



Universiteit  
Leiden  
The Netherlands

## The clustering of $H\beta + [O III]$ and $[O II]$ emitters since $z \sim 5$ : dependencies with line luminosity and stellar mass

Khostovan, A.A.; Serrano Goncalves Sobral, D.; Mobasher, B.; Best, P.N.; Smail, I.; Matthee, J.; ... ; Stott, J.P.

### Citation

Khostovan, A. A., Serrano Goncalves Sobral, D., Mobasher, B., Best, P. N., Smail, I., Matthee, J., ... Stott, J. P. (2018). The clustering of  $H\beta + [O III]$  and  $[O II]$  emitters since  $z \sim 5$ : dependencies with line luminosity and stellar mass. *Monthly Notices Of The Royal Astronomical Society*, 478(3), 2999-3015. doi:10.1093/mnras/sty925

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/69563>

**Note:** To cite this publication please use the final published version (if applicable).

# The clustering of H $\beta$ + [OIII] and [OII] emitters since $z \sim 5$ : dependencies with line luminosity and stellar mass

A. A. Khostovan<sup>1\*</sup>†, D. Sobral<sup>2,3</sup>, B. Mobasher<sup>1</sup>, P. N. Best<sup>4</sup>, I. Smail<sup>5,6</sup>,  
J. Matthee<sup>3</sup>, B. Darvish<sup>7</sup>, H. Nayyeri<sup>8</sup>, S. Hemmati<sup>9</sup>, J.P. Stott<sup>2,10</sup>

<sup>1</sup>Department of Physics & Astronomy, University of California, Riverside, United States of America

<sup>2</sup>Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK

<sup>3</sup>Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

<sup>4</sup>SUPA, Institute for Astronomy, Royal Observatory of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

<sup>5</sup>Centre for Extragalactic Astrophysics, Department of Physics, Durham University, Durham DH1 3LE, UK

<sup>6</sup>Institute for Computational Cosmology, Durham University, Durham DH1 3LE, UK

<sup>7</sup>Cahill Center for Astronomy and Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA

<sup>8</sup>Department of Physics & Astronomy, University of California, Irvine, CA 92697, USA

<sup>9</sup>Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA

<sup>10</sup>Sub-department of Astrophysics, Department of Physics, University of Oxford, Oxford OX1 3RH, UK

## ABSTRACT

We investigate the clustering properties of  $\sim 7000$  H $\beta$ + [OIII] and [OII] narrowband-selected emitters at  $z \sim 0.8 - 4.7$  from the High- $z$  Emission Line Survey. We find clustering lengths,  $r_0$ , of  $1.5 - 4.0 h^{-1}$  Mpc and minimum dark matter halo masses of  $10^{10.7-12.1} M_{\odot}$  for our  $z = 0.8 - 3.2$  H $\beta$ + [OIII] emitters and  $r_0 \sim 2.0 - 8.3 h^{-1}$  Mpc and halo masses of  $10^{11.5-12.6} M_{\odot}$  for our  $z = 1.5 - 4.7$  [OII] emitters. We find  $r_0$  to strongly increase both with increasing line luminosity and redshift. By taking into account the evolution of the characteristic line luminosity,  $L^*(z)$ , and using our model predictions of halo mass given  $r_0$ , we find a strong, redshift-independent increasing trend between  $L/L^*(z)$  and minimum halo mass. The faintest H $\beta$ + [OIII] emitters are found to reside in  $10^{9.5} M_{\odot}$  halos and the brightest emitters in  $10^{13.0} M_{\odot}$  halos. For [OII] emitters, the faintest emitters are found in  $10^{10.5} M_{\odot}$  halos and the brightest emitters in  $10^{12.6} M_{\odot}$  halos. A redshift-independent stellar mass dependency is also observed where the halo mass increases from  $10^{11} M_{\odot}$  to  $10^{12.5} M_{\odot}$  for stellar masses of  $10^{8.5} M_{\odot}$  to  $10^{11.5} M_{\odot}$ , respectively. We investigate the interdependencies of these trends by repeating our analysis in a  $L_{\text{line}} - M_{\text{star}}$  grid space for our most populated samples (H $\beta$ + [OIII]  $z = 0.84$  and [OII]  $z = 1.47$ ) and find that the line luminosity dependency is stronger than the stellar mass dependency on halo mass. For  $L > L^*$  emitters at all epochs, we find a relatively flat trend with halo masses of  $10^{12.5-13} M_{\odot}$  which may be due to quenching mechanisms in massive halos which is consistent with a transitional halo mass predicted by models.

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

## 1 INTRODUCTION

Our current understanding of galaxy formation and evolution implies that galaxies formed hierarchically and inside dark matter halos, such that the baryon clustering traces the

underlying dark matter distribution (see [Benson 2010](#) for a review and references therein). We thus expect a galaxy-halo connection for which the evolving properties of galaxies are tied into the changes of their host halos. A detailed investigation of the dark matter halo properties of galaxies and their evolution is then crucial in setting constraints on current models of galaxy formation.

Previous theoretical studies have looked into the galaxy-

\* NASA Earth and Space Science Fellow

† E-mail: akhostov@gmail.com

halo connection in several ways. One such method is by using semi-analytical models that identify dark matter halos from  $N$ -body simulations and populating them with galaxies based on analytic relations of the underlying baryon evolution (see [Baugh 2006](#) and [Somerville & Davé 2015](#) for reviews). Another method is using halo occupation distribution (HOD) models that use probability distributions of how many galaxies reside in halos with a specific mass (see [Cooray & Sheth 2002](#) for a review). A similar approach is abundance matching, which works by assigning the most massive galaxies to the most massive halos (e.g., [Behroozi et al. 2010](#); [Guo et al. 2010](#); [Moster et al. 2010](#)), although there are several caveats in this technique such as the scatter of stellar mass for a given halo mass and the contribution of satellite galaxies (e.g., [Contreras et al. 2015](#)).

On the observational side, large, wide-field, spectroscopic surveys (e.g., SDSS: [York et al. 2000](#), 2dFGRS: [Colless et al. 2001](#), DEEP2: [Davis et al. 2003](#), PRIMUS: [Coil et al. 2011](#), GAMA: [Driver et al. 2011](#)) in the last two decades have made it possible to investigate the clustering properties of galaxies as a function of different types (e.g., colors, luminosities, star formation rates, and stellar masses). For example, studies have found that red, passive galaxies are more clustered than blue, active galaxies (e.g., [Norberg et al. 2002](#); [Zehavi et al. 2005](#); [Coil et al. 2008](#); [Zehavi et al. 2011](#); [Guo et al. 2013](#)). In terms of stellar continuum luminosities (e.g.  $B$ -band luminosity), there is evidence for a luminosity-dependency with halo mass such that brighter galaxies tend to populate more massive halos (e.g., [Marulli et al. 2013](#); [Guo et al. 2014](#); [Harikane et al. 2016](#)).

There are a number of observational studies that have investigated the dependence of clustering strength/dark matter halo mass on stellar mass (e.g., [Meneux et al. 2008](#), [2009](#); [Wake et al. 2011](#); [Lin et al. 2012](#); [Mostek et al. 2013](#); [McCracken et al. 2015](#)). The connection between dark matter halo and stellar mass also forms the basis of abundance matching (e.g., [Behroozi et al. 2013b](#); [Skibba et al. 2015](#); [Harikane et al. 2016](#)). However, recent studies have shown this to be more complicated with the relation between the stellar mass and halo mass also being a function of other properties. For example, [Matthee et al. \(2017\)](#) used the hydrodynamical EAGLE simulation to investigate the scatter in the stellar-halo mass relation and came to the conclusion that either the scatter is mass dependent or it depends on more complex halo properties. [Contreras et al. \(2015\)](#) studied the galaxy-halo connection using two independent  $N$ -body simulations and found a monotonic increasing trend between halo mass and galaxy properties, such as stellar mass, although they find a considerable scatter for a given halo mass. A recent observational study by [Coil et al. \(2017\)](#) using the combined PRIMUS and DEEP2 surveys concluded that there is a wide range of stellar masses for a given halo mass and found that the relationship is also very much dependent on the specific star formation rate.

Other studies have also explored the dependencies on halo mass based on star-formation rates (SFRs) and specific SFRs (sSFRs). Recent measurements using  $H\alpha$  (tracing the instantaneous SFR) up to  $z \sim 2$  find that the clustering signal strongly increases with increasing line luminosity ([Sobral et al. 2010](#); [Stroe & Sobral 2015](#); [Cochrane et al. 2017](#)). Surprisingly, [Sobral et al. \(2010\)](#) found that the dependency is also redshift-independent in terms of  $L/L^*(z)$ ,

