM towards individual galaxies and then divide them in two sub-samples characterized by a transmission above or below the theoretical prescription. We create average spectra of combined VUDS and VANDELS data for each set of galaxies in seven redshift bins.

**Results.** The resulting spectra clearly exhibit the variance of the IGM transmission that can be seen directly from high redshift galaxy observations. Computing the optical depth based on the IGM transmission, we find an excellent agreement with QSOs results. In addition, our measurements seem to suggest that there is a large dispersion of redshift where complete Gunn-Peterson Trough happens, depending on the line of sight.

**Key words.** Extragalactic astronomy – Spectroscopy – High redshift – Intergalactic medium

## 1. Introduction

The light coming from distant sources (like Quasi Stellar Objects and galaxies) is absorbed by the gaseous hydrogen systems that are lying along the line of sight. At increasing redshift, this effect, called the Inter-Galactic Medium (IGM) absorption can be so important that all the light at wavelengths blueward the

Lyman  $\alpha$  [ $\text{Ly}\alpha$ ] at 1216Å becomes invisible to us. Such absorption has been the topic of numerous studies and this is mainly connected with the growth of the Large Scale Structure as predicted by the hierarchical structure formation scenario and it is tightly connected with the epoch of Reionization (e.g. Cen et al. 1994).

Twenty five years ago, Madau (Madau 1995, hereafter *M95*) proposed a theoretical model able to reproduce the shape of the extinction curve as a function of redshift. Using this model, they conclude that the IGM transmission,  $\text{Tr}(\text{Ly}\alpha)$ , decreases with increasing redshift and that the scatter, at a given redshift should be large, e.g., from 20% to 70% with an average of 40% at  $z=3.5$ . A decade later, Meiksin (Meiksin 2006, hereafter *M06*) produced

\* Based on observations made with ESO Telescopes at the La Silla or Paranal Observatories under programme ID(s) 194.A-2003

† Based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, under Large Program 185.A-0791

\*\* Oliver Le Fèvre, 1960-2020

an update of this model using the  $\Lambda$  Cold Dark Matter cosmological predictions. More recently, Inoue et al. (2014) developed a new model of transmission. Their model predicts a weaker average absorption in the range  $z=3-5$  while it becomes stronger at  $z>6$ . On the observational side, studies of the IGM have been provided almost exclusively by Quasi Stellar Objects (QSOs). Faucher-Giguère et al. (2008b) used 86 high-resolution quasar spectra with a high signal-to-noise ratio to provide reference measurements of the dispersion over  $2.2 < z < 4.6$ . Dall’Aglio et al. 2008 used 40 quasars bright quasars to produce the measurements of the IGM optical depth between  $z=2.5$  and  $z=4.5$  while Becker et al. (2013) has extended the redshift range up to 5.5 with more than 6000 quasars. They all found that the transmission is indeed decreasing with increasing redshift.

Until a few years ago, no observational studies had been made of the evolution of the IGM transmission from galaxy samples mainly because of the lack of large spectroscopic samples with high signal-to-noise ratios at high redshift that probed a wavelength range between the Lyman-limit at  $912\text{\AA}$  and  $\text{Ly}\alpha$ . Hence, the IGM transmission towards extended galaxies had not yet been compared with the well known results for point-like QSO samples. In two recent papers we were able to compute the IGM transmission for a large sample of galaxies using the VIMOS Ultra Deep Survey [VUDS; Le Fèvre et al. 2015; Thomas et al. 2017, hereafter T17] and the VANDELS survey [McLure et al. 2018; Pentericci et al. 2018; Thomas et al. 2020, hereafter T20]. In these studies we showed that not only we can measure the IGM transmission using galaxy spectra but that it is crucial for understanding how galaxies are selected with Lyman Break techniques (Steidel et al. 1995) which highly depends on the IGM.

In this letter, we assembled more than 2500 galaxies from VUDS and VANDELS at  $z>2.7$  with measured IGM absorption (using the SPARTAN tool) to create averaged IGM spectra in different redshift bins to look at the IGM extinctions and visually see the effect of the variance of IGM on galaxy data. We describe the VUDS and VANDELS galaxy sample and selection in Sect 2. The fitting method with the SPARTAN tool (Thomas 2021) and the recipe we use to compute the IGM are described in Sect 3. The results are presented in Sect 4 and Sect. 5. We discuss the results in Sect 6 and conclude in Sect. 7.

All magnitudes are given in the AB system (Oke & Gunn 1983) and we use a cosmology with  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $h = 0.7$ .

## 2. Data

In this paper, we use deep spectra collected by the VUDS (Le Fèvre et al. 2015) and VANDELS (McLure et al. 2018; Pentericci et al. 2018). We briefly describe here both surveys and refer the reader to the survey description papers for more in-depth details. Both surveys have been carried out with the now-decommissioned Visible Multi-Object Spectrograph [VIMOS] instrument installed at the Nasmyth focus of Unit Telescope 3 of the Very Large Telescope.

