

The clustering of typical Ly α emitters from $z \sim 2.5\text{--}6$: host halo masses depend on Ly α and UV luminosities

A. A. Khostovan¹, ¹ ¹ D. Sobral,^{2,3} B. Mobasher,¹ J. Matthee^{1,3,4}, ¹ R. K. Cochrane¹, N. Chartab¹, ¹ M. Jafariyazani,¹ A. Paulino-Afonso^{2,6,7}, S. Santos² and J. Calhau²

¹Department of Physics and Astronomy, University of California, Riverside, USA

²Department of Physics, Lancaster University, Lancaster LA1 4YB, UK

³Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

⁴Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland

⁵SUPA, Institute for Astronomy, Royal Observatory of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

⁶Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, P-1349-018 Lisboa, Portugal

⁷Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, P-1749-016 Lisboa, Portugal

Accepted 2019 July 30. Received 2019 June 22; in original form 2018 October 31

ABSTRACT

We investigate the clustering and halo properties of ~ 5000 Ly α -selected emission-line galaxies (LAEs) from the Slicing COSMOS 4K (SC4K) and from archival NB497 imaging of SA22 split in 15 discrete redshift slices between $z \sim 2.5$ and 6. We measure clustering lengths of $r_0 \sim 3\text{--}6 h^{-1}$ Mpc and typical halo masses of $\sim 10^{11} M_\odot$ for our narrowband-selected LAEs with typical $L_{\text{Ly}\alpha} \sim 10^{42\text{--}43}$ erg s $^{-1}$. The intermediate-band-selected LAEs are observed to have $r_0 \sim 3.5\text{--}15 h^{-1}$ Mpc with typical halo masses of $\sim 10^{11\text{--}12} M_\odot$ and typical $L_{\text{Ly}\alpha} \sim 10^{43\text{--}43.6}$ erg s $^{-1}$. We find a strong, redshift-independent correlation between halo mass and Ly α luminosity normalized by the characteristic Ly α luminosity, $L^*(z)$. The faintest LAEs ($L \sim 0.1 L^*(z)$) typically identified by deep narrowband surveys are found in $10^{10} M_\odot$ haloes and the brightest LAEs ($L \sim 7 L^*(z)$) are found in $\sim 5 \times 10^{12} M_\odot$ haloes. A dependency on the rest-frame 1500 Å UV luminosity, M_{UV} , is also observed where the halo masses increase from 10^{11} to $10^{13} M_\odot$ for $M_{\text{UV}} \sim -19$ to -23.5 mag. Halo mass is also observed to increase from $10^{9.8}$ to $10^{12} M_\odot$ for dust-corrected UV star formation rates from ~ 0.6 to $10 M_\odot \text{ yr}^{-1}$ and continues to increase up to $10^{13} M_\odot$ in halo mass, where the majority of those sources are active galactic nuclei. All the trends we observe are found to be redshift independent. Our results reveal that LAEs are the likely progenitors of a wide range of galaxies depending on their luminosity, from dwarf-like, to Milky Way-type, to bright cluster galaxies. LAEs therefore provide unique insight into the early formation and evolution of the galaxies we observe in the local Universe.

Key words: galaxies: evolution – galaxies: haloes – galaxies: high-redshift – galaxies: star formation – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

The current state of galaxy formation and evolution theory suggests that galaxies formed with the assistance of their host dark matter haloes, where deep gravitational potential wells allowed for the accretion of cold gas to form the galaxies and fuel star formation activity (see Benson 2010 and Somerville & Davé 2015 for a

review). The era between the ‘end’ of the epoch of reionization ($z \sim 6$) and cosmic noon (peak of cosmic star-formation activity; $z \sim 2$) constitutes an important time period in the history of the Universe. It is within that time (~ 2 Gyr) that galaxies rapidly evolved with their typical star formation rates increasing by an order of magnitude (e.g. Madau & Dickinson 2014; Khostovan et al. 2015). Since galaxies reside and evolve within dark matter haloes, the host haloes likely play a fundamental role in the overall evolution of galaxies. How do galaxies and their host haloes co-evolve?

Addressing such a fundamental question requires large samples of high- z galaxies with well-understood selection functions. Since galaxies reside in dark matter haloes, their spatial clustering

* E-mail: akhostov@gmail.com

† NASA Earth and Space Science Fellow

‡ Zwicky Fellow

directly traces the host dark matter haloes, although this requires a few assumptions (e.g. halo mass and bias functions, occupation distributions; see Cooray & Sheth 2002 for a review).

One class of high- z star-forming galaxies that can be used to investigate clustering properties are Ly α emitters (LAEs). These are typically young, low-mass (e.g. Gawiser *et al.* 2006; Finkelstein *et al.* 2007; Guaita *et al.* 2011; Hagen *et al.* 2016; Shimakawa *et al.* 2017; Hao *et al.* 2018; Kusakabe *et al.* 2019), compact galaxies (e.g. Malhotra *et al.* 2012; Kobayashi *et al.* 2016; Paulino-Afonso *et al.* 2018) with high rest-frame equivalent widths (e.g. Malhotra & Rhoads 2002; Hu *et al.* 2010; Ciardullo *et al.* 2012; Zheng *et al.* 2014; Hashimoto *et al.* 2017), low metallicities, and high ionization states (e.g. Finkelstein *et al.* 2011; Erb *et al.* 2016; Trainor *et al.* 2016).

Samples of LAEs selected via narrowband surveys provide an efficient and robust window for probing the high- z Universe (e.g. Rhoads *et al.* 2000; Ouchi *et al.* 2008; Nilsson *et al.* 2009). Narrowband surveys have the added advantage of enabling large samples of galaxies by directly observing emission lines associated with star formation and active galactic nucleus (AGN) activity using specially designed photometric filters. Because the filter widths are quite narrow [in the range of 50–200 Å in full width at half-maximum (FWHM)], emission-line galaxies selected in narrowband surveys have reliable redshifts within 1–2 per cent error (albeit with typical 5–10 per cent contamination; see e.g. Sobral *et al.* 2018a) and are prime sources for future spectroscopic follow-up studies.

Previous narrowband surveys have searched for LAEs at $z \sim 2$ –7 (e.g. Cowie & Hu 1998; Rhoads *et al.* 2000; Gronwall *et al.* 2007; Ouchi *et al.* 2008; Matthee *et al.* 2015; Santos, Sobral & Matthee 2016; Konno *et al.* 2018; Sobral *et al.* 2018a), but only a few studies have investigated their clustering properties. The earliest work on LAE clustering was done by Ouchi *et al.* (2003), which observed 87 LAEs at $z = 4.86$ in the 0.15 deg² Subaru Deep Field and reported the first angular correlation functions (ACFs) and clustering lengths for LAEs ($r_0 = 3.5 \pm 0.3 h^{-1}$ Mpc). Subsequent narrowband surveys have allowed for measurements of LAE clustering properties at $z \sim 2$ –7 (e.g. Shimasaku *et al.* 2004; Kovač *et al.* 2007; Ouchi *et al.* 2010; Zheng *et al.* 2016; Hao *et al.* 2018; Ouchi *et al.* 2018). Guaita *et al.* (2010) presented the first $z \sim 2.1$ measurement using a sample of 250 LAEs in the 0.36 deg² ECDF-S field and found $r_0 = 4.8 \pm 0.9$ Mpc h^{-1} and a typical halo mass of $\sim 3.2 \times 10^{11} M_\odot$. The recent $z \sim 2$ measurements of Kusakabe *et al.* (2018) presented the latest constraints using a sample of ~ 1250 LAEs in four separate fields for a total survey area of 1 deg² and found $r_0 = 2.38_{-0.39}^{+0.34}$ Mpc h^{-1} and a typical halo mass of $4 \times 10^{10} M_\odot$. Both surveys cover the same redshift, but report significantly different results. In respect to Guaita *et al.* (2010), the survey area of Kusakabe *et al.* (2018) is about three times larger and two times fainter in line flux. Given that the Ly α luminosity functions are steep (faint-end slopes of $\alpha \sim -1.8$; e.g. Konno *et al.* 2016; Sobral *et al.* 2018a), deeper samples will be dominated by faint LAEs, which highlights the importance of investigating the clustering properties in terms of their Ly α luminosities.

Although the past two decades have produced a handful of LAE clustering measurements, only a few have focused on how the measured clustering properties are correlated to the physical properties of LAEs (e.g. line luminosity, stellar mass, star formation rates). Ouchi *et al.* (2003) showed that the brightest ($z \sim 4.8$) LAEs are more clustered than the faint LAEs. Bielby *et al.* (2016) presented clustering measurements of $z \sim 3.1$ LAEs in bins of R -band limiting magnitude (corresponding to rest-frame 1500Å UV

continuum luminosity, M_{UV}) and found $r_0 \sim 3 h^{-1}$ Mpc at $R_{\text{lim}} = 27.5$ mag and $r_0 \sim 4.5 h^{-1}$ Mpc at $R_{\text{lim}} = 26$ mag, such that LAEs with high UV continuum luminosities (a tracer of star formation) are more strongly clustered. Recently, Kusakabe *et al.* (2018) found a weak correlation between Ly α line luminosity limit and clustering length/halo mass. This suggests that the Ly α and UV continuum luminosity, which to first order trace star formation activity, correlate with halo mass. However, sample selection effects and varying survey depths prohibit detailed analysis of the evolution and origin of such trends.

Besides Ly α studies, narrowband surveys focused on H α (Sobral *et al.* 2010; Cochrane *et al.* 2017, 2018), [O III] (Khostovan *et al.* 2018), and [O II] (Khostovan *et al.* 2018) emission-line-selected galaxies have also found strong trends between the physical properties of star-forming/active galaxies and their host halo properties. These surveys reveal strong correlations between halo mass and line luminosity (proxy for star formation rate) up to $z \sim 2.23$ (H α), $z \sim 3.3$ ([O III]), and $z \sim 4.7$ ([O II]). Taking into account the evolution in typical line luminosity, such studies found that these trends are redshift independent suggesting that the host halo and residing galaxy co-evolve in unison over cosmic time (Cochrane *et al.* 2017; Khostovan *et al.* 2018).

Although much work has been done on quantifying the clustering/halo properties and its relation to the physical properties of star-forming galaxies, not much focus has been applied to such analysis with Ly α -selected samples, which allow for studying whether these trends hold up to the epoch of reionization. In this paper, we use the Slicing COSMOS 4K (SC4K) survey and archival NB497 imaging to investigate the clustering properties of LAEs in 15 discrete redshift slices in the range $z \sim 2.5$ –5.8 with a total of ~ 5000 LAEs within the 2 deg² COSMOS and 1.3 deg² SA22 fields, corresponding to comoving volumes between 1 and 6×10^6 Mpc³ and total comoving volume of $\sim 6 \times 10^7$ Mpc³ for the full survey.

The paper is organized as follows: in Section 2, we describe the sample of LAEs. In Section 3, we present how we generate our random samples, the methodology in measuring the ACFs, the clustering length, and halo masses, as well as corrections for cosmic variance and a discussion on contamination. In Section 4, we show our methodology in measuring the UV continuum luminosity and slope, as well as the UV star formation rates. In Section 5, we present our main results with discussion regarding its interpretation. We present our main conclusions and final remarks in Section 6.

Throughout this paper we assume a Lambda cold dark matter cosmology with $H_0 = 70$ km s^{−1} Mpc^{−1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$. Our reported values for halo masses and r_0 are in units of $M_\odot h^{-1}$ and Mpc h^{-1} , respectively, unless otherwise specified. Our measurements of star formation rates assume a Salpeter IMF.

2 LY α SAMPLE

2.1 Slicing COSMOS 4K at $2.5 < z < 6$

Our sample is drawn from the publicly available Slicing COSMOS in 4K survey (SC4K; Sobral *et al.* 2018a; Paulino-Afonso *et al.* 2018), which contains 3908 LAEs. The survey uses Subaru imaging from 12 intermediate bands (Capak *et al.* 2007; Taniguchi *et al.* 2007, 2015) in the ~ 2 deg² COSMOS field (Capak *et al.* 2007; Scoville *et al.* 2007) that are reanalysed following the procedures outlined in Sobral *et al.* (2018a). The SC4K survey also includes imaging using four narrowband filters: NB392 ($z = 2.2$; Matthee

Table 1. List of the filters corresponding to rest-frame 1500 Å for each intermediate and narrowband sample. The central wavelength and widths are the rest-frame parameters of the corresponding filters. All photometry used in measuring β_{UV} are shown in the β_{UV} filters column with the number of filters used shown as N_{filters} . Note that β_{UV} for the combined intermediate band samples is measured as listed in each individual filter. The sample sizes of the combined samples include a 30 per cent completeness limit cut on each individual intermediate band sample as measured by Sobral et al. (2018a).

Sample	z	N_g	Filter	Eff. wave. (Å)	FWHM (Å)	β_{UV} filters	N_{filters}
IA427	2.51 ± 0.08	748	V_{540}	1538.97	282.55	$BVgri$, IA464–IA827, NB711, NB816	20
IA464	2.81 ± 0.09	313	V_{540}	1414.50	260.02	$VrizY$, IA484–IA827, NB711, NB816	17
IA484	2.99 ± 0.09	713	r'_{645}	1619.17	295.38	$VrizYJ$, IA505–IA827, NB711, NB816	17
IA505	3.17 ± 0.09	484	r'_{645}	1551.84	283.10	$VrizYJ$, IA527–IA827, NB711, NB816	16
IA527	3.33 ± 0.10	642	r'_{645}	1487.06	271.28	$VrizYJ$, IA574–IA827, NB711, NB816	15
IA574	3.74 ± 0.11	98	i'_{790}	1672.97	297.20	$rizYJ$, IA624–IA827, NB711, NB816	13
IA624	4.13 ± 0.12	143	i'_{790}	1538.93	273.38	$izYJ$, IA679–IA827, NB711, NB816	11
IA679	4.58 ± 0.14	80	z_{915}	1637.86	251.20	$izYJH$, IA709–IA827, NB711, NB816	11
IA709	4.82 ± 0.13	63	z_{915}	1568.56	240.58	$izYJH$, IA738–IA827, NB816	9
IA738	5.06 ± 0.13	79	z_{915}	1506.92	231.12	$izYJH$, IA767–IA827, NB816	8
IA767	5.33 ± 0.15	33	z_{915}	1449.94	222.38	$zYJH$, IA827, NB816	6
IA827	5.78 ± 0.14	36	Y_{1029}	1511.68	151.97	$zYJHK$	5
NB497	3.10 ± 0.02	1198	r'_{645}	1576.82	287.65	$grizJ$	5
NB711	4.86 ± 0.03	78	z_{915}	1564.14	239.90	$izYJH$, IA738–IA827	9
NB816	5.71 ± 0.04	192	Y_{1029}	1532.06	154.02	$zYJH$	4
IA427–IA484	$2.75^{+0.33}_{-0.33}$	1577	–	–	–	–	–
IA505–IA527	$3.25^{+0.18}_{-0.17}$	1074	–	–	–	–	–
IA574–IA624	$3.94^{+0.32}_{-0.31}$	185	–	–	–	–	–
IA679–IA738	$4.82^{+0.37}_{-0.38}$	192	–	–	–	–	–
IA767–IA827	$5.56^{+0.37}_{-0.38}$	53	–	–	–	–	–

et al. 2017; Sobral et al. 2017), NB501 ($z = 3.1$; Matthee et al. 2017), NB711 ($z = 4.8$; Sobral et al. 2018a), and NB816 ($z = 5.7$; Santos et al. 2016). The NB392 and NB501 observations were conducted using the Wide Field Camera on the Isaac Newton Telescope, while all other narrowband and intermediate band observations are from archival Subaru imaging. We restrict our analysis to the samples with $z \gtrsim 2.5$ and also those samples for which the image-to-image variation is negligible and spatially contiguous¹. This includes all 12 intermediate bands and the NB711 and NB816 narrowband samples.

We refer the reader to Sobral et al. (2018a) for details regarding the sample selection. In brief, initial emission-line galaxy candidates were selected by applying a rest-frame equivalent width cut of 25 and 50 Å for narrowbands and intermediate bands, respectively, along with a nebular excess significance cut of $\Sigma > 3$. A combination of spectroscopic redshifts, photometric redshifts, and colour–colour diagnostics were used to select potential LAEs. These candidates were then visually checked to remove any contaminants arising from artefacts not removed in the catalogue generation (e.g. diffraction patterns, edge effects resulting in poor S/N) and sources that have their narrow or intermediate band photometry boosted by the presence of a bright halo from a nearby star in the image. In total, a final sample of 3702 LAEs spanning between $z \sim 2.5$ and 6 were selected based on the 12 intermediate bands and two narrowband

¹Although SC4K has a total of four narrowband filters, we restrict ourselves to the NB711 and NB816 samples. This is because the NB392 and NB501 images are not spatially contiguous (i.e. chip gaps), which requires special care when generating the random/mock samples (see Section 3.1). The other issue with the NB392 and NB501 images is the inhomogeneous depths. Image-to-image depth variation has to also be carefully taken into account to properly generate the random samples. Because of these limitations, we decided to exclude them from the final sample of LAEs used in this study.

filters used in this study. In total, 102 of the 3702 LAEs have spectroscopic confirmation.