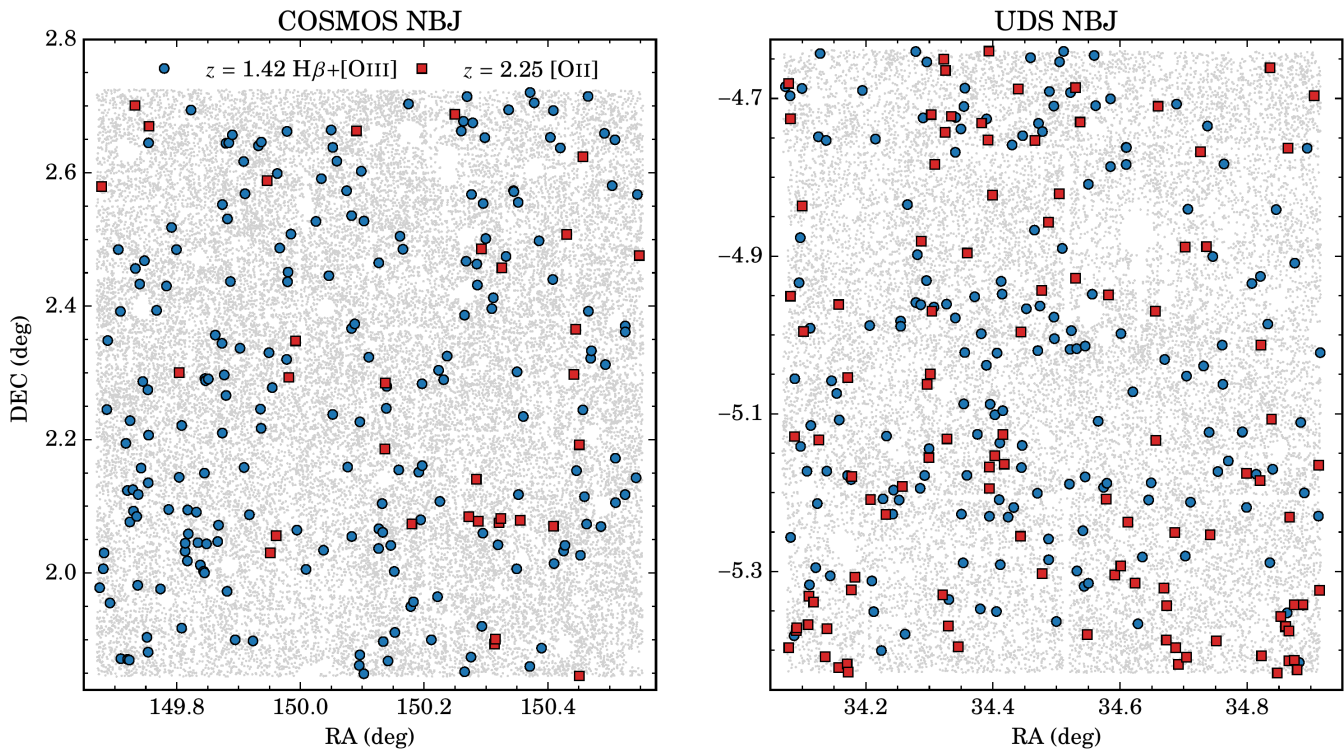
with  $L^*$  being the characteristic  $H\alpha$  luminosity at each redshift, equivalent to a characteristic SFR ( $SFR^*$ , [Sobral et al. 2014](#)). These studies also find that the trend may flatten for emitters with line luminosities  $\gtrsim L^*$  where emitters seem to reside in  $\sim 10^{13-13.5} M_\odot$  halos. This is consistent with the typical halo masses of AGN-selected samples ([Hickox et al. 2009](#); [Mendez et al. 2016](#)) with recent spectroscopic studies finding that the AGN fraction increases with line luminosity such that emission line-selected galaxies with  $L \gg L^*$  are primarily AGNs ([Sobral et al. 2016](#)). [Dolley et al. \(2014\)](#) used a  $24\mu\text{m}$ -selected sample between  $0.2 < z < 1.0$  and found a dependency between total infrared luminosity and halo mass. Using the DEEP2 samples, [Mostek et al. \(2013\)](#) found that the clustering amplitude for  $z \sim 1$  blue galaxies strongly increases with SFR and decreasing sSFR while the red population showed no significant correlation with SFR and sSFR.

The trends highlighted above are based on samples of the nearby Universe and a handful of  $z \sim 1-2$  studies. When and how these trends formed is important for our understanding of how halos and galaxies coevolve and also helps to constrain galaxy evolution models. In order to effectively study the clustering properties of galaxies, we require samples that are well-defined in terms of selection criteria, cover a range of redshifts to trace the evolving parameters over cosmic time, cover multiple and large comoving volumes to reduce the effects of cosmic variance, span a wide range in physical properties to properly subdivide the samples (e.g., line luminosity bins), and have known redshifts.

In this study, we use a sample of  $H\beta + [\text{OIII}]$  and  $[\text{OII}]$  emission line-selected galaxies from [Khostovan et al. \(2015\)](#) to study the clustering properties and dependencies with line luminosity and stellar mass up to  $z \sim 5$  in 4 narrow redshift slices per emission line. Since our samples are emission line-selected, this gives us the advantage of knowing the redshifts of our sources within  $\sigma_z = 0.01-0.03$  (based on the narrowband filter used) and forms a simple selection function, which is usually not the case with previous clustering studies using either broadband filters or spectroscopic surveys. Our samples are also large enough ( $\sim 7000$  sources) to properly subdivide to study the dependency of galaxy properties on the clustering strength and spread over the COSMOS and UDS fields ( $\sim 2 \text{ deg}^2$ ) to reduce the effects of cosmic variance.

This paper is structured as follows: in §2, we describe our samples and the mock random samples used in the clustering measurements. In §3 we present our methodology of measuring the angular correlation function, discuss the effects of contamination, describe how we corrected for cosmic variance, present our measurements of the spatial correlation function, and describe our model to convert the clustering length to minimum dark matter halo mass. In §4 we analyze the results for the full sample measurements in terms of the clustering length and halo masses. In §5 we look at the individual dependencies with halo mass starting with line luminosity and followed by stellar mass. We then show the dependency with halo mass in a line luminosity-stellar mass grid space. In §6 we present our interpretations of the results. We present our main conclusions in §7.

Throughout this paper we assume  $\Lambda$ CDM cosmology with  $H_0 = 70 \text{ km s}^{-1}$ ,  $\Omega_m = 0.3$ , and  $\Omega_\Lambda = 0.7$ . All stellar masses reported assume a Chabrier initial mass function.



**Figure 1.** The full COSMOS and UDS on-sky coverages with the NBJ filter. Shown in *blue circles* and *red squares* are the  $z = 1.42$   $H\beta+[OIII]$  and  $z = 2.25$   $[OII]$  emitters, respectively. The *grey dots* are all sources in the raw catalog used to select emission-line galaxies and clearly outline the masked regions which are associated with bright stars and artifacts. We refer the reader to Sobral et al. (2013) for a detailed description of how the masked regions were identified. The spatial distribution shown for both the  $H\beta+[OIII]$  and  $[OII]$  emitters already shows, visually and qualitatively, signatures of a non-random distribution. To properly quantify the clustering signal, we need to produce random samples that carefully take into account masked regions as outlined above.

## 2 SAMPLE

### 2.1 Emission-Line Galaxy Sample

In this study, we use the large sample of  $H\beta+[OIII]$  and  $[OII]$  selected emission-line galaxies from the narrow-band High- $z$  Emission Line Survey (HiZELS; Geach et al. 2008; Sobral et al. 2009, 2012, 2013) presented by Khostovan et al. (2015). Our samples are distributed over the COSMOS (Scoville et al. 2007) and UDS (Lawrence et al. 2007) fields with a combined areal coverage of  $\sim 2 \text{ deg}^2$  which equates to comoving volume coverages of  $\sim 10^6 \text{ Mpc}^3$ . The sample consists of 3475  $H\beta+[OIII]$  emitters at narrow redshift slices of  $z = 0.84, 1.42, 2.23,$  and  $3.24$  and 3298  $[OII]$  emitters at  $z = 1.47, 2.25, 3.34,$  and  $4.69$ . There are 223 and 219 spectroscopically confirmed  $H\beta+[OIII]$  and  $[OII]$  emitters, respectively, drawn from the UDSz Survey (Bradshaw et al. 2013; McLure et al. 2013), Subaru-FMOS measurements (Stott et al. 2013), Keck/DEIMOS and MOSFIRE measurements (Nayyeri et al., in prep), PRIMUS MUlti-object Survey (PRIMUS; Coil et al. 2011), and VIMOS Public Extragalactic Redshift Survey (VIPERS; Garilli et al. 2014). Recent Keck/MOSFIRE measurements of  $z = 1.47 - 3.34$  emitters are also included as well as recent VLT/VIMOS measurements for UDS sources (Khostovan et al., in prep).

The selection criteria used is explained in detail in Khostovan et al. (2015). In brief,  $H\beta+[OIII]$  and  $[OII]$  emitters are selected based on a combination of spectroscopic

measurements, photometric redshifts, and color-color selections (in order of priority) from the HiZELS narrowband color excess catalog of Sobral et al. (2013). Sources that have detections in multiple narrowband filters were also included in the final sample as the multiple emission line detections are equivalent to spectroscopic confirmation (e.g., the detection of  $[OII]$  in NB921 and  $H\alpha$  in NBH, see Sobral et al. 2012;  $[OIII]$  in NBH and  $H\alpha$  in NBK, Suzuki et al. 2016; see also Matthee et al. 2016 and Sobral et al. 2017 for dual NB-detections of  $Ly\alpha$  and  $H\alpha$  emitters at  $z = 2.23$ ).

Stellar masses of the sample were measured by Khostovan et al. (2016) using the SED fitting code of MAGPHYS (da Cunha et al. 2008), which works by balancing the stellar and dust components (e.g., the amount of attenuated stellar radiation is accounted for in the infrared). The level of AGN contamination was assessed by Khostovan et al. (2015) to be on the order of  $\sim 10 - 20\%$  using the  $1.6\mu\text{m}$  bump as a proxy via the color excesses in the *Spitzer* IRAC bands. Overall, the sample covers a wide range in physical properties with stellar masses between  $10^{8-11.5} M_{\odot}$ ,  $EW_{\text{rest}}$  between  $10 - 10000 \text{ \AA}$ , and line luminosities between  $10^{40.5-43.0} \text{ erg s}^{-1}$ , providing a wealth of different types of “active” galaxies (star-forming + AGN; Khostovan et al. 2016). This is important when investigating the connection between physical and clustering properties of galaxies.

A unique advantage of narrowband surveys in terms of clustering studies is knowing the redshift distribution of each line (emission line-selected) which removes any redshift

projections. Figure 1 shows the spatial distribution of the NBJ samples ( $\text{H}\beta + [\text{OIII}]$   $z = 1.42$  and  $[\text{OII}]$   $z \sim 2.25$ ) where, visually, it is clear that sources in both samples have a non-random, spatial clustering.

## 2.2 Random Sample

When looking for a clustering signal, an equivalent and consistent random catalog is required to test for a non-random spatial distribution within the sample. If all the sources within the sample are consistent with a random spatial distribution, then no spatial correlation would exist within the errors. Therefore, the methodology of creating the random sample has to be consistent with the real dataset in terms of depth, survey geometry, and masked regions (see Figure 1).