VUDS (P.I. O. Le Fèvre) is a European Southern Observatory [ESO] large program designed to study the first billion years of galaxy evolution. It is based on the observation of ten thousand sources up to  $z = 6.5$  selected primarily via the photometric redshift method performed with the LePhare software (Arnouts et al. 1999; Ilbert et al. 2006). Each target was observed for a total of  $\sim 14\text{h}$  in the wavelength range  $3500 \leq \lambda \leq 9350\text{\AA}$  using the two low resolution grism of the instrument at  $R\sim 240$ .

The data reduction was carried out using the VIPGI software (Scodreggio et al. 2005). As the targets are in well known fields (VIMOS VLT Deep Survey [VVDS] 2h field, Cosmic evolution Survey [COSMOS] and Extended Chandra Deep Field South [ECDFS]), a large amount of multi-wavelength photometric data are available to complement the spectroscopic information. We direct the reader to Le Fèvre et al. (2015) for a full description of the survey.

The second sample comes from the VANDELS survey (P.I. R. McLure & L. Pentericci). VANDELS is an ESO public spectroscopic survey aiming at the exploration of the high redshift Universe at  $1 < z < 7$ . More than 2100 galaxies have been observed using the VIMOS medium resolution grism at  $R\sim 600$  from  $4800$  to  $10000\text{\AA}$  and the exposure time varied from  $20\text{h}$  up to  $80\text{h}$ . As it was the case for VUDS, most targets were selected using the photometric redshift technique. The data reduction was carried out using the EASYLIFE package (Garilli et al. 2012). Targets were selected in the two widely observed Ultra deep survey [UDS] and Chandra deep field south [CDFS] fields which allows the association of the spectroscopy to a large amount of multi-wavelength photometric data. Readers can find more information about the survey in the two papers presenting it: McLure et al. (2018) and Pentericci et al. (2018).

For both surveys, all the redshift estimations were performed using the EZ software (Garilli et al. 2010), a template cross-correlating code, and they both use an equivalent redshift flag system to assess the quality of the redshift measurements. In this scheme, galaxies with a redshift-flag 2,3,4 are those with the most reliable redshift measurements, with a probability to be correct of 75%, 95%, and 100%, respectively.

In this study, we are interested in galaxies with IGM transmissions higher and lower than the mean theoretical transmission. As defined in T17 and T20, this quantity is measured between rest-frame  $1070\text{\AA}$  and  $1170\text{\AA}$ . Such requirement imposes a lower redshift limit of  $z=3.7$  for VANDELS (taking a lower wavelength of  $\sim 5000\text{\AA}$ ) and  $z=2.7$  for VUDS (lower wavelength at  $\sim 3900\text{\AA}$ ). Finally, we select only galaxies with a redshift flag of 2, 3 or 4 which ensures to have the best quality redshift. This leads to a combined sample of 2502 galaxies including 296 galaxies from VANDELS and 2206 galaxies from VUDS. As the goal is to analyse galaxy spectra, it is worth mentioning that VANDELS spectra have been rebinned to match VUDS wavelength grid.

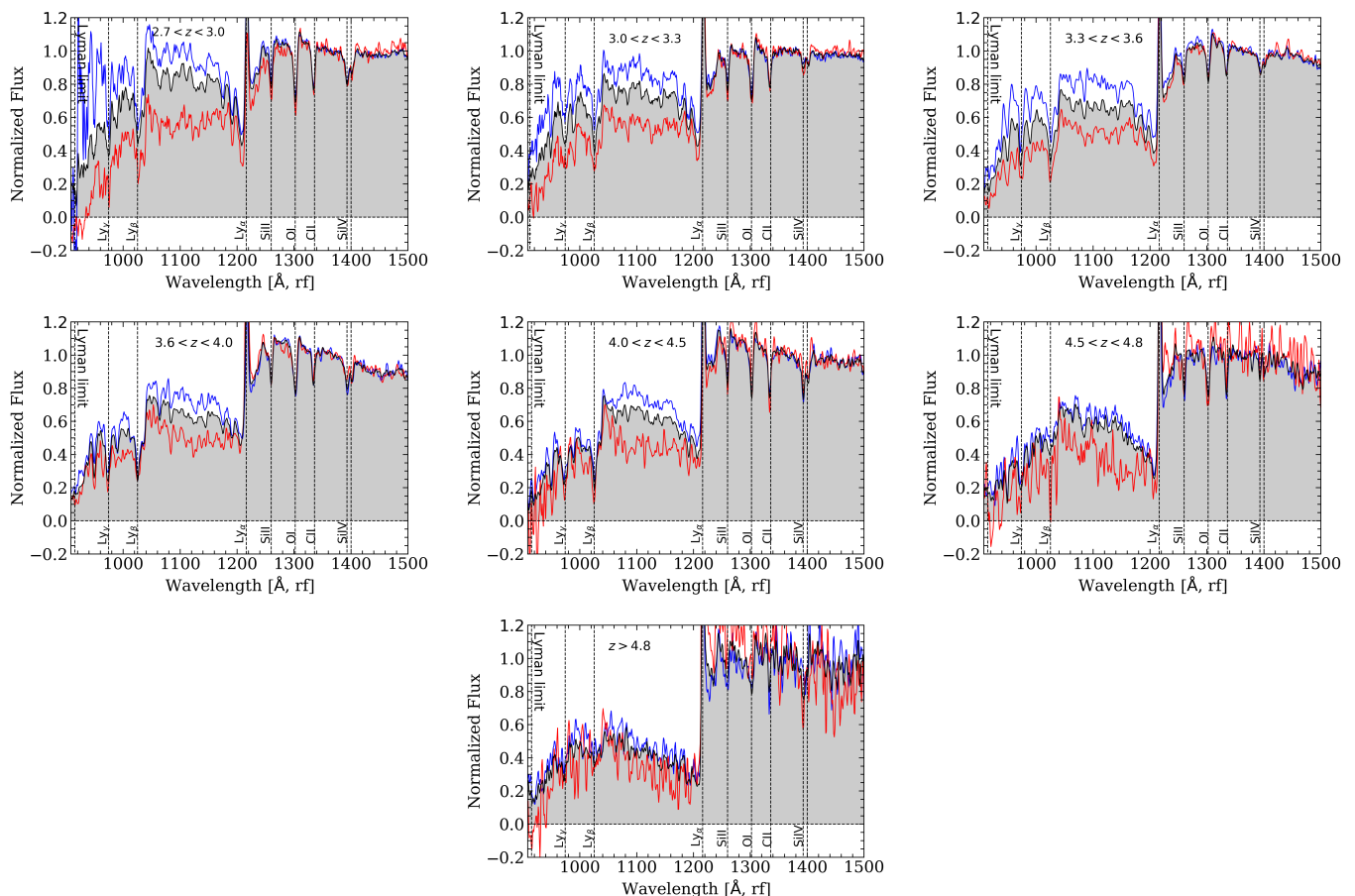
## 3. Method: The SPARTAN tool and IGM estimation