Table 1 highlights the redshifts and sample sizes of all the LAE redshift slices. To take advantage of larger sample sizes, especially at the high- z end (e.g. IA767 and IA827), we combine the intermediate band samples to form five larger samples in redshift bins as described at the bottom of Table 1. The choice of combinations was based on maximizing the sample size, with the different completeness limits per individual sample taken into account. This also includes minimizing the redshift widths of the final combined sample so as to remove possible cosmic evolutionary effects when using the samples in our clustering measurements.

2.2 SA22 NB497 at $Z = 3.12$

In addition to the SC4K sample, we also use a sample of 1198 $z = 3.1$ LAEs observed in the 1.38 deg² SA22 field using archival NB497 imaging (Matsuda et al. 2004; Yamada et al. 2012). The observations were done using Suprime-Cam on the Subaru 8.2 m telescope and consisted of seven contiguous, homogeneous pointings, which also covered the large SSA22 protocluster identified by Steidel et al. (1998). The data and source selection were done independently of Yamada et al. (2012) and followed the methodology of Matthee et al. (2017). CFHT MegaCam ugr photometry from the CFHTLS survey was used in the source selection as opposed to the original Subaru SuprimeCam BV photometry used in Yamada et al. (2012). All LAE candidates were selected with a rest-frame equivalent width cut of 25 Å and with a similar colour–colour selection criteria used for $z \sim 3$ LBG candidates (Hildebrandt et al. 2009). Details regarding the source selection are presented in an upcoming paper (Matthee et al., in preparation). Of the 1198 $z = 3.1$ LAEs, 54 of them are spectroscopically confirmed from observations by Yamada et al. (2012) and Saez et al. (2015).

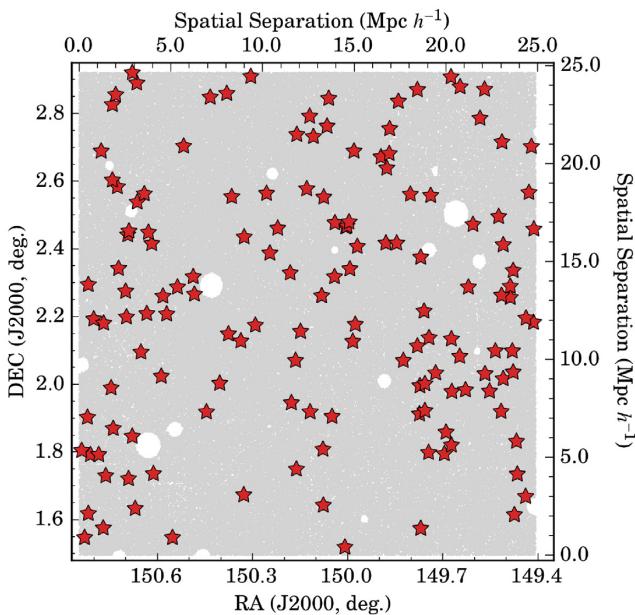


Figure 1. An example of the on-sky distribution using the $z = 4.13$ IA624 sample of LAEs shown as red stars. Specific regions are masked throughout the field (shown as empty regions within the grey shaded areas) to account for bright stars and various artefacts. These masked regions are taken into account when generating the random catalogues. The flux depths throughout the images are homogeneous which decreases the effects of image-to-image variations in the clustering measurements.

3 CLUSTERING MEASUREMENTS

3.1 Random sample

We adopt the approach used in Khostovan et al. (2018), which followed Sobral et al. (2010) to generate random samples. For each narrowband and intermediate band sample, we use the corresponding masked regions maps (see Sobral et al. 2018a) to remove parts of the survey where the imaging is poor in quality or affected by diffraction patterns around bright stars. Fig. 1 shows an example of the masked regions overlaid with our $z = 4.13$ IA624 LAE sample. The bluer intermediate bands (IA427–IA574) also include an extra $\sim 0.02 \text{ deg}^2$ masking due to the smaller u -band imaging area, which is necessary to distinguish LAEs from other line emitters. Each sample is homogeneous in depth (see Sobral et al. 2018a for details) which allows us to exclude the effects of variable depth in generating the mock random samples. In total, our random samples consist of $\sim 10^6$ mock sources per corresponding intermediate/narrowband sample, for which a subset is selected randomly when making measurements of the ACFs (see Section 3.2).

3.2 Angular correlation function

The two-point correlation function is a statistical tool that traces the clustering properties of a given sample by comparing the angular (or spatial) distribution to a random distribution (Peebles 1980). The angular two-point correlation function is typically defined as:

$$dP_{12} = \mathcal{N}(1 + w(\theta_{12})) d\Omega_1 d\Omega_2, \quad (1)$$

where dP_{12} is the probability of finding two galaxies at positions Ω_1 and Ω_2 with an angular separation of θ_{12} for a complete sample with number density, \mathcal{N} . In the case that no angular/spatial correlation

exists, then $w(\theta) = 0$ and the probability reduces into a uniform distribution. Therefore, a non-zero correlation function indicates that galaxies are clustered on the sky.

To measure the ACF, we use the methodology proposed by Sobral et al. (2010) and modified by Khostovan et al. (2018). This methodology uses the Landy & Szalay (1993, LS) estimator to measure the observed correlation function described as:

$$w(\theta) = 1 + \left(\frac{N_R}{N_D} \right)^2 \frac{DD(\theta)}{RR(\theta)} - 2 \left(\frac{N_R}{N_D} \right) \frac{DR(\theta)}{RR(\theta)} \quad (2)$$

where DD , DR , and RR are the number of data–data, data–random, and random–random galaxy pairs, respectively, and N_R and N_D are the number of random and data (real) galaxies, respectively. The errors are described as:

$$\Delta w(\theta) = \frac{1 + w(\theta)}{\sqrt{DD(\theta)}} \quad (3)$$

and are assumed to be Poisson errors. Norberg et al. (2009) showed that Poisson statistics underestimates the ‘true’ errors and suggests using bootstrapping for sample sizes of the order of 10^7 sources. Khostovan et al. (2018) find no significant difference between the two error assessments for sample sizes similar to the one used in this study (see Appendix C of Khostovan et al. 2018).

We use the following fitting function:

$$w(\theta) = A_w \left(\theta^\beta - \frac{\sum RR\theta^\beta}{\sum RR} \right) \quad (4)$$

where the second term is the integral constraint (IC), which takes into account the underestimation of $w(\theta)$ at large angular separations and A_w and β are the clustering amplitude and slope, respectively. Since our sample sizes are not large enough to constrain the slope, we assume $\beta = -0.8$ (fiducial; see Peebles 1980) for all our measurements.

As described in Khostovan et al. (2018), we measure the correlation function by randomly selecting a starting bin centre between 1 and 5 arcsec (e.g. 8–40 kpc at $z \sim 2.5$, 6–30 kpc at $z \sim 5.7$), constant bin sizes of $\log_{10} \Delta\theta \sim 0.1$ –0.3 arcsec, and a maximum angular separation of 7200 arcsec (e.g. 58 Mpc at $z \sim 2.5$, 43 Mpc at $z \sim 5.7$). The random sample is drawn from our large mock catalogue mentioned in Section 3.1. We randomly select the number of mock galaxies to be 10–500 times the number of observed LAEs. With both the real and random samples, we measure $w(\theta)$ via the LS estimator as described in equation (2). We then fit our power-law model as described in equation (4) to measure the clustering amplitude.

This whole process is iterated 2000 times while varying the bin sizes, centres, and random sample sizes per redshift slice. The reported clustering amplitude measurements and their associated errors are based on the median of the A_w distribution drawn from all realizations with the errors being a combination of the scatter (random) in the distribution and also the median (systematic) error of all 2000 realizations added in quadrature. The benefit of this approach (see also Sobral et al. 2010), as highlighted by Khostovan et al. (2018), is that it takes into account the systematic effects due to bin selection (e.g. centres, widths), which are especially important for small sample sizes (< 100 sources).

3.3 Real space correlation function

The real space correlation function, $\xi(r)$, measures the spatial clustering of galaxies and is typically described as a power law of the form $\xi(r) = (r/r_0)^\gamma$, where r_0 is the clustering length and γ is

the slope of the correlation function. The ACF (see Section 3.2) is a projection of the spatial correlation function and typical clustering studies relate the two using the Limber approximation (Limber 1953). Although this works for typical redshift surveys, Simon (2007) showed quantitatively that the approximation fails, especially at large angular separations and when the width of the redshift distributions become similar to a delta function². As a consequence, at large angular separations the slope of the ACF changes from $\gamma + 1$ to γ , such that $w(\theta)$ is a rescaled version of $\xi(r)$. Various narrowband studies have observed this rescaling at a wide range of redshifts (e.g. Gawiser et al. 2007; Guaita et al. 2010; Sobral et al. 2010; Cochrane et al. 2017; Khstovyan et al. 2018; Kusakabe et al. 2018; Ouchi et al. 2018).

To properly measure the spatial correlation function for our samples requires that we use the exact form of the Limber equation. We follow the methodology of Khstovyan et al. (2018) which uses the exact Limber equation as defined by Simon (2007):

$$w(\theta) = \frac{r_0^{-\gamma}}{1 + \cos \theta} \int_0^\infty \int_{\bar{r}\sqrt{2(1-\cos\theta)}}^{2\bar{r}} \frac{2p(\bar{r} - \Delta)p(\bar{r} + \Delta)}{R^{-\gamma-1}\Delta} dR d\bar{r}$$

$$\Delta = \sqrt{\frac{R^2 - 2\bar{r}^2(1 - \cos\theta)}{2(1 + \cos\theta)}} \quad (5)$$

where p describes the sample redshift distribution (essentially the filter profile) in units of comoving distance, R is the distance between two sources, and \bar{r} is the mean spatial position of two sources. As discussed in Section 3.2, our samples are not large enough to constrain the slopes of the correlation functions. Therefore, we fix the slopes such that $\beta = -0.8$ and $\gamma = -1.8$ (fiducial³; $\beta = \gamma + 1$). We use the exact filter profiles associated with each narrow/intermediate band sample and use equation (5) along with our observed measurements of $w(\theta)$ to measure the clustering length, r_0 , per each iteration as described in Section 3.2.

The redshift distributions of the combined intermediate samples are the combination of all the filter profiles associated with each respective intermediate band sample. We also weight the redshift distributions by the number of LAEs specific to an intermediate band sample. For example, the $z \sim 2.8$ sample consists of 1577 LAEs for which 634, 286, and 657 emitters are from the IA427, IA464, and IA484 samples, respectively, with the 30 per cent completeness limits included per sample as reported by Sobral et al. (2018a). Since the number of LAEs is then not homogeneous per intermediate band sample, the final redshift distribution for the combined sample is weighted by the number of emitters in each intermediate sample.

When we measure the clustering properties in bins of galaxy properties (e.g. line luminosity, UV continuum), we are essentially selecting a subsample from the full redshift distribution. Therefore, to properly measure the spatial correlation function, we must augment the weighting of the redshift distributions which is done using the same approach as described above to reflect the relative contribution of each individual intermediate band sample to the

²Simon (2007) showed that using the Limber approximation, the clustering amplitude is related to the spatial clustering length as $A_w \propto r_0^{1-\gamma}/(2\Delta r)$, with Δr being the comoving filter width. As redshift distributions become narrower ($\Delta r \rightarrow 0$), the approximation breaks down and produces unphysical measurements for A_w .

³The fiducial slopes were determined by Totsuji & Kihara (1969) and have remained the same for the past 50 yr as it is found to best represent the angular and spatial correlation functions for a wide range of galaxy samples

combined sample for a given specific bin of galaxy properties. For example, the full $z \sim 2.8$ sample has a relative contribution of 40, 18, and 42 per cent from the individual IA427, IA464, and IA484-selected LAEs, respectively. If we look at a specific bin in Ly α luminosity that consists of a total of 383 LAEs, the relative contribution changes to 28, 19, and 53 per cent for each respective sample. Therefore, we augment the weighted redshift distributions to reflect the changing contributions of each individual intermediate band sample when making our measurements. We follow this approach for all measurements of the spatial correlation function.

3.4 Dark matter halo model

The spatial clustering of galaxies is related to the overall dark matter distribution as:

$$b_{\text{eff}}^2 = \frac{\xi_{\text{gg}}(r)}{\xi_{\text{mm}}(r)}, \quad (6)$$

where b_{eff} is the effective galaxy bias, ξ_{gg} and ξ_{mm} are the galaxy–galaxy and matter–matter spatial correlation functions, respectively. The effective galaxy bias is related to the halo occupation distribution by:

$$b_{\text{eff}}(z) = \frac{\int_{M_{\min}}^\infty b_h(M, z)n_h(M, z)\langle N_g(M, z) \rangle dM}{\int_{M_{\min}}^\infty n_h(M, z)\langle N_g(M, z) \rangle dM}, \quad (7)$$

where b_h and n_h are the halo bias and mass functions for a given halo mass, M , respectively, $N_g(M, z)$ is the galaxy-halo occupation function, and M_{\min} is the minimum dark matter halo mass. The effective halo mass can then be calculated as:

$$M_{\text{eff}} = \frac{\int_{M_{\min}}^\infty M n_h(M, z)\langle N_g(M, z) \rangle dM}{\int_{M_{\min}}^\infty n_h(M, z)\langle N_g(M, z) \rangle dM} \quad (8)$$

for a given sample at a specific redshift.

There are numerous prescriptions for the galaxy-halo occupation that ranges from simple one-to-one occupation to three parameter models (e.g. Kravtsov et al. 2004), to five parameter models (e.g. Zheng et al. 2005), and can be as complex as 12 parameter models (e.g. Geach et al. 2012; although such models apply assumptions to reduce the number of fitted parameters). The one-to-one occupation model assumes a single galaxy resides in each halo, while more complex models take into account multiple galaxies occupying a single halo (central + satellite galaxies).

Since many of our samples are not large enough to properly constrain multiparameter halo occupation distribution models, we assume a simple one-to-one occupation model ($\langle N_g(M, z) \rangle = 1$) where each LAE is a central galaxy hosted by a dark matter halo by a minimum halo mass, M_{\min} . This enables us to adopt a consistent approach throughout this work, even where sample sizes are small. We use the COLOSSUS package (Diemer 2018) in order to measure ξ_{mm} at the redshifts corresponding to our samples. The effective bias is measured at $r = 8 h^{-1}$ Mpc (comoving) which corresponds to the regime for which the linear matter power spectrum dominates. The Tinker et al. (2010) halo bias and Tinker et al. (2008) halo mass functions are used for b_h and n_h in equation (7), respectively.

Throughout this paper, we refer to the effective halo mass as ‘halo mass’ unless otherwise stated.

