

# The VLT LBG redshift survey – V. Characterizing the $z = 3.1$ Lyman $\alpha$ emitter population

R. M. Bielby,<sup>1\*</sup> P. Tummuangpak,<sup>1,2</sup> T. Shanks,<sup>1</sup> H. Francke,<sup>3</sup> N. H. M. Crighton,<sup>4</sup>  
E. Bañados,<sup>5,6</sup> Jorge González-López,<sup>5</sup> L. Infante<sup>5</sup> and A. Orsi<sup>7</sup>

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

<sup>2</sup>Department of Physics, Khon Kaen University, Muang, Khon Kaen, 40002, Thailand

<sup>3</sup>Joint ALMA Observatory/European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago, Chile

<sup>4</sup>Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Victoria 3122, Australia

<sup>5</sup>Institute of Astrophysics, Center for Astro-Engineering, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile

<sup>6</sup>Max Planck Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany

<sup>7</sup>Centro de Estudios de Física del Cosmos de Aragón, Plaza San Juan 1, Planta-2, E-44001 Teruel, Spain

Accepted 2015 December 10. Received 2015 December 10; in original form 2014 December 15

## ABSTRACT

We present a survey of  $z \sim 3$  Ly $\alpha$  emitters (LAEs) within the fields of the VLT Lyman break galaxies (LBG) redshift survey. The data encompass five independent survey fields co-spatial with spectroscopic LBG data and covering a larger total area than previously analysed for LAE number counts and clustering. This affords an improved analysis over previous work by minimizing the effects of cosmic variance and allowing the cross-clustering analysis of LAEs and LBGs. Our photometric sample consists of  $\approx 600$  LAE candidates, over an area of  $1.07 \text{ deg}^2$ , with equivalent widths of  $\gtrsim 65 \text{ \AA}$  and a flux limit of  $\approx 2 \times 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1}$ . From spectroscopic follow-up, we measured a success rate of  $78 \pm 18$  per cent. We find the *R*-band continuum luminosity function to be  $\sim 10$  times lower than the luminosity function of LBGs at this redshift, consistent with previous studies. Exploiting the large area of the survey, we estimate the LAE auto-correlation function and find a clustering length of  $r_0 = 2.86 \pm 0.33 h^{-1} \text{ Mpc}$ , low compared to the  $z \sim 3$  LBG population, but somewhat higher than previous LAE measurements. This corresponds to a median halo mass of  $M_{\text{DM}} = 10^{11.0 \pm 0.3} h^{-1} M_{\odot}$ . We present an analysis of clustering length versus continuum magnitude and find that the measurements for LAEs and LBGs are consistent at faint magnitudes. Our combined data set of LAEs and LBGs allows us to measure, for the first time, the LBG–LAE cross-correlation, finding a clustering length of  $r_0 = 3.29 \pm 0.57 h^{-1} \text{ Mpc}$  and a LAE halo mass of  $10^{11.1 \pm 0.4} h^{-1} M_{\odot}$ . Overall, we conclude that LAEs inhabit primarily low-mass haloes, but form a relatively small proportion of the galaxy population found in such haloes.

**Key words:** galaxies: evolution – galaxies: high-redshift – galaxies: luminosity function, mass function – large-scale structure of Universe.

## 1 INTRODUCTION

Two key methods for efficiently selecting high-redshift galaxies are via their Lyman break feature (Lyman break galaxies – LBGs) and via their Lyman  $\alpha$  emission (Lyman  $\alpha$  emitters – LAEs). Selecting high-redshift galaxies through their strong emission in the Ly $\alpha$  feature using the narrow-band imaging method has come to be a very effective technique to isolate high-redshift galaxies. There are many observations that have been made to uncover galaxies with

strong Ly $\alpha$  emission at various redshifts (e.g. Hu & McMahon 1996; Steidel et al. 1996; Cowie & Hu 1998; Hu, Cowie & McMahon 1998; Ouchi et al. 2003, 2008, 2010; Hayashino et al. 2004; Gawiser et al. 2007; Gronwall et al. 2007; Rauch et al. 2008; Bañados et al. 2013).

At redshift  $z \sim 3$ , there are a number of photometric and spectroscopic LAE samples (e.g. Steidel et al. 2000; Fynbo et al. 2003; Hayashino et al. 2004; Matsuda et al. 2005; Venemans et al. 2007; Nilsson et al. 2007; Gawiser et al. 2007; Gronwall et al. 2007; Rauch et al. 2008; Ouchi et al. 2008; McLinden et al. 2014). Gronwall et al. (2007) surveyed the Extended *Chandra* Deep Field-South (ECDFS), identifying  $\approx 160$  photometrically selected  $z = 3.1$

\* E-mail: richard.bielby@durham.ac.uk

LAEs over an area of  $0.28 \text{ deg}^2$  and to a Ly $\alpha$  luminosity limit of  $L_{\text{Ly}\alpha} \gtrsim 10^{42} \text{ erg s}^{-1}$ . Their colour constraints correspond to a rest-frame equivalent width limit of  $\text{EW} \gtrsim 20 \text{ \AA}$ . By measuring the Ly $\alpha$  and rest-frame UV continuum luminosity functions in combination with the equivalent width distribution, they surmised that LAEs contain a non-negligible amount of dust and are therefore not of a primordial origin (but that they do represent a young galaxy population.). Ouchi et al. (2008) presented a survey of  $\approx 350$  photometrically identified (with 41 spectroscopically confirmed) LAEs over  $\approx 1 \text{ deg}^2$  in the *Subaru/XMM*–Newton Deep Survey (SXDS) field to a limiting Ly $\alpha$  luminosity of  $L_{\text{Ly}\alpha} \gtrsim 10^{42} \text{ erg s}^{-1}$  (and a rest-frame equivalent width limit of  $\gtrsim 65 \text{ \AA}$ ). Combining their  $z = 3.1$  data with higher redshift samples, they find that LAEs (at a given equivalent width and luminosity limit) are more common at earlier epochs and are a tracer of either low-extinction (when compared to LBGs) or younger stellar populations.

Although LAEs in general have been widely studied, their clustering properties remain comparatively sparsely studied, due to the need for large survey fields. At  $z \approx 3$ , Hayashino et al. (2004), Gawiser et al. (2007) and Ouchi et al. (2010) have each presented measurements of the LAE angular correlation function. Hayashino et al. (2004) do not present a quantification of their clustering result, whilst Gawiser et al. (2007) and Ouchi et al. (2010) measure clustering lengths of  $r_0 = 2.52^{+0.6}_{-0.7} h^{-1} \text{ Mpc}$  and  $r_0 = 1.70^{+0.39}_{-0.46} h^{-1} \text{ Mpc}$ , respectively. These two measurements are based on differing equivalent width constraints of  $\text{EW} > 20 \text{ \AA}$  and  $\text{EW} > 65 \text{ \AA}$ , respectively, with the lower EW cut giving a larger clustering length (although the two measurements only differ by  $\approx 1\sigma$ ). At higher redshifts, LAE clustering measurements have been obtained at  $z \sim 4$  (Kovač et al. 2007; Ouchi et al. 2010),  $z \sim 5$  (Ouchi et al. 2003; Shimasaku et al. 2004; Shioya et al. 2009), and  $z \sim 6$ – $7$  (Ouchi et al. 2010). In each case, these clustering measurements are based on single observational fields of at most  $1 \text{ deg}^2$ , whilst in some cases (e.g. Hayashino et al. 2004, at  $z = 3.1$ ) they are actively targeted on high-density regions. Indeed, cosmic variance can be a significant source of error in LAE clustering studies (Shimasaku et al. 2004; Francke 2009) and measurements based on single fields can prove unrepresentative of the population as a whole. The largest of these works thus far remains that of Ouchi et al. (2010), which presented LAE clustering results at  $z = 3.1$ ,  $z = 3.7$ ,  $z = 5.7$ , and  $z = 6.6$ . An apparent evolution in the LAE clustering was observed, with  $r_0$  (and therefore the clustering bias) increasing with increasing redshifts, implying a constant host halo mass with redshift of  $\approx 10^{11} h^{-1} \text{ Mpc}$ .

Here we present observations of LAEs in the deep imaging fields of the VLT LBG Redshift Survey (VLRS) aimed at studying the clustering properties of LAE galaxies. Our LAE sample extends over five independent observational fields each measuring  $\sim 0.5 \times 0.5$  in area, thus combatting the effects of cosmic variance. These fields were selected based on the presence of bright quasars at  $z \gtrsim 3$  and so are not specifically targeted at galaxy overdensities as with some previous works (e.g. Hayashino et al. 2004). We target the

$z \approx 3$  LAE population in order to overlap with our  $z \approx 3$  spectroscopic LBG sample in these fields, allowing cross-analysis between the two populations. The primary aim of the VLRS has been to probe the intergalactic medium (IGM) of  $z \sim 3$  galaxies by observing galaxies close quasi-stellar object (QSO) sightlines (Crighton et al. 2011; Tummuangpak et al. 2014). At present the survey consists of nine fields in total, with spectroscopic observations of LBGs taken with VLT VISIBLE Multi-Object Spectrograph (VIMOS). The primary survey is described in detail by Bielby et al. (2011, 2013). By adding narrow-band data to our survey data, we aim to enhance the scope of the survey, probing the volume for both LAEs and other sources of Ly $\alpha$  photons (e.g. Ly $\alpha$  blobs and Ly $\alpha$  emission from and around QSOs).

In this paper, we present the narrow-band data and analysis of LAE galaxies within the VLRS. The data cover five independent survey fields, giving the significant advantage of reducing the potential impact of cosmic variance on our results compared to past LAE surveys. We discuss our deep imaging data in these fields and the selection of LAEs at  $z \approx 3.1$  using the NB497 narrow-band filter in Section 2. We then report on spectroscopic follow up observations and our optimized selection criteria based on these in Section 3. In Section 4, we present narrow-band number counts and the rest-frame UV continuum luminosity function for our LAE sample, whilst in Section 5 we present an analysis of the clustering of the LAEs. Finally, we present a discussion of our results in Section 6 and our conclusions in Section 7.

Throughout this work, we adopt a cosmology given by  $(\Omega_m, \Omega_\Lambda, h, \sigma_8) = (0.25, 0.75, 0.73, 0.83)$ , where  $h$  is the Hubble constant in units of  $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . Distances and volumes are given in comoving coordinates unless otherwise stated. All magnitudes are given in the AB system unless otherwise stated.

## 2 IMAGING DATA

The VLRS consists of nine deep fields (each measuring between  $0.25 \text{ deg}^2$  and  $1 \text{ deg}^2$ ) centred on bright  $z \gtrsim 3$  QSOs and containing both deep broad-band photometry and spectroscopy (Bielby et al. 2011, 2013). We have added to these data by using the *Subaru* Suprime-Cam instrument (Miyazaki et al. 2002) to obtain deep narrow-band (497 nm) imaging in five of the VLRS fields, which allows for identification of LAEs at  $z \sim 3.1$  (details in Section 3.2). The observations were taken on 2009 September 19.

Suprime-Cam is a mosaic CCD camera with 10 CCDs, measuring  $2048 \times 4096$  pixels, which covers a  $34 \times 27 \text{ arcmin}^2$  field of view with a pixel scale of  $0.20 \text{ arcsec}$ . We observed  $\sim 0.5 \times 0.5$  fields centred on the bright quasars QSO B2359+068, LBQS 0301–0035, QSO J0124+0044, PKS 2126–158, and LBQS 2231–0015 (details are in Table 1). These fields were observed with the narrow-band [O III] NB497 filter ( $4977 \text{ \AA}$ , full width at half maximum (FWHM)  $77 \text{ \AA}$ ; Hayashino et al. 2003). Individual exposure times were 1200 s for each frame and the seeing was generally sub-arcsec

**Table 1.** Details of narrow-band imaging (NB497) from the *Subaru*/Suprime-Cam.

Central QSO	R.A. (J2000)	Dec.	QSO redshift	$T_{\text{exp}}$ (s)	Seeing (arcsec)	$M_{\text{ZP}}$	$1\sigma$ depth
QSO B2359+068	00:01:40.57	+07:09:54.1	3.234	6000	0.61	31.94	28.38
QSO J0124+0044	01:24:03.78	+00:44:32.7	3.807	4800	0.55	31.93	27.47
LBQS 0301–0035	03:03:41.05	–00:23:21.9	3.175	7200	0.60	31.93	28.57
PKS 2126–158	21:29:12.18	–15:38:41.0	3.268	6000	0.96	31.89	27.90
QSO B2231–0015	22:34:08.99	+00:00:01.7	3.027	6000	0.87	31.93	27.77

**Table 2.** Details of broad-band imaging observations from Bielby et al. (2011, 2013).

Central QSO	Band	$T_{\text{exp}}$ (ks)	Seeing (arcsec)	$1\sigma$ depth	Instrument
QSO B2359+068	<i>B</i>	7.2	1.45	28.23	KPNO/MOSAIC
	<i>R</i>	6.0	1.15	27.76	
QSO J0124+0044	<i>B</i>	2.8	1.50	27.99	KPNO/MOSAIC
	<i>V</i>	3.1	1.40	27.74	
LBQS 0301–0035	<i>B</i>	6.4	1.28	28.30	KPNO/MOSAIC
	<i>R</i>	4.8	1.19	27.66	
PKS 2126–158	<i>B</i>	7.8	1.60	28.44	CTIO/MOSAIC2
	<i>R</i>	6.4	1.50	28.27	
QSO B2231–0015	<i>B</i>	13.2	1.01	27.28	WFCam (INT)
	<i>R</i>	19.2	1.01	26.98	

( $\sim 0.6$ – $0.7$  arcsec). For the purposes of photometric calibration, we observed the standard star LTT 9491 (R.A. 23:19:34.98, Dec.  $-17:05:29.8$  J2000).

In order to select LAE galaxies, we also require *B* and *V* or *R* deep broad-band imaging, which are available from our previous work in these fields. The full details of the broad-band data and its reduction are provided by Bouché & Lowenthal (2004) and Bielby et al. (2011, 2013), whilst a summary of the broad-band data is given in Table 2.

## 2.1 Data reduction

To reduce the Suprime-Cam raw narrow-band data, we used the pipeline software SDFRED (the Suprime-Cam Deep field REDuction package,<sup>1</sup> Ouchi et al. 2004a) which comprises IRAF, SExtractor (Bertin & Arnouts 1996), and the mosaic-CCD data reduction software (Yagi et al. 2002). The package includes bias subtraction, flat fielding, distortion+atmospheric dispersion corrections, matching the PSF size, sky subtraction, masking vignetting caused by the auto-guider (AG), masking bad pixels, image alignments and scaling, and mosaicking. The reduction procedure is briefly described below.

The images were first bias subtracted using the median value of the overscan region on each line of a given file (the Suprime-CAM CCDs are noted to have very little bias pattern, so only the subtraction of the overscan median is required.) and the overscan regions were then trimmed from the images. The flat-field was then created using a total of 25 dithered object frames from across the five observed fields, with objects and regions vignettted by the AG probe masked. With this flat-field applied to the individual object frames, the astrometric distortion correction was performed (correcting for the telescope optics and the differential atmospheric dispersion) using a fifth-order polynomial transformation (Miyazaki et al. 2002). The point spread function (PSF) was then matched across the images by applying a Gaussian smoothing kernel to individual images. The sky background was then determined by interpolating over a mesh pattern. The image was divided into a grid of  $64 \times 64$  pixel squares (corresponding to  $12.9 \times 12.9$  arcsec<sup>2</sup>). A bilinear interpolation was used to determine the global sky background from this grid and the result was subtracted from the individual science frames.

An initial astrometry solution was calculated using SDFRED, whereby the images were matched internally (i.e. using one of

the science frames as a reference) and matched to a reference stellar catalogue. The images from each field were then co-added into stacked final images.

We applied the cosmic ray rejection using the rejected-mean algorithm, CRREJECT, from IRAF.

## 2.2 Object detection and photometry

Extraction of sources from the images was performed using SExtractor (Bertin & Arnouts 1996). SExtractor was run in dual image mode using the narrow-band image as the detection image for each of the broad-band images and the narrow-band image itself. Photometric zero-points were calculated using the narrow-band observations of LTT 9491. As there is no NB497 narrow-band standard photometry available for the standard star, we assumed the narrow-band magnitude to be equal to the quoted *V*-band magnitude of  $V = 14.06$  (the  $B - V$  colour of this star is  $+0.03$ ). The zero-point magnitude for our observations given a 1 s exposure was  $M_{\text{ZP}} = 24.34$ , which equates to  $M_{\text{ZP}} = 32.04$  for our 1200 s exposures. Taking into account atmospheric absorption for each field, the airmass ranged between  $\approx 1.1$  and 1.5. Taking an extinction coefficient of  $k = 0.12$  gives a range of extinction corrections of  $\approx 0.13$ – $0.19$ . The resulting individual zero-points are given in Table 1.