The estimation of the IGM transmission is done using the SPARTAN (Spectroscopic And photometRic fitting Tool for Astronomical aNalysis) software (Thomas 2021). It is a Python software created to be able to fit both photometry and spectroscopy in a common environment. We provide in this section the relevant aspects for this paper. SPARTAN uses a  $\chi^2$  minimization technique that consist on a comparison of the observations to theoretical galaxy models (see T20 for more details).

In order to select galaxies on the basis of their estimated IGM transmission, we need to measure it. This is done through the two-step template fitting method described in T20 which allows for the reduction of the degeneracy between dust extinction and IGM absorption. The first step is done to estimate the dust attenuation parameter  $E(B-V)$ , based on the SPARTAN fitting of all the photometric data point only. The second step is done fixing the dust extinction (from the first pass) on the spectral data which provides the IGM transmission. The fitting is

**Table 1.** Details of the stacked spectra displayed in Fig. 1. For each stack we provide the number of individual spectra entering the stacking of the spectrum ( $N_X$ ) and the average redshift of all these galaxies ( $\langle z_X \rangle$ ). Quantities with ‘all’, ‘more’, ‘less’ refer to the stack with all the galaxies (in grey in Fig.1), galaxies with the more transmitting IGM (blue) and galaxies with the less transmitting IGM (red). The last but one column shows the flux difference between the more IGM transmitted galaxies and the less IGM transmitted galaxies, computed at 1070-1170Å while the last column shows the fractional difference. It is worth mentioning that  $N_{all} \neq N_{more} + N_{less}$  as we are considering object  $0.5\sigma$  above ( $N_{more}$ ) or below ( $N_{less}$ ) the average.  $N_{all}$  include as well all galaxies with an IGM absorption equal to the theoretical average. Finally, the Signal to noise ratio (SNR), is the one computed in the stacked spectra in the IGM region at 1070-1170Å.

Redshift bin	$\langle z_{all} \rangle$	$N_{all}$	$\langle z_{more} \rangle$	$N_{more}$	$SNR_{more}$	$\langle z_{less} \rangle$	$N_{less}$	$SNR_{less}$	$\Delta \langle F \rangle_{1070-1170}$	$\frac{\Delta \langle F \rangle}{\langle F_{all} \rangle}$
$2.7 < z < 3.0$	2.8583	749	2.8643	333	17.8	2.8548	197	11.2	0.37	0.45
$3.0 < z < 3.3$	3.1392	546	3.1225	188	17.3	3.1540	183	13.4	0.31	0.41
$3.3 < z < 3.6$	3.4315	412	3.4285	102	15.8	3.4361	181	11.6	0.29	0.44
$3.6 < z < 4.0$	3.8133	305	3.7901	80	13.6	3.8146	105	8.2	0.23	0.36
$4.0 < z < 4.5$	4.2198	274	4.2545	113	14.2	4.1693	53	7	0.28	0.44
$4.5 < z < 4.8$	4.6209	122	4.6157	58	8.8	4.6472	16	2.7	0.24	0.42
$z > 4.8$	5.0411	94	5.0707	45	5	5.0464	18	2.6	0.14	0.34