3.5 Cosmic (sample) variance

One of the major systematic uncertainties that we take into account in our measurements is the effect of cosmic or sample variance

Table 2. The clustering properties for the full population of LAEs per sample. Shown are the redshifts, filter names, number of LAEs per sample, the corresponding survey area and comoving volume in deg² and Mpc³, respectively, the characteristic line luminosity ($L^*(z)$), the median Ly α luminosity, the clustering amplitude measured from the observed ACFs, the exact clustering length, r_0 , measured using equation (5), and the effective halo mass measured using our model described in Section 3.4. Each sample presented is within the ~ 2 deg² COSMOS field, except for the 1.38 deg² SA22 NB497 sample. All $L^*(z)$ measurements are from Sobral et al. (2018a) except for the narrowband samples. The NB816 L^* is measured by Santos et al. (2016). We use the redshift evolution of $L^*(z)$ from the SC4K samples measured in Sobral et al. (2018a) to measure $L^*(z)$ for the NB497 and NB711 samples. Our NB711 $L^*(z)$ measurement is consistent with Shioya et al. (2009) which measured $\log_{10} L^* = 42.9^{+0.5}_{-0.3}$ erg s⁻¹.

z	Filter	N_{gal}	Area (deg ²)	Volume (10 ⁶ Mpc ³)	$\log_{10} L^*(z)$ (erg s ⁻¹)	Med. $\log_{10} L$ (erg s ⁻¹)	A_w (arcsec ^{-0.8})	r_0^{exact} (Mpc h ⁻¹)	\log_{10} Halo mass (M _⊙ h ⁻¹)
2.51	IA427	748	1.94	4.0	$42.76^{+0.07}_{-0.07}$	42.70	$6.47^{+1.25}_{-1.20}$	$4.13^{+0.42}_{-0.42}$	$11.59^{+0.16}_{-0.16}$
2.81	IA464	313	1.94	4.2	$42.83^{+0.36}_{-0.19}$	43.06	$15.51^{+3.10}_{-2.93}$	$7.24^{+0.76}_{-0.74}$	$12.28^{+0.13}_{-0.13}$
2.99	IA484	713	1.94	4.3	$42.64^{+0.06}_{-0.05}$	42.86	$4.86^{+1.13}_{-1.08}$	$3.85^{+0.46}_{-0.45}$	$11.30^{+0.20}_{-0.19}$
3.17	IA505	484	1.94	4.3	$42.80^{+0.09}_{-0.07}$	42.92	$4.46^{+1.42}_{-1.09}$	$3.62^{+0.54}_{-0.50}$	$11.14^{+0.25}_{-0.23}$
3.33	IA527	642	1.94	4.5	$42.68^{+0.07}_{-0.06}$	42.86	$5.94^{+1.13}_{-1.13}$	$3.89^{+0.43}_{-0.42}$	$11.20^{+0.18}_{-0.18}$
3.74	IA574	98	1.96	4.9	$43.03^{+0.18}_{-0.15}$	43.12	$23.39^{+7.97}_{-6.13}$	$9.56^{+1.50}_{-1.26}$	$12.31^{+0.19}_{-0.16}$
4.13	IA624	143	1.96	5.2	$42.83^{+0.17}_{-0.15}$	43.02	$15.59^{+5.09}_{-4.68}$	$7.49^{+1.17}_{-1.22}$	$11.89^{+0.20}_{-0.21}$
4.58	IA679	80	1.96	5.5	$43.15^{+0.16}_{-0.15}$	43.26	$37.35^{+10.52}_{-10.51}$	$10.81^{+1.79}_{-1.66}$	$12.21^{+0.20}_{-0.18}$
4.82	IA709	63	1.96	5.1	$42.98^{+0.17}_{-0.14}$	43.22	$24.38^{+11.39}_{-8.69}$	$8.26^{+1.79}_{-1.87}$	$11.81^{+0.28}_{-0.29}$
5.06	IA738	79	1.96	5.1	$43.30^{+0.23}_{-0.19}$	43.42	$19.68^{+8.13}_{-5.54}$	$7.79^{+1.61}_{-1.10}$	$11.67^{+0.27}_{-0.19}$
5.33	IA767	33	1.96	5.5	$43.30^{+0.28}_{-0.20}$	43.55	$39.53^{+19.24}_{-18.98}$	$12.74^{+2.50}_{-2.62}$	$12.21^{+0.23}_{-0.24}$
5.79	IA827	36	1.96	4.9	$43.35^{+0.24}_{-0.19}$	43.60	$76.99^{+25.06}_{-24.01}$	$15.56^{+2.51}_{-2.71}$	$12.34^{+0.18}_{-0.20}$
3.10	NB497	1198	1.38	1.0	42.77	42.22	$8.95^{+1.54}_{-1.55}$	$3.11^{+0.30}_{-0.29}$	$10.89^{+0.18}_{-0.17}$
4.86	NB711	78	1.96	1.2	43.15	43.05	$17.85^{+10.81}_{-7.43}$	$4.57^{+1.24}_{-1.33}$	$10.97^{+0.42}_{-0.45}$
5.71	NB816	172	1.96	1.8	$43.25^{+0.09}_{-0.06}$	42.82	$19.18^{+3.83}_{-3.83}$	$5.04^{+0.55}_{-0.56}$	$10.87^{+0.17}_{-0.17}$
2.75	–	1577	1.94	12.5	–	–	$2.89^{+0.63}_{-0.59}$	$4.50^{+0.54}_{-0.51}$	$11.63^{+0.18}_{-0.17}$
3.25	–	1074	1.94	8.8	–	–	$3.17^{+0.85}_{-0.74}$	$3.75^{+0.52}_{-0.50}$	$11.17^{+0.23}_{-0.22}$
3.94	–	185	1.96	10.1	–	–	$10.41^{+4.26}_{-3.52}$	$7.62^{+1.56}_{-1.48}$	$11.97^{+0.26}_{-0.25}$
4.82	–	192	1.96	15.7	–	–	$12.71^{+4.36}_{-3.98}$	$9.24^{+1.94}_{-1.93}$	$11.96^{+0.26}_{-0.26}$
5.56	–	53	1.96	10.4	–	–	$44.55^{+19.63}_{-15.12}$	$16.16^{+3.80}_{-3.52}$	$12.43^{+0.26}_{-0.24}$

which arises from the limited survey area. Sobral et al. (2010) measured the effects of cosmic variance on the clustering amplitude using their H α sample at $z = 0.84$ by randomly sampling areas between 0.05 and 0.5 deg² in their 1.3 deg² coverage of the COSMOS field. They find that the uncertainty in the clustering amplitude as a function of area scales as $20 \times \Omega^{-0.35}$, with Ω being the survey size in deg². For the case of our ~ 2 deg² survey (see Table 2 for the survey size per redshift slice), the uncertainty in the clustering amplitude is ~ 16 per cent of A_w , which corresponds to ~ 11 per cent of the clustering length, r_0 . We incorporate these systematic errors by adding them in quadrature to the measured uncertainties.

3.6 Contamination

Contamination is typically assumed to cause an underestimation of the observed clustering signal where contaminants are randomly distributed in the field. Quantitatively, the clustering signal, A_w , will be underestimated by a factor of $(1 - f)^2$, with f being the contamination fraction. This translates to a factor of $(1 - f)^{2|\gamma|}$ for r_0 (the clustering length).

As discussed in Khostovan et al. (2018), the effect of contamination is not as straightforward in narrowband samples since the contaminants are other emission-line-selected galaxies. In the case of this study, our contaminants will be primarily low- z interlopers, such as [O II], [O III], and H α emitters. These low- z interlopers

exhibit non-random clustering (e.g. Shioya et al. 2008; Sobral et al. 2010; Stroe & Sobral 2015; Cochrane et al. 2017; Khostovan et al. 2018) and therefore can either cause an overestimation or underestimation of the clustering signal. A benefit of SC4K is that H α and [O III] become low- z interlopers starting in IA679 and IA505, respectively, and are small in numbers due to volume such that their effects are negligible.

Sobral et al. (2018a) investigated the contamination fraction for the SC4K sample using the available spectroscopic measurements. Of the 132 sources with spectroscopic redshifts, 112 were confirmed to be LAEs suggesting a contamination fraction of ~ 15 per cent, which is typical of large-area Ly α narrowband surveys. Sobral et al. (2018a) also investigated whether this contamination was dependent on redshift, Ly α luminosity, and rest-frame EW and found that it is constant around 10–20 per cent. Using the simple $(1 - f)^2$ factor, a contamination fraction of 15 per cent would increase A_w by ~ 38 per cent and r_0 by ~ 20 per cent, but with the assumption that these contaminants are randomly distributed, which, as discussed above, should not be the case. A detailed investigation of the non-random clustering of contaminants and their effect on the underlying measurements is needed for clustering studies, in general; although, we note this is beyond the scope of this paper. We instead omit from correcting the clustering measurements for contamination effects, but cite the numbers above as the maximum effect contaminants can have on the clustering signal, assuming randomly distributed contaminants.

4 REST-FRAME UV PROPERTIES

4.1 Determining M_{UV} and β_{UV}

The typical shape of the SED of star-forming galaxies at $1300\text{\AA} < \lambda < 2800\text{\AA}$ can be fit by a power law of the form $f_\lambda \propto \lambda^{\beta_{\text{UV}}}$, where f_λ is the flux density, typically in units of $\text{erg s}^{-1} \text{cm}^{-2} \text{\AA}^{-1}$, and β_{UV} is the UV spectral slope. Since the cross-section of dust grains effectively absorbs UV light, the amount of dust attenuation can be measured using the UV slopes (e.g. Calzetti, Kinney & Storchi-Bergmann 1994; Meurer, Heckman & Calzetti 1999), although it should be noted that redder UV slopes ($\beta_{\text{UV}} > -2$) can also signify galaxies with mature, evolved stellar populations. We expect this degeneracy in interpreting β_{UV} to be negligible as our samples are emission line selected and are then typically dominated by populations of star-forming galaxies. Investigating this degeneracy will be part of a forthcoming work (Santos et al., in preparation).

We measure β_{UV} by fitting the power law described above using the available photometry in the rest-frame range of $1300\text{\AA} < \lambda < 2800\text{\AA}$. We measure the 1500\AA UV continuum absolute magnitudes (M_{UV}) by:

$$M_{\text{UV}} = m_{\text{UV}} - 5 \log_{10} \left(\frac{d_L}{10\text{pc}} \right) + 2.5 \log_{10}(1+z) \quad (9)$$

where m_{UV} is the observed UV magnitude and d_L is the luminosity distance. Table 1 shows the corresponding observer-frame photometric band associated with m_{UV} used to measure M_{UV} , the rest-frame effective wavelength and FWHM of the filter, and the observer-frame filters used in measuring β_{UV} . Although the filters cover 1500\AA within the FWHM, their effective wavelengths are off-centred by a maximum of $\sim 170\text{\AA}$ which causes an offset in our measurements of M_{UV} . We calculate the maximum offsets to be ~ -0.06 , -0.12 , and -0.18 mag for UV spectral slopes of $\beta = -1.5$, -1 , -0.5 and for a 170\AA offset towards redder wavelengths. In principle, the offset can be taken into account by applying the correction, $-2.5(\beta_{\text{UV}} + 2)\log_{10}(\lambda/1500\text{\AA})$, to M_{UV} . Since our LAEs typically have blue spectral slopes, especially for the higher redshift samples, the offsets are negligible. For the case of our sources with redder spectral slopes ($\beta_{\text{UV}} > -2$), we find that the uncertainties in M_{UV} are larger than the offsets.

Since our samples are Ly α selected, we are prone to detect low stellar mass sources for which the stellar continuum is below the survey detection limit compared to continuum-selected surveys. In such cases, we apply a lower limit to M_{UV} using the 3σ detection limit of the photometry. The lack of stellar continuum also means that we are not able to measure β_{UV} for a subset of our sources. There are also sources for which the uncertainties in β_{UV} are quite high due to weak stellar continuum measurements. To take these effects into account, we take all measurements of β_{UV} that have an S/N ($|\beta_{\text{UV}}/\Delta\beta_{\text{UV}}| > 3$) and measure the median. This is then used as the median stacked spectral slope for those that have S/N < 3 or no β_{UV} measurement.

4.2 Star formation rates of LAEs

Typically, narrowband surveys measure star formation rates using the observed or dust-corrected emission line luminosity in conjunction with a star formation rate calibration (e.g. Ly et al. 2011; Sobral et al. 2013, 2014). In the case of LAEs, measuring star formation rates using the Ly α line introduces several caveats. Even though Ly α traces the ionizing radiation of star formation activity, it is easily scattered. This increases the likelihood of being

absorbed by dust and also decreases the surface brightness, such that Ly α photons are spread out over a larger area (see Dijkstra 2017 and Sobral & Matthee 2019 for discussion).

To measure the star formation rates of our LAEs, we instead use the UV continuum luminosities, M_{UV} , as described in Section 4.1, which traces the population of short-lived, massive O , B , and A type stars corresponding to a star formation activity time-scale of ~ 100 Myr. We assume the Kennicutt (1998) SFR(UV) calibration:

$$\text{SFR(UV)} = 1.4 \times 10^{-28} \left(\frac{L_v}{\text{erg s}^{-1} \text{Hz}^{-1}} \right) M_{\odot} \text{yr}^{-1}, \quad (10)$$

where L_v is the UV luminosity per unit frequency. This calibration is valid for the range 1500 – 2800\AA , where L_v is consistently flat (assuming $\beta_{\text{UV}} \sim -2$) and assumes a Salpeter IMF. We assume the Meurer et al. (1999) calibration to correct M_{UV} for dust attenuation:

$$A_{\text{UV}} = 4.43 + 1.99\beta_{\text{UV}}, \quad (11)$$

where A_{UV} is the UV dust extinction and β_{UV} is the UV spectral slope described in Section 4.1.

5 RESULTS AND DISCUSSION

5.1 Clustering properties of LAEs

In this section, we present our results on the clustering properties of our LAE samples. The ACFs are shown for all our narrowband, intermediate band, and combined samples in Figs 2, C1, and C2, respectively. Each bin in angular separation encompasses the median measurement of $w(\theta)$ for all 2000 realizations and the errors includes the scatter in $w(\theta)$ and the median Poisson error described in equation (3). In this respect, we are taking into account the effects of selecting some arbitrary fixed bin size and bin width in making our final ACF measurements. We overlay the fits based on the measured r_0 in equation (5) and shown in Table 2 and find that it is consistent with the median ACFs. Note that, as described above, the measured r_0 is based on the distribution of 2000 r_0 measurements that correspond to each individual ACF.

As described in Section 3.2, we use the exact form of the Limber equation as outlined by Simon (2007) to fit the ACFs and find that it best represents the observed measurements, especially at higher angular separations where the deviation from a simple power law occurs. This is more pronounced in the $z \sim 3.1$ NB497 ACF at angular separations greater than 600arcsec , which corresponds to comoving separations of $\sim 3.4\text{ Mpc }h^{-1}$, as shown in the top panel of Fig. 2. Previous narrowband studies have also observed deviations from the simple power law form at high angular separations (e.g. Sobral et al. 2010; Cochrane et al. 2017; Khstovyan et al. 2018).

Fig. 2 includes the observed ACFs from various clustering studies of LAEs (Ouchi et al. 2003; Gawiser et al. 2007; Murayama et al. 2007; Shioya et al. 2009; Ouchi et al. 2010; Bielby et al. 2016; Ouchi et al. 2018). We find that, in general, our ACFs and fits are in agreement with the literature. Note that the ACFs shown are not corrected for cosmic variance effects, which would severely affect $\sim \text{arcmin}^2$ measurements (e.g. Ouchi et al. 2003; Gawiser et al. 2007; see Table B1 for survey size). This is probably why the $z \sim 4.86$, 543 arcmin^2 Subaru Deep Survey measurements of Ouchi et al. (2003) are systematically above our measurements, but still within 1σ .

5.1.1 Effects of the 1-halo term

Typically, ACFs trace two distinct clustering regimes. The first is the galaxy–galaxy angular correlation within a single dark matter

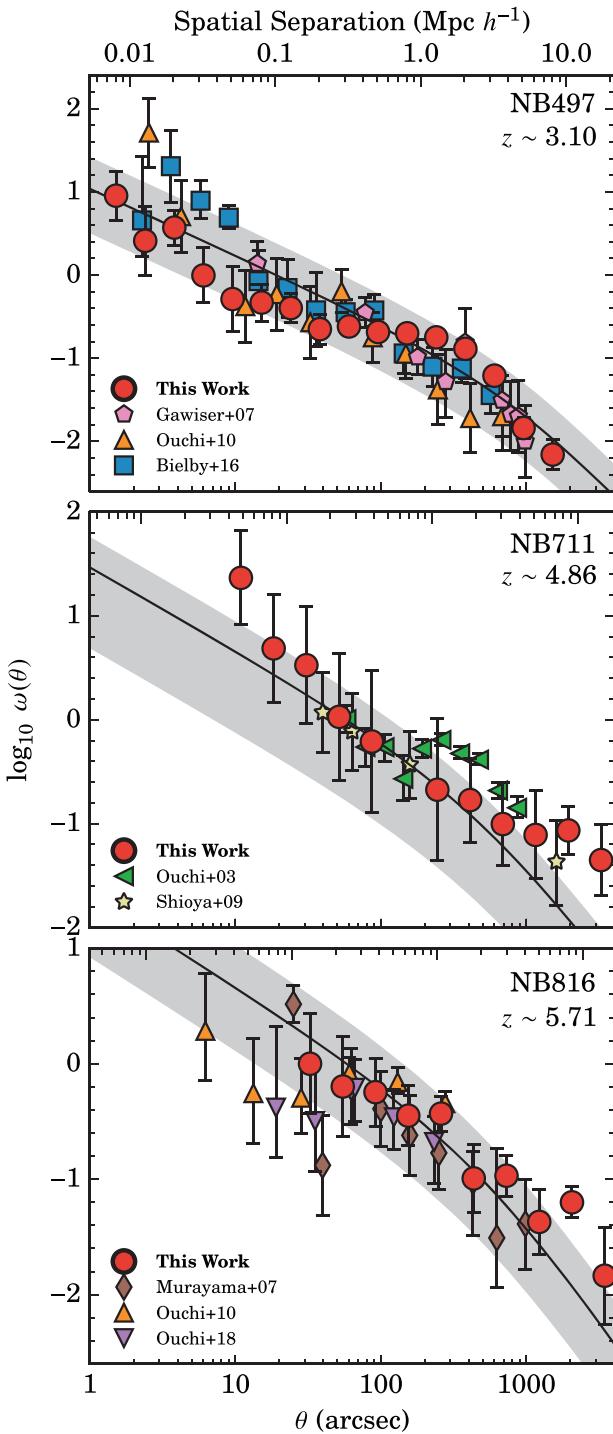


Figure 2. The ACFs for each narrowband sample. The red circles are the median observed measurements for $w(\theta)$ based on all 2000 iterations of measuring $w(\theta)$ with varying bin sizes and centres. The black line shows the best-fitting model as described in equation (5) with the 1σ uncertainty represented as the grey. The spatial axis shown in each panel corresponds to the spatial separation for a given angular separation at the redshift of the samples shown. We only detect the 1-halo term at $\theta < 10$ arcsec ($r < 53$ kpc h^{-1}) for our $z \sim 3.1$ sample, which suggests that our sample does contain satellites. We ruled out the large overdensity in the SA22 field as the source of the 1-halo term (see Appendix A). The $z \sim 4.86$ ACF also shows hints of a 1-halo term although this is based on one bin. Overlaid are the ACFs from various narrowband surveys that are consistent with the redshifts of our samples.

halo, referred to as the 1-halo term. The second is the galaxy–galaxy angular correlation, with galaxies residing in separate dark matter haloes, which is referred to as the 2-halo term. The 1-halo term signal is observed at low angular separations as a deviation from a simple power-law ACF model towards higher $w(\theta)$ (enhanced clustering) and traces the clustering properties of both central and satellite galaxies, while the 2-halo term is observed at larger angular separations.