Aperture magnitudes were measured using an aperture diameter of 3.0 arcsec (approximately twice the seeing FWHM of the broad-band data). Total magnitudes were measured using MAG\_AUTO.  $1\sigma$  magnitude depths were estimated using the errors calculated in 3.0 arcsec diameter apertures in SExtractor.

These initial catalogues were used for the selection of targets for the VLT VIMOS observations described in the following section.

## 2.3 Improved astrometry, object detection and photometry

The initial astrometric solution was discovered to be inaccurate during the analysis of the VIMOS spectroscopic data (as described in Section 3). We therefore re-calibrated the astrometry and produced improved photometric catalogues for all of our fields.

This re-calibration of the astrometric solution was performed using SCAMP (Bertin 2006) with each stacked narrow-band image in conjunction with the associated VLRS broad-band images, thus providing a consistent astrometric solution across the narrow- and broad-band images. Since we applied the geometric transformation to all images, the object positions in each broad-band image were matched with objects in the *NB* band. The distortion correction in

<sup>1</sup> <http://www.naoj.org/Observing/DataReduction/>

**Table 3.** Details of the VIMOS LAE observations.

Field	R.A. (J2000)	Dec.	Constraints	$N_{\text{LAE}}$ (Sel/obs)	Seeing (arcsec)	Integration time (h)
QSO B2359+068	00:02:11.457	+07:15:32.25	$(R - m_{\text{NB497}}) > 0.2$ $(B - m_{\text{NB497}}) > 0.9$	304/37	1.0	2.0
QSO J0124+0044	01:24:36.236	+00:51:07.19	$(V - m_{\text{NB497}}) > 1.1$ $(B - m_{\text{NB497}}) > 1.4$	141/45	1.0	4.3
LBQS 0301-0035	03:03:10.208	-00:16:20.98	$(R - m_{\text{NB497}}) > 0.8$ $(B - m_{\text{NB497}}) > 1.4$	214/46	1.0	3.3

the reduction process corrected the geometric distortion and we obtained good astrometry to better than  $\pm 0.2$  arcsec rms over the image. Astrometry was made based on USNO at ESO catalogue and  $\sim 1000$  stars identified in the stacked images. The position of USNO objects was approximately uniformly distributed over the entire stacked images. The absolute coordinates of our objects were obtained from these USNO objects. The images were then transformed based on the SCAMP solution using SWARP (Bertin et al. 2002).

We then produced two sets of catalogues based on these distortion-corrected images, both sets using SExtractor in dual image with one set using the narrow-band images as the detection image and the second set using a  $\chi^2$  combined image as the detection image. The  $\chi^2$  combined image was produced using SWARP and in each field combines the narrow-band and  $B$  and  $R$  (or  $V$ ) images in an optimal way for source detection (see Bertin et al. 2002, for details).

### 3 SPECTROSCOPIC OBSERVATIONS

#### 3.1 Overview

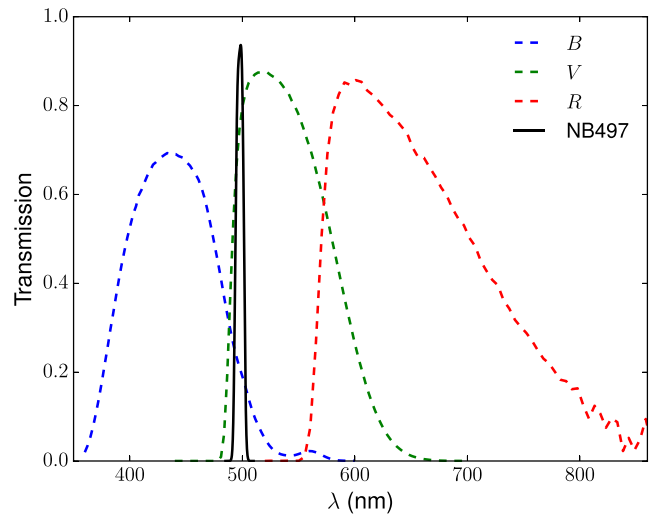
We have made spectroscopic follow-up observations of the LAE candidates with the VIMOS (Le Fèvre et al. 2003) on the Very Large Telescope (VLT). The VIMOS focal plane is divided into four quadrants, each measuring  $7 \times 8$  arcmin<sup>2</sup> (with gaps of  $\sim 2$  arcmin between adjacent detectors). We observed with the HR\_Blue grism, allowing us to observe  $\sim 20$ – $30$  objects per detector with a spectral coverage of  $4150 < \lambda < 6000$  and a resolution of  $R \sim 2050$  (given our 1 arcsec slit-width).

VIMOS Mask Preparation was made by using VIMOS Mask Preparation Software (VMMPs)<sup>2</sup> and the observations were conducted at the end of August to the beginning of 2011 September (part of the observing run ESO-ID 086.A-0520 B, P.I. H. Francke).

The observations were taken in three of our five LAE fields: QSO B2359+068; QSO J0124+0044; and LBQS 0301-0035. Details of the observations are given in Table 3.

#### 3.2 Initial photometric target selection

LAEs were selected for the VLT VIMOS spectroscopic observations using the narrow-band 497 nm and broad-band  $B$ ,  $V$  and  $R$  band data already described. The transmission curves for each of these are shown in Fig. 1. For this selection we used a narrow-band magnitude cut of  $m_{\text{NB497}} \leq 26$  in combination with  $(B - m_{\text{NB497}})$  and  $(V - m_{\text{NB497}})$  or  $(R - m_{\text{NB497}})$  colour cuts. These colour cuts



**Figure 1.** The transmission curve of the filters. The solid line indicates the narrow-band filter, NB497. The dashed lines indicate broad-band filters,  $B$ ,  $V$ , and  $R$  from left to right, respectively.

were tailored individually to each field in order to optimize the slit allocations in the VIMOS fields targeted. The colour cuts used in each field are given in Table 3, along with the total number of objects selected (across the whole  $\approx 0.5 \times 0.5$  field in each case) and the number of these that were targeted in the VIMOS masks (i.e. over an area of  $\approx 16 \times 16$  arcmin<sup>2</sup>).

Different groups use different selection criteria to search for LAE candidates photometrically. For example, narrow-band and combined broad-band observations (e.g.  $(B + R)/2$ ), one narrow-band filter and one broad-band filter, and one narrow-band filter and multiple individual broad-band filters. We choose the latter approach, using a two-colour approach which applied  $R - \text{NB497}$  colour and  $B - \text{NB497}$  colour, similar to those in Fynbo et al. (2003), Nilsson et al. (2007), Ouchi et al. (2008).

Following Ouchi et al. (2008), we started from a colour cut of  $(V - m_{\text{NB497}}) = 1.2$ , which equates to a cut in constant rest-frame equivalent width of  $\text{EW}_0 \approx 45 \text{ \AA}$  (or  $\text{EW}_0 \approx 65 \text{ \AA}$  when taking the Ly $\alpha$  Gunn–Peterson trough into account – see Ouchi et al. 2008). For our fields containing  $R$ -band data, but no  $V$ -band data, we took into account the mean  $(V - R)$  colour difference in the  $22 < m_{\text{NB497}} < 25$  galaxy population giving an equivalent  $R$ -band cut is  $(R - m_{\text{NB497}}) = 0.8$ . In addition to this, we also applied a cut using the  $B$  band of  $B - \text{NB497} > 1.3$ . In order to maximize the slit allocations in the VIMOS masks, we then allowed these colour constraints to be relaxed. The resulting colour cuts used for the VIMOS target selection in each field are given in Table 3. These colour cuts were combined with a magnitude limit of  $m_{\text{NB497}} \leq 26$ .

<sup>2</sup> <http://www.eso.org/sci/observing/phase2/SMGuidelines/VMMPs.html>



**Table 4.** Galaxies with confirmed Ly $\alpha$  emission from the VLT VIMOS observations. The top, middle, and bottom entries correspond to the galaxies in the fields of the quasars B2359+0653, J01214+0044, and Q0301–0035, respectively.

ID	RA (J2000)	Dec.	$z$	$m_{497}$	$B - m_{497}$ (AB)	$R - m_{497}$
VLRS J000223.50+071321.0	0.597932	7.222500	3.0939	23.48	2.44	1.50
VLRS J000243.85+071021.5	0.682742	7.172648	3.1079	24.49	2.75	4.40
VLRS J000222.05+071010.6	0.591878	7.169614	3.0543	24.81	1.96	2.91
VLRS J000221.03+070915.6	0.587649	7.154351	3.1187	24.69	2.11	3.34
VLRS J000150.11+072144.3	0.458817	7.362311	3.1112	25.57	–	–
VLRS J000200.93+071939.3	0.503911	7.327593	3.0994	25.46	–	1.43
VLRS J000143.31+071209.4	0.430497	7.202622	3.1015	25.55	–	1.40
VLRS J000154.47+070912.2	0.476987	7.153391	3.1052	24.17	3.99	2.28
ID	RA (J2000)	Dec.	$z$	$m_{497}$	$B - m_{497}$ (AB)	$V - m_{497}$
VLRS J012457.44+004648.0	21.239342	0.780015	3.1030	24.70	1.68	3.29
VLRS J012451.71+004502.5	21.215461	0.750693	3.0520	25.18	2.15	1.30
VLRS J012450.74+004352.3	21.211416	0.731202	3.0811	24.55	–	0.24
VLRS J012445.90+004317.0	21.191265	0.721403	3.0706	25.23	1.81	1.12
VLRS J012509.24+005710.7	21.288526	0.952986	3.1228	24.84	3.05	–
VLRS J012503.13+005634.7	21.263065	0.942975	3.1104	23.71	3.22	3.21
VLRS J012408.25+005656.9	21.034377	0.949161	3.0748	24.15	3.02	4.67
VLRS J012416.71+005622.2	21.069636	0.939499	3.0672	25.20	3.21	4.52
VLRS J012413.08+005529.5	21.054527	0.924861	3.0734	25.54	0.44	0.86
VLRS J012408.69+005341.3	21.036244	0.894817	3.0930	24.94	2.60	3.74
VLRS J012428.28+005249.8	21.117851	0.880499	3.0761	24.00	–	1.53
VLRS J012406.88+004750.3	21.028689	0.797311	3.0792	24.39	2.04	2.39
VLRS J012417.96+004733.7	21.074841	0.792696	3.0786	24.25	3.22	3.07
ID	RA (J2000)	Dec.	$z$	$m_{497}$	$B - m_{497}$ (AB)	$R - m_{497}$
VLRS J030326.84–001805.8	45.861860	–0.301630	3.0951	25.96	–	0.14
VLRS J030337.40–001959.3	45.905844	–0.333158	3.1322	25.67	0.54	0.39
VLRS J030319.11–002139.7	45.829661	–0.361044	3.0646	25.89	–	–
VLRS J030325.62–002350.5	45.856789	–0.397361	3.0757	25.28	2.96	–
VLRS J030326.31–002400.2	45.859633	–0.400077	3.0541	24.82	1.48	1.05
VLRS J030333.31–001047.6	45.888794	–0.179914	3.0746	24.57	2.07	–
VLRS J030256.04–001007.2	45.733499	–0.168692	3.0802	24.32	2.47	2.51
VLRS J030248.36–001022.7	45.701519	–0.172993	3.0943	25.52	–	–0.01
VLRS J030241.28–001140.6	45.672003	–0.194622	3.1031	25.28	3.09	–
VLRS J030244.58–001341.8	45.685771	–0.228293	3.0730	23.87	1.52	1.21
VLRS J030254.08–001414.6	45.725362	–0.237397	3.0854	24.08	3.67	2.24
VLRS J030244.33–001911.9	45.684716	–0.319997	3.1251	24.37	1.41	0.62
VLRS J030257.61–002058.5	45.740049	–0.349601	3.0939	24.76	–	3.26
VLRS J030257.31–002301.0	45.738794	–0.383629	3.1019	25.62	1.90	0.82

### 3.3 Data reduction

The reduction of the spectroscopic data was performed with the VIMOS pipeline `ESOREX` packages.<sup>3</sup> The main procedure includes creating master calibration data, reducing science frames, and extracting objects. Following the pipeline manual, we first created a master-bias with the recipe `VMBIAS`. An output master-bias was then used in the reduction of the flat-field, arc lamp, and scientific exposures. The next step was using the recipe `VMMOSCALIB` to process flat-field exposures and arc lamp exposures. We used the `VMMOSSCIENCE` recipe to process science frames with the cosmic ray rejection applied at this point in the process. We then combined the images for each field using the `IRAF IMCOMBINE` task.

The object extraction is made by applying an optimal extraction algorithm (Horne 1986). The wavelength calibration was performed using the input wavelength calibration and sky lines, and sky background subtracted. We found diagonal stray light interfered with the wavelength calibration and sky subtraction in quadrants 3 and 4. To fix this, we isolated the brightest part of this light from the flat lamp, fitted the smooth pattern of the flat and subtracted it out.

### 3.4 Lae identification

We found significant emission within the wavelength range of the narrow-band filter in eight of the 23 candidates in the QSO B2359+0653 field, 13 of the 45 in the QSO J0124+0044 field, and 14 of the 46 in the LBQS 0301–0035 field. These are listed in full in Table 4, whilst the reduced 1D spectra (with 2D spectra inset) are shown in Figs A1–A3 for the three separate fields. In each figure,

<sup>3</sup> <http://www.eso.org/sci/software/cpl/esorex.html>

the solid black line shows the reduced 1D spectrum, the shaded grey region shows the normalized NB497 filter response and the inset image shows the 2D spectrum within  $\pm 60 \text{ \AA}$  of the peak of the NB497 filter response curve.

Given all our detections are single line detections, the possibility exists that some of these could be [O II] emission from faint  $z \approx 0.33$  galaxies. The spectral resolution of our observations is marginally sufficient to discern the double peaked emission inherent in the O II double lines (the resolution is  $\sim 2.5 \text{ \AA}$  while the [O II] doublet is about  $2.7 \text{ \AA}$ ). Although the resolution is very close to the peak separation, even in the event of the doublet being smoothed out in the spectrum by the instrument response, the presence of the [O II] doublet would lead to a significantly broad line. We have analysed each emission line, plotting the doublet separation over each and evaluating the likelihood of each line being [O II] emission. In most cases ( $\gtrsim 90$  per cent), the emission shows no sign of being a doublet (i.e. there is no double peak and the single peak emission is not broad enough to be the two [O II] lines convolved into one via the instrument response). For the remaining  $\lesssim 5$  per cent, we cannot discount the possibility of the emission being low-redshift [O II], but at the same time, none are clear-cut cases of [O II] emission. These ambiguous  $\lesssim 5$  per cent are broad emission that show no significant detections of double peaks within the spectral noise. Indeed, higher signal-to-noise ratio and resolution observations would be needed to fully discern any double peaked nature to these broader lines. In addition, we note that even if double-peaked emission were detected, the complex nature of the escape of Ly $\alpha$  photons from galaxies can also produce double emission peaks. In all we find no strong evidence of any double emission in any of the detected emission lines and conclude that  $\gtrsim 95$  per cent of these lines are Ly $\alpha$  emission at  $z \approx 3.06$ .

The success rate of the observations was somewhat low in part due to the relaxed constraints used to select candidates. Additionally, the slit mask for the second quadrant in the QSO B2359+0653 field was not properly aligned in its mount leading to no detections in that quadrant. However, in addition to this, a poor astrometric solution was found to have contributed to the low success rate. The astrometric errors for the images created two issues: (1) in the QSO B2359+0653 field, a number of targets were not correctly aligned within the slits during the observations; and (2) inconsistencies between the astrometry between the narrow band and broad-bands led to incorrect broad-band photometry measurements. Apertures placed to measure the photometry in the broad-band images were thus offset from the intended targets in some cases, leading to incorrectly faint magnitudes being measured and objects with no emission being promoted into the selection criteria.

Following the spectroscopic observations, we thus recalculated matched astrometric solutions for all the data as described in Section 2.3.