**Fig. 1.** Average spectra with different IGM-attenuated sample. All spectra are given in the restframe (rf). We show 7 different redshift bins (from top left to bottom left):  $2.7 < z < 3.0$ ,  $3.0 < z < 3.3$ ,  $3.3 < z < 3.6$ ,  $3.6 < z < 4.0$ ,  $4.0 < z < 4.5$ ,  $4.5 < z < 4.8$  and  $z > 4.8$ . For each panel we show the averaged spectrum for the more transmitted galaxies in blue, the averaged spectrum of the less transmitted galaxies in red and the averaged spectrum of all the galaxies in grey

done using Bruzual & Charlot (2003) models with a Chabrier (2003) initial mass function. The stellar-phase metallicity ranges from sub-solar ( $0.2Z_{\odot}$  and  $0.4Z_{\odot}$ ) to solar ( $1.0Z_{\odot}$ ). We assume an exponentially delayed star formation history of the form  $SFR \propto t \times \tau^{-2} \times \exp(-t/\tau)$  with a time-scale parameter,  $\tau$ , ranging from 0.1 Gyr to 2.0 Gyr. The ages (corresponding to the time since onset of star formation) range from 0.01 Gyr to 4 Gyr. It is worth noting that this range of age is further limited by the age of the Universe at the redshift that is considered during the fit. For the photometric fitting, the  $E(B-V)_s$  parameter can vary from

0.0 to 0.39 (in 0.03 steps). As described above, this parameter is fixed during the second pass spectral fitting. Finally, the IGM prescription that is used from T17 in which we built IGM transmission templates around the mean of M06. At a given redshift, six empirical additional curves at  $\pm 0.5\sigma$ ,  $\pm 1.0\sigma$  and  $\pm 1.5\sigma$  were created. This allows us to use the IGM as a free parameter in our fit and explore a large range of IGM transmission. At  $z = 3$  the IGM transmission,  $Tr(Ly\alpha)$ , can range from 20% to 100% while at  $z = 5$ , it may vary from 5% to 50%. The  $Ly\alpha$  transmission is estimated directly on these models as the average of the selected

model between 1070Å and 1170Å (we refer the reader to T20 for examples of spectral fitting).

#### 4. The more and less IGM transmitted galaxies at $z > 2.7$

The objective of this paper is to look at the galaxies in different redshift bins which have an IGM transmission different from the theoretical mean. In order to achieve this goal, we create average spectra of galaxies selected based on their IGM transmission computed from the SPARTAN fitting. This has been done in seven redshift bins,  $2.7 \leq z < 3.0$ ,  $3.0 \leq z < 3.3$ ,  $3.3 \leq z < 3.6$ ,  $3.6 \leq z < 4.0$ ,  $4.0 \leq z < 4.5$ ,  $4.5 \leq z < 4.8$  and  $z > 4.8$ . The choice of these seven bins was a trade-off between maximizing the number of points and keeping a high number of galaxies to achieve a SNR high enough for our study. Nevertheless, It is worth mentioning that the change of this binning, e.g. with less bins, does not change our results. The galaxies in each bin were separated into two sub-samples: the more transmitted galaxies, i.e. those with an IGM transmission measured at a level higher than the predicted mean of the M06 prescription (with a selected model at  $\geq 0.5\sigma$ , see previous section) and the less transmitted galaxies, i.e., with an IGM transmission measured at a level lower than the mean of the M06 prescription (with a selected model at  $\leq 0.5\sigma$ ). For each set of galaxies we have created an average spectrum using the *specstack*<sup>1</sup> tool (Thomas 2019b). The tool works as follows. For a given set of galaxies, we de-redshift all the individual spectra and normalise them in a region redward of the Ly $\alpha$  line (in our case 1345Å-1380Å due to the absence of spectral feature in that region). Then we regrid the spectra in a common wavelength grid. Finally, at a given wavelength, we compute the average of all the flux density in each given pixel weighted by the individual spectra SNR using a  $3\sigma$  clipping method. In the seven redshift bins, the resulting average spectra based on our 2502 galaxies from both VUDS and VANDELS are presented in Fig. 1 and their associated data are displayed in Table 1. The averaged spectra were computed between 912Å and 1500Å rest-frame. In each redshift bin we show three average spectra: an average spectrum made from the more IGM-transmitted galaxies, an average spectrum made from the less IGM-transmitted galaxies and finally an average spectrum made from all the galaxies available in the considered redshift bin independent of the IGM transmission<sup>2</sup>.