We find that most of our samples show no significant detection of the 1-halo term, which suggests that the LAEs in our sample are primarily centrals and have low/negligible satellite fractions. This could be due to selection bias as we are selecting LAEs with strong emission lines and are missing the faint, low-mass population that forms the bulk of the satellite population.

We detect a signature of a 1-halo term in the $z \sim 3.1$ NB497 sample at angular (comoving) separations of ~ 10 arcsec (~ 50 kpc h^{-1}), although we note that the observed ACFs are still consistent with the exact Limber equation fits. One possible reason for the detection of the 1-halo term could be the presence of a significant overdense region in the SA22 field (Steidel et al. 1998, 2000; Matsuda et al. 2004; Yamada et al. 2012) which, in principle, would cause elevated correlation function measurements at lower angular separations. We test this idea in Appendix A by masking the overdense region and repeating our measurements. We find no significant difference between the ACFs for the full SA22 field and the case where the overdense region is masked.

Another possibility is that the $z \sim 3.1$ NB497 sample is deep enough in line luminosity to observe satellite LAEs. We test this idea in Appendix A by applying varying line luminosity thresholds and find that the 1-halo term disappears at $L_{Ly\alpha} \gtrsim 0.4 L^*$, such that the satellite fraction is negligible beyond this threshold. Other clustering studies of emission-line galaxies (Cochrane et al. 2017) and LBGs (Harikane et al. 2018) also show that the satellite fractions are typically $\lesssim 5$ per cent, such that they are negligible. Since measurements of the satellite population is not the main focus of this paper, we defer further discussion but assume based on past works and our own observations that such a population has minimal effects on our measurements, particularly for the $L_{Ly\alpha} \gtrsim 0.4 L^*$ population.

5.1.2 Clustering length

With the observed ACFs, we measure the spatial clustering lengths using our approach highlighted in Section 3.3. Fig. 3 shows the redshift evolution of the clustering length, r_0 , for all our LAE samples without any cuts. Although there is a distinct difference between the NB- and IB (combined IB)-selected results, both show an increasing r_0 with increasing redshift and are shown in Table 2. Comparing the NB- and IB-samples yields an apparently different redshift evolution where by $z = 5.7$, the IB samples are three times as clustered compared to the NB samples. A similar result is seen when comparing the NB- and combined IB-samples. The main cause of this difference is due to sample selection and survey parameters as the narrowbands are ~ 0.2 – 0.8 dex deeper in Ly α luminosity than their corresponding intermediate bands, while covering a smaller volume (see Table 2).

The issue of sample selection effects on the clustering results become evident when comparing IB-to-IB samples. For example, the $z = 2.8$ IA464 sample is 0.2 and 0.4 dex shallower in depth in comparison to the $z = 2.5$ IA427 and $z = 3.0$ IA484 sample, respectively, and is found to be more clustered by a factor of two. This suggests that the clustering signal is dependent on

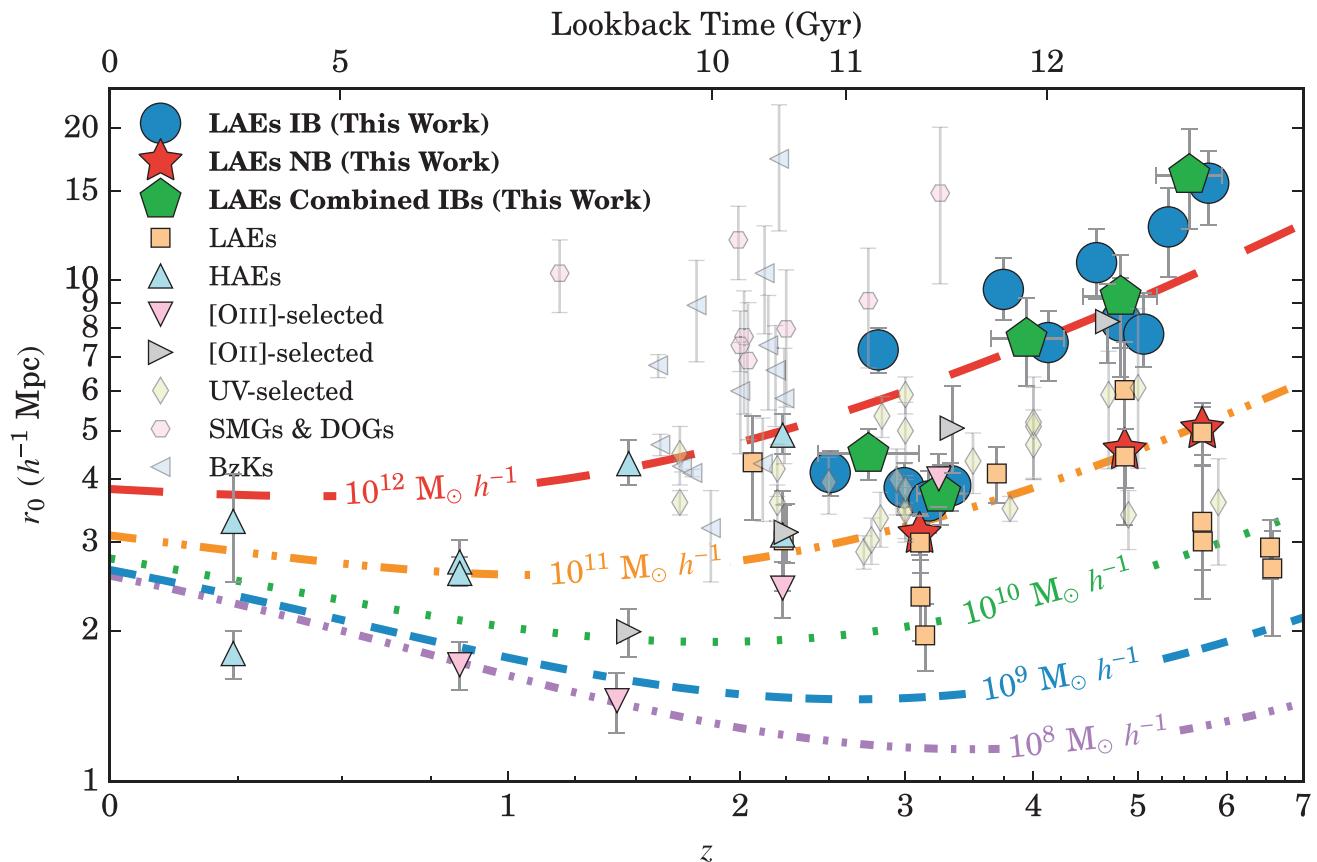


Figure 3. The redshift evolution of the clustering length, r_0 , for our intermediate band, combined, and narrowband-selected LAEs. We find that r_0 increases with increasing redshift up to $z \sim 6$ for all our samples. The systematic offset in r_0 between our intermediate and narrowband-selected samples is attributed to selection effects (e.g. depth of each sample). Comparison with narrowband-selected samples drawn from the literature also show widely varying r_0 (Ouchi et al. 2003; Gawiser et al. 2007; Murayama et al. 2007; Shioya et al. 2009; Guaita et al. 2010; Ouchi et al. 2010; Bielby et al. 2016; Ouchi et al. 2018). The redshift evolutions of host dark matter haloes are also shown for minimum halo masses in the range $10^{8\text{--}12} \text{ M}_\odot \text{ } h^{-1}$ based on our one-to-one halo occupation model. Our intermediate-band-selected LAEs are found to be hosted by $10^{11\text{--}12} \text{ M}_\odot$ haloes, while our narrowband-selected LAEs are hosted by $\sim 10^{11} \text{ M}_\odot$ haloes for all redshifts. We also compare to H α (Geach et al. 2008; Shioya et al. 2008; Sobral et al. 2010; Stroe & Sobral 2015; Cochrane et al. 2017), [O II] (Khstovyan et al. 2018), and [O III] (Khstovyan et al. 2018) narrowband studies and find that, for the overlapping redshift ranges, we are in agreement, suggesting that the various emission-line-selected galaxies reside in dark matter haloes with similar masses. Included are measurements from UV-selected samples (Foucaud et al. 2003; Ouchi et al. 2004; Adelberger et al. 2005; Kashikawa et al. 2006; Savoy et al. 2011; Bielby et al. 2013; Barone-Nugent et al. 2014; Durkalec et al. 2015, 2018), Sub-millimeter Galaxies (SMGs; Blain et al. 2004; Hickox et al. 2012; Wilkinson et al. 2017), Dust-Obscured Galaxies (DOGs; Brodwin et al. 2008; Toba et al. 2017), and star-forming BzK-selected galaxies (BzKs; Kong et al. 2006; Hayashi et al. 2007; Blanc et al. 2008; Hartley et al. 2008; McCracken et al. 2010; Lin et al. 2012; Ishikawa et al. 2015). Each selection type covers a specific subset of star-forming galaxies resulting in wide ranges of clustering lengths, which highlights the need to properly compare samples by accounting for sample biases.

Ly α luminosity and to properly compare clustering properties requires that we take this effect into account. We showed the importance of this effect in an earlier work for other emission-line-selected samples (Khstovyan et al. 2018).

Fig. 3 also includes the clustering lengths associated with minimum halo masses in the range $10^{8\text{--}12} \text{ M}_\odot$ host dark matter haloes as a function of redshift. We find that our IB-selected LAEs typically reside in host haloes with a minimum mass range of $\sim 10^{11\text{--}12} \text{ M}_\odot$ and the NB-selected LAEs show a consistent minimum host halo mass of $\sim 10^{11} \text{ M}_\odot$ for all redshifts observed. There is no redshift evolution observed in host halo mass, which suggests that galaxies observed as LAEs at different redshifts reside in haloes of similar mass. Table 2 highlights the effective host halo masses per sample.

We also include the r_0 measurements of NB-selected LAEs drawn from the literature in Fig. 3 (Ouchi et al. 2003; Murayama et al. 2007; Guaita et al. 2010; Ouchi et al. 2010; Bielby et al. 2016; Ouchi et al. 2018). Differences in measuring clustering lengths and halo masses in comparison to our approach are taken into account and described in Appendix B. Fig. 3 also includes H α (Shioya et al. 2008; Sobral et al. 2010; Stroe & Sobral 2015; Cochrane et al. 2017; Kashino et al. 2017), [O III] (Khstovyan et al. 2018), and [O II] emitters (Takahashi et al. 2007; Khstovyan et al. 2018).

We find an excellent agreement between our measurement of $r_0 = 3.11^{+0.30}_{-0.29} \text{ Mpc } h^{-1}$ for our $z \sim 3.1$ NB497 sample and $r_0 = 2.99 \pm 0.40 \text{ Mpc } h^{-1}$ from the VLT LBG redshift survey of Bielby et al. (2016). Both our work and Bielby et al. (2016) use a similar NB497 filter and are somewhat similar in survey parameters and

selection, although their sample size is smaller (~ 600 LAEs) and they apply a higher rest-frame equivalent width cut (65 Å). We find that the other $z \sim 3.1$ studies report a lower r_0 with a $> 1\sigma$ deviation with the ECDF-S MUSYC imaging survey of Gawiser et al. (2007) measuring an $r_0 = 2.34 \pm 0.43$ Mpc h^{-1} and Ouchi et al. (2010) measuring $r_0 = 1.96 \pm 0.30$ Mpc h^{-1} for LAEs in the SXDS field. The Ouchi et al. (2010) $z \sim 3.1$ sample is somewhat deeper than our NB497 sample with a limiting flux of $\sim 1.2 \times 10^{-17}$ erg s $^{-1}$ cm $^{-2}$. The Gawiser et al. (2007) sample is also somewhat deeper with a limiting flux of $\sim 1.5 \times 10^{-17}$ erg s $^{-1}$ cm $^{-2}$, such that the discrepancy is most likely due to the fainter LAEs being picked up by the two respective studies.

Our $z = 4.86$ NB711 r_0 measurement is found to be in agreement with the Shioya et al. (2009) 1.83 deg 2 COSMOS measurement of $r_0 = 4.44 \pm 0.59$ Mpc h^{-1} , despite the different source extraction and sample selection used by Sobral et al. (2018a). Ouchi et al. (2003) performed an earlier clustering analysis of LAEs in the 543 arcmin 2 Subaru Deep Field using a similar NB711 filter and reported an $r_0 = 6.03 \pm 1.49$ Mpc h^{-1} , which is within 1σ agreement with our results.

We also find an agreement within 1σ between our $z = 5.71$ NB816 r_0 measurement and that of the 1.95 deg 2 COSMOS measurement of Murayama et al. (2007), while the SXDS measurements of Ouchi et al. (2010) and the HSC SILVERRUSH measurements of Ouchi et al. (2018) are lower. The cause of the difference is likely due to survey depth (Ly α luminosity; e.g. the SXDS measurements are close to 1 mag deeper in terms of 5σ narrowband detection limits) and also cosmic variance.

Comparing our measurements to continuum-selected samples from the literature shows that LAEs and LBGs have similar r_0 , as shown in Fig. 3. Measurements from BzK-, SMG-, and DOG-selected samples show clustering lengths higher than our Ly α measurements at all redshifts. This is primarily due to selection effects as samples, such as DOGs, will select more massive, dustier populations in comparison to our LAE samples, which primarily select dust-free, low-mass systems.

5.2 Line luminosity

Motivated by the results of Sobral et al. (2010), Cochrane et al. (2017), and Khostovan et al. (2018), we investigate the trends between Ly α luminosity and host dark matter halo mass. Our measurements reported here are the first in the literature that cover a wide dynamic range in Ly α luminosity and redshift. Throughout the rest of this paper, we will use our combined intermediate band samples, along with our narrowband samples, to increase the sample statistics, while also ensuring that the redshift range per sample is small enough that any redshift evolution within each combined sample is negligible.