### 3.5 Selection efficiency and optimized selection criteria

The selection success rates are relatively low in our spectroscopic sample of LAE candidates (31 per cent). This is the result of the original criteria being comparatively flexible in order to allow the maximal number of slits placed per VIMOS field, whilst as discussed the astrometric calibrations were sub-optimal prior to the spectroscopic follow-up, causing errors in the photometry. In this section we therefore provide an optimized and uniform set of selection constraints, maximizing the numbers of successfully identified LAEs based on our updated photometric catalogues and the spectroscopic observations.

First of all, we now use a narrow-band magnitude cut of  $m_{\text{NB497}} \leq 25.0$  (as opposed to  $m_{\text{NB497}} \leq 26$  previously), as in all three fields there is significant scatter from the galaxy locus at fainter magnitudes than this due to the measurement uncertainties on the photometry. The converse of this is that the measured colours of many LAEs beyond this limit are scattered into the galaxy locus in the colour–magnitude diagram.

Secondly, we optimise the  $V - m_{\text{NB497}}$ ,  $R - m_{\text{NB497}}$  and  $B - m_{\text{NB497}}$  colour cuts at  $m_{\text{NB497}} \leq 25.0$  in order to maximize the numbers of LAEs found and minimize the number of interlopers given the spectroscopic observations. We also apply the constraint that the three colour criteria correspond to the same equivalent width constraint. For this we simulate LAE spectra by assuming a  $\beta = -1.4$  power-law slope (e.g. Berry et al. 2012) combined with a range of equivalent Ly $\alpha$  emission lines. We find that taking colour constraints of  $V - m_{\text{NB497}} \geq 1.2$ ,  $R - m_{\text{NB497}} \geq 1.0$  and  $B - m_{\text{NB497}} \geq 1.4$  selects 18 out of the 20 spectroscopically confirmed  $m_{\text{NB497}} \leq 25.0$  LAEs, whilst missing two of these. The selection also includes five non-detections from the spectroscopic observations that may be contamination or weakly emitting  $z \approx 3.1$  LAEs (likely scattered into the selection region due to noise on the photometry). These colour cuts are equivalent to a rest-frame equivalent width cut of  $\approx 45 \text{ \AA}$ , when assuming the simple  $\beta = -1.4$  continuum power-law slope. We do not, however, include the effect of the Gunn–Peterson trough and so, as discussed and modelled in Ouchi et al. (2008), our nominal  $\approx 45 \text{ \AA}$  selection corresponds to a  $\approx 65 \text{ \AA}$  cut when taking this into account.

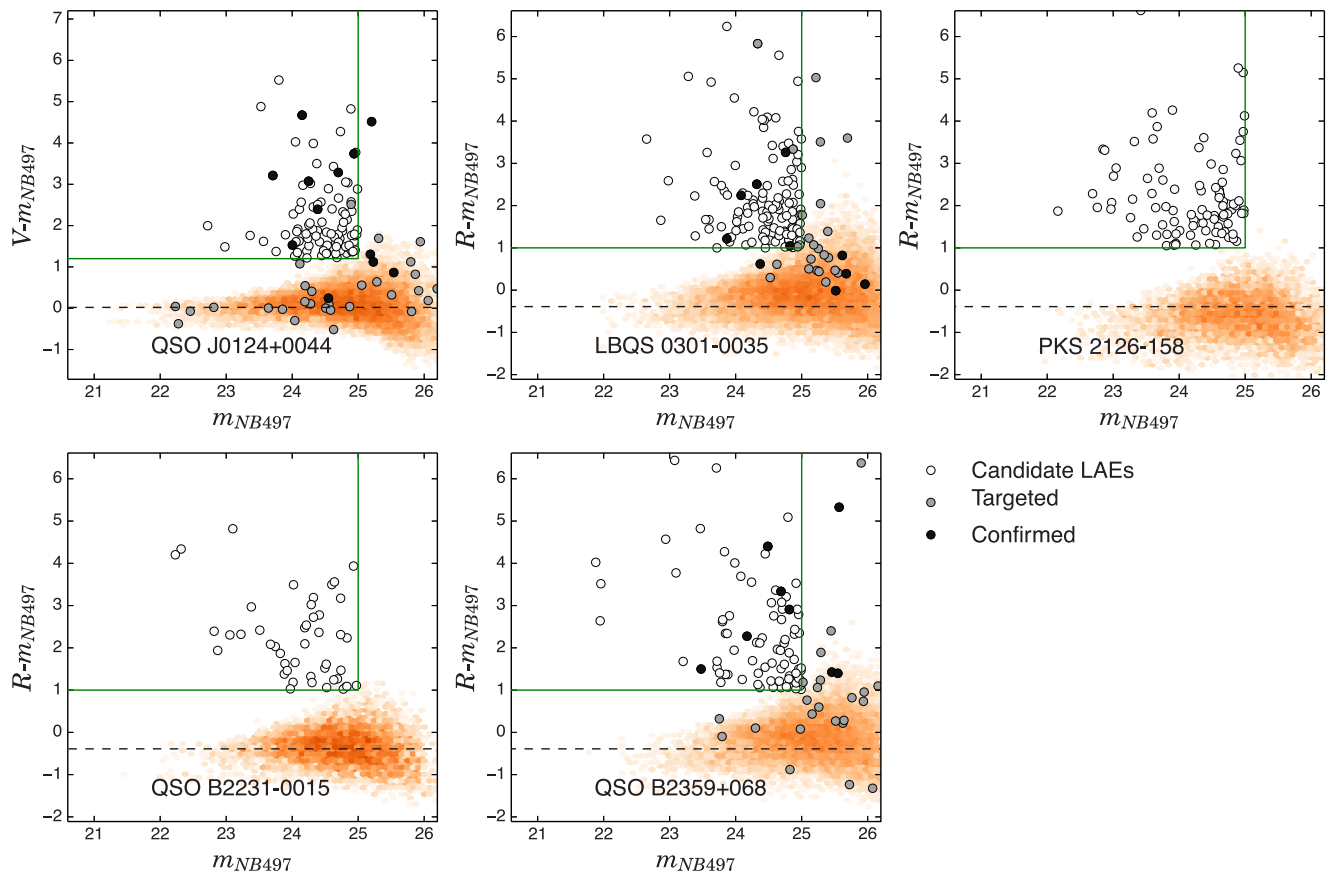
Additionally, we also reject objects with SExtractor flags greater than zero; mask low signal-to-noise ratio regions of the images (primarily the image edges) to avoid noise contamination in these regions; and (in order to minimize the contribution of artefacts in the images) apply elongation ( $e_{\text{SEX}}$ ) and FWHM limits. By inspection of the images, objects with a measured elongation of  $e_{\text{SEX}} > 2.8$  are exclusively artefacts (primarily cosmic ray hits not removed by CRREJECT). In addition, objects with measured FWHM less than the measured image FWHM are also rejected (on inspection these are also exclusively artefacts and primarily remnant cosmic rays.).

The selection criteria are summarized as follows:

- (i)  $20 < m_{\text{NB497}} < 25$ ;
- (ii)  $(V - m_{\text{NB497}}) > 1.2$  or  $(R - m_{\text{NB497}}) > 1.0$ ;
- (iii)  $(B - m_{\text{NB497}}) > 1.4$ ;
- (iv)  $\text{FLAG}_{\text{SEX}} = 0$ ;
- (v)  $e_{\text{SEX}} < 2.7$ ;
- (vi)  $\text{FWHM} \geq \text{FWHM}_{\text{image}}$ .

These colour cuts are shown by the solid lines in Figs 2 and 3. In each figure the galaxy population is shown by the 2D histogram, whilst selected LAE candidates are shown by open circles. The dashed horizontal lines show the position of the locus in  $(V - m_{\text{NB497}})$  or  $(R - m_{\text{NB497}})$  colour (Fig. 2) and  $(B - m_{\text{NB497}})$  colour (Fig. 3). The grey and black filled circles show LAE candidates observed with VLT VIMOS and are described further in Section 3. The seeing for each field (i.e.  $\text{FWHM}_{\text{image}}$ ) is given in Table 1. As discussed, we have produced catalogues based on object detection in the narrow-band images alone and also in a  $\chi^2$  combined image of the narrow-band and broad-band images. We have run our selection on both of these sets of catalogues. In the results that we report here, we use the  $\chi^2$ -image detected catalogues. We have run all our analysis on the narrow-band image detected catalogues and find consistent results within the estimated uncertainties.

The numbers and sky densities of candidates in each field given by these criteria are given in Table 5. Taking only those



**Figure 2.** Colour–magnitude diagrams showing  $V - m_{\text{NB497}}$  versus  $m_{\text{NB497}}$  for the QSO J0124+0044 field and  $R - m_{\text{NB497}}$  versus  $m_{\text{NB497}}$  for the LBQS 0301–0035, PKS 2126–158, QSO B2231–0015, and QSO B2359+068 fields. The shaded 2D histogram shows the general galaxy population. Grey and black filled circles show LAE candidates observed with VLT VIMOS, the grey showing those with no spectroscopic detection and the black showing those confirmed as  $z = 3.06$  LAEs. The region to the upper left of the solid horizontal and vertical bounding lines give the updated selection region, whilst the open circles show LAE candidates selected based on these bounds. Note that the objects identified as being targeted with VIMOS that lie within the main galaxy population are those that were subject to significant photometric errors due to the astrometric mis-alignment of images (see Section 3.4).

spectroscopically observed targets that fall within the updated selection criteria, we find a success rate of  $78 \pm 18$  per cent. The proportion of targets selected in this new sample that were also present in the original target selection is  $\approx 40$  per cent, thus calculating the success rate for our updated selection using the spectroscopic sample is arguably only representative of 40 per cent of the sample. However, the remaining 60 per cent of targets selected in the updated selection were missed in the original selection due to astrometric errors. The loss of targets due to the sub-optimal astrometric solution has no connection with intrinsic source properties (rather it is a function of the on-sky coordinates of the sources) and so should not introduce any biases between the 40 per cent that were selected and the 60 per cent that were not, so it would be reasonable to assume that the result for the 40 per cent that were included holds for the 60 per cent that were missed due to astrometric issues.

Comparing to the success rate in other works, Ouchi et al. (2008) identified line emitters from 60 per cent of their targets at this redshift, whilst Fynbo, Möller & Thomsen (2001) and Fynbo et al. (2003) reported the spectroscopic follow-up success rate of 75–90 per cent for  $z \sim 3$  LAE surveys.

The distribution of the selected LAE candidates in each field is shown in Fig. 4, with the same symbols as in previous plots (i.e. open circles for LAE candidates and filled grey and black circles

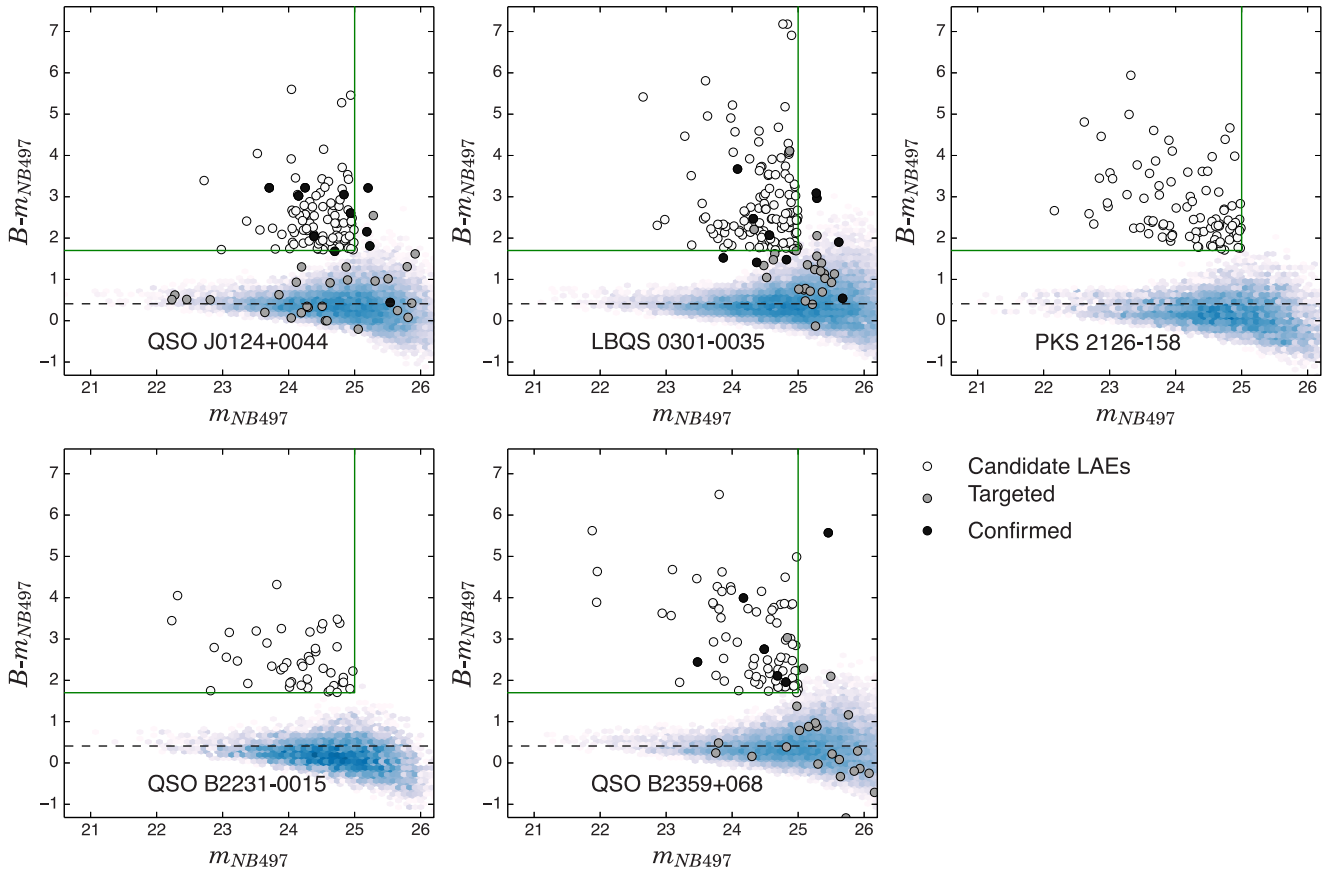
for candidates observed with VLT VIMOS). Filled stars in Fig. 4 denote the positions of background QSOs in the fields.

### 3.6 Redshift and spatial distribution of LAEs

The redshift distribution,  $n(z)$ , for the VLT VIMOS confirmed LAEs is shown in Fig. 5 (solid blue histogram). For comparison, the throughput curve of the NB497 filter is also shown (solid black curve).

The survey sky coverage is illustrated in Fig. 4, where the distributions of candidate LAEs (small grey circles) and confirmed LAEs (large black circles) are shown for each field. Background  $z \gtrsim 3.1$  QSOs are also shown (filled-stars). These five fields cover a total area (after masking) of  $1.07 \text{ deg}^2$ .

Integrating along the normalized VIMOS redshift distribution and combining it with this total sky area give an LAE survey volume of  $2.51 \times 10^5 h^{-3} \text{ Mpc}^3$ . Integrating instead the NB497 filter response curve gives a somewhat higher survey volume estimate of  $3.35 \times 10^5 h^{-3} \text{ Mpc}^3$ . In the analysis that follows, we take the volume given by the filter response curve given that the measured redshift distribution is somewhat under-sampled. For reference, a top-hat redshift distribution (between  $z = 3.05$  and  $z = 3.14$ ) gives a survey volume of  $4.69 \times 10^5 h^{-3} \text{ Mpc}^3$ .



**Figure 3.** Colour–magnitude diagrams showing  $B - m_{\text{NB497}}$  versus  $m_{\text{NB497}}$  for our five QSO fields: QSO J0124+0044, LBQS 0301–0035, PKS 2126–158, QSO B2231–0015, and QSO B2359+068. The shaded 2D histogram shows the general galaxy population. Grey and black filled circles show LAE candidates observed with VLT VIMOS, the grey showing those with no spectroscopic detection and the black showing those confirmed as  $z = 3.06$  LAEs. The region to the upper left of the solid horizontal and vertical bounding lines give the updated selection region, whilst the open circles show LAE candidates selected based on these bounds. Note that the objects identified as being targeted with VIMOS that lie within the main galaxy population are those that were subject to significant photometric errors due to the astrometric mis-alignment of images (see Section 3.4).