First of all, it is important to note that 70% of the average spectra are made of more than 100 individual spectra, the other 30% are at the highest redshift bins where we have less spectra. This results in the average spectrum corresponding to the less IGM transmitted galaxies having a higher level of noise. In each redshift bin, the average redshift of that comprise the three composite spectra are similar and it is therefore meaningful to compare them directly. Very importantly, Fig.1 shows that redward, of the Ly $\alpha$  line, the spectra of the three samples are very close to each other which means that the underlying galaxy populations are similar, more notably, the dust properties. Therefore, the strong variation of flux below the Ly $\alpha$  line depends primarily on the IGM extinction along the line of sight. At every redshift, the less transmitted sample is much dimmer at 1070Å-1170Å than the more transmitted one, except for the highest redshift spectra (see next section). Based on the measurement on the average spectra, the flux density at 1070Å-1170Å of less

transmitted galaxies is on average 40% fainter than that of more transmitted galaxies. This difference will have a very important influence on the selection of galaxies based on the Lyman-break technique (as shown in T17 at  $z \sim 3.2$  with *ugr* colour-colour diagram). However, we note that for the two highest redshift bins, the flux density of the average sample (in gray) is very similar to that of more transmitted sample. In these bins, there are fewer galaxies in the sample which is less transmitting, therefore the average tends to be closer to the more transmitted one. We investigate the potential origin of this asymmetry in section 6.1. For a better visualisation, the evolution of the flux itself in this region is displayed in Fig.2 where all the regions have been redshifted to the average redshift of the bin (see Table 1).

We compute the average flux of the rest-frame stacked spectra in the IGM transmission region (1070-1170Å), and the evolution of this average flux is shown in Fig.3. As seen in the stacks, the average flux is decreasing with redshift as expected due to the evolution of the IGM (M06 & M95). Nevertheless, the evolution seems to be different for the two samples. We compute the difference in average flux between the more and the less transmitted sample,  $\Delta F$ . These measurements are reported in the last but one column of Table 1. We see that the difference is decreasing with increasing redshift. It goes from 0.37 at  $z \sim 2.85$  down to 0.14 at  $z \sim 5.04$ , indicating that the difference in transmission between the two sets of galaxies is reducing as well. We also computed the fractional difference,  $\Delta F / \langle F_{all} \rangle$  and reported it in the last column of Table 1. It slightly decreases with redshift going from 0.45 at  $z \sim 2.85$  to 0.34 at  $z \sim 5.04$ . Nevertheless, the scatter of the points does not let us draw any conclusion from this evolution. These evolutions are well fitted by a simple linear function,  $f(z) = a \times z + b$  (valid for taking all the points of Fig.3 from  $z_{source} = 2.85$  and  $z_{source} = 5.04$ ):

$$\begin{aligned} \langle F \rangle_{less} &= -0.101[\pm 0.008] \times z + 0.866[\pm 0.029] \\ \langle F \rangle_{more} &= -0.181[\pm 0.019] \times z + 1.447[\pm 0.072] \end{aligned} \quad (1)$$

The fitting method has been carried out using the `curve_fit` function of the `scipy.optimize` package (Virtanen et al. 2020) which performs a  $\chi^2$  minimization. The errors on the parameters correspond to the standard deviation estimated from the covariance matrix<sup>3</sup>.

#### 5. Optical depth

QSO studies traditionally compute the HI optical depth from the IGM transmission defined as:

$$\tau_{eff} = -\ln \text{Tr}(\text{Ly}\alpha), \quad (2)$$

where the  $\text{Tr}(\text{Ly}\alpha)$  is the Lyman  $\alpha$  transmission computed directly from the IGM template. We apply this transformation to each of our set of galaxies and the results are reported in Table. 2. It is also worth mentioning that QSO studies often use the redshift of the absorbers in the Lyman  $\alpha$  forest defined as:

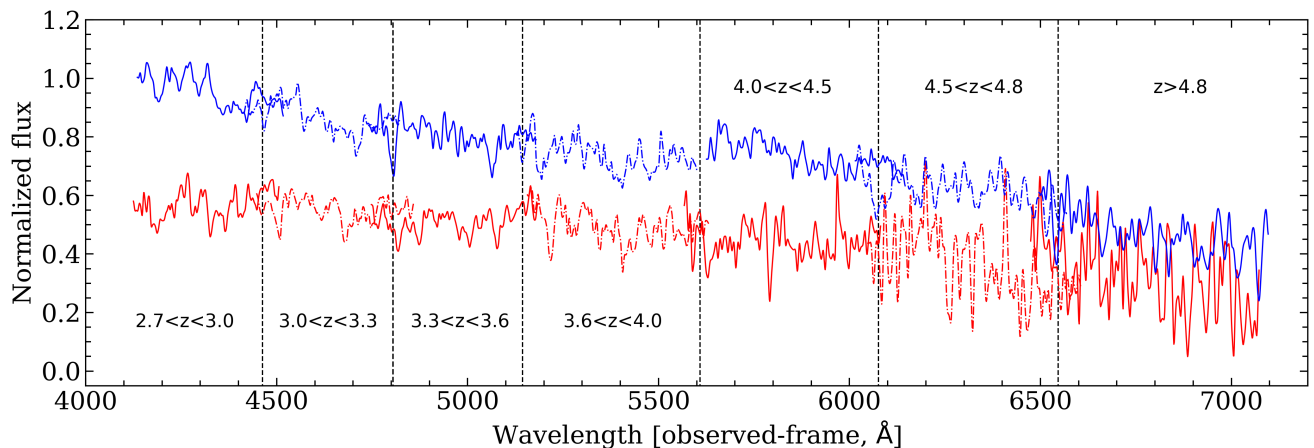
$$\lambda_{\text{Ly}\alpha}(1 + z_{abs}) = \lambda_0(1 + z_s), \quad (3)$$