We create our random samples on an image-by-image basis in order to take into account the different survey depths.<sup>1</sup>As we also want to investigate the dependency with line luminosity and stellar mass (see §5), we populate each image using the line luminosity functions of [Khostovan et al. \(2015\)](#). For each image, we calculate the total effective area which takes into account the masked areas. We then integrate the [Khostovan et al. \(2015\)](#) luminosity functions down to the  $3\sigma$  detection limit of each image to calculate the total number of sources expected within the image area. This is then rescaled up by a factor of  $10^5$  such that each random sample generated has a total of  $\sim 10^6$  mock sources for each field. Figure 1 shows the masked regions of the NBJ images for both the COSMOS and UDS fields that are taken into account when generating the random samples.

## 3 MEASURING THE CLUSTERING OF $\text{H}\beta + [\text{OIII}]$ AND $[\text{OII}]$ EMITTERS

### 3.1 Angular Correlation Function

The two-point angular correlation function (ACF;  $w(\theta)$ ) is defined as:

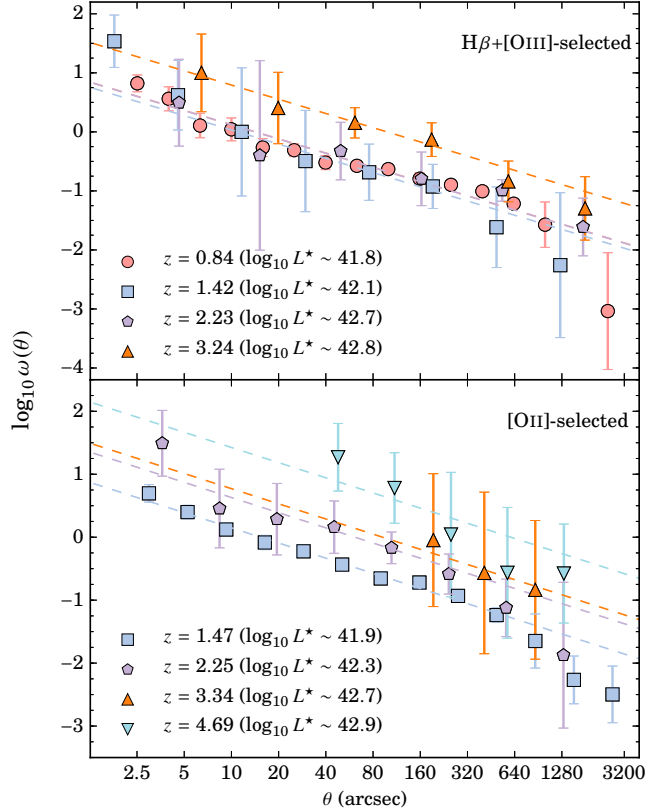
$$dP_{12} = N^2[1 + w(\theta)]d\Omega_1 d\Omega_2 \quad (1)$$

where  $P_{12}$  is the excess probability of finding two galaxies (galaxy 1 and galaxy 2) within a solid angle,  $\Omega$ , at a given angular separation,  $\theta$ , and with a mean number density  $N$ . Galaxies are randomly distributed for the case of  $w(\theta) = 0$  while a non-zero  $w(\theta)$  corresponds to a non-random distribution. We use the [Landy & Szalay \(1993, LS\)](#) estimator to measure the two-point angular correlation function as it has been shown to be the most reliable and has the best edge corrections when compared to other major estimators ([Kerscher et al. 2000](#)). The LS estimator is defined as:

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

where  $w(\theta)$  is the angular correlation function,  $DD$  is the number of data-data pairs,  $RR$  is the number of random-random pairs,  $DR$  is the number of data-random pairs,  $\theta$  is

<sup>1</sup> Refer to Table 2 of [Sobral et al. \(2013\)](#) for information regarding the depth of each image.



**Figure 2.** The angular correlation function based on the median of all the 2000 realizations per sample with the corresponding Limber approximation fits. All the data points are calculated using the LS estimator. The fits shown are constrained to angular separations for which the ACF is best described as a power law with slope,  $\beta = -0.8$ . There is evidence for an evolution in the clustering amplitude, but we stress the point that the clustering signal is sensitive to the range of physical properties (e.g., luminosity and stellar mass), which we explore in §5.

the angular separation, and  $N_R$  and  $N_D$  are the total number of random and data sources, respectively. The error associated with the LS estimator is defined as:

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

which assumes Poisson error.

Due to our small sample sizes in comparison to other clustering studies (e.g., SDSS), binning effects could introduce uncertainties in measuring the ACFs. This is basically a signal-to-noise problem where due to the small sample sizes, the way one bins can affect the measured data-data and data-random pairs. For example, bin sizes that are too small will result in bins of data-data pairs (signal) that are not sufficiently populated such that the random-random pairs (noise) will dominate the measured  $w(\theta)$ .

To take this into account, we measure the ACF 2000 times assuming Poisson errors as described in Equation 3 with varying bin centers and sizes. For each ACF, we apply a random bin size ( $\Delta \log \theta = 0.05 - 0.25$  dex) with  $\theta_{\min} = 1.0''$  to  $5.0''$  (randomly selected per ACF) and  $\theta_{\max} = 3100''$ . Each realization draws 10 - 100 times the number of real sources from the random sample discussed in Section 2.2

and the number of data-data, random-random, and data-random pairs are measured. We then fit a power law of the form:

$$w(\theta) = A_w \left( \theta^\beta - IC \right)$$

$$IC = \frac{\sum RR \theta^\beta}{\sum RR} \quad (4)$$

with  $A_w$  as the clustering amplitude and  $\beta$  as the power-law slope. The second equation is the integral constraint (IC; Roche et al. 2002) that takes into account the limited survey area. We note that the integral constraint has a marginal effect on our measurements of  $r_0$  as HiZELS coverage is  $> 1 \text{ deg}^2$ . The final measurements and errors for  $A_w$  and the clustering length ( $r_0$ ; see §3.5) are based on the distributions of values from the 2000 ACFs. In this way, we take into account the effects associated with binning.

Table 1 shows our  $A_w$  and  $\beta$  measurements. We find that our measurements are reasonably consistent (within  $\sim 1\sigma$ ) with  $\beta \sim -0.80$ . We also fit Equation 4 with a fixed  $\beta = -0.80$  (fiducial value in clustering studies) and use these measurements throughout the rest of the paper.

Figure 2 shows the median  $w(\theta)$  for the 2000 realizations and the fits for the best-fitted  $A_w$ . We find signs of the 1-halo term (small-scale clustering/contribution of satellite galaxies) at angular separations  $< 20''$  ( $\sim 150 \text{ kpc}$ ) for the  $z = 0.84$  H $\beta$ + [OIII] sample. This is the deepest of all the H $\beta$ + [OIII] samples and probably includes faint, dwarf-like systems that can be potential satellites (the sample includes sources with stellar masses down to  $10^{8.5} M_\odot$ ). The deviation from the power law fit seen for the lowest angular separation bin in the  $z = 1.42$  H $\beta$ + [OIII] correlation function is consistent with the 1-halo term, but this is quite weak (within  $1\sigma$  deviation). We find no significant detection of the 1-halo term in the [OII] samples. One possible cause for the 1-halo term is the presence of large overdense regions that can increase the satellite fraction. For example, there is a  $\sim 10 \text{ Mpc}$ -scale structure at  $z = 0.84$  that contains several X-ray confirmed clusters/groups and large filaments within the COSMOS field (e.g., Sobral et al. 2011, Darvish et al. 2014) but we defer from a detailed analysis of the satellite fractions as it is beyond the scope of this work.

### 3.2 Bootstrapping or Poisson Errors?

There are three main error estimators that are typically employed in clustering studies: bootstrapping, jackknifing, and Poisson. In the case that Poisson errors are assumed (as is the case with this study), then the errors are defined as shown in equation 3. Norberg et al. (2009) studied these three estimators to see how reliably each measures the ‘true’ errors of the ACFs. They found that bootstrapping overestimates the errors by  $\sim 40$  percent and jackknifing fails at small-scales but can reproduce the errors at large-scales, while Poisson errors were found to underestimate the errors.

One characteristic of the results of Norberg et al. (2009) is that the sample size used in their simulations is comparable to that of SDSS ( $10^{5-6}$  sources). Since Poisson errors become significantly smaller for larger sample sizes, it then would become apparent that Poisson errors could severely underestimate the ‘true’ errors of the ACFs. This may not

be entirely true for our sample sizes, which are typically between  $10^{2-3}$  sources. We test this by using our  $z = 3.24$  H $\beta$ + [OIII] sample (179 sources) for which the bin size and centers were fixed and calculated the ACFs assuming Poisson errors and also bootstrapping with 2000 realizations. We find that the errors on average are similar such that Poisson errors for small sample sizes are comparable to bootstrapping errors. Note that, as described in §3, we assume Poisson errors for each individual ACF but also take into account binning effects by repeating our measurements of the ACF with varying bin sizes and centers such that our final measurements are based on the distributions of these realizations.

### 3.3 Effects of Contamination

The issue of contamination can be marginal or quite significant and is based on many factors such as the sample selection. Clustering studies typically consider the contaminants in a sample to be randomly distributed, such that the clustering amplitude is underestimated by a factor of  $(1-f)^2$ , with  $f$  being the contamination fraction. For the clustering length,  $r_0$ , this results in an underestimation by a factor of  $(1-f)^{2/|\beta|}$ .

The level of contamination was briefly investigated in Khostovan et al. (2015) and was found to be on the order of  $\sim 10$  percent for the lowest redshift samples. This would result in a 23 percent increase in  $A_w$  and a 12 percent increase in  $r_0$ . Note that this assumes that the contaminants are randomly distributed and, hence, lowers the clustering strength, which may not be true for narrowband surveys. For our samples, contaminants could be due to galaxies with misidentified emission lines. For example, a source at  $z = 1.47$  that is misidentified as [OII] in the NB921 filter could actually be a  $z = 0.84$  [OIII] emitter or a  $z = 0.40$  H $\alpha$  emitter. Because galaxies selected by nebular emission lines are shown to be clustered as well (see below and Sobral et al. 2010 and Cochrane et al. 2017 for H $\alpha$ ), the effects could possibly be negligible and not follow the typical  $(1-f)^2$  correction factor. Therefore, we do not correct our measurements due to contamination.