**Table 5.** Number of LAE candidates from our selection in each field.

Field	$N_{\text{LAE}}$	$\rho_{\text{LAE}}$ ( $\text{arcmin}^{-2}$ )
QSO B2359+068	124	0.160
QSO J0124+0044	139	0.186
LBQS 0301–0035	170	0.201
PKS 2126–158	140	0.158
QSO B2231–0015	70	0.119
All fields	643	0.167

### 3.7 Selection completeness

In order to estimate the selection completeness, we first estimate the percentage completeness as a function of source magnitude of the individual survey images. To do so, we simulate sources in the images in set magnitude bins. We attempt to detect the simulated sources using `SEXTRACTOR`, identically to the extraction process for the original data. For each magnitude bin, we simulate 100 sources, with the number recovered giving the fractional completeness in each bin. We avoid source confusion by only placing simulated sources in empty regions with no genuine sources found within 2.5 times the image FWHM.

We applied these detection rates with simulated LAE spectra with a range of equivalent widths and assuming a spectral slope of  $\beta = -1.4$  (e.g. Berry et al. 2012). The results of this simulation for each field are shown in Fig. 6. Each panel shows one of our five individual fields, with the shaded regions showing the range of our LAE selection in  $R$ -EW space. The contour curves show the 20 per cent, 40 per cent, 60 per cent, and 80 per cent completeness levels.

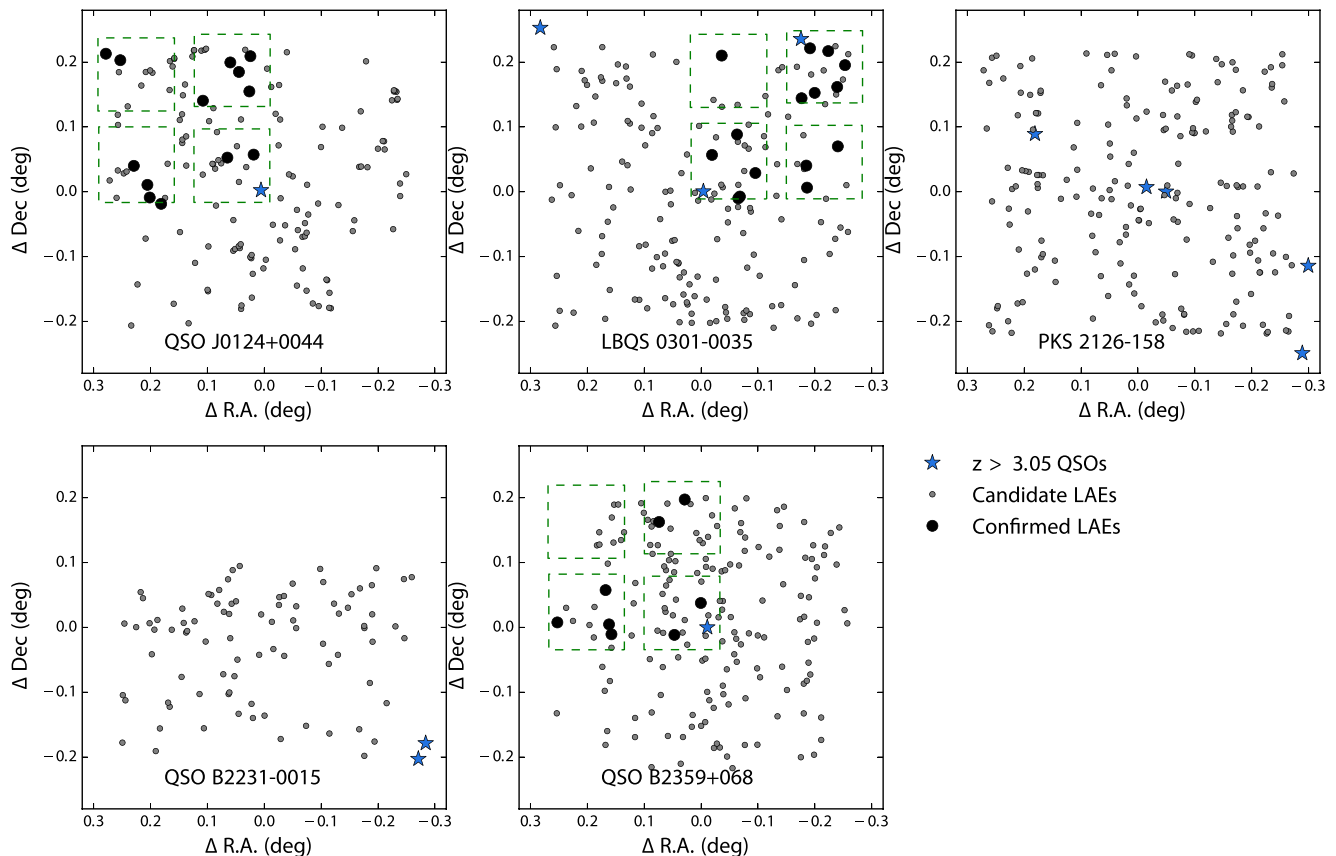
Overall, we find good consistency across our five fields in terms of the completeness profile. At fainter continuum luminosities (i.e.  $R \gtrsim 25.5$ ), we find that we increasingly miss the high equivalent width range of our selection with all five fields falling to  $\approx 20$  per cent completeness at  $R \approx 27$ .

## 4 NUMBER COUNTS AND LUMINOSITY FUNCTION

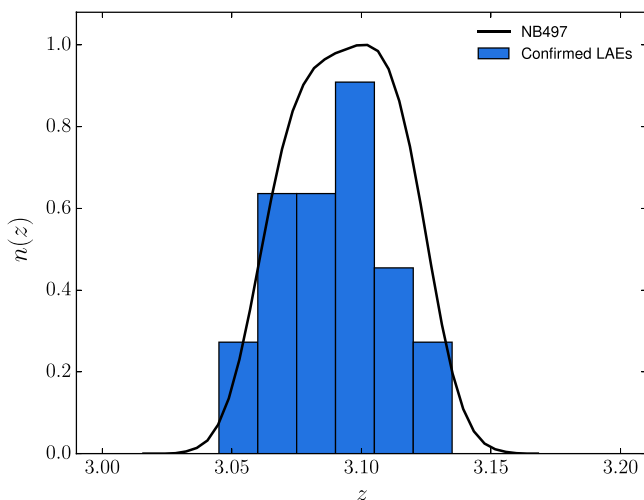
### 4.1 Narrow-band number counts

Using the updated selection criteria given in the previous section, we calculate the galaxy number counts for the LAE sample as a function of narrow-band magnitude. This is shown in Fig. 7, where the filled black circles show our results. The error bars show field-to-field error estimates on the points. For reference, we also





**Figure 4.** The distribution of candidate (grey filled circles) and confirmed (black filled circles) LAEs in the five survey fields. Known QSOs at the LAE redshifts or higher (i.e.  $z > 3.05$ ) are also plotted for reference (stars). The dashed green boxes show the areal coverage of the VIMOS quadrants for the spectroscopic observations (note that the fields PKS 2126–158 and QSO B2231–0015 fields have not been observed spectroscopically). The central QSO in the QSO B2231–0015 field is not plotted as it has a redshift lower than that of the LAE selection ( $z_{\text{QSO}} = 3.027$ ).



**Figure 5.** Redshift distribution for LAE spectroscopic identification from VIMOS (normalized histogram) and the NB497 filter with wavelength converted to redshift (solid curve). The NB497 filter is normalized such that the peak is unity.

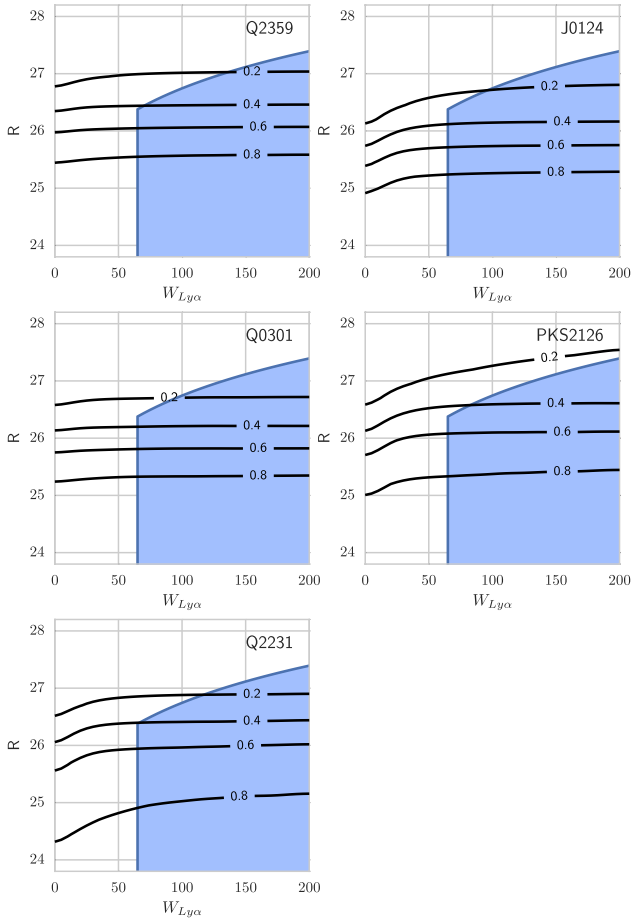
show the NB497 number counts for all sources in our fields (grey circles).

For comparison, we show the  $z \approx 3.1$  narrow-band number counts published by Gronwall et al. (2007, diamonds) and Ouchi et al.

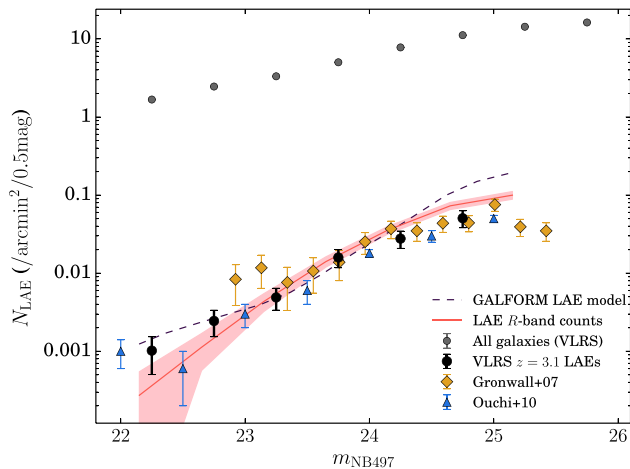
(2008, triangles). Gronwall et al. (2007) conducted a  $0.28 \text{ deg}^2$  survey at  $z \approx 3.1$  in the ECDFS. Their selection is based on a measured colours derived from the NB5000 filter and a broad-band image constructed using a combination of images taken using the  $B$ - and  $V$ -band filters. Their criteria were thus:  $m_{\text{NB5000}} - m_{B+V} < 1.03$ , with a narrow-band magnitude limit of  $m_{\text{NB5000}} \leq 25.4$ . This gave a sample of 162 galaxies with rest-frame equivalent widths of  $> 20 \text{ \AA}$  (approximately  $80 \text{ \AA}$  in the observer’s frame). The Ouchi et al. (2008) selection criteria are based on cuts of  $V - m_{\text{NB503}} > 1.2$  and  $B - V > 0.5$ . They detected 356 photometric LAEs, with rest-frame equivalent widths of  $\geq 65 \text{ \AA}$ , over a  $\approx 1 \text{ deg}^2$  area.

A direct comparison of the Gronwall et al. (2007) and (Ouchi et al. 2008) number counts of  $z \sim 3.1$  LAEs is shown by Ciardullo et al. (2012). Ciardullo et al. (2012) extended the study of Gronwall et al. (2007) by re-imaging the ECDF-S with a  $57 \text{ \AA}$  FWHM nearly top-hat filter centred at  $\approx 5010 \text{ \AA}$ . They found a total number of 360  $z \approx 3.1$  LAEs, a subset of which are also selected by Gronwall et al. (2007). They then reproduced the luminosity function and concluded that their result is statistically identical to values in Gronwall et al. (2007) and Ouchi et al. (2008), once differences in the filter FWHM and EW selection constraints were accounted for.

The VLRS LAE number counts are consistent with the number counts of Gronwall et al. (2007) at the  $\approx 1\sigma$  level, and marginally higher than the counts of Ouchi et al. (2008) (although still only at the  $\approx 1\text{--}2\sigma$  level). Our plotted number counts include a factor of 0.78 to account for the number of non-detections when our sample was



**Figure 6.** Estimated selection completeness in the five observed fields based on simulated sources in the survey images. The shaded region shows the extent of our selection criteria in  $R$  –EW space, whilst the contour curves show the 20 per cent, 40 per cent, 60 per cent, and 80 per cent completeness levels.



**Figure 7.** The number counts of photometrically detected LAEs based on the selection criteria given in Section 3.5 (filled black circles). The diamonds and triangles show number counts for comparable LAE selections by Gronwall et al. (2007) and Ouchi et al. (2008), respectively. Small filled grey circles show the number counts of all objects detected in the narrow-band imaging. The solid curve shows model LAE number counts from the GALFORM galaxy formation model. The shaded curve shows the LAE  $R$ -band counts adapted from Fig. 8 with a shift of  $\Delta m = -1.6$  to overlay the narrow-band counts.

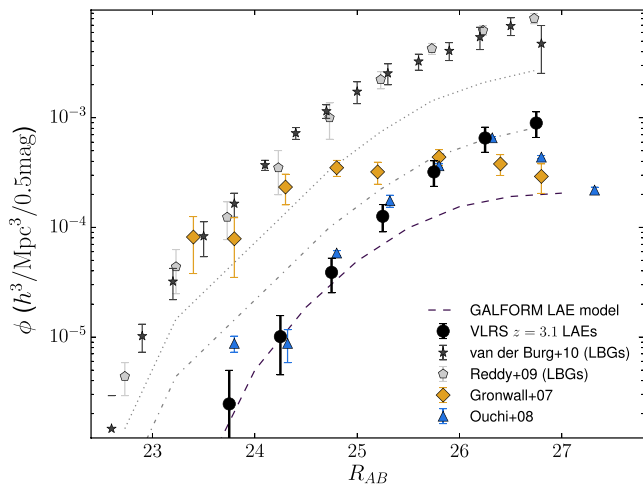
observed spectroscopically. This is perhaps a conservative estimate for the success of our selection criteria in identifying LAEs, given that some fraction of those non-detections could have been LAEs below the detection threshold of the spectroscopic observations. In the most extreme case, i.e. if all the non-detections were in fact genuine LAEs, the VLRS LAE number counts shown in Fig. 7 could be increased by a factor of  $\approx 25$  per cent. This would perhaps cause some tension with the counts of Ouchi et al. (2008) at the  $2\text{--}3\sigma$  level.

For context in comparing the individual results, the equivalent width limits for each survey are: 65 Å (this paper), 20 Å (Gronwall et al. 2007), and 65 Å (Ouchi et al. 2008). Ciardullo et al. (2012) suggest that the lower equivalent width limit of Gronwall et al. (2007) should measure  $\approx 10$  per cent higher numbers of LAEs compared to Ouchi et al. (2008), and hence our own sample. In addition to the different EW cuts, given the narrower filter used by Gronwall et al. (2007), one would expect to measure lower counts by a factor of  $\frac{2}{3}$  given the smaller volume coverage this entails. Ciardullo et al. (2012) do not explicitly state that the difference in filter width is included in their prediction of a difference of 10 per cent in the numbers of LAEs between the samples, however their estimate would seem somewhat low if it is not (for example when compared to the equivalent width distribution of LBGs measured by Shapley et al. 2003).