with  $z_s$  the redshift of the source and  $\lambda_{\text{Ly}\alpha} = 1215.67\text{Å}$ . To be able to compare our measurements to the QSO's literature, we transform all our redshifts in this section to  $z_{abs}$  using  $\lambda_0 = 1120\text{Å}$ , the middle of the region where we measure  $\text{Tr}(\text{Ly}\alpha)$ .

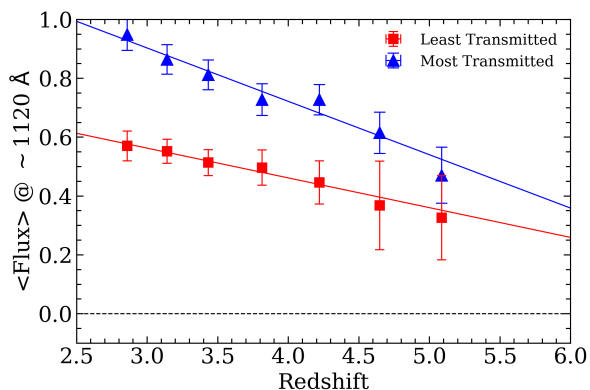
<sup>1</sup><https://specstack.readthedocs.io/en/latest/>

<sup>2</sup>The evolution of the Ly $\alpha$  transmission itself is shown in T20.

<sup>3</sup>[https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve\\_fit.html](https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html)



**Fig. 2.** Evolution of the region where the  $\text{Ly}\alpha$  transmission is computed (1070-1170Å) in the different redshift bins of Fig.1 but this time in the observed frame (the redshift of the spectra are set to the average redshift of the redshift bin). The vertical dashed lines show the separation between each redshift bin (in the middle of the overlap). The color coding is the same as Fig.1.



**Fig. 3.** Evolution of the averaged flux in the 1070-1170Å region in the averaged spectra for the more (in blue) and the less (red) IGM-transmitted average spectra. The two solid lines are linear fits to the evolutions whose functional shapes are given in equation 1.

**Table 2.** Optical depth from the IGM transmission in each of our galaxy subsamples. Quantities with ‘all’, ‘more’, ‘less’ refer to the stack with all the galaxies (in grey in Fig.1), galaxies with the more transmitting IGM (blue) and galaxies with the less transmitting IGM (red). Redshift have been transformed using eq.3.

$z_{abs}^{all}$	$\tau_{eff}^{all}$	$z_{abs}^{more}$	$\tau_{eff}^{more}$	$z_{abs}^{less}$	$\tau_{eff}^{less}$
2.55	$0.21 \pm 0.1$	2.56	$0.01 \pm 0.05$	2.55	$0.54 \pm 0.04$
2.81	$0.30 \pm 0.1$	2.79	$0.06 \pm 0.02$	2.83	$0.61 \pm 0.04$
3.08	$0.41 \pm 0.12$	3.07	$0.13 \pm 0.03$	3.09	$0.70 \pm 0.05$
3.43	$0.56 \pm 0.15$	3.41	$0.23 \pm 0.09$	3.44	$0.86 \pm 0.07$
3.80	$0.64 \pm 0.15$	3.84	$0.41 \pm 0.09$	3.76	$0.96 \pm 0.05$
4.17	$0.87 \pm 0.19$	4.17	$0.64 \pm 0.07$	4.20	$1.23 \pm 0.06$
4.56	$1.16 \pm 0.22$	4.59	$0.91 \pm 0.12$	4.57	$1.61 \pm 0.1$

We compare our estimation to QSO’s data with data from Dall’Aglio et al. (2008), Faucher-Giguère et al. (2008), Becker et al. (2013), Becker et al. (2015) and galaxy based optical depth estimation from Monzon et al. (2020). Our measurement from all our galaxies are in excellent agreement with each of these sample over the whole redshift range. At the high redshift end of our sample, the scatter of Becker et al. (2015) also reaches our less/more transmitted galaxies. We fit the evolution of the optical

depth for each of our sample with an analytical function of the form:

$$\tau_{eff}(z) = A \times (1 + z)^\gamma \quad (4)$$

The fit is carried out using the same algorithm as Sect. 4 using the scipy package. We find, for the three sample:

- All galaxies:  $A = 2.655 \times 10^{-3} \pm 0.001$ ,  $\gamma = 3.536 \pm 0.170$
- More transmitted:  $A = 9.540 \times 10^{-6} \pm 6.397 \times 10^{-6}$ ,  $\gamma = 6.717 \pm 0.413$
- less transmitted:  $A = 0.026 \pm 0.006$ ,  $\gamma = 2.354 \pm 0.147$

It is worth mentioning that the result of this fit is valid for the redshift range we consider here from  $z_{source}=2.85$  to  $z_{source}=5.04$ .