### 3.4 Cosmic Variance

Cosmic variance can greatly affect the clustering measurements. If the areal coverage is small ( $\lesssim \text{arcmin}^2$  scales), then the measured clustering amplitude and subsequent results can vary considerably, especially if the region probed is a significant overdense region or a void. Therefore, it is important that the clustering measurements are done on large fields ( $\gtrsim 1 \text{ deg}^2$ ).

Sobral et al. (2010) measured the effects of cosmic variance for the HiZELS H $\alpha$   $z = 0.84$  sample (734 emitters) on the clustering amplitude. This was done by measuring  $A_w$  (fixed  $\beta = -0.80$ ) for randomly sized regions between  $0.05 \text{ deg}^2$  to  $0.5 \text{ deg}^2$  with the larger areas randomly sampled 100 times ( $0.3 - 0.5 \text{ deg}^2$ ) and the smaller areas randomly sampled 1000 times. We refer the reader to Figure 3 of Sobral et al. (2010) where they show that the uncertainty in  $A_w$  (in percentage) is related to the area covered and is best fit with a power-law of the form  $20 \times \Omega^{-0.35}$ , with  $\Omega$  representing the area in units of  $\text{deg}^2$ .

We note that the HiZELS coverage at the time of Sobral et al. (2010) was only  $1.3 \text{ deg}^2$  in the COSMOS field and used only  $J$ -band coverage. In this paper we are using the current HiZELS coverage (all four narrowband filters in  $zJHK$ ) which includes both the COSMOS and UDS fields for a combined areal coverage of  $\sim 2 \text{ deg}^2$  (Sobral et al. 2013). This corresponds to a decreased uncertainty of  $\sim 16\%$  due to cosmic variance in the measurement of  $A_w$ . We incorporate this uncertainty by adding  $\sim 16\%$  of  $A_w$  in quadrature to the error from the fit. For the clustering length,  $r_0$ , we propagate the error from  $A_w$  and find that the error in  $r_0$  is increased by  $\sim 11\%$ .

### 3.5 Real-Space Correlation

The two-point (real-space) correlation function is a useful tool in measuring the physical clustering of galaxies and is best described, empirically, by  $\xi = (r/r_0)^\gamma$ , with  $r_0$  being the clustering length. One key requirement in measuring the two-point correlation function is the redshift distribution of the sample. The benefit of narrowband surveys is that the redshift distribution of the sample is easily derived from the narrowband filter profile (e.g., Sobral et al. 2010; Stroe & Sobral 2015) such that it is equivalent to taking a narrow redshift slice of  $\sigma_z \sim 0.01 - 0.03$  (depending on the central redshift; see Table 2 in Khostovan et al. 2015).

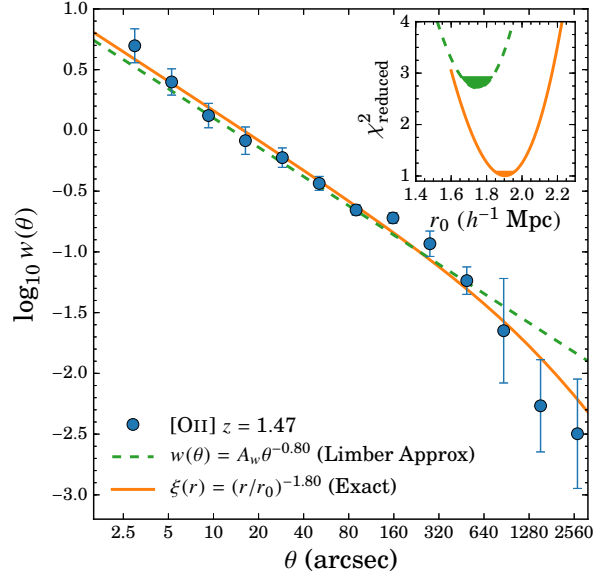
Traditionally, the Limber approximation (Limber 1953) is used to relate the real-space correlation to the angular correlation function. Simon (2007) showed that the approximation works for surveys that use broad filters and for small angular separations but fails for narrow filters and large angular separations. They find that for large angular separations and very narrow filters,  $\omega(\theta)$  becomes a rescaled version of  $\xi(r)$  where the slope of  $w(\theta)$  changes from  $\gamma + 1$  to  $\gamma$ . Sobral et al. (2010) used the exact Limber equation proposed by Simon (2007) and found that, for their sample of  $z = 0.84$  H $\alpha$  emitters, the break-down in the Limber approximation occurs at angular separations  $\sim 600''$  with an  $r_0 = 2.6 \pm 0.3 h^{-1} \text{ Mpc}$  measured from the approximation and  $r_0 = 2.7 \pm 0.3 h^{-1} \text{ Mpc}$  from the exact equation.

We adopt the exact equation presented by Simon (2007) and used by Sobral et al. (2010) to relate the real-space and angular correlation functions and calculate  $r_0$ . The relation is described as:

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

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

where  $p$  is the filter profile in radial comoving distance, which is written as the mean spatial position of two sources,  $r_1$  and  $r_2$ , such that  $\bar{r} = (r_1 + r_2)/2$  with  $R$  being the distance between the two sources using the law of cosines. We refer the reader to Simon (2007) for a detailed description regarding the derivation of this equation. The filter profile, which traces the underlying redshift distribution of the sample, is assumed to be a Gaussian function. We fit for the true filter profile based on the transmission curves of the actual narrowband filters. Table A1 shows a comparison between the



**Figure 3.** The angular correlation function for the  $z = 1.47$  [OII] sample. Shown are the observed  $w(\theta)$  measurements as in Figure 2 with the corresponding Limber approximation and exact Limber equation fits. We use the full range of angular separations for both fits, even though the Limber approximation is found to fail at  $\theta \sim 500''$ . The exact equation results in a reduced  $\chi^2 \approx 1$  compared to  $\approx 2.8$  when using the Limber approximation. The clustering lengths are  $r_{0,\text{exact}} = 1.90 \pm 0.21$  compared to  $r_{0,\text{limber}} = 1.75 \pm 0.21$  (errors corrected for cosmic variance). The errors shown in the  $\chi^2$  distribution are only based on the fits. The results shown here signify the importance of the exact Limber equation when using narrowband samples for large angular separations.

properties of the Gaussian and true filters in terms of redshifts. The power law slope of the spatial correlation function is also shown in Equation 5 and is assumed to be  $\gamma = -1.8$  ( $\gamma = \beta - 1$ ). We use Equation 5 to fit  $r_0$  to our measurements of  $w(\theta)$ .

Figure 3 shows the comparison between the Limber approximation (assuming a single power law to describe  $w(\theta)$  as shown in Equation 4) and the exact Limber equation as described in Equation 5 for the  $z = 1.47$  [OII] sample. We find that the Limber approximation breaks down at angular separations of  $\sim 500''$ . As discussed in Simon (2007) and in Appendix A, the point for where the Limber approximation fails is dependent on the filter width (the width of the redshift distribution) and the transverse distance (central redshift).

Also shown on Figure 3 is the reduced  $\chi^2$  measurements of the fits. We find that the exact equation has a reduced  $\chi^2$  of  $\approx 1$  in comparison to 2.8 for the Limber approximation-based fit with  $r_{0,\text{exact}} = 1.90 \pm 0.21 h^{-1} \text{ Mpc}$  compared to  $r_{0,\text{limber}} = 1.75 \pm 0.21 h^{-1} \text{ Mpc}$  (errors include cosmic variance contribution; see §3.4). Although both methods produce measurements that are consistent within  $1\sigma$  (errors dominated by cosmic variance), our results shown on Figure 3 highlights the importance of using the exact Limber equation to measure the clustering length since it can compensate for the rescaling of the ACF due to the effects of using nar-

**Table 1.** The clustering properties for our  $H\beta+[OIII]$  and  $[OII]$  samples. The power-law slope,  $\beta$ , in the ACF is shown and corresponds to the clustering amplitude,  $A_{w,free}$ , which corresponds to when  $\beta$  is a free-parameter in the fit. All other measurements shown have  $\beta$  fixed to  $-0.8$ , which corresponds to  $\gamma = -1.8$  in the real-space two-point correlation function.  $r_{0,exact}$  is the clustering length measured using the exact Limber equation as defined in Equation 5. Dark matter halo masses are measured using our  $r_0$ -halo mass models.

Clustering Properties for Full Sample						
$z$	$N_D$	$\beta$	$A_{w,free}$ (arcsec)	$A_{w,\beta=-0.8}$ (arcsec)	$r_{0,exact}$ (Mpc $h^{-1}$ )	$\log_{10} M_{min}$ ( $M_{\odot} h^{-1}$ )
$H\beta+[OIII]$ Emitters						
0.84	2477	$-0.69^{+0.03}_{-0.03}$	$5.19^{+1.32}_{-1.22}$	$11.53^{+2.33}_{-2.33}$	$1.71^{+0.19}_{-0.19}$	$11.18^{+0.33}_{-0.33}$
1.42	371	$-0.79^{+0.07}_{-0.04}$	$7.47^{+3.58}_{-3.24}$	$8.32^{+2.18}_{-2.08}$	$1.45^{+0.20}_{-0.20}$	$10.70^{+0.40}_{-0.40}$
2.23	270	$-0.81^{+0.15}_{-0.12}$	$11.10^{+12.42}_{-6.57}$	$10.42^{+2.80}_{-2.62}$	$2.43^{+0.31}_{-0.31}$	$11.61^{+0.22}_{-0.22}$
3.24	179	$-0.78^{+0.04}_{-0.03}$	$42.28^{+13.22}_{-13.56}$	$48.70^{+10.71}_{-10.83}$	$4.01^{+0.49}_{-0.49}$	$12.08^{+0.17}_{-0.17}$
$[OII]$ Emitters						
1.47	3285	$-0.83^{+0.02}_{-0.04}$	$10.06^{+2.66}_{-2.21}$	$11.61^{+2.34}_{-2.34}$	$1.99^{+0.22}_{-0.22}$	$11.46^{+0.23}_{-0.24}$
2.25	137	$-0.78^{+0.05}_{-0.03}$	$25.51^{+9.08}_{-9.18}$	$29.99^{+7.24}_{-7.00}$	$3.14^{+0.43}_{-0.41}$	$12.03^{+0.21}_{-0.20}$
3.34	35	$-0.79^{+0.23}_{-0.06}$	$53.67^{+41.66}_{-44.95}$	$57.49^{+22.49}_{-24.67}$	$5.06^{+1.08}_{-0.94}$	$12.37^{+0.28}_{-0.24}$
4.69	18	$-0.83^{+0.04}_{-0.04}$	$208.50^{+116.82}_{-91.58}$	$139.44^{+53.69}_{-44.63}$	$8.25^{+1.54}_{-1.44}$	$12.62^{+0.22}_{-0.20}$

rowband filters. Throughout the rest of this paper, we refer to  $r_0$  as the clustering length measured using Equation 5.