Although our main aim has been to check the representiveness of our LAE counts as a prelude to using the LAE for clustering studies, we now briefly compare our counts to a recent theoretical model. Shown in Fig. 7 therefore is the Orsi, Lacey & Baugh (2012) model for LAEs updated to the latest variant of the GALFORM semi-analytical model of galaxy formation (Lacey et al. 2015). This model for LAEs makes use of a Monte Carlo  $\text{Ly}\alpha$  radiative transfer code to compute the escape fraction of LAEs at different redshifts. The GALFORM model is calibrated to match a number of observational data, mostly at  $z = 0$  (e.g. optical and NIR LFs,  $M_{\text{BH}} - M_{\text{bulge}}$  relation, morphological fractions), but also sub-mm, FIR counts and UV luminosity functions at high redshift. A radiation transfer model was then used to model the  $\text{Ly}\alpha$  escape fraction, based on an outflow model which is itself dependent on the individual galaxy properties. This radiation transfer and outflow model (which in this case consists of an expanding thin shell of material) was tuned to match the luminosity function of Ouchi et al. (2008). Given the agreement between our own results and those of Ouchi et al. (2008), it is no surprise to find the model successfully reproduces our own narrow-band number counts at  $m_{\text{NB497}} \lesssim 24.2$ , although the agreement worsens at fainter magnitudes (as it does with the Ouchi et al. 2008 results). Even to the extent that they agree, the tuning of the model to the luminosity function of Ouchi et al. (2008) means that the observed LAE number counts are therefore not a test of the model. This and the disagreement at faint fluxes must be borne in mind when comparing the LAE clustering data with the model in Section 5.2.

## 4.2 Continuum luminosity function

We estimated the  $R$ -band continuum luminosity function of our sample of LAEs. As in Gronwall et al. (2007), we use the filter curve to define the survey volume used for the luminosity function calculation, which gives a volume of  $3.35 \times 10^5 h^{-3} \text{Mpc}^3$ . The resulting continuum luminosity function evaluated across all five fields (and scaled in accordance with the estimated 78 per cent success rate of our sample) is shown in Fig. 8 (filled black circles). Plotted points are corrected for the incompleteness as a function



**Figure 8.** The  $R$ -band luminosity function of the VLSR  $z = 3.1$  LAEs (filled circles). For comparison, we also show the equivalent LAE luminosity functions from Gronwall et al. (2007, diamonds) and Ouchi et al. (2008, triangles). Also shown is the observed  $r$ -band luminosity functions of  $z \sim 3$  LBGs from Reddy & Steidel (2009, pentagons) and from the CFHT fields of van der Burg et al. (2010, stars). The dotted and dash–dotted curves show the Reddy & Steidel (2009) luminosity function reduced by a factor of 3 and a factor of 10, respectively, whilst the dashed curve shows the GALFORM model result.

of  $R$  magnitude estimated in Section 3.7. The error bars give the uncertainty based on field-to-field estimates using our five imaging fields in combination with the uncertainty introduced by the estimate of the same success rate.

We show our results in comparison to those of Gronwall et al. (2007, diamonds) and Ouchi et al. (2008, triangles). Our results are consistent with those of Ouchi et al. (2008) within the calculated errors to a magnitude of  $R \approx 26$ . In contrast, our luminosity function (and that of Ouchi et al. 2008) is only consistent with that of Gronwall et al. (2007) at magnitudes of  $R \gtrsim 25$ . At  $R \lesssim 25$ , the Gronwall et al. (2007) luminosity function is significantly elevated above our own and approaches the volume densities of LBGs at the same redshift – shown by the filled stars (van der Burg et al. 2010) and pentagons (Reddy & Steidel 2009). Indeed, whilst Ouchi et al. (2008) report their LAE luminosity function to be  $\approx \frac{1}{10}$  of that of the Reddy & Steidel (2009) LBG luminosity function (represented by the dash–dotted curve in Fig. 8), Gronwall et al. (2007) find that their LAE  $M_{UV}$  luminosity function is equivalent to  $\approx \frac{1}{3}$  of the Reddy & Steidel (2009) LBG luminosity function (represented by the dotted curve). Our own LAE  $M_{UV}$  luminosity function is consistent with LAEs (of equivalent width  $\gtrsim 65 \text{ \AA}$ ) being  $\approx 10$  times less common than LBGs, given magnitudes of  $R \lesssim 26$  ( $M_{UV} \lesssim -19.6$ ). The difference between the Gronwall et al. (2007) and our own results would appear to be driven by the differing equivalent width cuts used. This is supported by the distribution function of equivalent widths in LBGs as presented by Shapley et al. (2003), whereby there are approximately three times as many galaxies with equivalent widths in the range  $\gtrsim 20 \text{ \AA}$  (equivalent to the Gronwall et al. 2007), than in our range  $\gtrsim 65 \text{ \AA}$  at magnitudes of  $R \lesssim 25.5$ .

We note that as with the number counts in the previous section, we have applied a correction on the assumption that our sample successfully identifies LAEs in 78 per cent of cases. In the case of this being overly pessimistic, the presented luminosity function could be boosted by up to  $\approx 25$  per cent. Within the plotted error bars, this does not change any of the comparisons discussed above.

Again, we also present the GALFORM model results (which are calculated based on our own selection criteria), shown by the dashed curve in Fig. 8. Our observations and the model are consistent within the observational uncertainties at  $R \lesssim 25$ ; however the model under-predicts the numbers of galaxies at magnitudes fainter than this limit (contrasting with the narrow-band counts in which the model over-predicts the faint end.). We note that when the specific selection constraints of Gronwall et al. (2007) are applied to the GALFORM model, the model under-predicts the continuum luminosity function at both the bright and faint ends, whilst when the same is done with the Ouchi et al. (2008) selection constraints, the result is the same as seen with our own data. These comparisons between our results and previous observational results and the model will be instructive when analysing the different clustering results.

Returning to Fig. 7, we show the continuum number counts (shifted by  $\Delta m = -1.6$ , pale red shaded curve) in comparison to the narrow-band number counts. In both the narrow-band and continuum number counts (and the continuum luminosity function), we see evidence for a turn over at faint magnitudes. Indeed, Ouchi et al. (2008), for example, measure  $L_{Ly\alpha}^* \approx 5 \times 10^{42} \text{ erg s}^{-1}$  for their  $z \approx 3.1$  LAEs, which corresponds to a narrow-band magnitude of  $m_{NB} \sim 24.5$ . Our results support the assertion that such LAE samples reaching depths of  $m_{NB} \approx 25$  are probing the knee of both the NB and continuum luminosity functions. Briefly comparing the narrow-band and continuum number counts, we find that the continuum number counts appear steeper, hinting at an evolution in the EW distribution as a function of galaxy brightness.

## 5 LAE AUTO-CORRELATION FUNCTION

We now measure the clustering properties of the photometric LAE sample to derive halo masses and other properties.

### 5.1 Clustering estimator

We estimate the angular auto-correlation function,  $w(\theta)$ , using the Landy–Szalay estimator (Landy & Szalay 1993), which is given by

$$w(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)} \quad (1)$$

where  $DD(\theta)$ ,  $DR(\theta)$ , and  $RR(\theta)$  are the numbers of galaxy–galaxy pairs, galaxy–random pairs and random–random pairs as a function of  $\theta$ . For each field, we generated uniform random points with the same area as our masked LAE samples, with 20 times the number of LAE candidates in each field. Fig. 9 shows the distribution of the random points used in each field for the clustering calculation, showing the extent of the fields and the masking implemented.

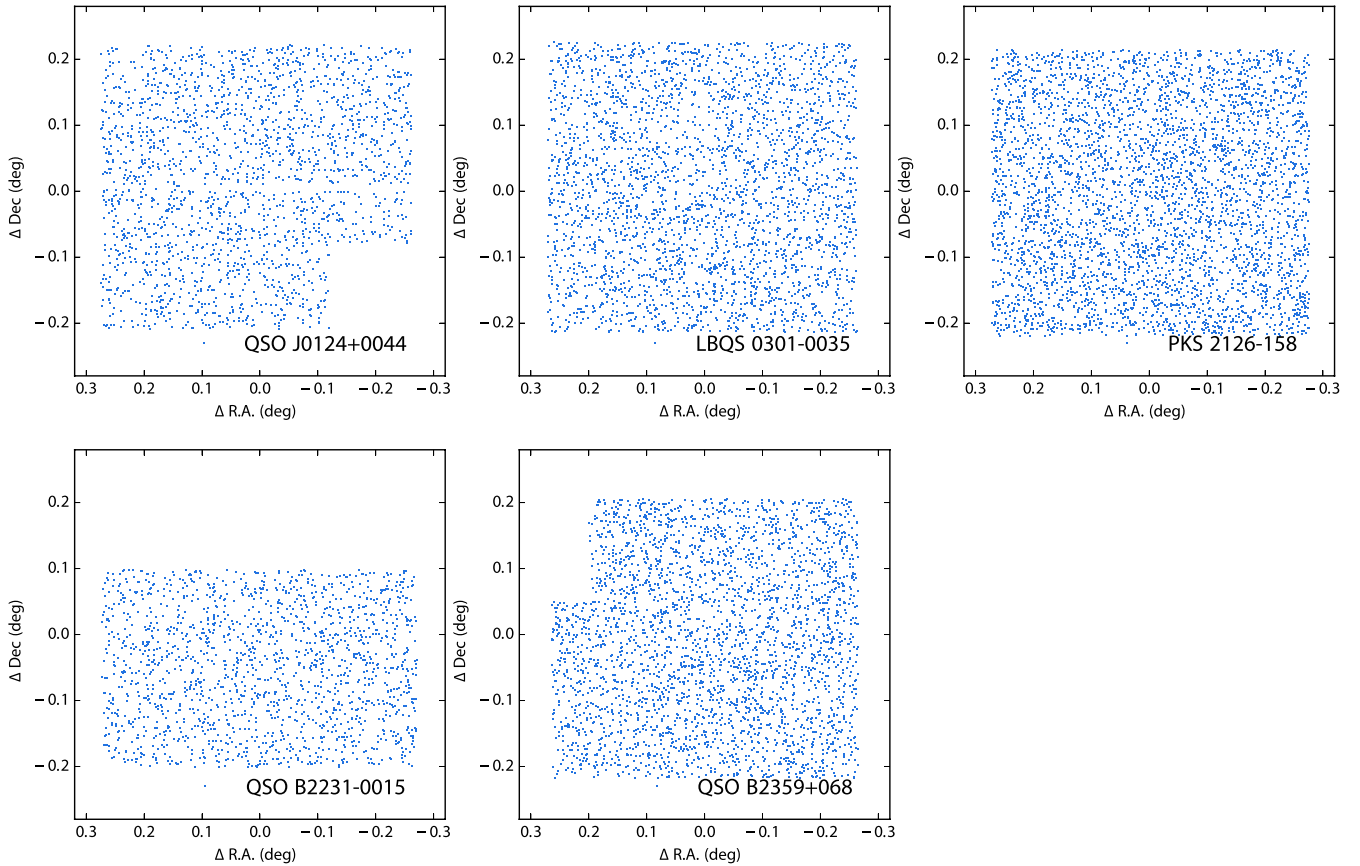
We use two error estimators here: simple Poisson error estimates; and field-to-field error estimates. The Poisson estimate is given by

$$\sigma_{\text{Poi}}(\theta) = \frac{1 + w(\theta)}{\sqrt{DD(\theta)/2}}. \quad (2)$$

The field-to-field error estimate is given by the error on the mean of the measurement across the fields and is calculated using:

$$\sigma_{\text{FF}}(\theta) = \sqrt{\frac{1}{N} \frac{1}{N-1} \sum_{i=1}^N [w_i(\theta) - \bar{w}(\theta)]^2}, \quad (3)$$

where  $N$  is the number of fields (i.e.  $N = 5$ ),  $w_i(\theta)$  is a measurement from the  $i$ th field and  $\bar{w}(\theta)$  is the mean value.



**Figure 9.** The distribution of the random catalogues used in the correlation function analysis highlighting the masked regions in each of the fields.

## 5.2 Angular clustering results

Fig. 10 shows the results of the clustering analysis (filled black circles). We use field-to-field errors everywhere except at  $\theta < 1$  arcmin, where Poisson estimates are used. The plotted points include an amplitude correction for contamination in the photometrically selected sample, which takes the form:

$$A_i = \frac{A_o}{(1 - f_c)^2} \quad (4)$$

where  $A_i$  and  $A_o$  are the intrinsic and observed correlation amplitude, respectively, and  $f_c$  is the contamination fraction (i.e.  $f_c = 0.22$ ). The plotted points have all been adjusted upwards by a factor of  $1/(1 - f_c)^2 = 1.64$ . The points are also corrected for the integral constraint, which accounts for the effect of finite field sizes. This is estimated following the method outlined in Bielby et al. (2014) and takes a value of  $\mathcal{I} = 0.024$ .

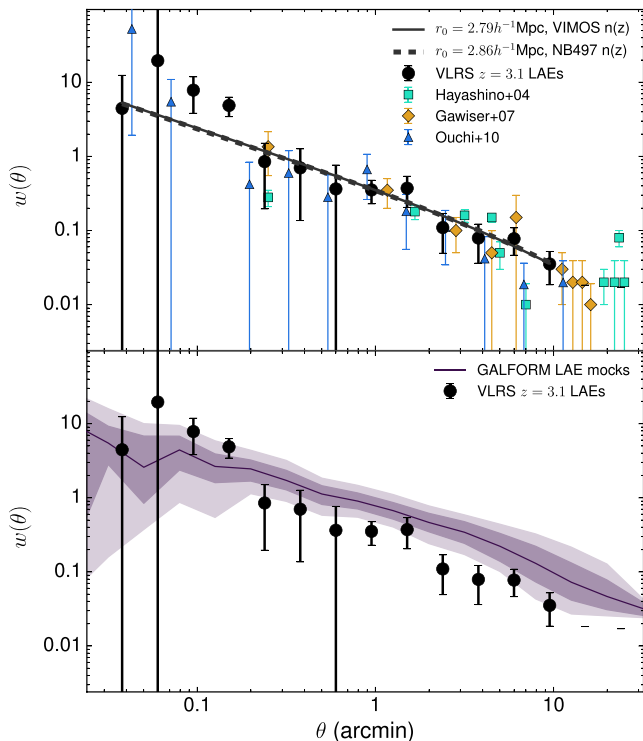
We compare our results in the top panel of Fig. 10 with several measurements from other authors. Squares show the result from Hayashino et al. (2004) which is obtained from 283  $z = 3.1$  LAE candidates with a narrow-band magnitude limit of  $m_{\text{NB}497} < 25.8$  and  $\text{EW}_{\text{rest}} \geq 38 \text{ \AA}$  observed in the SSA22a field using the *Subaru* Telescope. This was a further observation of SSA22a of Steidel et al. (2000) who found 72 LAEs. However, the observed area in Hayashino et al. (2004) is about 10 times larger than in Steidel et al. (2000). Both Steidel et al. (2000) and Hayashino et al. (2004) concluded that no significant clustering of  $z = 3.1$  LAEs has been shown inside the SSA22a area. We also compared our results with Gawiser et al. (2007) (represented by diamonds) who studied the clustering properties of 162 LAEs at  $z = 3.1$  at a narrow-band

completeness limit of  $m_{5000} \leq 25.4$  which were observed in the deep narrow-band MUSYC survey. Their  $\text{EW}_{\text{rest}}$  limit is  $\approx 20 \text{ \AA}$ . Ouchi et al. (2010) used 356  $z = 3.1$  LAEs from Ouchi et al. (2008) who carried out a narrow-band survey in the  $1 \text{ deg}^2$  *Subaru/XMM-Newton* Deep Survey (SXDS) to measure the correlation function at  $m_{\text{NB}5003} \leq 25.3$  (triangles).

Comparing our results with each measurement, we found that our results agree with Ouchi et al. (2010) within their error bars. At larger scales, Hayashino et al. (2004) and Gawiser et al. (2007) show slightly higher clustering amplitude than ours, but are also in agreement with our results within the error estimates. Given the different equivalent width, and flux limits of each survey, this may be evidence that the clustering amplitude of LAEs is particularly insensitive to the  $\text{Ly}\alpha$  sample selection. In other words, LAEs trace the density field in the same way regardless of their selection. Care should be taken when comparing the  $w(\theta)$  measurements of samples with differing filter widths however. Whilst Hayashino et al. (2004), Ouchi et al. (2010), and our own sample use a filter with FWHM  $\approx 75 \text{ \AA}$ , the Gawiser et al. (2007) LAE sample uses a filter with FWHM  $\approx 50 \text{ \AA}$ . A more reliable comparison can be made when comparing the  $r_0$  measurements that follow, which take into account the filter profile. Overall, our results provide the most accurate measurement of the LAE clustering signal to date given our large total area and the use of multiple fields, which minimize any cosmic variance effects.