We show the host halo mass in terms of Ly α luminosity in Fig. 4. For all redshift samples, we find that host halo mass increases with increasing Ly α luminosity. Including the Ly α luminosity threshold $z = 2.2$ (Kusakabe et al. 2018) and $z = 3.1$ (Gawiser et al. 2007; Ouchi et al. 2010) literature measurements, along with our $z \sim 2$ –3 samples, we find that halo mass increases from $10^{9.7\text{--}12.8}$ M $_{\odot}$ between $10^{41.7\text{--}43.6}$ erg s $^{-1}$ in Ly α luminosity. We find similar results when looking at the higher redshift samples, in conjunction with luminosity threshold measurements from the literature. The main difference between the redshift samples is an offset in Ly α luminosity with the high- z measurements shifted to higher Ly α luminosities. This could reflect the cosmic evolution in the Ly α luminosity functions. If so, this could be taken into account

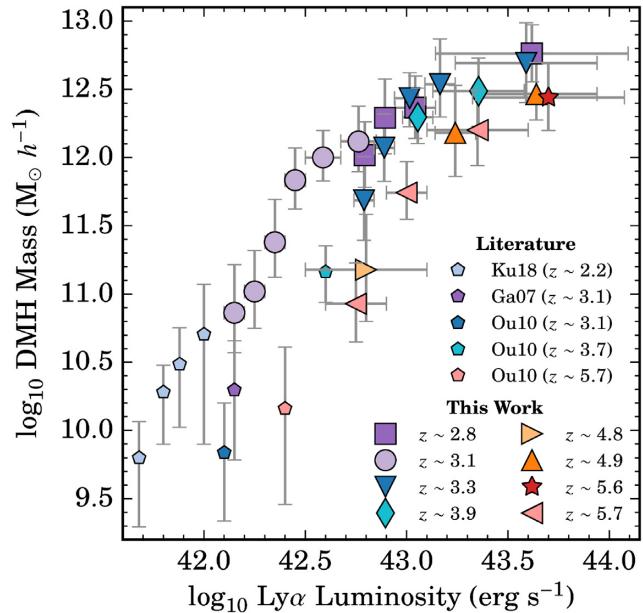


Figure 4. Halo mass in terms of observed Ly α luminosity. For each redshift sample, we see that halo mass increases with increasing line luminosity. Between $z \sim 2$ and 3, our measurements show an increase in halo mass from $10^{9.7}$ to $10^{12.8}$ M $_{\odot}$ for Ly α luminosities in the range $10^{41.7\text{--}43.6}$ erg s $^{-1}$. Similar trends are also seen at $z > 3$, but are shifted to higher line luminosities in comparison to the $z \sim 2$ –3 samples.

in order to investigate the evolution of clustering/halo properties of LAEs.

Fig. 5 shows host halo mass in terms of Ly α luminosity normalized by the characteristic line luminosity, $L^*(z)$. The measurements of $L^*(z)$ used are shown in Table 2 and are taken from Sobral et al. (2018a), which used the same SC4K sample we use in this study. Since we combine our intermediate band samples, we carefully take into account the variation in $L^*(z)$ between each individual intermediate band sample by first applying the corresponding $L^*(z)$ and then binning in terms of $L/L^*(z)$.

We find a strong, increasing trend between host halo mass and $L/L^*(z)$ from $z \sim 2.8$ to 5.7, covering 1.3 Gyr of cosmic history from the end of reionization to the peak of cosmic star formation. Our faintest LAEs ($L \sim 0.25L^*(z)$) are observed to reside in $10^{10.8}$ M $_{\odot}$ haloes and our brightest LAEs ($L \sim 7L^*(z)$) reside in $10^{12.8}$ M $_{\odot}$ haloes. The typical L^* galaxy is observed to be found in $\sim 10^{12}$ M $_{\odot}$ host dark matter haloes. Surprisingly, these are found to be redshift-independent suggesting that LAEs of the same $L^*(z)$ type at different redshifts reside in similar halo masses.

Fig. 5 also includes Ly α luminosity threshold measurements drawn from the literature at various redshifts (Gawiser et al. 2007; Ouchi et al. 2010; Kusakabe et al. 2018). Due to the nature of these measurements, they help to constrain the faint end of Fig. 5 and are primarily single measurements per redshift, except for Kusakabe et al. (2018), which made five measurements (although we only show four as their deepest measurement is poorly constrained). The literature measurements, along with our own observations, show significantly strong, redshift-independent trends between Ly α luminosity and effective halo mass.

To quantify the trends seen in Fig. 5, we fit two different models: a single power law and a piecewise power law with the pivot point

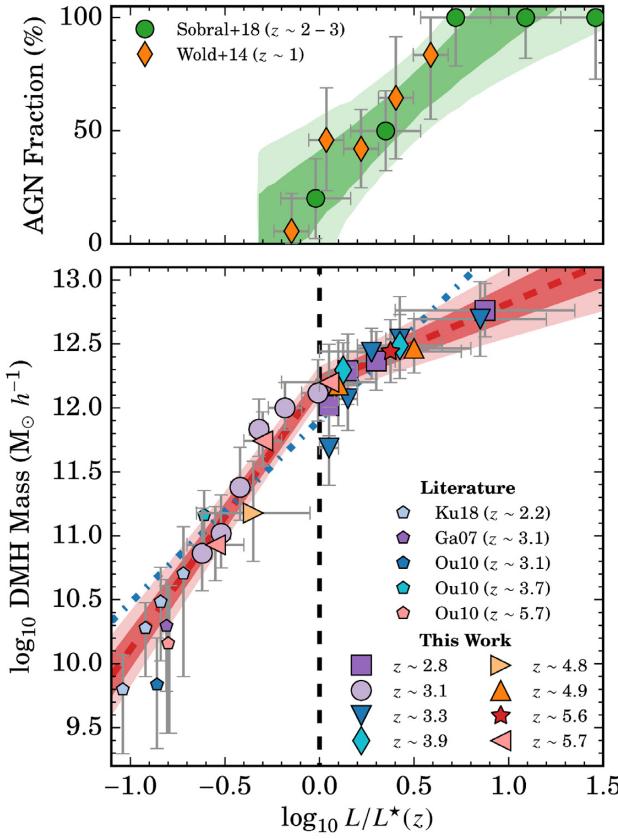


Figure 5. Host halo mass and Ly α luminosity normalized by the characteristic line luminosity, $L^*(z)$. We find a strong, redshift-independent trend between host halo mass and the $L^*(z)$ normalized line luminosity similar to previous narrowband works for H α -, [O II]-, and [O III]-selected emission-line galaxies (Sobral et al. 2010; Cochrane et al. 2017; Khostovan et al. 2018). These are quantified by a single (blue line) and piecewise (red line) power-law model, with the 1 σ and 2 σ regions shown in dark and light red, respectively. We find that the observed trends become shallower at $L > L^*(z)$, which may be a signature of a transitional halo mass where it becomes increasingly improbable that a star-forming galaxy resides in higher host halo masses. The continuous, shallower increase can also be a sign of AGN contribution at the brightest Ly α luminosities. Recent work by Sobral et al. (2018b) find a strong increase in the AGN fractions with $L/L^*(z)$ and that by $L \sim 2L^*(z)$, the AGN fraction is ≈ 50 per cent.

at L^* . The best-fitting single power law is

$$\frac{M_{\text{halo}}}{M_\odot/h} = 10^{11.91^{+0.05}_{-0.05}} \left(\frac{L}{L^*(z)} \right)^{1.44^{+0.14}_{-0.12}} \quad (12)$$

with a slope near unity. Although the single power law seems to represent the observations around $L \sim L^*(z)$, there is a deviation towards lower and higher line luminosities. Based on this deviation, we use a piecewise power law that is separated at $L^*(z)$ with a best fit of:

$$\frac{M_{\text{halo}}}{M_\odot/h} = 10^{12.19^{+0.06}_{-0.06}} \begin{cases} \left(\frac{L}{L^*(z)} \right)^{2.08^{+0.12}_{-0.12}} & L < L^* \\ \left(\frac{L}{L^*(z)} \right)^{0.63^{+0.12}_{-0.12}} & L > L^* \end{cases} \quad (13)$$

where the slopes above and below $L^*(z)$ are quite different. The best fits show a steeply increasing halo mass with line luminosity up to $L^*(z)$ with a slope of 2.08 ± 0.12 followed by a slowly increasing

halo mass at brighter line luminosities with a slope of 0.63 ± 0.12 and a typical halo mass of $10^{12.19 \pm 0.06} M_\odot$ at $L^*(z)$.

5.2.1 What causes the trend change at $L > L^*(z)$?

The slope change that is seen in Fig. 5 could be due to a change in the nature of the population of LAEs (e.g. Ly α emission is no longer driven by star formation but by AGN activity). This would result in the fraction of star-forming galaxies to decrease with increasing luminosity. Above $10^{12} M_\odot$, the star formation efficiency decreases due to accelerated gas accretion caused by the deeper gravitational potentials of higher mass haloes resulting in fewer star-forming galaxies with increasing halo mass (e.g. Dekel & Birnboim 2006; Bower et al. 2017). This idea of a transitional or characteristic halo mass has been observed for H α -, [O III]-, and [O II]-selected emitters between $z \sim 0.4$ and 5 (Sobral et al. 2010; Khostovan et al. 2018) and by studies of star-forming and passive galaxies (e.g. Hartley et al. 2013; Doherty et al. 2014).

To understand whether AGN contribution could be causing a trend change at $L > L^*(z)$, we include the $z \sim 2$ –3 AGN fraction measurements of Sobral et al. (2018b) and $z \sim 1$ measurements of Wold, Barger & Cowie (2014) in the top panel of Fig. 5. About 20 per cent of $z \sim 1$ –3 LAEs are found to be AGNs around L^* and by $2L^*(z)$, half of the population of LAEs are AGNs. Calhau et al., submitted, found a strong correlation between the fraction of X-ray-detected AGNs and Ly α luminosity. Matthee et al. (2017) found that $z \sim 2.3$ LAEs are about 50 per cent X-ray AGNs at $> 10^{44}$ erg s $^{-1}$ (see also Konno et al. 2016). The halo masses measured for our $> 2L^*(z)$ samples are also consistent with previous AGN clustering studies (e.g. halo masses of $\gtrsim 10^{12.5} M_\odot$; Hickox et al. 2009; Koutoulidis et al. 2013; Allevato et al. 2016; Mendez et al. 2016; Hale et al. 2018). Overall, we find that the brightest LAEs in our samples show properties that are consistent with AGN populations at high z .

5.3 Rest-frame UV continuum

In the previous section, we found that the line luminosity of LAEs correlates with the host halo, regardless of redshift, such that the brightest LAEs reside in the most massive haloes. Here we explore how the host halo mass can depend on the rest-frame UV properties, specifically the 1500 Å UV continuum luminosity (M_{UV}) and the UV-measured star formation rate. Our method of measuring both properties is described in Section 4.

Fig. 6 shows how the observed (not corrected for dust) M_{UV} and the host halo mass are correlated. We find a strong trend where the host halo mass increases with increasing UV luminosity. The most UV-bright LAEs ($M_{\text{UV}} < -22$) are found to reside in $10^{13} M_\odot$ haloes and the fainter ones ($M_{\text{UV}} > -20$) are found in $< 10^{11.5} M_\odot$ haloes. We find a redshift-independent trend without the need to normalize by the characteristic UV luminosity, M_{UV}^* . Interestingly, $M_{\text{UV}}^*(z)$ has been observed by previous work to be constant within the redshift range of our samples (e.g. Oesch et al. 2010; Alavi et al. 2016). This indicates that the redshift-independent trends are not an artefact of the model-dependent Schechter parameters.

We also include M_{UV} -limit measurements from the literature which cover the faintest end of Fig. 6 (Ouchi et al. 2003; Gawiser et al. 2007; Murayama et al. 2007; Guaita et al. 2010; Ouchi et al. 2010). Presently, Bielby et al. (2016) is the only work that covered multiple M_{UV} -limit thresholds for which they measured halo masses. Their measurements cover the range $-18 < M_{\text{UV}} < -20$

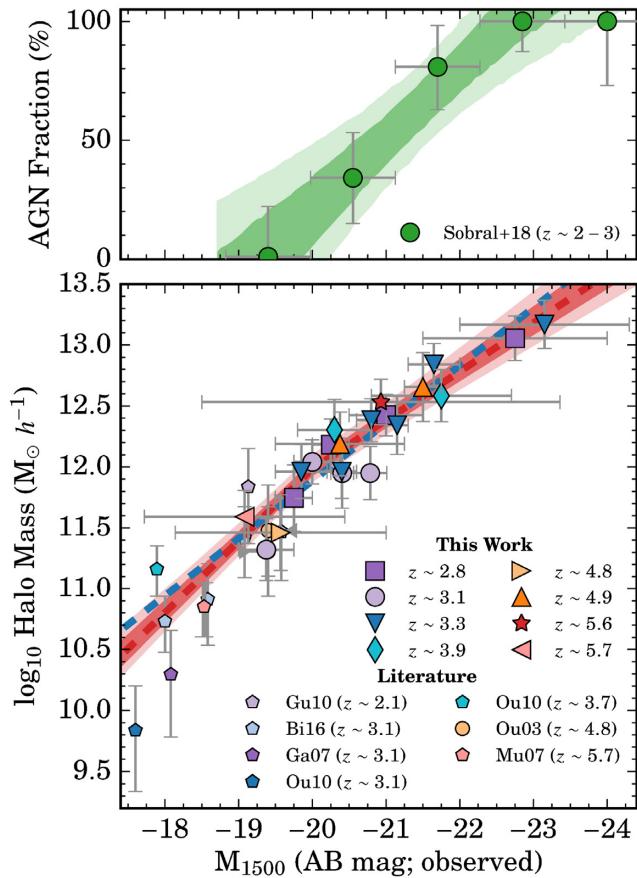


Figure 6. The host halo mass versus the rest-frame 1500 Å UV continuum luminosity for each redshift bin. We observe a strong, redshift-independent trend for the full range of UV luminosities observed where galaxies with the brightest continuum reside in massive haloes. Our best-fitting models are shown in blue and red, dashed lines corresponding to our single power law and piecewise power-laws fits, respectively. The 1σ and 2σ regions for our piecewise power-law models is highlighted as dark and light red regions, respectively. Included are the M_{UV} limit measurements from the literature. Bielby et al. (2016) covered multiple M_{UV} limits and also found a similar trend and even extend our observed trends down to $M_{UV} \sim -18$. The AGN fraction measurements of Sobral et al. (2018b) are shown in the top panel and show that our UV-bright LAEs are primarily AGNs.

and show an increasing trend between M_{UV} and halo mass in perfect agreement with the trends we observe with our samples. Furthermore, their $M_{UV} > -19$ measurements complement ours by showing that the trends seen at brighter UV luminosities continues down to $M_{UV} \sim -18$.

Using both our measurements and those from the literature, we fit a piecewise power law:

$$\frac{M_{halo}}{M_\odot/h} = \begin{cases} 10^{11.99^{+0.05}_{-0.06} - 0.40^{+0.03}_{-0.04} (M_{UV} + 20)} & M_{UV} > -20 \\ 10^{11.99^{+0.05}_{-0.06} - 0.60^{+0.10}_{-0.13} (M_{UV} + 20)} & M_{UV} < -20 \end{cases} \quad (14)$$

with the pivot at $M_{UV} = -20$ mag, which is consistent with a changing slope towards fainter UV luminosities as shown by the literature measurements. The different slopes are statistically significant ($>1\sigma$) and show a typical host halo mass of $\sim 10^{12} M_\odot$ at $M_{UV} \sim -20$ mag. This is similar to what we find for typical $L^*(z)$ galaxies as shown in Fig. 5, although the trend change is not as statistically significant. We also find the best-fitting single power

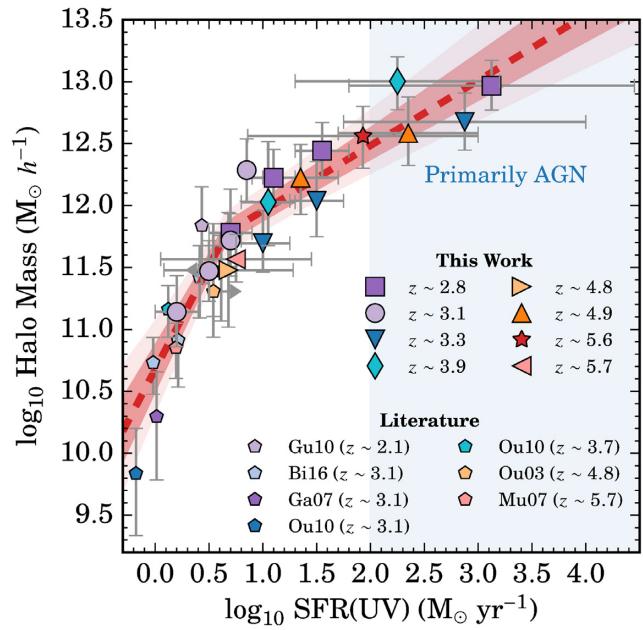


Figure 7. The host halo mass as measured in bins of dust-corrected rest-frame UV star formation rate. We find that an increasing, redshift-independent trend between increasing halo mass and increasing star formation rate. Our best-fitting model is shown as a red dashed line with the 1σ and 2σ regions highlighted as dark and light red regions, respectively. Included are the dust-corrected M_{UV} -limit literature measurements from various narrowband surveys. Bielby et al. (2016) covers multiple star formation rate bins and also shows a similar trend in comparison to our observations, although for a limited star formation rate range. Above $10 M_\odot \text{yr}^{-1}$ and halo masses of $10^{12} M_\odot$, the observed trends become shallower, similar to our observations of the halo mass–Ly α luminosity trends.

law of $\log_{10} M_{halo}/M_\odot/h = (-0.47 \pm 0.03)M_{UV} + (2.43 \pm 0.62)$, although we find a significant deviation at $M_{UV} < -19$ mag.