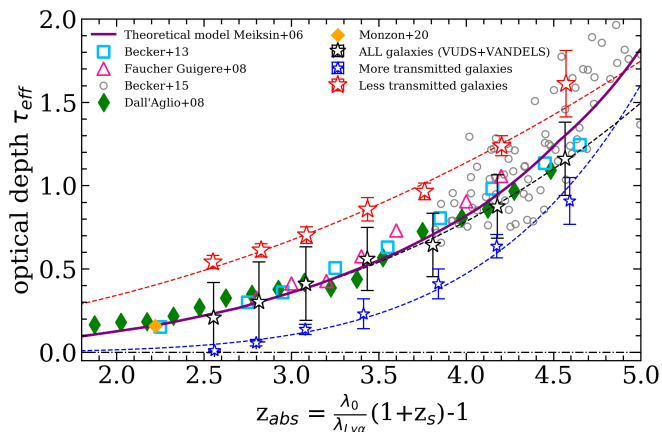
Finally, as we can see, the evolution of the optical depth for both more/less transmitted galaxies is well separated at any redshift which would indicate that some galaxies may become visible, below the  $\text{Ly}\alpha$  line, at very different redshift depending only on their IGM optical depth. At a fixed optical depth, the difference in redshift reaches  $\geq 1$ .

## 6. Discussion

### 6.1. Asymmetry in the distribution of more/less transmitting lines of sight.

As reported in section 4, we observe that in the highest redshift bins, the number of galaxies with a LOS classified as less transmitting is smaller than the the number of classified galaxies with a LOS classified as more transmitting. We investigate this asymmetry by looking at the two different surveys independently (see Table 3).

Taking into account the VANDELS survey alone the distribution of galaxies between less and more transmitting line of sights is not showing this asymmetry. On the other hand, when looking at VUDS data alone, the effect is present. This most probably indicates that the two pass fitting method is not completely efficient when applied to the VUDS data. This couple be explained by the the number of points used for the dust estimation using photometric SED-fitting (we use up to 20 bands for VANDELS



**Fig. 4.** Evolution of the optical depth for each sample of our galaxies and comparison with QSO's data. More IGM transmitted galaxies are shown by the blue stars, less transmitted galaxies are in red stars and black star represent the full sample of our galaxies. Dashed lines represent the fit to each sample and are in the same colors. We also show point from the literature with measurements from Dall'Aglio et al. 2008 in green diamonds, from Faucher-Giguère et al. 2008 in pink triangles, from Becker et al. 2013 in light blue square and from Becker et al. 2015 in empty gray circles. The orange point is based on galaxy data from Monzon et al. (2020). Finally, the purple plain line shows the Meiksin modelisation (shifted to the  $z_{abs}$ , see Eq. 3).

VANDELS	$N_{gal_{more}}$	$N_{gal_{less}}$
$4.0 < z < 4.5$	39	43
$4.5 < z < 4.8$	18	14
$z > 4.8$	12	14
VUDS	$N_{gal_{more}}$	$N_{gal_{less}}$
$4.0 < z < 4.5$	77	10
$4.5 < z < 4.8$	40	2
$z > 4.8$	33	4

**Table 3.** Number of galaxies with IGM transmission lower and higher than the theoretical mean from M06.

and around 10 for VUDS). This can be partially resolved by setting lower limits on the photometric fitting, i.e. forcing a lower dust attenuation as shown in T17. This would result in a more symmetrical distribution as for the VANDELS sources. The effect of such action would be to remove the slight jump that can be seen in Fig. 3 at  $z > 4$  and shift the measurements of the optical depth closer to the Meiksin model (see Fig. 4).

## 6.2. $\gamma$ values

As shown in the previous section, the fits of Fig. 4 show large differences in parameters range, especially for the  $\gamma$  parameters that ranges from 2.3546 for the galaxies that are less transmitted to 6.3974 for the more transmitting galaxies. In the Lyman alpha forest studies framework, the  $\gamma$  parameter is related to the evolution of the number density of Ly $\alpha$  forest clouds with redshift (Press et al. 1993). In the same paper, the evolution of the optical depth is fitted by a functional of the form  $\tau_{eff}(z) = A(1+z)^\beta$ , with  $\beta = \gamma + 1$ . In this case the  $\beta$  value can be used directly to estimate the evolution of the number of clouds with redshift given by:

$$\frac{dN}{dz} \propto B(1+z)^{\beta-1}. \quad (5)$$

In that framework, we must remove 1 to all our values of  $\gamma$  presented in Sect. 5. We see that for the less transmitted galaxies, the  $\gamma$  value is the lowest. This indicates a slower evolution with redshift. These line of sight are the one that are the most populated by clouds which leads to a lower transmission. This slower evolution can be explained by the fact that the strong increase of the optical depth at high- $z$  (Becker et al. 2015) is less prominent for these line of sights because they are already heavily populated by clouds. On the contrary, the more transmitted galaxies are fitted with a high  $\gamma$  value which indicates a strong evolution with redshift. For these galaxies, the lines of sight are less populated by clouds, leading to a higher IGM transmission. The strong evolution of the optical depth at high- $z$  translates in a rapid increase of the number of clouds along the line of sight with redshift for the more transmitted galaxies.