### 3.6 Clustering Length to Dark Matter Halo Mass

Our theoretical understanding of galaxy formation is that galaxies form with the assistance of the gravitational potentials of dark matter halos such that all galaxies reside in a halo. In effect, the spatial clustering of galaxies is then related to the clustering of dark matter. Matarrese et al. (1997) and Moscardini et al. (1998) used this link between galaxies and dark matter halos to predict the clustering length of a sample for a given minimum dark matter halo mass and redshift. In this section, we use the same methodology used to generate their predictions, but update to the latest cosmological prescriptions.

We first begin by measuring the matter-matter spatial correlation function using a suite of cosmological codes called Colossus (Diemer & Kravtsov 2015). This is calculated by taking the Fourier transform of the matter power spectrum, assuming an Eisenstein & Hu (1998) transfer function. We then calculate the effective bias by using the following equation:

$$b_{eff}(z) = \frac{\int_{M_{min}}^{\infty} b_h(M, z)n(M, z)dM}{\int_{M_{min}}^{\infty} n(M, z)dM} \quad (6)$$

where  $b_h(M, z)$  and  $n(M, z)$  are the halo bias and mass functions, respectively. The effective bias is defined as the integrated halo bias and mass functions above some minimum dark matter halo mass,  $M_{min}$ , and normalized to the number density of halos. We then relate the effective bias to the spatial correlation of galaxies by:

$$b_{eff}^2 = \xi_{gg}/\xi_{mm} \quad (7)$$

with  $\xi_{gg}$  and  $\xi_{mm}$  being the galaxy-galaxy and matter-matter spatial correlation functions, respectively.

We use the Tinker et al. (2010) halo bias prescription and the Tinker et al. (2008) halo mass function. The previous predictions of Matarrese et al. (1997) and Moscardini et al. (1998) used the Press & Schechter (1974)

halo mass function and Mo & White (1996) halo bias functions. Their assumed  $\Lambda$ CDM cosmology was also different ( $H_0 = 65 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.4$ , and  $\Omega_{\Lambda} = 0.6$ ) than the current measurements. We present a discussion regarding the uncertainties of assuming a bias and mass function in Appendix B.

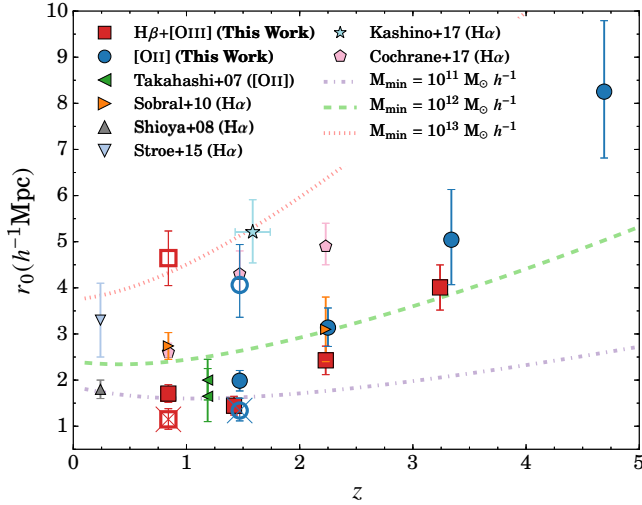
Note that our approach is very much similar to the methodology used in halo occupation distribution (HOD) modeling (e.g., Kravtsov et al. 2004). In comparison to the framework of HOD, we are assuming that all galaxies are centrals (only one galaxy occupies each host halo) and reside in halos with mass  $\geq M_{min}$ . This is an oversimplification in comparison to typical HOD models where we have only one free parameter (minimum dark matter halo mass), but we note that HOD modeling typically employs 3 - 5 free parameters (e.g., Kravtsov et al. 2004; Zheng et al. 2005) with even more complex models incorporating 8 free parameters (e.g., Geach et al. 2012). We instead resort to using our one parameter approach but caution the reader that directly comparing our results with minimum halo masses is inconsistent. Any study from the literature that is used to compare with our results in this paper have their minimum halo masses computed using their  $r_0$  measurements and our  $r_0$ -halo mass model.

## 4 CLUSTERING OF $H\beta+[OIII]$ AND $[OII]$ EMITTERS

Figure 4 shows the evolution of  $r_0$  for  $H\beta+[OIII]$  and  $[OII]$  emitters up to  $z \sim 3$  and  $\sim 5$ , respectively. These are the first measurements of the clustering length for  $H\beta+[OIII]$  and  $[OII]$  emission-line galaxies to be reported. Included are the  $r_0$  predictions for dark matter halos with minimum masses between  $10^{11} - 10^{13} M_{\odot}$  based on our model described in §3.6.

We find that, based on the full population of emitters in each sample,  $H\beta+[OIII]$  emitters tend to reside in  $\sim 10^{10.7} - 10^{12.1} M_{\odot}$  dark matter halos while the  $[OII]$  emitters are found to vary less with  $\sim 10^{11.5} M_{\odot}$  at  $z = 1.47$  to  $\sim 10^{12.6} M_{\odot}$  at  $z = 4.69$ , although these are driven by selection effects (e.g., highest redshift sample will be bi-





**Figure 4.** Shown is the evolution of the clustering length up to  $z \sim 5$ . Included are the predicted clustering lengths for minimum dark matter halo masses between  $10^{11-13} M_{\odot}$ . Although there is a clear sign of a redshift evolution in  $r_0$ , we stress the point that this is due to selection bias such that these measurements are sensitive to the range of physical properties, such as line luminosity. As a demonstration, we overlay the brightest (*open symbol*) and faintest (*open symbol with a cross*) line luminosity bins (see Table 2) with the symbol type and color consistent with that used for the full sample measurement. The brightest emitters are found to have  $r_0$  measurements  $\sim 2-3$  times that of the full sample and the faintest emitters with  $\sim 50\%$  lower  $r_0$  values.

ased towards higher line luminosities which, as shown in §5.1, leads to higher  $r_0$ ). In comparison to each other, all overlapping samples, except for the  $z = 1.47$  samples, have similar  $r_0$  measurements within  $1\sigma$  error bars. This then suggests that  $H\beta+[OIII]$ - and  $[OII]$ -selected galaxies reside in dark matter halos with similar masses. Included in Figure 4 are the  $H\alpha$  measurements of Shioya et al. (2008), Sobral et al. (2010), Stroe & Sobral (2015), Cochrane et al. (2017), and Kashino et al. (2017). The Sobral et al. (2010) measurement at  $z = 2.23$  is consistent with that of the  $H\beta+[OIII]$  and  $[OII]$  samples at the same redshift, suggesting that  $H\beta+[OIII]$ - and  $[OII]$ -selected emitters reside in dark matter halos with similar masses as  $H\alpha$ -selected emitters and can be tracing a similar underlying population of star-forming/active galaxies. We also include the  $z \sim 1.2$   $[OII]$  measurements of Takahashi et al. (2007). Although our closest sample in terms of redshift is at  $z = 1.47$ , we find that our measurements are in agreement.

Despite the agreement between  $H\alpha$ ,  $H\beta+[OIII]$ , and  $[OII]$  samples, we note that such a comparison is not entirely fair. An example is the  $H\alpha$  measurement of Stroe & Sobral (2015) and Shioya et al. (2008). Both cover the same redshift range of  $z = 0.24$ , but the Shioya et al. (2008) has a depth of  $\sim 10^{39.5} \text{ erg s}^{-1}$  in  $L_{H\alpha}$  while the Stroe & Sobral (2015) depth is  $\sim 10^{41.0} \text{ erg s}^{-1}$  and covers significantly larger volumes. This results in a factor of two difference in the  $r_0$  measured and almost two orders of magnitude difference in the minimum dark matter halo mass by these two studies which arises from the dependency of the clustering length with line luminosity (see §5.1).

As a demonstration of this same feature, we show

$r_0$  of the brightest (open symbols) and faintest (open symbols with a cross) galaxies in our  $H\beta+[OIII]$   $z = 0.84$  and  $[OII]$   $z = 1.47$  samples. We find that the most luminous (faintest) galaxies have higher (lower) clustering lengths relative to the full sample measurement. This suggests a line luminosity dependency not just in the  $H\alpha$  measurements, but also in the  $H\beta+[OIII]$  and  $[OII]$  measurements. Therefore, any comparison, as shown in Figure 4, needs to be interpreted with caution as each measurement for a full sample will be dependent on how wide a range of line luminosities is covered. For example, the  $r_0$  measured for the  $z = 4.69$   $[OII]$  sample is biased towards higher  $r_0$  values since the sample is biased towards the brightest  $[OII]$  emitters. To investigate the redshift evolution of the clustering and dark matter halo properties of galaxies, we need to then study its dependencies.

## 5 DEPENDENCIES BETWEEN GALAXY PROPERTIES AND DARK MATTER HALO

In this section we present our results on how the clustering evolution of  $H\beta+[OIII]$  and  $[OII]$  emitters depends on line luminosities and stellar masses.