We also include the AGN fraction measurements of Sobral et al. (2018b) in the top panel of Fig. 6. The AGN fraction is found to increase with UV luminosity, such that 50 per cent of LAEs are AGNs by $M_{UV} \sim -21$ mag. We find halo masses of $> 10^{12.5} M_\odot$ within the AGN-dominated regime, which is consistent with other clustering studies of AGNs (e.g. Hickox et al. 2009; Mendez et al. 2016). This is also similar to what we find for the halo mass dependency with Ly α luminosity such that the brightest LAEs in terms of Ly α and UV luminosity at high- z are consistent with AGNs.

5.4 Star formation rate

The results in Fig. 6 are based on the observed M_{UV} for which the UV luminosity is not corrected for dust. To explore how host halo mass depends directly on the star formation rate, we dust correct M_{UV} using the UV slope, β_{UV} , and use the Kennicutt (1998) calibration as described in Section 4.

Fig. 7 shows the host halo mass for each LAE sample in bins of UV star formation rates. We find that the host halo mass increases with increasing star formation rate at all redshifts. The trends observed are also found to be redshift independent, similar to the other trends with galaxy properties that have been noted in this paper. The range of dark matter halo masses shown varies greatly, with the least active galaxies ($SFR \sim 1.6 M_\odot \text{yr}^{-1}$) residing in $10^{11.2} M_\odot$ haloes and the most active ($SFR \sim 100 M_\odot \text{yr}^{-1}$) residing

in $10^{12.5} M_\odot$ haloes. SFRs $> 100 M_\odot \text{ yr}^{-1}$ primarily have their UV continuum emission powered by AGNs as we saw in Fig. 6 and therefore should be interpreted with caution. This region is highlighted in Fig. 7.

Included in Fig. 7 are the M_{UV} -limit measurements from $z \sim 2$ to 6 studies from the literature (Ouchi et al. 2003; Gawiser et al. 2007; Murayama et al. 2007; Guaita et al. 2010; Ouchi et al. 2010; Bielby et al. 2016). The measurements were redone to match with the assumptions made in this work (see Appendix B for details) and converted to SFR using equation (10). The typical β_{UV} slope for these samples are bluer than $\beta_{\text{UV}} \sim -2$, which implies zero to minimal dust attenuation (e.g. see fig. 2 of Ono et al. 2010). Because these measurements are M_{UV} (SFR)-limit studies, they help to constrain the least active end ($\text{SFR} \lesssim 1.6 M_\odot \text{ yr}^{-1}$) of Fig. 7.

We find that two trends are present in Fig. 7 where the halo mass increases rapidly from low SFR to $\sim 4.5 M_\odot \text{ yr}^{-1}$ and continues to increase with a shallower slope to higher SFRs. To quantify these trends, we fit our measurements and those from the literature with a piecewise power law. The best fit is:

$$\frac{M_{\text{halo}}}{M_\odot/h} = 10^{11.77^{+0.08}_{-0.08}} \begin{cases} \left(\frac{\text{SFR}}{4.5}\right)^{1.66^{+0.26}_{-0.23}} & \text{SFR} < 4.5 \frac{M_\odot}{\text{yr}} \\ \left(\frac{\text{SFR}}{4.5}\right)^{0.52^{+0.09}_{-0.06}} & \text{SFR} > 4.5 \frac{M_\odot}{\text{yr}} \end{cases} \quad (15)$$

with a typical halo mass of $10^{11.77^{+0.08}_{-0.08}} M_\odot$ at $\text{SFR} \sim 4.5 M_\odot \text{ yr}^{-1}$, which is the point for which we visually see a change in the trend in Fig. 7.

In comparison to the halo mass– $L_{\text{Ly}\alpha}$ trend we measured, there are many important similarities. The pivot point in the piecewise has similar halo masses and the slopes of both trends are very much similar. This could suggest that $L_{\text{Ly}\alpha}$ is indeed tracing the star formation activity, despite the many caveats surrounding using Ly α as a star formation indicator (see Dijkstra 2017 for a review).

The typical halo mass measured at $\text{SFR} = 4.5 M_\odot \text{ yr}^{-1}$ is consistent with the peak of star formation efficiency found in haloes of $\sim 10^{12} M_\odot$. This is similar to what we also find for the halo mass– $L_{\text{Ly}\alpha}$ results. The changing slope seen above $4.5 M_\odot \text{ yr}^{-1}$ is most likely due to the combined effects of a larger population of AGN and the existence of LAEs that are undergoing an intense period of star formation activity.

The observed trends suggest that the processes that govern star formation activity and the production of the Ly α line in LAEs are strongly tied to the host halo mass properties. The redshift independence reinforces the idea that this connection is independent of time such that haloes and their residing galaxies co-evolve with each other in unison. This would then suggest that one of the most important characteristics that governs the evolution of an LAE is the host dark matter halo mass.

5.4.1 Comparisons to UV-selected samples

Throughout this paper we have focused on LAEs and how they relate to their host halo properties. Here, we investigate how our results relate to UV-selected samples. We use a compilation of halo mass measurements from the literature that are selected as $z \sim 2$ BzK (Lin et al. 2012) and $z \sim 2$ –5 LBGs (Kashikawa et al. 2006; Hildebrandt et al. 2009; Harikane et al. 2016, 2018). We also include the recent $z \sim 3$ VUDS spectroscopic survey measurements (Durkalec et al. 2018).

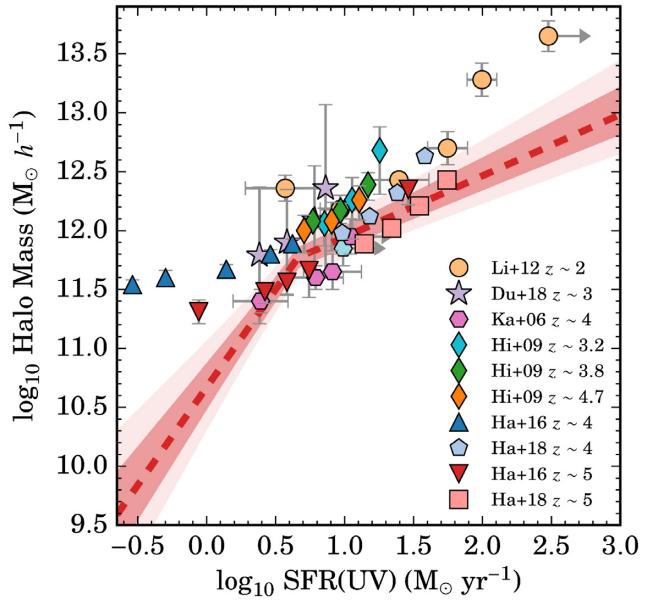


Figure 8. Comparison between our best-fitting halo mass–UV star formation rate relation for LAEs (shown as a dashed red line) and a compilation of measurements from $z \sim 2$ –5 UV-selected samples: BzK (Lin et al. 2012), LBGs (Kashikawa et al. 2006; Hildebrandt et al. 2009; Harikane et al. 2016, 2018), and the VUDS spectroscopic survey (Durkalec et al. 2018). The 1σ and 2σ regions of our model are highlighted in *dark* and *light* red, respectively. The literature measurements are in agreement with our model for $\text{SFR} \gtrsim 3 M_\odot \text{ yr}^{-1}$ ($\log_{10} \text{SFR} \gtrsim 0.5$), although with some scatter which can be due to the different methods used in measuring halo masses in each study (e.g. different clustering slopes, halo occupation model prescriptions, halo mass and bias models, wide redshift distributions). Below $3 M_\odot \text{ yr}^{-1}$, LBGs are measured to reside in higher mass haloes in comparison to LAEs.

Fig. 8 shows the comparison of our results with UV-selected samples drawn from the literature. All measurements from the literature confirm an increasing halo mass with increasing star formation rate. We find that the $z \sim 2$ BzK measurement of Lin et al. (2012) is somewhat higher than our measurements, with their most active galaxies found in more massive haloes. The VUDS measurements of Durkalec et al. (2018) are found to be within 1σ agreement down to $\sim 3 M_\odot \text{ yr}^{-1}$.

The majority of literature measurements shown in Fig. 8 are from LBG-selected samples at $z \sim 2$ –5. We find our measurements are in agreement with LBG studies at $\text{SFR} \gtrsim 3 M_\odot \text{ yr}^{-1}$, while a deviation is seen at lower SFRs. The typical LBG at $z \sim 4$ –5 with $\text{SFR} \gtrsim 3 M_\odot \text{ yr}^{-1}$ is found to reside in $10^{11.3-11.7} M_\odot$ haloes, while we find LAEs reside in significantly lower mass haloes with decreasing star formation rate in respect to LBGs. Our result is in agreement with Bielby et al. (2016), where they measured the LAE–LBG cross-correlation function at $z \sim 3$ and concluded that LAEs comprise the low-luminosity portion of LBGs that reside in low-mass haloes.

We note that our results and those from the literature shown in Fig. 8 are not directly comparable. This is due to the different methods used in measuring halo mass, ranging from the different clustering slopes and methodology in quantifying the correlation functions (e.g. incorporating redshift distributions, Limber approximation) to the halo model prescriptions used (e.g. halo occupation distributions, halo mass and bias functions). All these factors need to be taken into account for a proper quantitative comparison, although, qualitatively, we find that LAEs only seem to deviate

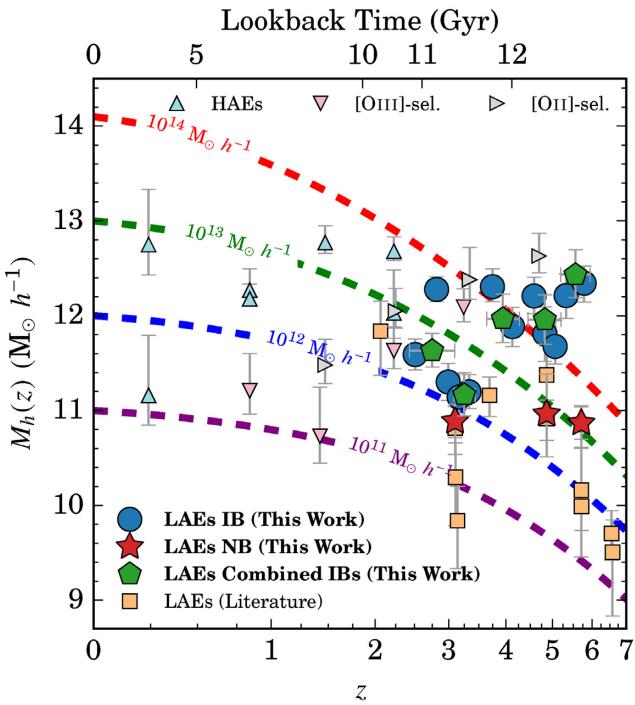


Figure 9. The present-day descendants of emission-line-selected galaxies. The dashed lines are the evolutionary tracks of present-day $10^{11-14} M_{\odot}$ host haloes assuming the halo mass accretion model of Behroozi et al. (2013). Our narrowband-selected LAEs are found to be progenitors of galaxies in present-day $10^{12-13} M_{\odot}$ haloes ('Milky Way-like'), while our intermediate band samples are in more massive present-day haloes ($>10^{13} M_{\odot}$ haloes; 'cluster-like'). Other LAE narrowband surveys show similar results such that Ly α emitters are the progenitors of a wide range of present-day galaxies, similar to what is found for H α -, [O II]-, and [O III]-selected galaxies.

from the halo properties of UV-selected samples at lower UV star formation rates.

5.5 Present-day descendants of LAEs

At all redshift slices, we find that LAEs cover a wide range in host halo mass and redshift-independent trends between halo mass and various galaxy properties. But what types of galaxies do LAEs evolve into in respect of the present-day? We can address this question using halo mass accretion models to predict the present-day halo mass of a host halo at a given redshift.

Fig. 9 shows the present-day halo masses of our full LAE samples measured using the halo mass accretion tracks of Behroozi, Wechsler & Conroy (2013). Given the wide range in halo masses we have observed throughout this study, we find that LAEs cover a wide range of present-day descendants from dwarf-like ($M_{\text{halo}}(z=0) \sim 10^{11} M_{\odot}$), to Milky Way-like ($M_{\text{halo}}(z=0) \sim 10^{12} M_{\odot}$), and galaxies residing in cluster-like environments ($M_{\text{halo}}(z=0) > 10^{13} M_{\odot}$) all the way to the richest clusters. Since we find the brightest LAEs are highly clustered and reside in $10^{13} M_{\odot}$ haloes, they are most likely the 'protoclusters' that are known to be the densest regions in the high- z Universe and evolve to become the massive, rich galaxy clusters we see in the local Universe (e.g. Franck & McGaugh 2016; Overzier 2016 and references therein). Our results suggest that bright LAEs are useful candidates in selecting potential overdense regions for further study on massive cluster formation.

The wide range in the type of descendants shows that LAEs are great tools in studying how galaxies formed and evolved to the ones we currently see in the local Universe. Other narrowband-selected samples overlaid in Fig. 9 also show a wide range in present-day descendants, suggesting that narrowband surveys, in general, can provide us with samples of star-forming galaxies at various redshifts to map out the evolutionary track of galaxies from high z to the present-day.

6 CONCLUSIONS

We present a comprehensive investigation of the clustering and halo properties of ~ 5000 LAEs and explore their dependences on Ly α luminosity, UV continuum luminosity, and UV star formation rate in multiple redshift slices in the range $z \sim 2.5-6$. We highlight the main results of this study here:

(i) The clustering lengths of the narrowband LAE samples are shown to increase from $r_0 \sim 3.1$ to $5.0 \text{ Mpc } h^{-1}$ in the range $z \sim 3.1-5.7$. The intermediate band and combined intermediate band samples show a more rapid increase in r_0 from $\sim 4.5 \text{ Mpc } h^{-1}$ at $z \sim 2.5$ to $\sim 16 \text{ Mpc } h^{-1}$ by $z \sim 5.8$. The typical halo masses of the narrowband samples are found to be $\sim 10^{11} M_{\odot}$, while the intermediate band samples range between $\sim 10^{11}$ and $10^{12} M_{\odot}$.

(ii) Host halo mass is found to increase with increasing Ly α luminosity at all redshifts probed.

(iii) Normalizing Ly α luminosity by $L^*(z)$ shows a redshift-independent trend with host halo mass. LAEs are found to reside in a wide range of host haloes ranging from $10^{9.75} M_{\odot}$ at $\sim 0.1 L^*(z)$ to $10^{12.1} M_{\odot}$ at $L^*(z)$ and $10^{12.8} M_{\odot}$ at $\sim 10 L^*(z)$.

(iv) We find a strong, redshift-independent trend between host halo mass and observed 1500Å UV continuum luminosity. LAEs with $M_{\text{UV}} \sim -18$ mag are found to reside in $10^{10.5} M_{\odot}$ haloes and $M_{\text{UV}} \sim -23$ mag in $10^{13} M_{\odot}$ haloes.

(v) We also find a strong, redshift-independent trend between host halo mass and dust-corrected UV star formation rate. We find that LAEs with $\text{SFR} \sim 1 M_{\odot} \text{ yr}^{-1}$ reside in $10^{10} M_{\odot}$ haloes and $\sim 100 M_{\odot} \text{ yr}^{-1}$ reside in $10^{12.5} M_{\odot}$ haloes.

(vi) For both Ly α luminosity and UV SFR, we observe sharp trend changes. In the case of Ly α luminosity, we find that the host halo mass scales as $(L/L^*(z))^{2.08}$ and $(L/L^*(z))^{0.63}$ for below and above $L^*(z)$, respectively. A similar trend is seen between halo mass and SFR with the trend change occurring at $\sim 4.5 M_{\odot} \text{ yr}^{-1}$. This is attributed to a changing population of LAEs where the brightest LAEs, in terms of line luminosity and SFRs, primarily have their emission powered by AGNs and not star formation activity.

(vii) We find that LAEs are progenitors of a wide range of present-day galaxies depending on their Ly α luminosity ranging from dwarf-like systems (similar to LMC and SMC) to galaxies in cluster-like environments, such that LAEs are great tools to understand the evolutionary path of galaxies from high z to the local Universe.