In the lower panel of Fig. 10 we compare our results to the predictions of the GALFORM model. The filled circles are the same as in the top panel, whilst the solid curve and shaded regions show the median clustering measured from 225 mocks and the  $1\sigma$  and





**Figure 10.** Top panel: the angular correlation function of the  $z = 3.1$  LAEs. The filled black circles show the result based on our VLRs LAE sample. Also shown are model fits to the VLRs LAE correlation function based on using the  $n(z)$  from our VIMOS spectroscopically observed sample (solid curve) and using an  $n(z)$  based on the NB497 filter profile (dashed curve). A number of literature results are also plotted: Ouchi et al. (2010) result is shown by the blue triangles; the Gawiser et al. (2007) result is shown by the orange diamonds; and the result of Hayashino et al. (2004) is shown by the cyan squares. Lower panel: the same VLRs LAE result (filled black circles) alongside the LAE clustering prediction from the GALFORM model. The shaded regions show the 68th and 95th percentile uncertainties based on the standard deviation of 225  $1.07 \text{ deg}^2$  mock catalogues.

$2\sigma$  uncertainties. Each mock covers a survey area of  $1.07 \text{ deg}^2$  (i.e. equivalent to our own survey area) and so the uncertainties given by the shaded regions give a further estimate of the uncertainties in our observations (and one that corroborates our estimates based on the data itself).

We find that the model results over-predict the observational measurement at the  $\approx 2\sigma$  level. Although the semi-analytical model is able to reproduce the narrow-band and continuum luminosity function of LAEs at bright magnitudes, it is worth noting that the clustering results will be dominated by the more numerous faint LAEs, which the model fails to reproduce. It may be that the discrepancy here is connected to the model number count disagreements at faint magnitudes. Taking our results alongside comparable over-predictions of clustering measurements by semi-analytical simulations (e.g. Guo et al. 2011, 2013; Bielby et al. 2014), it is clear that clustering measurements provide an important additional constraint on galaxy formation models. Indeed, resolving these discrepancies at faint magnitudes in the LAE numbers and clustering remains a challenge for the semi-analytical models.

### 5.3 Real-space clustering

We now parametrize the LAE clustering in terms of the real-space clustering length as measured via the angular two-point correlation

function  $w(\theta)$ , which is effectively a weighted projection of the spatial two-point correlation function  $\xi(r)$ . It is common to transform from the spatial to the angular correlation function using Limber’s approximation (Limber 1953); however this requires that the depth of the galaxy survey,  $\pi$  (i.e. distance probed along the line of sight), should be much greater than the on sky maximum separation,  $\sigma$  (e.g. Limber 1953; Simon 2007). For our survey volume, this is not the case – the depth of our survey is  $\pi \approx 60 h^{-1} \text{ Mpc}$  and the width is  $\sigma \approx 40 h^{-1} \text{ Mpc}$ . We therefore use the full analytical form in transforming from the real-space clustering form to the projected angular clustering.

As the underlying model for the real-space clustering, we assume a power-law with a slope of  $\gamma = 1.8$  (e.g. Peebles 1980):

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-1.8} \quad (5)$$

where  $r$  is the real-space separation between two galaxies and  $r_0$  is the clustering length parameter.

The spatial correlation function,  $\xi(r)$  can be related to  $w(\theta)$  by (e.g. Phillipps et al. 1978):

$$w(\theta) = \frac{\int_0^\infty dz_1 f(z_1) \int_0^\infty dz_2 f(z_2) \xi(r)}{\left[\int_0^\infty dz f(z)\right]^2} \quad (6)$$

where  $f(z)$  is the radial distribution of sources which is given by

$$f(z) \equiv \chi^2(z) \frac{d\chi(z)}{dz} n_c(z) \phi(z) \quad (7)$$

where  $\phi(z)$  is the selection function of the sample,  $n_c(z)$  is the comoving number density of the sources,  $\chi$  is the radial comoving distance, and  $r = r(\theta, z_1, z_2)$  is a comoving separation between two points at  $z_1$  and  $z_2$ .

$$r \equiv \sqrt{\chi^2(z_1) + \chi^2(z_2) - 2\chi(z_1)\chi(z_2) \cos \theta} \quad (8)$$

We assume that over the redshift range probed, the actual number density of LAEs,  $n_c(z)$ , is constant, whilst we have calculated solutions with each of the filter response curve and the redshift distribution of the VIMOS detected sources as the selection function,  $\phi(z)$ . We performed a minimized  $\chi^2$  fit to determine the clustering length,  $r_0$ , given our  $w(\theta)$  clustering measurement and using equation (6) to transform from the real-space power-law correlation function,  $\xi(r)$ , and the angular correlation function,  $w(\theta)$ . With the  $n(z)$  of the spectroscopically confirmed LAEs and a fixed slope of  $\gamma = 1.8$ , we derived a result of  $r_0 = 2.79 \pm 0.34 h^{-1} \text{ Mpc}$ . The resulting  $w(\theta)$  model is shown by the solid curve in Fig. 10. With the selection function given by the filter transmission curve, we found a best-fitting value of  $r_0 = 2.86 \pm 0.33 h^{-1} \text{ Mpc}$  – shown by the dashed curve in Fig. 10. Given that these two estimates of  $r_0$  are consistent with each other within the estimated uncertainties, it is clear that the selection function is not significant for the result. We note that the data points appear to rise above the fitted model at small separations ( $\theta \lesssim 0.15 \text{ arcmin}$ ), potentially indicating the beginning of the one-halo term. Indeed, the scale at which this turn-up in the correlation function is seen is comparable to that reported for  $z \approx 3$  LBGs by Hildebrandt et al. (2009).

Such a low clustering measurement is consistent with previous measurements of the clustering of  $z \approx 3$  LAEs. Those shown in Fig. 10 have reported clustering lengths of  $r_0 = 2.5^{+0.6}_{-0.7} h^{-1} \text{ Mpc}$  (Gawiser et al. 2007) and  $r_0 = 1.70^{+0.39}_{-0.46} h^{-1} \text{ Mpc}$  (Ouchi et al. 2010), whilst Hayashino et al. (2004) did not present a measurement of the clustering length. As discussed, the Ouchi et al. (2010) sample is based on a comparable EW cut as our own sample and so presents the most like-for-like comparison. Compared to their ‘best’

estimate of the clustering length, we measure a moderately higher  $r_0$ , at the  $\approx 2\text{--}3\sigma$  level. Both measurements are subject to uncertainties in contamination and they report a ‘maximum’ clustering length estimate of  $r_0 = 1.99^{+0.45}_{-0.55}$ . Similarly, our own measurement may be ‘over-corrected’ for contamination based on the spectroscopic non-detections, i.e. if the non-detections were in fact LAE too faint to be detected in the VIMOS observations. If this were the case, and we assume no contamination of our sample, then we would measure a clustering length of  $r_0 = 2.18 h^{-1} \text{Mpc} \pm 0.26$ , somewhat more consistent with the previous measurement.

The LAE clustering is low compared to both the normal star-forming and passive galaxy populations at  $z \sim 2\text{--}3$ . Indeed, at  $z \sim 3$ , the LBG population of star-forming galaxies is found to have clustering lengths of  $r_0 \sim 4 h^{-1} \text{Mpc}$  (e.g. Adelberger et al. 2005a; Bielby et al. 2011, 2013). Given literature results suggesting a strong link between galaxy stellar mass and median halo mass (e.g. Bielby et al. 2014; McCracken et al. 2015), such a low clustering length is strongly suggestive of LAEs being low-mass galaxies residing in low-mass dark matter haloes.

#### 5.4 Dark matter halo masses

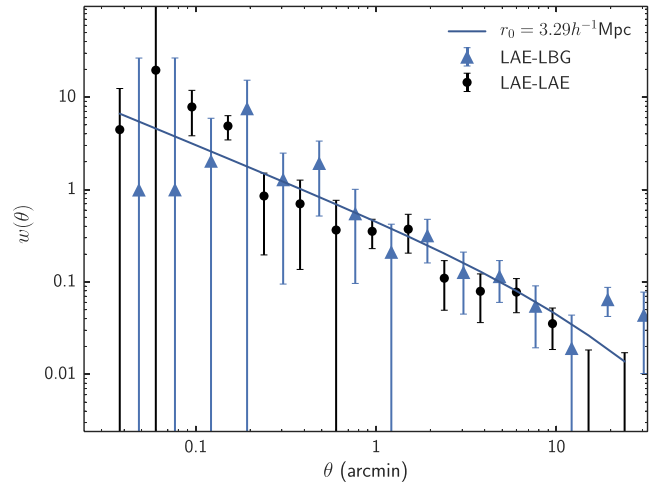
We now explicitly estimate median halo masses from our clustering results. Following Bielby et al. (2014), we calculated the mean masses of dark matter haloes within the galaxy samples by using the clustering results. We used the formalism developed by Mo & White (1996) which provides a relationship between the halo-bias to the mean halo mass via Sheth, Mo & Tormen (2001). Based on the measured clustering length derived using the filter transmission profile, we find a clustering bias of  $b = 2.13 \pm 0.22$  and halo mass  $M_{\text{DM}} = 10^{11.0 \pm 0.3} h^{-1} M_{\odot}$ .

Gawiser et al. (2007) reported a bias factor of  $b = 1.7 \pm 0.4$  and a median dark matter halo masses of  $M_{\text{DM}} = 10^{10.9 \pm 0.9} M_{\odot}$ , whilst Ouchi et al. (2010) reported a bias value of  $b = 1.7 \pm 0.8$  and a halo mass  $M_{\text{DM}} \approx 6.7^{+42.0}_{-6.7} \times 10^{10} M_{\odot}$ . We note that, although the cosmologies used in these papers and our own differ slightly, this only affects the halo mass values by a factor of  $\approx 10^{0.1}$ . Ouchi et al. (2010) concluded that the average dark halo mass of LAEs is  $\sim 10^{11 \pm 1} M_{\odot}$  at  $z=2\text{--}7$  and our own measurement is more consistent with this assertion than their own results at  $z = 3.1$ . As for the GALFORM model clustering result shown in Fig. 10, this corresponds to LAEs occupying dark matter haloes with a median mass of  $M_{\text{DM}} \approx 10^{11.3 \pm 0.3} h^{-1} M_{\odot}$ , approximately  $1\sigma$  higher than the measurement based on our observations.

The median halo mass for LBGs is estimated to be  $\sim 10^{12} M_{\odot}$ , about one order of magnitude larger than that of LAEs (Hamana et al. 2004; Ouchi et al. 2004b, 2005; Lee et al. 2006, 2009; Hildebrandt et al. 2009; McLure et al. 2010). Indeed, from our own LBG measurements, we estimate a median LBG halo mass at  $z \sim 3$  of  $M_{\text{DM}} = 10^{11.6 \pm 0.2} h^{-1} M_{\odot}$  with  $b = 2.37 \pm 0.21$  (Bielby et al. 2013), larger than the LAE measurements at the  $\approx 2\sigma$  level.

#### 5.5 LAE–LBG cross-correlation

We now turn to the LAE–LBG cross-correlation, combining the LAE candidates selected in this paper with the spectroscopically confirmed sample of LBGs presented by Bielby et al. (2013). Combining the cross-correlation of the two samples with their respective auto-correlation samples, we may infer the relationship between the dark matter density distributions that they each inhabit (see e.g. Ryan-Weber 2006; Tejos et al. 2014).



**Figure 11.** The LAE–LBG angular cross-correlation function calculated using the VLSR LAE and LBG samples (triangles). A fit to the data using a single power-law form (with a correlation length of  $r_0 = 3.29 \pm 0.57$ ) is shown by the solid curve. For comparison, the LAE auto-correlation is reproduced (circles).

From the Cauchy–Schwartz inequality (e.g. Cauchy 1821; Mitrinovic, Pecaric & Fink 2013), Adelberger et al. (2003) derived the following inequality for the auto- and cross-correlation functions of the two populations:

$$\xi_{ab}^2 \leq \xi_{aa} \xi_{bb}, \quad (9)$$

where  $\xi_{ab}$  is the cross-correlation and  $\xi_{aa}$  and  $\xi_{bb}$  are the respective auto-correlation functions. As discussed (in reference to the galaxy–HI cross-correlation) in Tejos et al. (2014), if the equality holds, the two samples are deemed to trace the same underlying mass distribution, and the relative biases can be used to infer the ratio between the dark matter halo populations that the two populations occupy.

We calculate the cross-correlation between the photometric LAE sample and a subset of the LBGs matching the criteria  $|z_{\text{LBG}} - 3.08| \leq 0.08$  (giving  $\approx 80$  LBGs). Although this leaves a relatively small number of LBGs, it optimises the cross-correlation signal by excluding poorly correlated pairs at large separations along the line of sight. For the calculation, we use the Landy–Szalay estimator:

$$w(\theta) = \frac{\langle D_a D_b(\theta) \rangle - \langle D_a R_b(\theta) \rangle - \langle R_a D_b(\theta) \rangle + \langle R_a R_b(\theta) \rangle}{\langle R_a R_b(\theta) \rangle} \quad (10)$$

where  $D_a D_b$  is the LBG–LAE pair count,  $D_a R_b$  is the LBG–random LAE pair count,  $D_a R_b$  is the random LBG–LAE pair count and  $R_a R_b$  is the random LBG–random LAE pair count (all normalized by the number of randoms per galaxy). We use random catalogues with 20 times the number of randoms as galaxies in each case.

The results are shown by the triangles in Fig. 11, alongside the LAE auto-correlation result (filled circles). We fit the cross-correlation result based on a single power-law form in real-space, i.e.  $\xi(r)$ , transformed to  $w(\theta)$  based on the formalism outlined in Section 5.3 (but now incorporating the two separate redshift distributions rather than just that of the LAE). The best-fitting result is shown by the solid curve in Fig. 11, corresponding to a correlation length of  $r_0 = 3.29 \pm 0.57 h^{-1} \text{Mpc}$  (we do not fit for the slope, but take a fixed value of  $\gamma = -1.8$ ). This compares to a LAE auto-correlation clustering length of  $r_0 = 2.86 \pm 0.33 h^{-1} \text{Mpc}$  and a LBG auto-correlation clustering length of  $r_0 = 3.83 \pm 0.24 h^{-1} \text{Mpc}$  (with a slope of  $\gamma = -1.6$ ; Bielby et al. 2013).

Evaluating equation (9) using the clustering amplitudes corresponding to  $r = 1 h^{-1}$  Mpc (i.e.  $r_0^y$ ), we find  $\xi_{ab}^2/\xi_{aa}\xi_{bb} = 1.28 \pm 0.46$ , i.e.  $\approx 1$ . As discussed by Adelberger et al. (2005b) and Tejos et al. (2014), this unity in the relation does indeed imply that the two photometric selections give samples of galaxies that consistently trace the same underlying dark matter distribution and we can therefore use this result to provide a second constraint on the typical halo mass that our LAE sample occupies. Similar to the auto-correlation case, the correlation function can now be related to the underlying dark matter correlation function by:

$$\xi_{\text{LAE-LBG}} = b_{\text{LAE}} b_{\text{LBG}} \xi_{\text{DM}} \quad (11)$$

where  $\xi_{\text{LAE-LBG}}$  is the LAE-LBG cross-correlation,  $b_{\text{LAE}}$  is the LAE bias,  $b_{\text{LBG}}$  is the LBG bias, and  $\xi_{\text{DM}}$  is the underlying dark matter clustering as in the previous section. Based on this and our cross-clustering results, we calculate a further measure of the LAE bias of  $b_{\text{LAE}} = 2.24 \pm 0.37$ . Using this to estimate a halo mass as in the previous section, we find a value of  $M_{\text{DM}} = 10^{11.1 \pm 0.4} h^{-1} M_{\odot}$ . The cross-correlation results corroborate the auto-correlation analysis with both dark matter halo estimates being consistent.

## 6 DISCUSSION

Our results, based on the largest  $z \sim 3$  narrow-band LAE sample yet surveyed and the first to combine multiple independent large fields, build on the results presented by others (e.g. Gawiser et al. 2007; Gronwall et al. 2007; Ouchi et al. 2008, 2010) to add to the overall picture of the nature of LAE galaxies at this epoch.

At first sight, LAEs at  $z \approx 3$  seem to be a high-density population with a sky density of  $\approx 0.1$  arcmin $^{-2}$  in a redshift range of width only  $\Delta z \approx 0.05$  to the various limits used in our survey. Taking the sky density of LBGs to be  $\approx 1.8$  arcmin $^{-2}$  in the  $2.5 < z < 3.5$  redshift range (Steidel et al. 2003), this corresponds to a roughly similar sky density. On the other hand, comparing galaxy numbers at fixed continuum luminosity (e.g. Ouchi et al. 2008, Fig. 8 in this paper) we confirm that LAE are a relatively rare component of the galaxy population with abundances  $\approx 10$  times lower than that of LBGs.