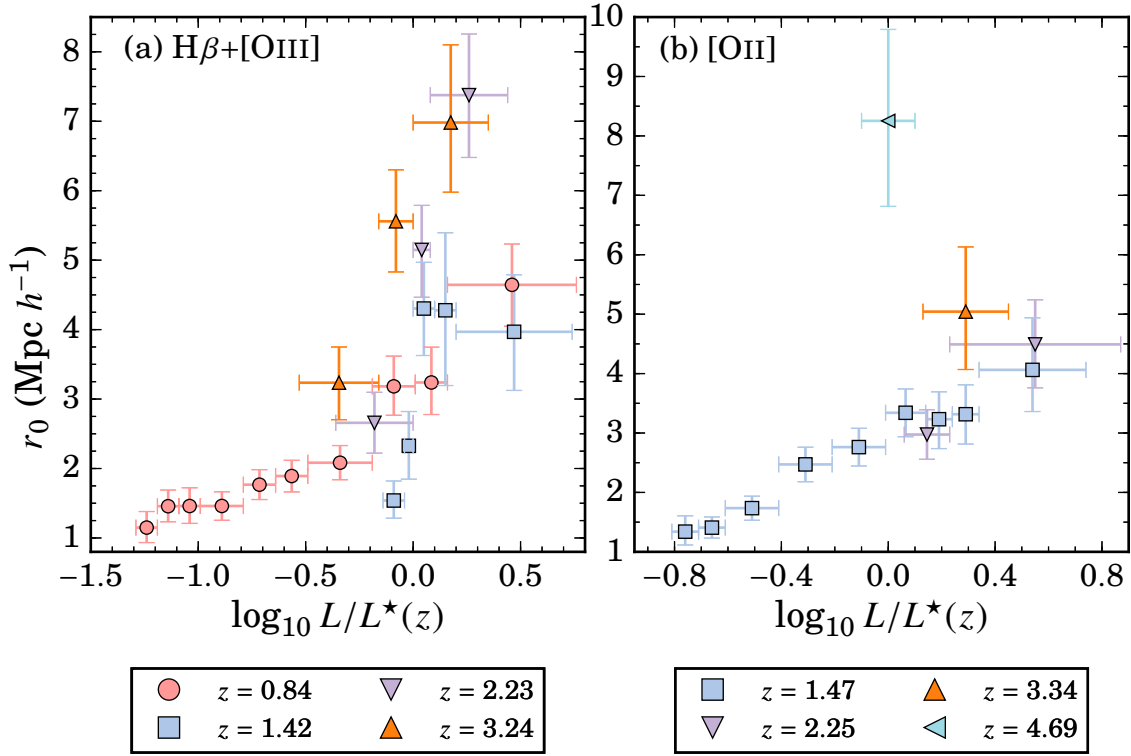
### 5.1 Observed Line Luminosity Dependency

As discussed in §4, the clustering properties of galaxies are tied to their physical properties such that an investigation of their dependencies is required to properly map out the clustering evolution and study the connection between dark matter halos and galaxies. In this section, we study how the clustering length is dependent on the observed line luminosities and link it to the dark matter halo properties.

Figure 5 shows the  $r_0$  dependency with line luminosity normalized by the characteristic line luminosity at the corresponding redshift,  $L/L^*(z)$ . The tabulated measurements are shown in Tables 2 and 3. The reason we show our measurements in terms of  $L/L^*(z)$  is so that we may investigate the clustering evolution of our samples independent of the cosmic evolution of the line luminosity functions. This was motivated by the results of Sobral et al. (2010) and Cochrane et al. (2017) for their  $H\alpha$  samples. Khostovan et al. (2015) showed that  $L^*(z)$  can evolve by a factor of  $\sim 11-12$  from  $z \sim 0.8-5$  for both  $H\beta+[OIII]$ - and  $[OII]$ -selected samples.

For each redshift slice, we find that  $r_0$  increases by a factor of  $\sim 2-4$  with increasing line luminosity. There is also a redshift evolution such that at a fixed  $L/L^*(z)$ ,  $r_0$  is increasing. For example, we find for our  $H\beta+[OIII]$  samples that the clustering length at  $L \sim L^*(z)$  is 3.2, 4.3, 5.2, and  $7.0 h^{-1} \text{ Mpc}$  at  $z = 0.84, 1.42, 2.23,$  and  $3.24$ , respectively, which corresponds to a factor of 2.2 increase in  $r_0$  within  $\sim 5$  Gyrs.

Our results suggest some redshift evolution in the clustering of galaxies as a function of line luminosity, but we must also take into account the intrinsic clustering evolution due to halos as shown in Figure 4. A reasonable way to assess if there is an evolution in the clustering properties is by investigating it in terms of halo masses and  $L/L^*(z)$  such that we take into account both the halo clustering (see Figure 4) and the line luminosity function evolutions. This



**Figure 5.** The clustering length measured in terms of  $L/L^*(z)$ . Studying the dependency of the clustering length with luminosity as a function of the ratio between line and characteristic luminosity removes the effects caused by the cosmic evolution in the luminosity functions. For each redshift slice we find that there is a strong correlation between the clustering length and  $L/L^*(z)$ . There is an evolution in the clustering length such that  $r_0$  increases with redshift at any given  $L/L^*(z)$ . For example, the clustering lengths at  $L \sim L^*(z)$  are 3.2, 4.3, 5.2, and 7.0  $h^{-1}$  Mpc for our  $H\beta+[OIII]$  samples at  $z = 0.84, 1.42, 2.23,$  and  $3.24$ . The same strong, increasing trend between  $r_0$  and  $L/L^*(z)$  is also seen for the  $[OII]$  sample.

relation was first studied by Sobral et al. (2010) for  $H\alpha$  emitters up to  $z = 2.23$  where they reported a strong, redshift-independent trend between halo mass and  $L/L^*(z)$ . Here we investigate if such a relation exists for our  $H\beta+[OIII]$  and  $[OII]$  emitters to even higher redshifts.

Figure 6 shows the line luminosity dependence on minimum dark matter halo masses (measured using our  $r_0$ -halo mass models as described in §3.6) with the measurements highlighted in Tables 2 and 3. We find that there is a strong relationship between line luminosity and halo mass for all redshift samples. More interestingly, we find no significant redshift evolution in the minimum dark matter halo mass such that galaxies reside in halos with similar masses independent of redshift at fixed  $L/L^*(z)$ . This is found for both  $H\beta+[OIII]$  and  $[OII]$ , as well as  $H\alpha$  studies (Geach et al. 2008; Shioya et al. 2008; Sobral et al. 2010; Cochrane et al. 2017) as shown in the bottom panel of Figure 6.

We quantify the observed trends by fitting both single and piecewise power laws to all measurements at all redshifts. The piecewise power laws are used in order to test the significance of a possible flattening of the observed, increasing trends for  $L > L^*(z)$ . Our single power law fits are:

$$M_{\min} = \begin{cases} 10^{12.48 \pm 0.07} \left(\frac{L}{L^*(z)}\right)^{1.77 \pm 0.21} & H\beta+[OIII] \\ 10^{12.87 \pm 0.06} \left(\frac{L}{L^*(z)}\right)^{1.17 \pm 0.14} & H\alpha \end{cases} \quad (8)$$

where we only show the measurements for  $H\beta+[OIII]$  and

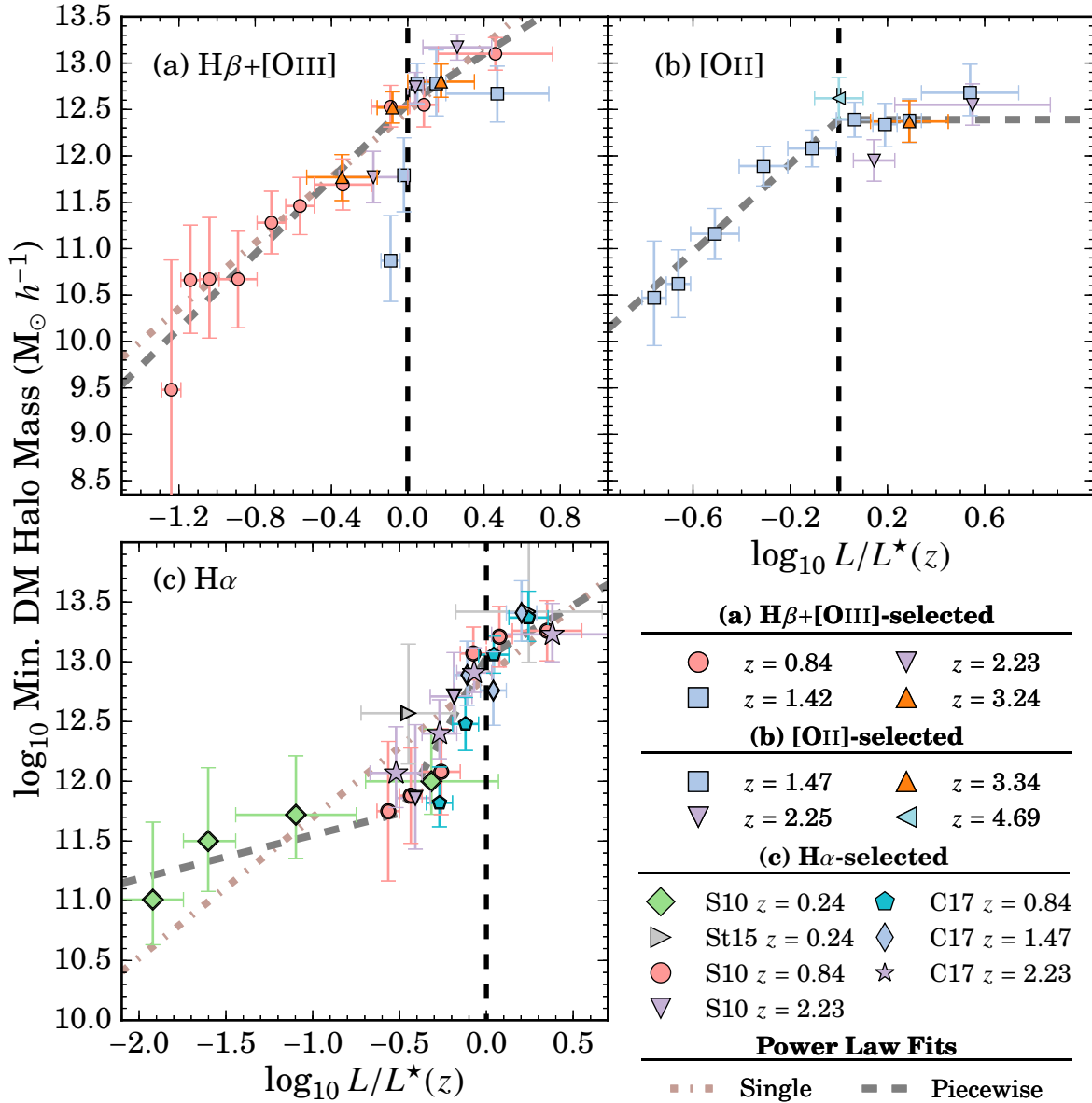
$H\alpha$  as the  $[OII]$  measurements show a clear deviation for  $L > L^*(z)$ . We find that the  $H\beta+[OIII]$  emitters show a steeper increasing trend in comparison to  $H\alpha$  but with a lower halo mass at  $L \sim L^*(z)$ .