Our results highlight the significant connection that host haloes and galaxies share from the end of the epoch of reionization to the peak of cosmic star formation. The redshift-independent trends with halo mass signify the co-evolution of galaxies and their host haloes and emphasize the importance of haloes in the overall evolution of galaxies. The results presented in this paper provide empirical relations that can be tested in simulations and provide new constraints in regard to the physics behind galaxy evolution. Our results also emphasize the importance of investigating clustering and halo properties in terms of various galaxy properties to

take selection effects into account when comparing samples from different surveys.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their useful comments and suggestions that helped improve this study. AAK acknowledges that this work was supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program – Grant NNX16AO92H. JM acknowledges support from the ETH Zwickly fellowship. RKC acknowledges funding from STFC via a studentship. APA acknowledges support from the Fundação para a Ciência e a Tecnologia FCT through the fellowship PD/BD/52706/2014 and the research grant UID/FIS/04434/2013. JC and SS both acknowledge their support from the Lancaster University PhD Fellowship.

We have benefited greatly from the publicly available programming language PYTHON, including the NUMPY, SCIPY, MATPLOTLIB, SCIKIT-LEARN, and ASTROPY packages, as well as the TOPCAT analysis program.

The SC4K samples used in this paper are all publicly available for use by the community (Sobral et al. 2018a). The catalogue is also available on the COSMOS IPAC website (<https://irsa.ipac.caltech.edu/data/COSMOS/overview.html>).

REFERENCES

- Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005, *ApJ*, 619, 697
- Alavi A. et al., 2016, *ApJ*, 832, 56
- Allevato V. et al., 2016, *ApJ*, 832, 70
- Barone-Nugent R. L. et al., 2014, *ApJ*, 793, 17
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Benson A. J., 2010, *Phys. Rep.*, 495, 33
- Bielby R. et al., 2013, *MNRAS*, 430, 425
- Bielby R. M. et al., 2016, *MNRAS*, 456, 4061
- Blain A. W., Chapman S. C., Smail I., Ivison R., 2004, *ApJ*, 611, 725
- Blanc G. A. et al., 2008, *ApJ*, 681, 1099
- Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, *MNRAS*, 465, 32
- Brodwin M. et al., 2008, *ApJ*, 687, L65
- Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, *ApJ*, 429, 582
- Capak P. et al., 2007, *ApJS*, 172, 99
- Ciardullo R. et al., 2012, *ApJ*, 744, 110
- Cochrane R. K., Best P. N., Sobral D., Smail I., Wake D. A., Stott J. P., Geach J. E., 2017, *MNRAS*, 469, 2913
- Cochrane R. K., Best P. N., Sobral D., Smail I., Geach J. E., Stott J. P., Wake D. A., 2018, *MNRAS*, 475, 3730
- Cooray A., Sheth R., 2002, *Phys. Rep.*, 372, 1
- Cowie L. L., Hu E. M., 1998, *AJ*, 115, 1319
- Dekel A., Birnboim Y., 2006, *MNRAS*, 368, 2
- Diemer B., 2018, *ApJS*, 239, 35
- Dijkstra M., 2017, preprint ([arXiv:1704.03416](https://arxiv.org/abs/1704.03416))
- Dolley T. et al., 2014, *ApJ*, 797, 125
- Durkalec A. et al., 2015, *A&A*, 583, A128
- Durkalec A. et al., 2018, *A&A*, 612, A42
- Erb D. K., Pettini M., Steidel C. C., Strom A. L., Rudie G. C., Trainor R. F., Shapley A. E., Reddy N. A., 2016, *ApJ*, 830, 52
- Finkelstein S. L. et al., 2011, *ApJ*, 729, 140
- Finkelstein S. L., Rhoads J. E., Malhotra S., Pirzkal N., Wang J., 2007, *ApJ*, 660, 1023
- Foucaud S., McCracken H. J., Le Fèvre O., Arnouts S., Brodwin M., Lilly S. J., Crampton D., Mellier Y., 2003, *A&A*, 409, 835
- Franck J. R., McGaugh S. S., 2016, *ApJ*, 833, 15
- Gawiser E. et al., 2006, *ApJ*, 642, L13
- Gawiser E. et al., 2007, *ApJ*, 671, 278
- Geach J. E., Smail I., Best P. N., Kurk J., Casali M., Ivison R. J., Coppin K., 2008, *MNRAS*, 388, 1473
- Geach J. E., Sobral D., Hickox R. C., Wake D. A., Smail I., Best P. N., Baugh C. M., Stott J. P., 2012, *MNRAS*, 426, 679
- Gronwall C. et al., 2007, *ApJ*, 667, 79
- Guaita L. et al., 2010, *ApJ*, 714, 255
- Guaita L. et al., 2011, *ApJ*, 733, 114
- Hagen A. et al., 2016, *ApJ*, 817, 79
- Hale C. L., Jarvis M. J., Delvecchio I., Hatfield P. W., Novak M., Smolčić V., Zamorani G., 2018, *MNRAS*, 474, 4133
- Hao C.-N., Huang J.-S., Xia X., Zheng X., Jiang C., Li C., 2018, *ApJ*, 864, 145
- Harikane Y. et al., 2016, *ApJ*, 821, 123
- Harikane Y. et al., 2018, *PASJ*, 70, S11
- Hartley W. G. et al., 2008, *MNRAS*, 391, 1301
- Hartley W. G. et al., 2013, *MNRAS*, 431, 3045
- Hashimoto T. et al., 2017, *A&A*, 608, A10
- Hayashi M., Shimasaku K., Motohara K., Yoshida M., Okamura S., Kashikawa N., 2007, *ApJ*, 660, 72
- Hayashino T. et al., 2004, *AJ*, 128, 2073
- Hickox R. C. et al., 2009, *ApJ*, 696, 891
- Hickox R. C. et al., 2012, *MNRAS*, 421, 284
- Hildebrandt H., Pielorz J., Erben T., van Waerbeke L., Simon P., Capak P., 2009, *A&A*, 498, 725
- Hu E. M., Cowie L. L., Barger A. J., Capak P., Kakazu Y., Trouille L., 2010, *ApJ*, 725, 394
- Ishikawa S., Kashikawa N., Toshikawa J., Onoue M., 2015, *MNRAS*, 454, 205
- Kashikawa N. et al., 2006, *ApJ*, 637, 631
- Kashino D. et al., 2017, *ApJ*, 843, 138
- Kennicutt R. C., Jr., 1998, *ARA&A*, 36, 189
- Khostovan A. A., Sobral D., Mobasher B., Best P. N., Smail I., Stott J. P., Hemmati S., Nayyeri H., 2015, *MNRAS*, 452, 3948
- Khostovan A. A. et al., 2018, *MNRAS*, 478, 2999
- Kobayashi M. A. R. et al., 2016, *ApJ*, 819, 25
- Kong X. et al., 2006, *ApJ*, 638, 72
- Konno A., Ouchi M., Nakajima K., Duval F., Kusakabe H., Ono Y., Shimasaku K., 2016, *ApJ*, 823, 20
- Konno A. et al., 2018, *PASJ*, 70, S16
- Koutoulidis L., Plonis M., Georgantopoulos I., Fanidakis N., 2013, *MNRAS*, 428, 1382
- Kováč K., Somerville R. S., Rhoads J. E., Malhotra S., Wang J., 2007, *ApJ*, 668, 15
- Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, *ApJ*, 609, 35
- Kubo M., Yamada T., Ichikawa T., Kajisawa M., Matsuda Y., Tanaka I., 2015, *ApJ*, 799, 38
- Kubo M., Yamada T., Ichikawa T., Kajisawa M., Matsuda Y., Tanaka I., Umehata H., 2016, *MNRAS*, 455, 3333
- Kusakabe H. et al., 2018, *PASJ*, 70, 4
- Kusakabe H. et al., 2019, *PASJ*, 71, 55
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Limber D. N., 1953, *ApJ*, 117, 134
- Lin L. et al., 2012, *ApJ*, 756, 71
- Ly C., Lee J. C., Dale D. A., Momcheva I., Salim S., Staudaher S., Moore C. A., Finn R., 2011, *ApJ*, 726, 109
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Malhotra S., Rhoads J. E., 2002, *ApJ*, 565, L71
- Malhotra S., Rhoads J. E., Finkelstein S. L., Hathi N., Nilsson K., McLinden E., Pirzkal N., 2012, *ApJ*, 750, L36
- Matsuda Y. et al., 2004, *AJ*, 128, 569
- Matthee J., Sobral D., Santos S., Röttgering H., Darvish B., Mobasher B., 2015, *MNRAS*, 451, 400
- Matthee J., Sobral D., Best P., Smail I., Bian F., Darvish B., Röttgering H., Fan X., 2017, *MNRAS*, 471, 629
- McCracken H. J. et al., 2010, *ApJ*, 708, 202
- Mendez A. J. et al., 2016, *ApJ*, 821, 55
- Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, 521, 64

- Murayama T. et al., 2007, *ApJS*, 172, 523
 Nestor D. B., Shapley A. E., Steidel C. C., Siana B., 2011, *ApJ*, 736, 18
 Nilsson K. K., Tapken C., Møller P., Freudling W., Fynbo J. P. U., Meisenheimer K., Laursen P., Östlin G., 2009, *A&A*, 498, 13
 Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, *MNRAS*, 396, 19
 Oesch P. A. et al., 2010, *ApJ*, 725, L150
 Ono Y., Ouchi M., Shimasaku K., Dunlop J., Farrah D., McLure R., Okamura S., 2010, *ApJ*, 724, 1524
 Ouchi M. et al., 2003, *ApJ*, 582, 60
 Ouchi M. et al., 2004, *ApJ*, 611, 685
 Ouchi M. et al., 2008, *ApJS*, 176, 301
 Ouchi M. et al., 2010, *ApJ*, 723, 869
 Ouchi M. et al., 2018, *PASJ*, 70, S13
 Overzier R. A., 2016, *A&AR*, 24, 14
 Paulino-Afonso A. et al., 2018, *MNRAS*, 476, 5479
 Peebles P. J. E., 1980, *The Large-Scale Structure of the Universe*, Princeton University Press
 Rhoads J. E., Malhotra S., Dey A., Stern D., Spinrad H., Jannuzzi B. T., 2000, *ApJ*, 545, L85
 Saez C. et al., 2015, *MNRAS*, 450, 2615
 Santos S., Sobral D., Matthee J., 2016, *MNRAS*, 463, 1678
 Savoy J., Sawicki M., Thompson D., Sato T., 2011, *ApJ*, 737, 92
 Scoville N. et al., 2007, *ApJS*, 172, 1
 Shimakawa R. et al., 2017, *MNRAS*, 468, 1123
 Shimasaku K. et al., 2004, *ApJ*, 605, L93
 Shioya Y. et al., 2008, *ApJS*, 175, 128
 Shioya Y. et al., 2009, *ApJ*, 696, 546
 Simon P., 2007, *A&A*, 473, 711
 Sobral D., Best P. N., Geach J. E., Smail I., Cirasuolo M., Garn T., Dalton G. B., Kurk J., 2010, *MNRAS*, 404, 1551
 Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, *MNRAS*, 428, 1128
 Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, *MNRAS*, 437, 3516
 Sobral D. et al., 2017, *MNRAS*, 466, 1242
 Sobral D., Santos S., Matthee J., Paulino-Afonso A., Ribeiro B., Calhau J., Khostovan A. A., 2018a, *MNRAS*, 476, 4752
 Sobral D. et al., 2018b, *MNRAS*, 477, 2817
 Sobral D., Matthee J., 2019, *A&A*, 623, A157
 Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51
 Steidel C. C., Adelberger K. L., Dickinson M., Giavalisco M., Pettini M., Kellogg M., 1998, *ApJ*, 492, 428
 Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2000, *ApJ*, 532, 170
 Stroe A., Sobral D., 2015, *MNRAS*, 453, 242
 Takahashi M. I. et al., 2007, *ApJS*, 172, 456
 Taniguchi Y. et al., 2007, *ApJS*, 172, 9
 Taniguchi Y. et al., 2015, *PASJ*, 67, 104
 Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
 Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878
 Toba Y. et al., 2017, *ApJ*, 835, 36
 Topping M. W., Shapley A. E., Steidel C. C., 2016, *ApJ*, 824, L11
 Totsuji H., Kihara T., 1969, *PASJ*, 21, 221
 Trainor R. F., Strom A. L., Steidel C. C., Rudie G. C., 2016, *ApJ*, 832, 171
 Webb T. M. A., Yamada T., Huang J.-S., Ashby M. L. N., Matsuda Y., Egami E., Gonzalez M., Hayashimo T., 2009, *ApJ*, 692, 1561
 Wilkinson A. et al., 2017, *MNRAS*, 464, 1380
 Wold I. G. B., Barger A. J., Cowie L. L., 2014, *ApJ*, 783, 119
 Yamada T., Nakamura Y., Matsuda Y., Hayashimo T., Yamauchi R., Morimoto N., Kousai K., Umemura M., 2012, *AJ*, 143, 79
 Zheng Z. et al., 2005, *ApJ*, 633, 791
 Zheng Z.-Y., Wang J.-X., Malhotra S., Rhoads J. E., Finkelstein S. L., Finkelstein K., 2014, *MNRAS*, 439, 1101
 Zheng Z.-Y., Malhotra S., Rhoads J. E., Finkelstein S. L., Wang J.-X., Jiang C.-Y., Cai Z., 2016, *ApJS*, 226, 23

APPENDIX A: EFFECTS OF THE SA22 OVERDENSE REGION

Previous work on the SA22 field has identified and extensively studied a significant overdense region (protocluster) comprised of LAEs and LBGs at $z \sim 3.1$ (e.g. Steidel et al. 1998, 2000; Hayashino et al. 2004; Webb et al. 2009; Nestor et al. 2011; Yamada et al. 2012; Kubo et al. 2015, 2016; Saez et al. 2015; Topping, Shapley & Steidel 2016). Fig. A1 shows a 0.12 deg² cut-out centred on the position of the overdense region highlighted by the source isodensity levels from Hayashino et al. (2004). We include the distribution of our $z \sim 3.1$ LAEs and find that it traces the underlying overdensity. The main question that needs to be addressed is how does the overdense region affect our observed ACFs as shown in the top panel of Fig. 2?

As described in Section 3.2, there is a signature of a 1-halo term in the NB497 ACFs which would suggest a satellite fraction of LAEs. We investigate if this is due to the presence of the significant overdense region discussed above by measuring the ACFs using the same methodology used throughout this paper and masking the 0.12 deg² field shown in Fig. A1. Note that we also mask the random maps as well to ensure a consistent survey geometry.

Fig. A2 shows the comparison between the ACFs for the full field and for the case where the overdense region is masked. We find that there is no difference between the two correlation functions, such that the overdensity does not significantly affect the overall clustering measurements. Hayashino et al. (2004) also found the overdense region to have a weak clustering signal relative to the whole SA22 field.

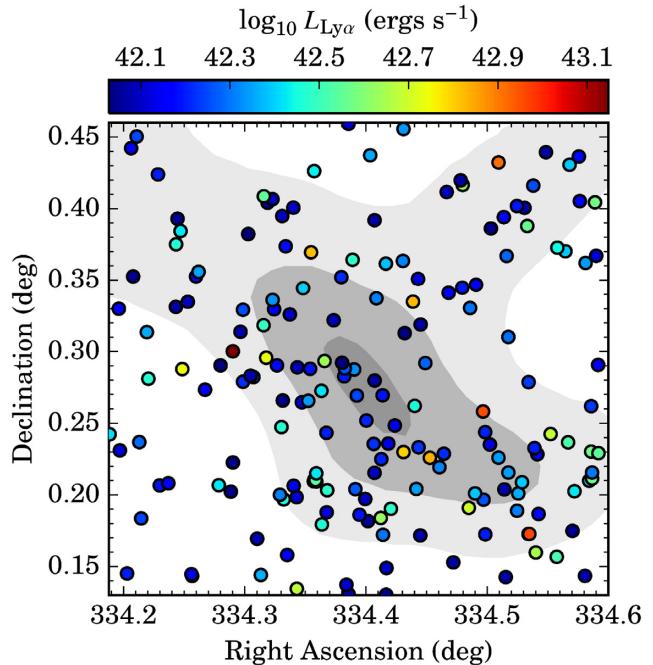


Figure A1. A 24 arcmin \times 18 arcmin cut-out of the SA22 field centred on the $z \sim 3.1$ protocluster. The circles represent the angular positions of our NB497-selected LAEs with the colour scale related to their line luminosities. The grey shaded regions are the source isodensity levels from Hayashino et al. (2004) and compiled by Saez et al. (2015) that highlight the location of the SA22 overdense region. The $z = 3.1$ LAEs that seem to primarily populate the overdense region (darker grey regions) have line luminosities of $\lesssim 10^{42.4}$ erg s⁻¹, which are the faintest LAEs in the sample.