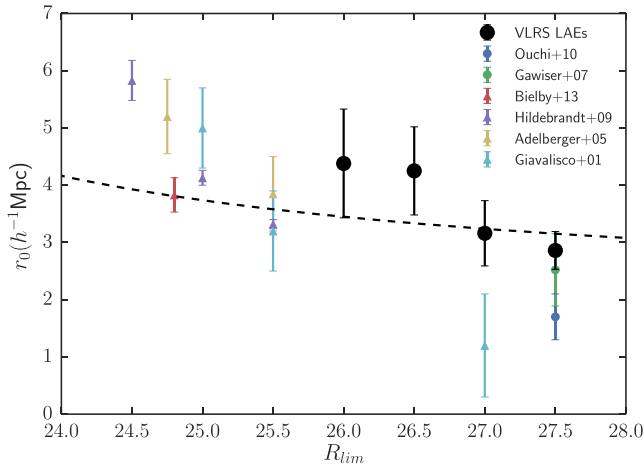
Given their relative rareness, one might expect that LAEs would occupy rare high-mass dark matter haloes based on a simple abundance matching approach. Our clustering results show a somewhat higher clustering signature than the previous best study of LAEs at this redshift (Ouchi et al. 2010); however the clustering results still clearly show that the LAEs are relatively poorly clustered compared to average star-forming galaxies at  $z \approx 3$  (e.g. Adelberger et al. 2005a; Bielby et al. 2013; Durkalec et al. 2015). Indeed, our measurement of LAE clustering implies a median halo mass of  $M_{\text{DM}} \approx 10^{11} h^{-1} M_{\odot}$ , for which the number densities are predicted to be  $\sim 0.008 h^3 \text{Mpc}^{-3}$  (integrating the mass function over the mass range  $10^{11 \pm 0.1} h^{-1} \text{Mpc}$ ) in a  $\Lambda$  cold dark matter Universe (Jenkins et al. 2001; Murray, Power & Robotham 2013). It therefore follows (given our measured number density of LAEs of  $\approx 2 \times 10^{-3} h^3 \text{Mpc}^{-3}$ ) that LAEs possess a relatively low occupation number (i.e. number of galaxies per halo) of  $\ll 1$  and trace the dark matter distribution relatively sparsely (corroborating the results of Ouchi et al. 2010). It is common to link a galaxy sample's properties as far as possible to its host halo properties, but it is apparent that the galaxy properties that govern LAE selection (i.e. Ly $\alpha$  equivalent width) are unlikely to be well correlated with the halo mass. Indeed, our results support the hypothesis that LAEs presence within dark matter haloes is stochastic in nature (Nagamine et al. 2010; Ouchi et al. 2010).

Our measured halo mass of  $M_{\text{DM}} = 10^{11.0 \pm 0.3} h^{-1} M_{\odot}$  is consistent with estimated LAE halo masses across a wide redshift range of  $2 < z < 7$  (i.e. the full range of redshifts probed by available LAE samples). Indeed, our measurement brings a tighter agreement with available measurements at other redshifts than the previous best measurement at  $z \approx 3$  (i.e. Ouchi et al. 2010). As such, this enhances the assertion made by Ouchi et al. (2010) that LAEs exist as a stage in the evolution of at least some galaxies upon the host dark halo reaching a mass of  $M_{\text{DM}} \approx 10^{11} h^{-1} M_{\odot}$ .

The LAE sample as selected using the NB497 filter lies within the redshift range of LBGs selected based on their  $U - B$  (or  $U - G$ ) colours. It is worth considering then how the galaxies identified by these two selections relate to each other. LBGs by definition are galaxies that possess a Lyman break; however this is not intrinsic to the galaxies themselves and so LAEs are also LBGs in the sense that they will have a Lyman break of some form, although it still may be too weak to satisfy the LBG cut (Cooke et al. 2014). Nevertheless, since the LAE number density at fixed continuum luminosity is  $\approx 10$  times lower than that of the LBGs, the LAE selection appears simply as an extreme subset of the LBG selected population. Indeed, 40 per cent of the LBG population has been shown to exhibit Ly $\alpha$  in emission, with  $\approx 25$  per cent of LBGs having equivalent widths of  $\gtrsim 20 \text{\AA}$  (Shapley et al. 2003) and a tail in Ly $\alpha$  equivalent width up to  $\sim 200 \text{\AA}$ . Clearly, at least some LAEs would fall into an LBG selected sample and the numbers seem to approximately tally given that Gronwall et al. (2007) reported number densities of  $\gtrsim 20 \text{\AA}$  equivalent width LAEs were approximately a third of the number densities of LBGs. Indeed, fig. 8 of Shapley et al. (2003) shows that  $\approx 12$  per cent of their LBG sample have equivalent widths of  $\gtrsim 60 \text{\AA}$ , which ties in very well with the number densities we find for our sample (as also noted by Ouchi et al. 2008 for their own  $z \approx 3.1$  EW  $\gtrsim 60 \text{\AA}$  LAE sample.).

Our clustering analysis corroborates this picture of the LAE selected sample being an extreme subsample of the LBG population, whereby they show lower clustering due to our LAE sample probing a low-luminosity subset of the LBG population (at least compared to the mostly  $R \lesssim 25$  LBG samples from which clustering measurements have been made.). This is consistent with the UV luminosity dependence of the observed LBG clustering, whereby LBG samples selected with fainter UV luminosities are observed to have lower clustering signals (e.g. Giavalisco & Dickinson 2001; Lee et al. 2006, 2009; Yoshida et al. 2008). In Fig. 12, we show the clustering results taken from a number of LBG (triangles) and LAE (circles) surveys including our own LBG clustering measurement and measurements of the clustering of our LAE sample when limited by a range of continuum magnitudes (filled black circles). The literature  $z \approx 3$  LAE results shown are from Gawiser et al. (2007) and Ouchi et al. (2010), and the  $z \sim 3$  LBG samples are from Giavalisco & Dickinson (2001), Adelberger et al. (2005a), Hildebrandt et al. (2009), Bielby et al. (2013). The dashed curve shows the predicted  $r_0$ -magnitude dependence assuming a 1:1 ratio between UV continuum luminosity and halo mass, and using the Mo & White (1996) formalism to derive clustering lengths.

We find a tentative continuum luminosity dependence in our LAE results, albeit only at the  $\approx 1\sigma$  level. Taking our own LBG clustering results (which include the five fields used for the LAE study in this paper), our results are in broad agreement since our LBG limit is  $R_{AB} \approx 24.8$  and our lowest LAE continuum limit in Fig. 12 is  $R_{AB} \approx 27$  or a factor of  $\approx 8$  in UV luminosity. Assuming that luminosity is approximately proportional to mass, given the above halo model this represents a factor of  $\approx 1.5$  in bias. This compares to the observed factor of  $1.35 \pm 0.2$  given the measured  $r_0$  for LBG and LAE



**Figure 12.** We show the clustering results as a function of limiting continuum magnitude for our LAE sample (large circles) and a range of  $z \approx 3$  LAE (circles) and LBG (triangles) samples. The dashed curve shows the predicted luminosity dependence based on a 1:1 relationship between luminosity and halo mass.

that we have measured above. The LAE clustering length of  $r_0 = 2.86 \pm 0.33 h^{-1}$  Mpc is consistent with the decreasing trend of LBG clustering length measured by Hildebrandt et al. (2009). The VLRs LAE and LBG samples show an  $R_{\text{lim}}-r_0$  dependence consistent with this simple model. At brighter magnitude cuts, however, a steeper dependence is exhibited by the results of Adelberger et al. (2005a) and Hildebrandt et al. (2009).

Larger LAE data sets are required to test if at the  $R_{AB} \approx 25$  LBG limit, LAEs and LBGs have the same clustering amplitude, but at this point it seems that there is consistency with a model where LAEs can be treated as a subset of LBGs and where halo mass is proportional to UV luminosity in this combined population. The LBGs have a spatial number density that is consistent with their halo masses implied from clustering but the LAE number density is too low, implying that they can only randomly sample their potential halo hosts. From the cross-correlation results and the equality of the ratio between the cross-correlation and the respective auto-correlations, it is also evident that, although the fainter LAEs only randomly sample smaller mass haloes, both the LBG and LAE populations trace the underlying dark matter distribution in a linear way but with significantly different bias.

## 7 CONCLUSIONS

In this paper, we have presented NB497 band photometric observations within five of our VLRs fields, taken using *Subaru* Suprime-Cam. Combining these narrow-band images with  $B$ ,  $V$  and  $R$  deep imaging, we have made a photometric selection of  $z \approx 3.1$  LAEs with an equivalent width limit of  $\approx 65 \text{ \AA}$ . We have also made spectroscopic follow-up observations of LAE candidates in three of the five fields, using VLT VIMOS. Two factors mean that this survey improves on previous work on the study of LAEs at comparable epochs: first our choice of five independent survey fields minimizes the potential effects of cosmic variance on our results compared to previous work; and secondly that our fields contain significant  $z \approx 3$  LBG data that have allowed us to perform a first cross-correlation analysis of the two populations.

The major findings of our study are summarized as follows.

(1) Our final LAE selection, based on the combination of  $B - m_{\text{NB497}}$  with either  $V - m_{\text{NB497}}$  or  $R - m_{\text{NB497}}$ , gives a success rate of  $78 \pm 18$  per cent based on our spectroscopic observations. The equivalent width limit of our final selection is  $\approx 65 \text{ \AA}$ , whilst the flux limit is  $\approx 2 \times 10^{-17} \text{ erg cm}^2 \text{ s}^{-1}$  (equivalent to  $L \approx 10^{42} \text{ erg s}^{-1}$ ). A catalogue of the photometric LAEs from the five VLRs fields presented here is available at the Strasbourg Astronomical Data Center (CDS, <http://cds.u-strasbg.fr/>).

(2) The spectroscopic observations produced 35 confirmed LAEs at  $z \approx 3.1$ , with 23 of these falling within our final optimized selection criteria. We found no strong evidence for any contamination from low-redshift  $\text{O II}$  emitters (although  $\lesssim 5$  per cent of the detections remain ambiguous given the observational constraints). The only significant ‘contamination’ of the selection is from objects with no discernible emission in the observed spectra, pointing to either low  $\text{Ly}\alpha$  flux objects and/or interlopers (both likely entering the selection due to photometric uncertainties).

(3) We have calculated the NB497 number counts for our LAE sample (corrected for contamination) and find it to be consistent within the uncertainties with the previous observations of Gronwall et al. (2007, equivalent width limit  $\approx 20 \text{ \AA}$ ) and Ouchi et al. (2008, equivalent width limit  $\approx 65 \text{ \AA}$ ). Our selection is shown to probe the knee of the LAE  $\text{Ly}\alpha$  luminosity function based on an observed turnover in the narrow-band number counts at  $m_{\text{NB}} \sim 24.5$ .

(4) We derived the  $R$ -band/rest-frame UV continuum luminosity functions of our sample of LAEs, again corrected for contamination. Our luminosity function lies at a level  $\approx 10$  times lower than that of the  $z \approx 3$  LBG luminosity function, consistent with the LAE population being a small subset of the LBG population. At  $R \lesssim 26$  ( $M_{\text{UV}} \lesssim -19.6$ ), our results are consistent with the fraction of strong emitters being constant or marginally increasing towards fainter continuum luminosities, consistent with the results of Ouchi et al. (2008). Over a comparable magnitude range, we find our UV continuum luminosity function is significantly lower (by a factor of  $\gtrsim 3$ ) than that measured by Gronwall et al. (2007); however this appears consistent with their lower equivalent width threshold of  $\approx 20 \text{ \AA}$ .

(5) We have measured the angular correlation function for our  $z \sim 3.1$  LAE photometric sample in our five fields. Our results (corrected for contamination) are significantly lower than the observed clustering of LBGs at the same redshift. We measure a clustering length of  $r_0 = 2.86 \pm 0.33 h^{-1}$  Mpc (assuming a slope of  $\gamma = 1.8$ ), which corresponds to a clustering bias of  $b = 2.13 \pm 0.22$  and a median halo mass of  $10^{11.0 \pm 0.3} h^{-1} M_{\odot}$ . This indicates that LAEs reside in low-mass dark matter haloes, but given their number densities they have a low occupation number – i.e. only a small fraction of such haloes actually host an LAE.

(6) We measure the LBG–LAE cross-correlation function for the first time and find the results to be in agreement with the auto-correlation analysis, providing a useful consistency check on our results. From the cross-correlation analysis, we find a bias for the LAE population of  $b = 2.24 \pm 0.37$  and a typical dark matter halo mass of  $M_{\text{DM}} = 10^{11.1 \pm 0.4} h^{-1} M_{\odot}$ .

(7) Thus our results for LAE number counts, clustering and cross-clustering are in agreement with a view that LAE, as selected here, are a low-luminosity subset of LBGs that inhabit low-mass dark matter haloes with a small occupancy. Comparison to current semi-analytical models shows good agreement in the narrow-band counts, but poorer agreement in the LAE clustering properties. It



will be interesting to see how future changes to semi-analytical models can better match the observational results.

## ACKNOWLEDGEMENTS

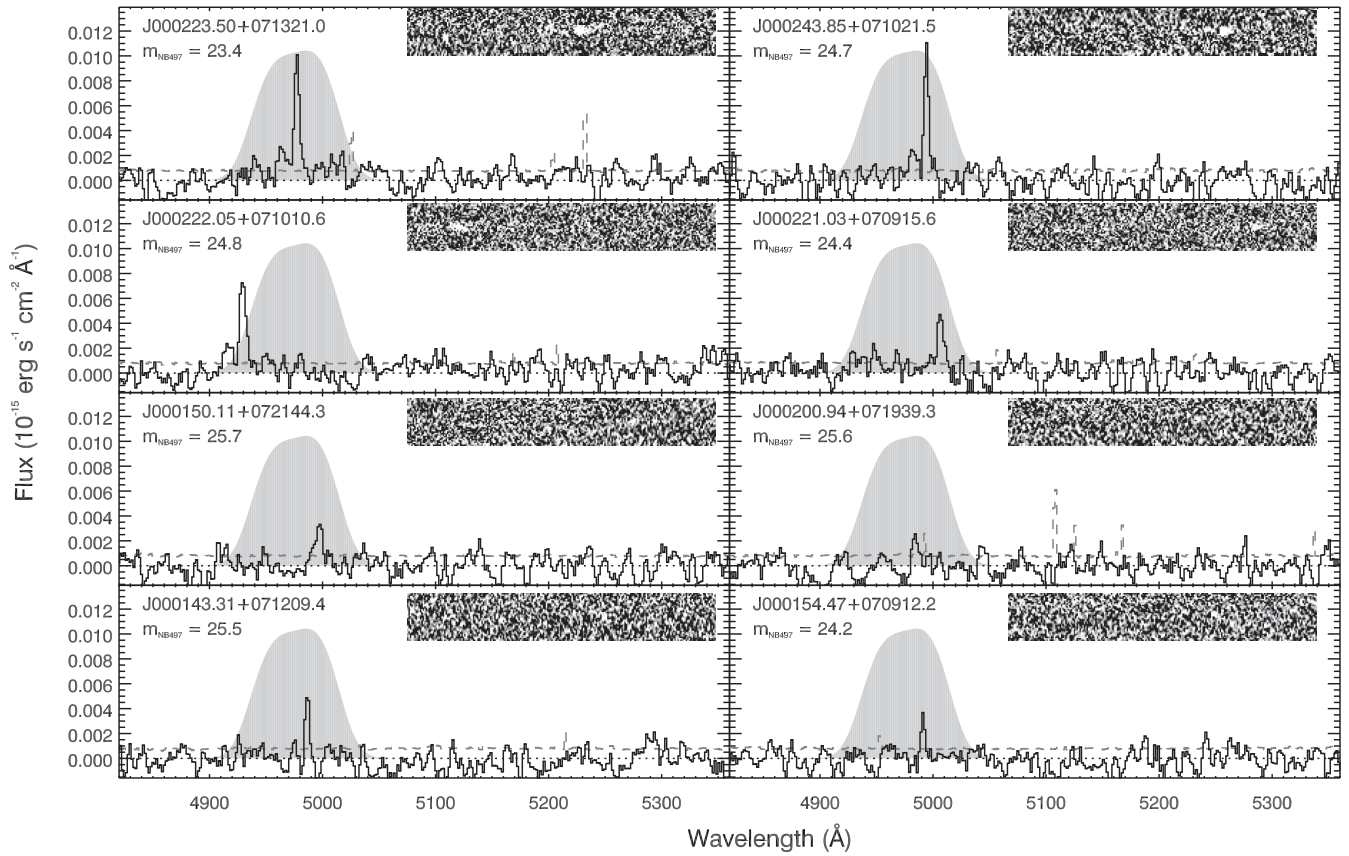
RMB and TS acknowledge the funding of their work through the UK STFC research grant ST/L00075X/1. PT acknowledges financial support from the Royal Thai Government. Leopoldo Infante and Jorge González-López obtained partial support from CATA, Conicyt Basal program. The plots in this publication were primarily produced using Matplotlib (Hunter 2007, <http://www.matplotlib.org>). We finally thank Professor Tomoki Hayashino for providing access to the NB497 filter for our observations.