## 7. Conclusions

Using data from both VUDS and VANDELS programs, we presented in this letter the effect of the IGM at redshift  $z > 2.7$ . Using averaged spectra we could observe the effect of the IGM for lines of sights that are more and less transmitting than the theoretical average prescription. This leads to two important results:

- Firstly, the variance of the IGM is clearly visible from galaxy data. The flux for the sub-sample with the more transmitting lines of sight is on average 63% higher in the 1070-1170Å region than the sub-sample with smaller IGM transmission. It has been already reported from the quasar point of view, e.g. Songaila (2004). From the theoretical point of view, the large variance of the IGM was already shown by Madau (1995) but not emphasized by more recent works. We defer a comprehensive comparison between our galaxy data with simulations to a future study.
- The estimation of the optical depth from the IGM transmission allows us to compare our data to multiple QSO studies and shows an excellent agreement. Finally, the observed variance of the IGM described above could lead to different full-opacity redshift for different categories of lines of sight. The evolution of the optical depth from both more and less attenuated galaxies suggests that galaxies could start to be visible below Lyman alpha at very different redshift; at  $z_s \sim 6.16$  for less transmitted galaxies and at  $z_s \sim 6.80$  for the more transmitted galaxies. Nevertheless, QSOs studies at  $z > 5$  indicate that the evolution of the optical depth would become steeper as we increase redshift.

*Acknowledgements.* We wish to thank the anonymous referee for a careful reading of our manuscript and a insightful comments and corrections. We thank the ESO staff for their continuous support for the VANDELS survey, particularly the Paranal staff, who helped us to conduct the observations, and the ESO user support group in Garching. VUDS Data products are made available at the CESAM data center, Laboratoire d'Astrophysique de Marseille, France. RA acknowledges support from FONDECYT Regular Grant 1202007. All the plots have been made using the Photon software (Thomas 2019a) and dfitspy (Thomas 2019c) was used for fits file handling. We would like to dedicate this paper to the memory of Olivier Le Fèvre, PI of the VUDS survey and co-I of the VANDELS survey.

## References

- Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540  
Becker, G. D., Bolton, J. S., Madau, P., et al. 2015, MNRAS, 447, 3402  
Becker, G. D., Hewett, P. C., Worseck, G., & Prochaska, J. X. 2013, MNRAS, 430, 2067  
Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

- Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, *ApJ*, 437, L9  
Chabrier, G. 2003, *PASP*, 115, 763  
Dall'Aglio, A., Wisotzki, L., & Worseck, G. 2008, *A&A*, 491, 465  
Faucher-Giguère, C.-A., Prochaska, J. X., Lidz, A., Hernquist, L., & Zaldarriaga, M. 2008, *ApJ*, 681, 831  
Garilli, B., Fumana, M., Franzetti, P., et al. 2010, *PASP*, 122, 827  
Garilli, B., Paoro, L., Scodreggio, M., et al. 2012, *PASP*, 124, 1232  
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *A&A*, 457, 841  
Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, *MNRAS*, 442, 1805  
Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2015, *A&A*, 576, A79  
Madau, P. 1995, *ApJ*, 441, 18  
McLure, R. J., Pentericci, L., Cimatti, A., et al. 2018, *MNRAS*, 479, 25  
Meiksin, A. 2006, *MNRAS*, 365, 807  
Monzon, J. S., Prochaska, J. X., Lee, K.-G., & Chisholm, J. 2020, *AJ*, 160, 37  
Oke, J. B. & Gunn, J. E. 1983, *ApJ*, 266, 713  
Pentericci, L., McLure, R. J., Garilli, B., et al. 2018, *A&A*, 616, A174  
Press, W. H., Rybicki, G. B., & Schneider, D. P. 1993, *ApJ*, 414, 64  
Scodreggio, M., Franzetti, P., Garilli, B., et al. 2005, *PASP*, 117, 1284  
Songaila, A. 2004, *AJ*, 127, 2598  
Steidel, C. C., Pettini, M., & Hamilton, D. 1995, *AJ*, 110, 2519  
Thomas, R. 2019a, *Astrophysic Source Code Library*, [record ascl:1901.007]  
Thomas, R. 2019b, *Astrophysic Source Code Library*, [record ascl:1904.018]  
Thomas, R. 2019c, *The Journal of Open Source Software*, 4, 1259  
Thomas, R. 2021, *Astronomy and Computing*, 34, 100427  
Thomas, R., Le Fèvre, O., Le Brun, V., et al. 2017, *A&A*, 597, A88  
Thomas, R., Pentericci, L., Le Fèvre, O., et al. 2020, *A&A*, 634, A110  
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261