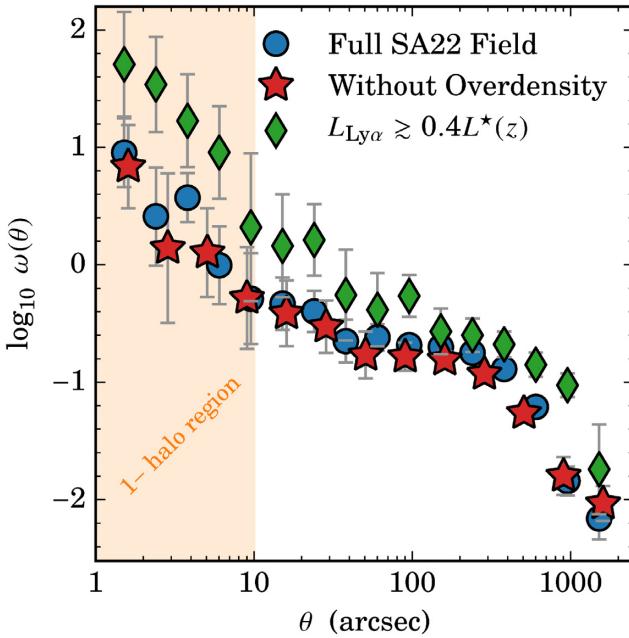


Figure A2. Comparison of the ACF for the Full SA22 field (blue circles) and for the case where the overdensity is masked (red stars). We find that both ACFs are consistent with each other, such that the $z \sim 3.1$ protocluster is not affecting our clustering measurements and that the 1-halo term still exists in both cases (region highlighted in orange). The green diamonds show the ACF for the Full SA22 field with the added condition of a line luminosity threshold $\gtrsim 0.4L^*$. We find that above this threshold, the deviation from a simple power law associated with the detection of a 1-halo term disappears, suggesting that the satellite fraction of galaxies is close to zero above $0.4L^*$.

We next investigate if the cause of the 1-halo term is due to flux depth of the sample. Fig. A1 shows the spatial and luminosity distribution of the $z \sim 3.1$ LAEs where concentrations of faint LAEs are found within the overdense region. To test if the 1-halo term arises from faint LAEs that comprise the satellite population, we measure the ACFs for the full SA22 field with varying line luminosity thresholds. Fig. A2 shows the 1-halo term disappears

by $L_{\text{Ly}\alpha} \gtrsim 0.4L^*$, which is consistent with the spatial/luminosity distribution shown in Fig. A1.

Although a detailed analysis of the satellite fractions is beyond the scope of this work, we find that all our samples do not exhibit signatures of population of satellite galaxies except for the NB497 sample, which we find to also have negligible satellite fraction effects above $0.4L^*$. If the satellite fractions do affect our measurements, then it would be primarily isolated to the lowest line luminosity ranges of our samples and would result in overestimations of their host dark matter haloes. Lastly, the most comprehensive study of satellite fractions of emission-line-selected galaxies was done by Cochrane et al. (2017) for their $z = 0.4$, 0.8, 1.5, and 2.2 NB-selected H α samples where they find that, on average, 3–5 per cent of the sample are satellites. In comparison to LBG samples, the recent 100 deg 2 HSC survey found satellite fractions of $\lesssim 5$ per cent between $z \sim 4$ and 6 (Harikane et al. 2018). Throughout the rest of this paper, we neglect the effects of the satellite population.

APPENDIX B: CLUSTERING MEASUREMENTS FROM THE LITERATURE

We have used several narrowband studies for the purpose of comparison throughout this paper. Due to our unique approach and the varying assumptions between each measurement drawn from the literature, we have to be careful about how we are comparing our measurements to the literature. To resolve this issue, we have to normalize the methodology between our clustering measurements and those from the literature. We achieve this by taking the observed ACFs from each narrowband study, fit equation (5) to measure r_0 and use the narrowband filter attributed to that study as the proxy for the redshift distribution, and include the errors associated with cosmic variance using the empirical relation measured by Sobral et al. (2010).

Table B1 shows our recalculations of r_0 for each narrowband study used in this paper for comparison purposes. We also include the measured halo masses in Table B1, which are based on the same assumptions described in Section 3.4. In comparison to the measurements reported in each study, the error bars we measure

Table B1. Clustering measurements of LAEs from the literature. Shown are the referred narrowband studies, redshifts per each sample, the corresponding narrowband filter identification, the survey area, the r_0 reported in the respective study, our measurement of r_0 based on our assumptions and methodology using the observed ACFs, and the host halo mass using our own r_0 -halo mass model.

Study	Redshift	Filter	Area (deg 2)	r_0 , reported (Mpc h $^{-1}$)	r_0 , measured (Mpc h $^{-1}$)	\log_{10} Halo mass (M $_\odot$ h $^{-1}$)
Hao et al. (2018)	2.23 ± 0.03	NB393	0.34	2.56 ± 0.33	3.02 ± 0.53	$11.32^{+0.29}_{-0.40}$
Bielby et al. (2016)	3.10 ± 0.03	NB497	1.07	2.86 ± 0.33	2.99 ± 0.40	$10.81^{+0.23}_{-0.28}$
Gawiser et al. (2007)	3.11 ± 0.02	NB4990	0.27	$2.52^{+0.56}_{-0.70}$	$2.34^{+0.43}_{-0.43}$	$10.30^{+0.36}_{-0.51}$
Guaita et al. (2010)	2.07 ± 0.02	NB3727	0.36	$4.80^{+0.90}_{-0.90}$	$4.33^{+1.01}_{-1.01}$	$11.84^{+0.31}_{-0.46}$
Murayama et al. (2007)	5.71 ± 0.04	NB816	1.95	—	$4.96^{+0.71}_{-0.71}$	$10.85^{+0.20}_{-0.25}$
Ouchi et al. (2003)	4.86 ± 0.03	NB711	0.15	—	$6.03^{+1.49}_{-1.49}$	$11.37^{+0.30}_{-0.43}$
Ouchi et al. (2010)	3.14 ± 0.03	NB503	0.98	$1.70^{+0.39}_{-0.46}$	$1.96^{+0.30}_{-0.30}$	$9.84^{+0.50}_{-0.36}$
Ouchi et al. (2010)	3.69 ± 0.03	NB570	0.96	$2.74^{+0.58}_{-0.72}$	$4.11^{+0.52}_{-0.52}$	$11.16^{+0.19}_{-0.22}$
Ouchi et al. (2010)	5.71 ± 0.04	NB816	1.03	$3.12^{+0.33}_{-0.36}$	$3.29^{+0.99}_{-0.99}$	$10.16^{+0.45}_{-0.70}$
Ouchi et al. (2010)	6.55 ± 0.05	NB921	0.90	$2.31^{+0.65}_{-0.85}$	$2.93^{+0.39}_{-0.39}$	$9.70^{+0.24}_{-0.28}$
Ouchi et al. (2018)	5.72 ± 0.05	NB816	13.8	$3.01^{+0.35}_{-0.35}$	$3.01^{+0.37}_{-0.37}$	$9.99^{+0.22}_{-0.25}$
Ouchi et al. (2018)	6.58 ± 0.05	NB921	21.2	$2.66^{+0.49}_{-0.70}$	$2.66^{+0.50}_{-0.71}$	$9.50^{+0.34}_{-0.67}$
Shioya et al. (2009)	4.86 ± 0.03	NB711	1.83	$4.40^{+1.30}_{-1.50}$	$4.44^{+0.59}_{-0.59}$	$10.91^{+0.20}_{-0.23}$

are typically larger than that reported in the literature due to the inclusion of cosmic variance effects.

We detail on a few of these studies as we had to apply specific corrections/extensions. For the Ouchi et al. (2003) study, only a measurement of the clustering amplitude and slope was reported for which we have extended this work by making measurements of r_0 and halo mass. The ACFs of Murayama et al. (2007) did not include an IC correction. Since the survey size and sample is essentially the same as our NB816 sample (note that we used the archival NB816 images which are the same used in Murayama et al. 2007), we use our IC to correct their ACFs. For the Shioya et al. (2009), we report their measurement of $4.4^{+1.3}_{-1.5}$ Mpc h^{-1} which

assumes a slope of $\gamma = -1.90 \pm 0.22$. In our recalculation of their measurement, we keep γ fixed on -1.80 , which is consistent with their measured slope.

APPENDIX C: INTERMEDIATE BAND ANGULAR CORRELATION FUNCTIONS

Shown in Figs C1 and C2 are the ACFs of our intermediate band and combined intermediate band samples. We show the median ACFs based on our 2000 iterations in measuring $w(\theta)$ in each figure and also include the best-fitting model as described in equation (5).

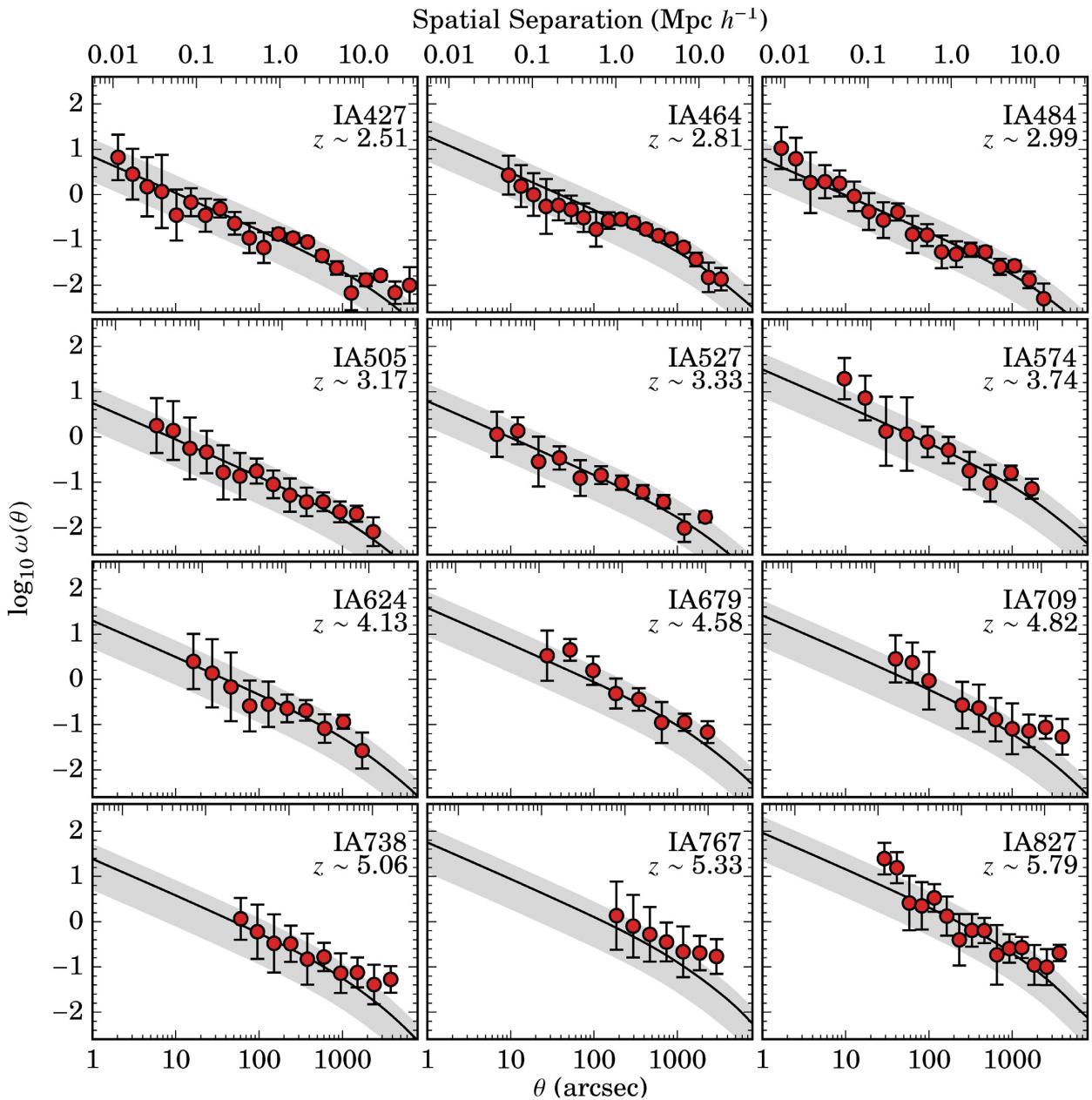


Figure C1. The ACFs for each intermediate band sample. The red circles are the median observed measurements for $w(\theta)$ based on all 2000 iterations. The black line shows the best-fitting model as described in equation (5) with the 1σ uncertainty represented as the grey. The spatial axis shown in each panel corresponds to the spatial separation for a given angular separation at the redshift of the samples shown. We find no significant detections of the 1-halo term, implying that satellite fractions are negligible for these LAEs. Note that the IB samples are biased towards the brightest LAEs (typically $L > L^*(z)$ galaxies).

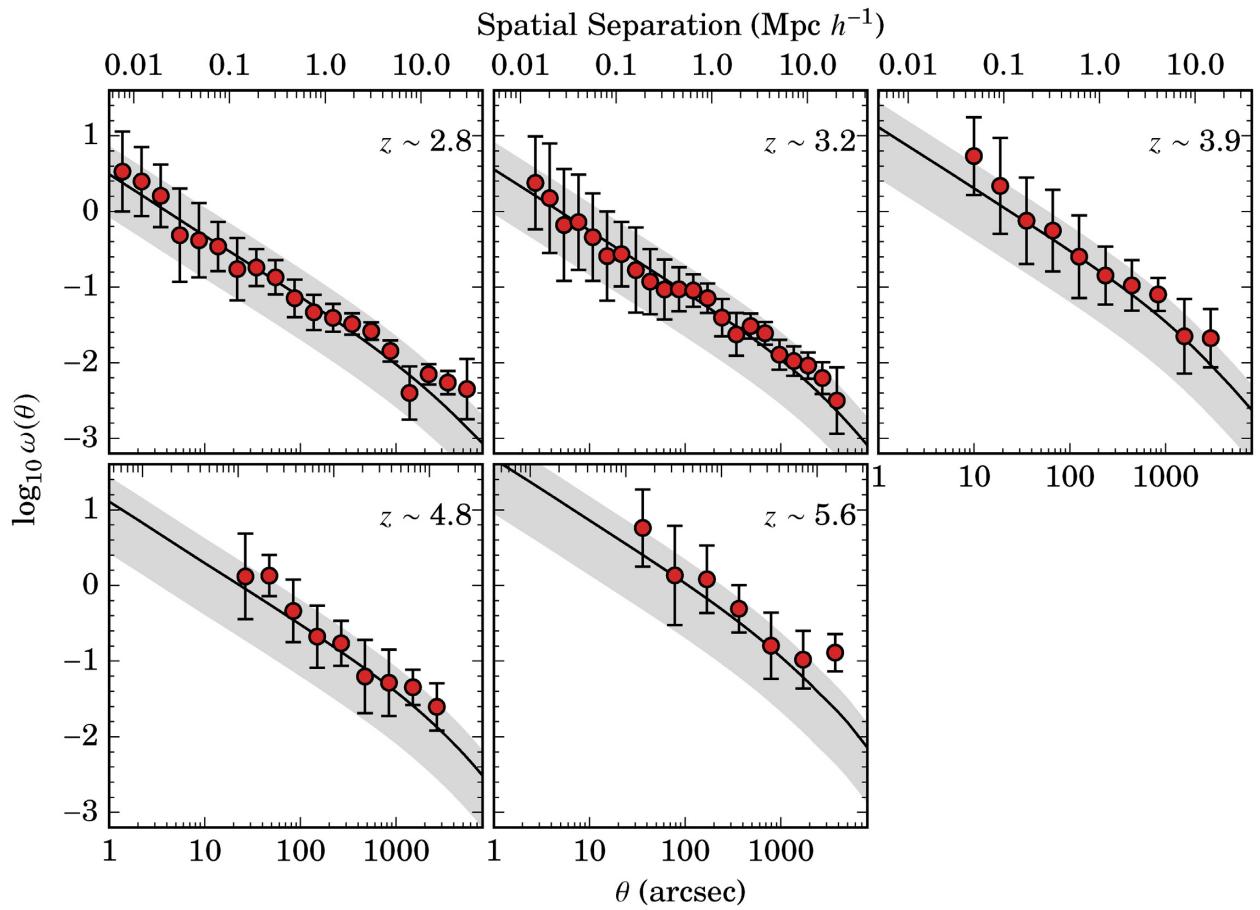


Figure C2. Same description as in Fig. C1 but for the combined intermediate samples. The main importance of the combined IB samples is the large increase in sample sizes, especially at $z > 4$ where the individual IB samples contain $\sim 30\text{--}100$ LAEs each. As found in Fig. C1, we find no significant detection of a 1-halo term such that the satellite fraction in these LAE samples are negligible.

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