Figure 6 shows a clear deviation from a single power law trend at  $L \sim L^*(z)$  for the  $[OII]$  samples. There is some signature of such a deviation in our  $H\beta+[OIII]$  and also the  $H\alpha$  samples from the literature where the slope of the trends becomes shallower. We fit piecewise power laws split at  $L \sim L^*(z)$  and find:

$$H\beta+[OIII] : \\ M_{\min} = 10^{12.56 \pm 0.11} \begin{cases} \left(\frac{L}{L^*(z)}\right)^{2.02 \pm 0.32} & L < L^* \\ \left(\frac{L}{L^*(z)}\right)^{1.35 \pm 0.47} & L > L^* \end{cases} \quad (9)$$

$$[OII] : \\ M_{\min} = 10^{12.39 \pm 0.08} \begin{cases} \left(\frac{L}{L^*(z)}\right)^{2.37 \pm 0.31} & L < L^* \\ \left(\frac{L}{L^*(z)}\right)^{0.003 \pm 0.003} & L > L^* \end{cases} \quad (10)$$

$$H\alpha : \\ M_{\min} = 10^{13.04 \pm 0.08} \begin{cases} \left(\frac{L}{L^*(z)}\right)^{0.36 \pm 0.20} & L < 0.3L^* \\ \left(\frac{L}{L^*(z)}\right)^{2.61 \pm 0.36} & 0.3L^* < L < L^* \\ \left(\frac{L}{L^*(z)}\right)^{0.87 \pm 0.43} & L > L^* \end{cases} \quad (11)$$

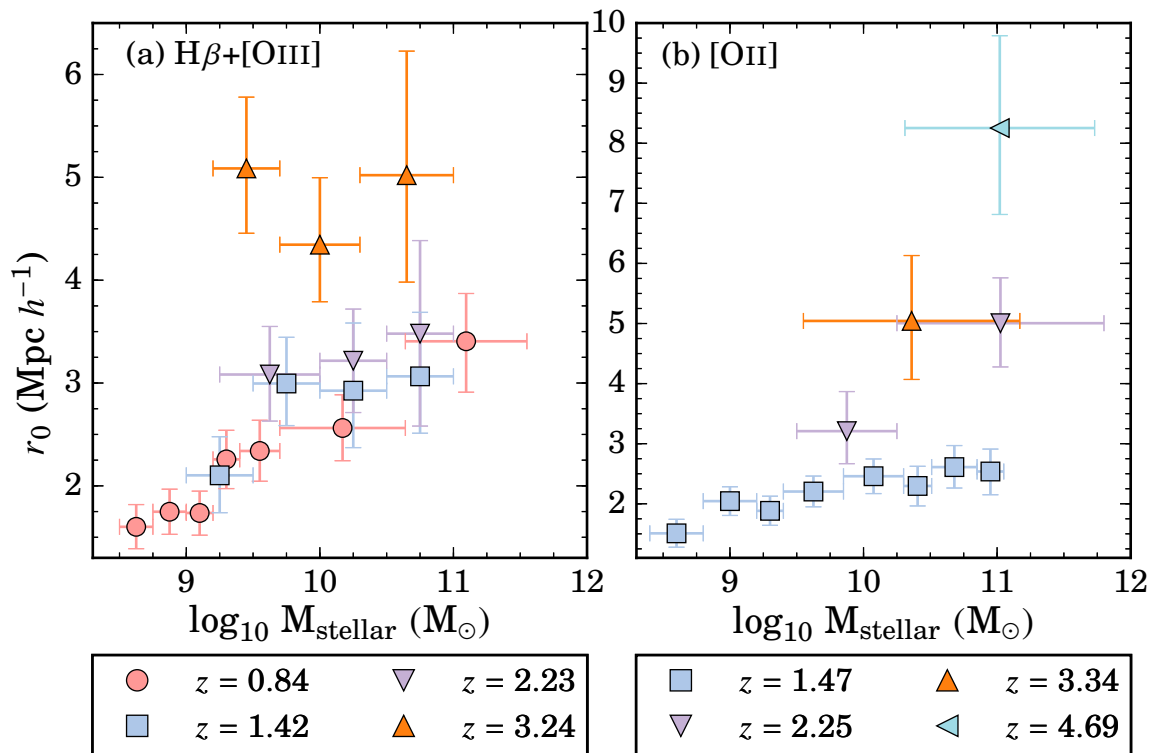


**Figure 6.** The dependency between  $L/L^*(z)$  versus minimum halo mass for our  $H\beta+[OIII]$  and  $[OII]$  samples. We find a strong correlation between line luminosity and dark matter halo mass and find no redshift evolution in  $L/L^*(z)$  such that galaxies at redshifts as high as  $z \sim 5$  for a given  $L/L^*(z)$  reside in halos of similar mass as galaxies at  $z \sim 1$ . As a comparison, we also include the  $H\alpha$  measurements at  $z = 0.24$  from Shioya et al. (2008) (recomputed by Sobral et al. (2010, S10)) and Stroe & Sobral (2015, St15),  $z = 0.84$  from Sobral et al. (2010), and  $z = 2.23$  from Geach et al. (2008) (recomputed by Sobral et al. (2010)). The latest  $H\alpha$  results of Cochrane et al. (2017, C17) are also included at  $z = 0.84, 1.47,$  and  $2.23$ . The consensus from  $H\alpha$  studies is a strong dependency between line luminosity and halo mass. For  $L > L^*$  emitters, we find a flat trend with halo mass consistent with  $10^{12.5} M_\odot$  for  $[OII]$  emitters and a shallower increasing trend for  $H\alpha$  and  $H\beta+[OIII]$  emitters, although the scatter in the measurements are  $\sim 0.5$  dex which can also be consistent with a flat trend.

where only the  $H\alpha$  measurements includes a second split at  $L \sim 0.3L^*$  which is only constrained by the  $z \sim 0.24$   $H\alpha$  measurements of Shioya et al. (2008). Therefore, we cannot state that the trend is redshift-independent below  $0.3L^*$  for  $H\alpha$ -selected emitters due to lack of measurements at different redshifts.

Equations 9 – 11 show a steep, increasing trend up to  $L \sim L^*$  followed by significantly shallower slopes beyond  $L^*$ . The  $H\beta+[OIII]$  fit shows the steepest slope of  $1.35 \pm 0.47$  beyond  $L^*$ , but we note that the spread in our halo mass measurements are quite large ( $\sim 0.7$  dex) such that a flat

slope can also be consistent with the measurements. The fits confirm a near constant halo mass for  $L > L^*(z)$  such that emission line-selected galaxies ( $H\alpha$ ,  $H\beta+[OIII]$ , and  $[OII]$ ) with different line luminosities  $> L^*$  reside in halos with similar masses regardless of redshift. This suggests that the mechanisms and processes causing this flattening of the line luminosity-halo mass relation is possibly the same in  $H\alpha$ ,  $H\beta+[OIII]$ , and  $[OII]$  emitters for all redshift slices probed. The flattening/shallower slope could also be due to the lower number density of  $10^{12.5-13.0} M_\odot$  halos given the comoving volume of our survey.



**Figure 7.** The clustering length as measured per stellar mass bins. We find that for both  $H\beta+[OIII]$  and  $[OII]$  emitters the clustering length increases with increasing stellar mass. Our results show that  $r_0$  also increases with redshift for a fixed stellar mass. In comparison to the line luminosity dependency, we find that the increasing trend with stellar mass is weaker but we note that this could be a result of the line luminosity dependency or vice versa. This is because for each stellar mass bin there is a wide range of line luminosities. We explore this inter-dependency in §5.3.

Our results also imply that there is a simple, redshift-independent relationship between the emission line luminosities of galaxies and their host halos once accounting for the evolution in  $L^*$  (Sobral et al. 2010). This has implications for theoretical studies that use photoionization codes along with semi-analytical modeling to study the connection between nebular emission lines and dark matter halo properties (e.g., Orsi et al. 2014).

The results reported in Equations 8 – 11 and shown in Figure 6 do not take into account the errors in  $L^*(z)$ . The errors for each sample are listed in Tables 2 and 3. We find that the errors are on the order of 0.05 dex for the lowest redshift samples and  $\sim 0.20$  dex for the highest redshift samples. Taking into account this error does not significantly remove the redshift dependency that we see in Figure 6, but may change the measurements shown in Equations 8 – 11.

### 5.2 Stellar Mass Dependency

In principle, the mass of a halo regulates the inflow of cold gas that is used to fuel star formation activity inside galaxies with the peak in star formation activity found to occur in  $10^{12} M_\odot$  halos (e.g., Peacock & Smith 2000; Seljak 2000; Moster et al. 2010; Behroozi et al. 2013a). It is then expected that there is a dependency between the stellar mass of a galaxy and its host halo mass, which forms the main basis of the abundance matching technique (e.g., Vale & Ostriker

2004). In this section we explore the stellar - halo mass relationship.