## REFERENCES

- Adelberger K. L., Steidel C. C., Shapley A. E., Pettini M., 2003, *ApJ*, 584, 45
- Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005a, *ApJ*, 619, 697
- Adelberger K. L., Shapley A. E., Steidel C. C., Pettini M., Erb D. K., Reddy N. A., 2005b, *ApJ*, 629, 636
- Bañados E., Venemans B., Walter F., Kurk J., Overzier R., Ouchi M., 2013, *ApJ*, 773, 178
- Berry M. et al., 2012, *ApJ*, 749, 4
- Bertin E., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, *ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV*. Astron. Soc. Pac., San Francisco, p. 112
- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
- Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, *ASP Conf. Ser. Vol. 281, Astronomical Data Analysis Software and Systems XI*. Astron. Soc. Pac., San Francisco, p. 228
- Bielby R. M. et al., 2011, *MNRAS*, 414, 2
- Bielby R. et al., 2013, *MNRAS*, 430, 425
- Bielby R. M. et al., 2014, *A&A*, 568, A24
- Bouché N., Lowenthal J. D., 2004, *ApJ*, 609, 513
- Cauchy A. L., 1821, *Cours d'analyse de l'École Royale Polytechnique*. Editrice Clueb Bologna
- Ciardullo R. et al., 2012, *ApJ*, 744, 110
- Cooke J., Ryan-Weber E. V., Garel T., Díaz C. G., 2014, *MNRAS*, 441, 837
- Cowie L. L., Hu E. M., 1998, *AJ*, 115, 1319
- Crighton N. H. M. et al., 2011, *MNRAS*, 414, 28
- Durkalec A. et al., 2015, *A&A*, 583, A128
- Francke H., 2009, *New Astron. Rev.*, 53, 47
- Fynbo J. U., Möller P., Thomsen B., 2001, *A&A*, 374, 443
- Fynbo J. P. U., Ledoux C., Möller P., Thomsen B., Burud I., 2003, *A&A*, 407, 147
- Gawiser E. et al., 2007, *ApJ*, 671, 278
- Giavalisco M., Dickinson M., 2001, *ApJ*, 550, 177
- Gronwall C. et al., 2007, *ApJ*, 667, 79
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, *MNRAS*, 428, 1351
- Hamana T., Ouchi M., Shimasaku K., Kayo I., Suto Y., 2004, *MNRAS*, 347, 813
- Hayashino T. et al., 2003, *Publ. Natl. Astron. Obser. Jpn*, 7, 33
- Hayashino T. et al., 2004, *AJ*, 128, 2073
- Hildebrandt H., Pielorz J., Erben T., van Waerbeke L., Simon P., Capak P., 2009, *A&A*, 498, 725
- Horne K., 1986, *PASP*, 98, 609
- Hu E. M., McMahon R. G., 1996, *Nature*, 382, 231
- Hu E. M., Cowie L. L., McMahon R. G., 1998, *ApJ*, 502, L99
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, *MNRAS*, 321, 372
- Kovač K., Somerville R. S., Rhoads J. E., Malhotra S., Wang J., 2007, *ApJ*, 668, 15
- Lacey C. G. et al., 2015, preprint ([arXiv:1509.08473](https://arxiv.org/abs/1509.08473))
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Le Fèvre O. et al., 2003, in Iye M., Moorwood A. F. M., eds, *SPIE Conf. Ser. Vol. 4841, Society of Photo-Optical Instrumentation Engineers Conf. Ser.* p. 1670, doi:10.1117/12.460959
- Lee K.-S., Giavalisco M., Gnedin O. Y., Somerville R. S., Ferguson H. C., Dickinson M., Ouchi M., 2006, *ApJ*, 642, 63
- Lee K.-S., Giavalisco M., Conroy C., Wechsler R. H., Ferguson H. C., Somerville R. S., Dickinson M. E., Urry C. M., 2009, *ApJ*, 695, 368
- Limber D. N., 1953, *ApJ*, 117, 134
- Matsuda Y. et al., 2005, *ApJ*, 634, L125
- McCracken H. J. et al., 2015, *MNRAS*, 449, 901
- McLinden E. M., Rhoads J. E., Malhotra S., Finkelstein S. L., Richardson M. L. A., Smith B., Tilvi V. S., 2014, *MNRAS*, 439, 446
- McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D. P., Targett T. A., Ellis R. S., 2010, *MNRAS*, 403, 960
- Mitrinovic D. S., Pecaric J., Fink A. M., 2013, *Classical and New Inequalities in Analysis*. Springer Science & Business Media. Dordrecht, Netherlands
- Miyazaki S. et al., 2002, *PASJ*, 54, 833
- Mo H. J., White S. D. M., 1996, *MNRAS*, 282, 347
- Murray S. G., Power C., Robotham A. S. G., 2013, *Astron. Comput.*, 3, 23
- Nagamine K., Ouchi M., Springel V., Hernquist L., 2010, *PASJ*, 62, 1455
- Nilsson K. K. et al., 2007, *A&A*, 471, 71
- Orsi A., Lacey C. G., Baugh C. M., 2012, *MNRAS*, 425, 87
- Ouchi M. et al., 2003, *ApJ*, 582, 60
- Ouchi M. et al., 2004a, *ApJ*, 611, 660
- Ouchi M. et al., 2004b, *ApJ*, 611, 685
- Ouchi M. et al., 2005, *ApJ*, 635, L117
- Ouchi M. et al., 2008, *ApJS*, 176, 301
- Ouchi M. et al., 2010, *ApJ*, 723, 869
- Peebles P. J. E., 1980, *The Large-Scale Structure of the Universe*. Princeton Univ. Press, Princeton, NJ
- Phillipps S., Fong R., Fall R. S. E. S. M., MacGillivray H. T., 1978, *MNRAS*, 182, 673
- Rauch M. et al., 2008, *ApJ*, 681, 856
- Reddy N. A., Steidel C. C., 2009, *ApJ*, 692, 778
- Ryan-Weber E. V., 2006, *MNRAS*, 367, 1251
- Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, *ApJ*, 588, 65
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Shimasaku K. et al., 2004, *ApJ*, 605, L93
- Shioya Y. et al., 2009, *ApJ*, 696, 546
- Simon P., 2007, *A&A*, 473, 711
- Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, *ApJ*, 462, L17
- Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2000, *ApJ*, 532, 170
- Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, *ApJ*, 592, 728
- Tejos N. et al., 2014, *MNRAS*, 437, 2017
- Tummuangpak P., Bielby R. M., Shanks T., Theuns T., Crighton N. H. M., Francke H., Infante L., 2014, *MNRAS*, 442, 2094
- van der Burg R. F. J., Hildebrandt H., Erben T., 2010, *A&A*, 523, A74
- Venemans B. P. et al., 2007, *A&A*, 461, 823
- Yagi M., Kashikawa N., Sekiguchi M., Doi M., Yasuda N., Shimasaku K., Okamura S., 2002, *AJ*, 123, 66
- Yoshida M., Shimasaku K., Ouchi M., Sekiguchi K., Furusawa H., Okamura S., 2008, *ApJ*, 679, 269

## APPENDIX A: VLT VIMOS 1D AND 2D SPECTRA OF CONFIRMED LAES

The figures show the observed-frame spectra taken using VLT VIMOS with the HR\_Blue grism as part of our spectroscopic follow-up observations. All the spectra shown have been identified as LAEs with significant emission within the NB497 filter passband.



**Figure A1.** VIMOS HR\_red spectra (black solid curve) for confirmed LAEs in the QSO B2359+0653 field. The grey filled region shows the transmission profile of the NB497 filter. In addition, the insets show the corresponding 2D spectra showing only the wavelength region covered by the filter. The dashed lines show the noise estimated from the standard deviation of the background signal across the slit.

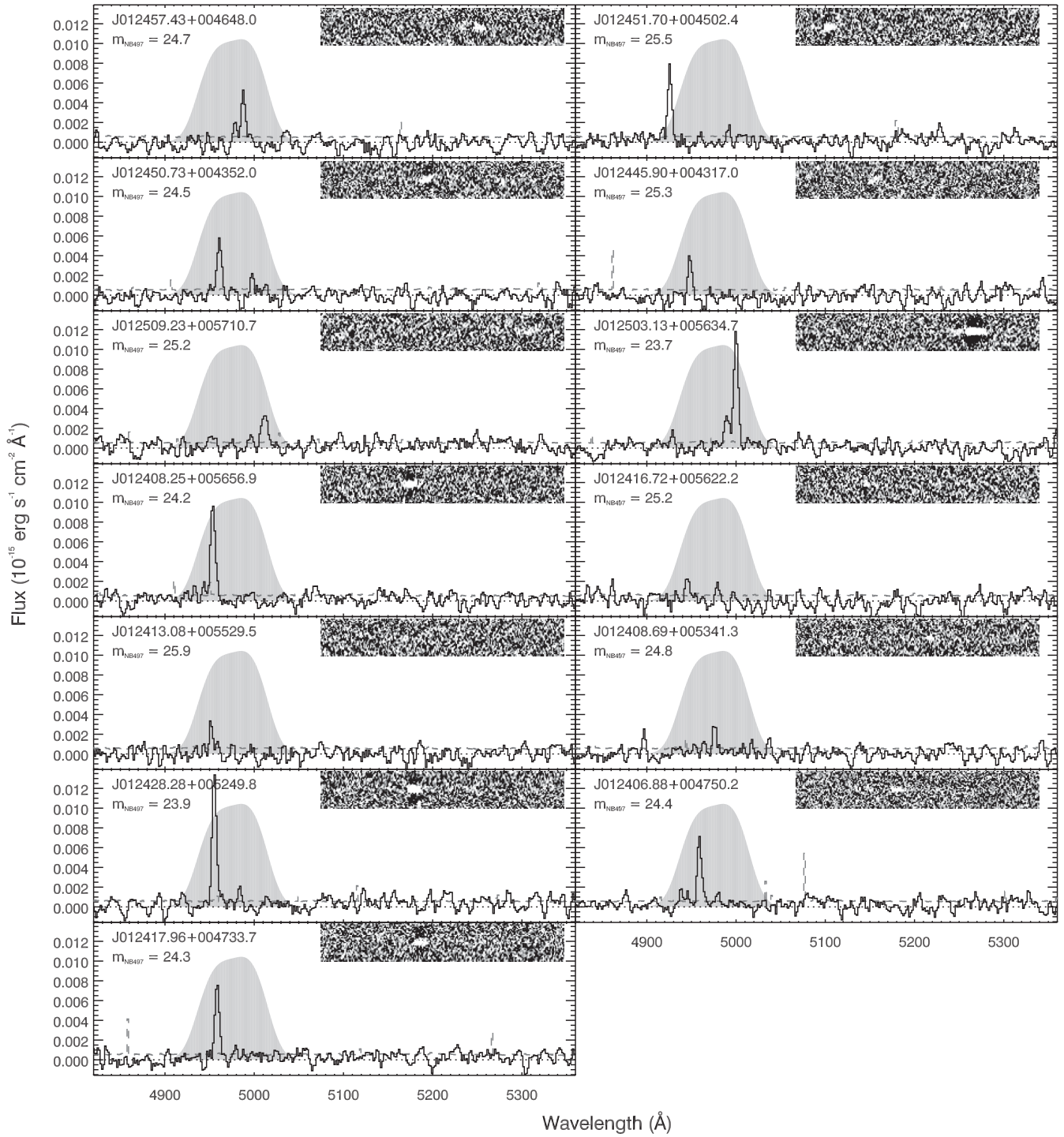
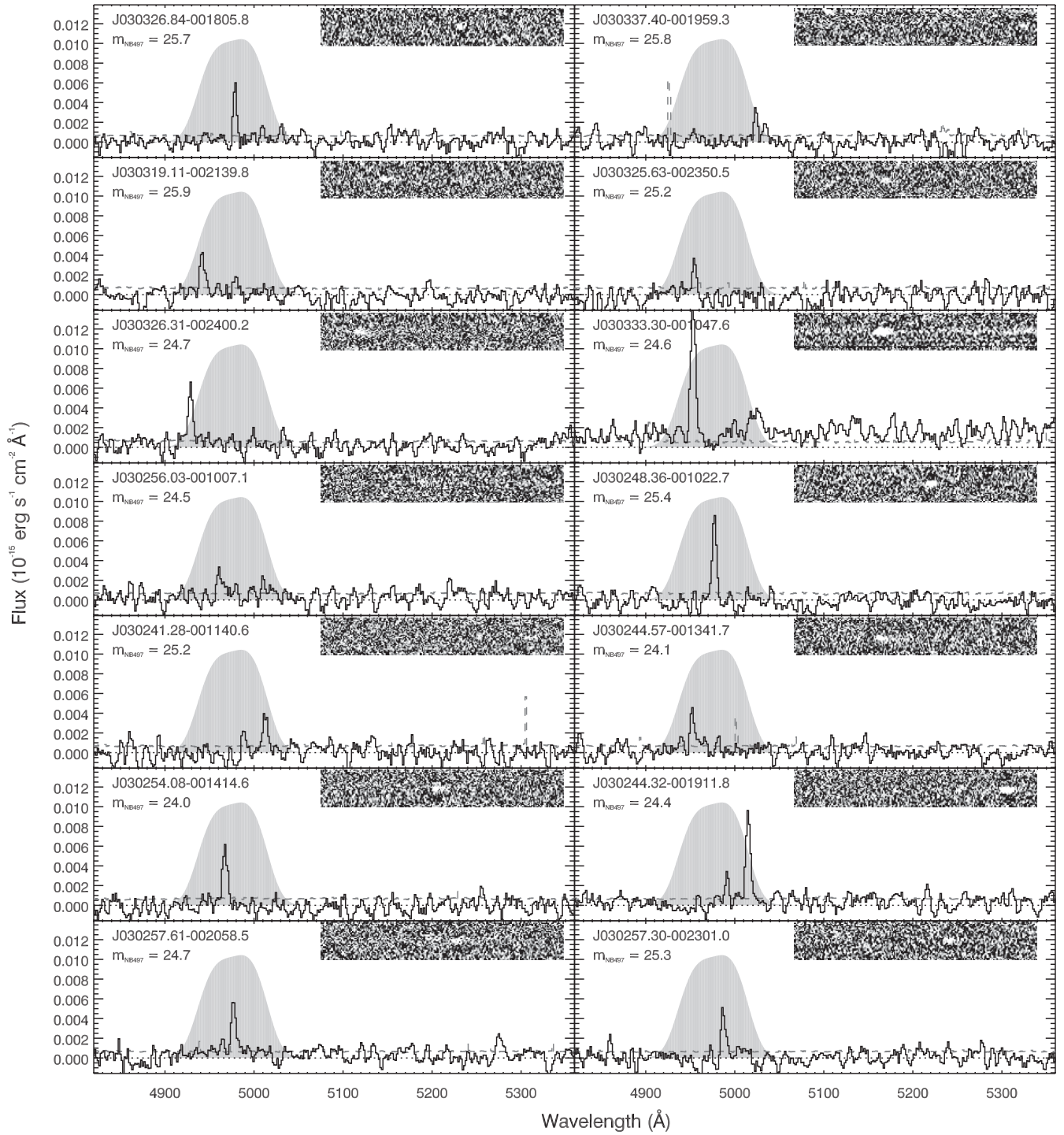


Figure A2. As for Fig. A1, but for QSO J0124+0044 field.



**Figure A3.** As for Fig. A1, but for LBQS Q0301–0035 field.

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