Figure 7 shows  $r_0$  per stellar mass bin for our samples of emission line-selected galaxies and listed in Tables 2 and 3. Similar to the results found in §5.1 for the line luminosity dependency, we find an increase in  $r_0$  with increasing stellar mass although not as pronounced as the line luminosity dependency, especially for the high- $z$  samples. The  $H\beta+[OIII]$   $z = 0.84$  shows an increase of a factor of  $\sim 2$  for the full range of stellar mass observed, while the  $z > 1$  show an increase of a factor ranging between 1.2 – 1.5. The  $[OII]$   $z = 1.47$  shows that the clustering length increases by a factor of  $\sim 1.7$ , which is weaker when compared to the line luminosity dependency.

We also find a strong redshift evolution for a fixed stellar mass. For example, the  $H\beta+[OIII]$  samples show a clustering length of 2.6, 2.9, 3.2, and 4.4  $Mpc h^{-1}$  for  $z = 0.84, 1.42, 2.23,$  and  $3.24$ , respectively at a fixed stellar mass of  $10^{10} M_\odot$ . To test if there is a redshift evolution we apply the same approach as was done with the line luminosity dependency by investigating the clustering evolution in terms of halo and stellar mass. We use our models as described in §3.6 to convert  $r_0$  to minimum halo mass.

Figure 8 shows the dependency between stellar and minimum halo mass for all redshift slices. We find a strong dependency between stellar and halo mass for  $z = 0.84$   $H\beta+[OIII]$  emitters between stellar masses of  $10^{8.5} M_\odot$  and  $10^{9.8} M_\odot$ . There is also a hint of a dependency for  $z = 1.42$   $H\beta+[OIII]$  emitters in the stellar mass range of  $\sim 10^{9.0} M_\odot$  to

**Table 2.** Clustering Properties of the  $H\beta+[OIII]$  samples as a function of line luminosities and stellar masses. We include  $L^*(z)$  for each sample as measured by [Khostovan et al. \(2015\)](#). All measurements assume a fixed  $\gamma = -1.8$ . The minimum dark matter halo masses are measured from the  $r_0$  measurements in conjunction with our  $r_0$ -halo mass models. All measurements are corrected for cosmic variance by adding in quadrature 11% of  $r_0$  in the total error cited.

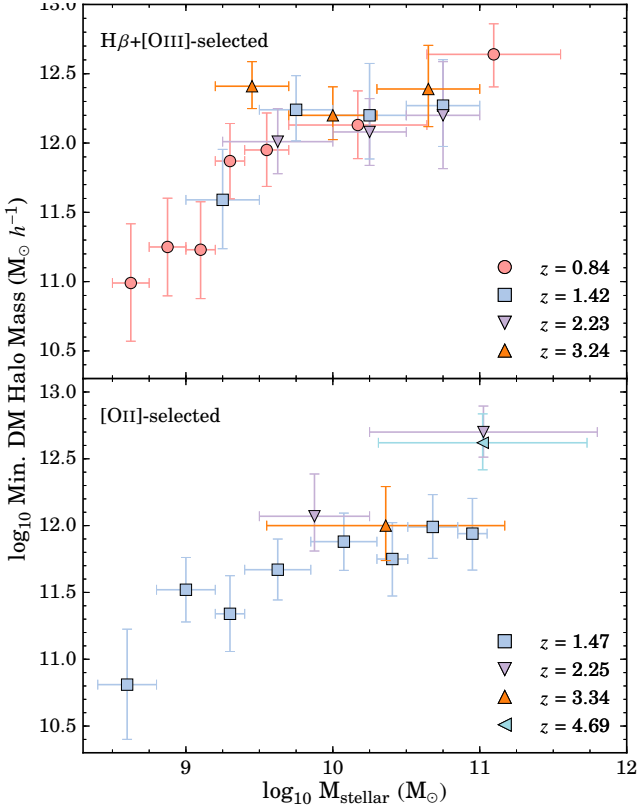
Subsample	$N_D$	$r_{0,\text{exact}}$ (Mpc $h^{-1}$ )	$\log_{10} M_{\text{min}}$ ( $M_{\odot} h^{-1}$ )
$H\beta+[OIII]$ $z = 0.84$ ( $\log_{10} L^* = 41.79^{+0.03}_{-0.05}$ )			
$40.50 < \log_{10} L_{\text{line}} < 40.60$	188	$1.15^{+0.23}_{-0.22}$	$9.48^{+1.40}_{-1.30}$
$40.60 < \log_{10} L_{\text{line}} < 40.70$	175	$1.46^{+0.23}_{-0.22}$	$10.66^{+0.59}_{-0.57}$
$40.70 < \log_{10} L_{\text{line}} < 40.80$	150	$1.46^{+0.26}_{-0.25}$	$10.67^{+0.67}_{-0.63}$
$40.80 < \log_{10} L_{\text{line}} < 41.00$	279	$1.46^{+0.20}_{-0.21}$	$10.67^{+0.52}_{-0.52}$
$41.00 < \log_{10} L_{\text{line}} < 41.15$	538	$1.77^{+0.22}_{-0.21}$	$11.28^{+0.34}_{-0.34}$
$41.15 < \log_{10} L_{\text{line}} < 41.30$	404	$1.89^{+0.23}_{-0.23}$	$11.46^{+0.31}_{-0.31}$
$41.30 < \log_{10} L_{\text{line}} < 41.60$	492	$2.08^{+0.25}_{-0.24}$	$11.69^{+0.28}_{-0.27}$
$41.60 < \log_{10} L_{\text{line}} < 41.80$	131	$3.18^{+0.44}_{-0.42}$	$12.53^{+0.23}_{-0.22}$
$41.80 < \log_{10} L_{\text{line}} < 41.95$	51	$3.24^{+0.51}_{-0.46}$	$12.55^{+0.26}_{-0.24}$
$41.95 < \log_{10} L_{\text{line}} < 42.55$	61	$4.64^{+0.59}_{-0.60}$	$13.10^{+0.17}_{-0.18}$
$8.50 < \log_{10} M < 8.75$			
$8.75 < \log_{10} M < 9.00$	368	$1.60^{+0.22}_{-0.21}$	$11.15^{+0.32}_{-0.32}$
$9.00 < \log_{10} M < 9.20$	483	$1.75^{+0.22}_{-0.22}$	$11.35^{+0.28}_{-0.28}$
$9.20 < \log_{10} M < 9.40$	391	$1.74^{+0.21}_{-0.22}$	$11.33^{+0.27}_{-0.28}$
$9.40 < \log_{10} M < 9.70$	294	$2.26^{+0.28}_{-0.28}$	$11.89^{+0.26}_{-0.26}$
$9.70 < \log_{10} M < 10.64$	271	$2.34^{+0.30}_{-0.29}$	$11.96^{+0.26}_{-0.26}$
$10.64 < \log_{10} M < 11.55$	213	$2.56^{+0.32}_{-0.32}$	$12.11^{+0.20}_{-0.19}$
$H\beta+[OIII]$ $z = 1.42$ ( $\log_{10} L^* = 42.06^{+0.06}_{-0.05}$ )			
$41.92 < \log_{10} L_{\text{line}} < 42.02$	74	$3.41^{+0.46}_{-0.49}$	$12.55^{+0.21}_{-0.22}$
$41.92 < \log_{10} L_{\text{line}} < 42.02$	191	$1.54^{+0.28}_{-0.25}$	$10.87^{+0.49}_{-0.44}$
$42.02 < \log_{10} L_{\text{line}} < 42.06$	63	$2.33^{+0.49}_{-0.48}$	$11.79^{+0.40}_{-0.39}$
$42.06 < \log_{10} L_{\text{line}} < 42.16$	58	$4.30^{+0.67}_{-0.68}$	$12.78^{+0.22}_{-0.22}$
$42.16 < \log_{10} L_{\text{line}} < 42.26$	25	$4.28^{+1.12}_{-1.08}$	$12.78^{+0.36}_{-0.35}$
$42.26 < \log_{10} L_{\text{line}} < 42.80$	34	$3.97^{+0.82}_{-0.84}$	$12.67^{+0.30}_{-0.31}$
$9.00 < \log_{10} M < 9.50$			
$9.50 < \log_{10} M < 10.00$	96	$2.10^{+0.38}_{-0.36}$	$11.54^{+0.33}_{-0.31}$
$10.00 < \log_{10} M < 10.50$	99	$3.00^{+0.45}_{-0.41}$	$12.14^{+0.21}_{-0.19}$
$10.50 < \log_{10} M < 11.00$	60	$2.93^{+0.66}_{-0.55}$	$12.11^{+0.31}_{-0.26}$
$10.50 < \log_{10} M < 11.00$	53	$3.06^{+0.62}_{-0.55}$	$12.18^{+0.28}_{-0.25}$
$H\beta+[OIII]$ $z = 2.23$ ( $\log_{10} L^* = 42.66^{+0.13}_{-0.13}$ )			
$42.30 < \log_{10} L_{\text{line}} < 42.66$	136	$2.66^{+0.44}_{-0.44}$	$11.77^{+0.28}_{-0.28}$
$42.66 < \log_{10} L_{\text{line}} < 42.74$	56	$5.15^{+0.64}_{-0.68}$	$12.74^{+0.16}_{-0.17}$
$42.74 < \log_{10} L_{\text{line}} < 43.10$	57	$7.38^{+0.88}_{-0.90}$	$13.17^{+0.14}_{-0.14}$
$9.25 < \log_{10} M < 10.00$			
$10.00 < \log_{10} M < 10.50$	120	$3.08^{+0.47}_{-0.45}$	$11.89^{+0.26}_{-0.25}$
$10.50 < \log_{10} M < 11.00$	66	$3.22^{+0.50}_{-0.50}$	$11.97^{+0.27}_{-0.27}$
$10.50 < \log_{10} M < 11.00$	41	$3.48^{+0.91}_{-0.90}$	$12.08^{+0.35}_{-0.34}$
$H\beta+[OIII]$ $z = 3.24$ ( $\log_{10} L^* = 42.83^{+0.19}_{-0.17}$ )			
$42.30 < \log_{10} L_{\text{line}} < 42.67$	68	$3.24^{+0.51}_{-0.53}$	$11.77^{+0.24}_{-0.25}$
$42.67 < \log_{10} L_{\text{line}} < 42.83$	67	$5.56^{+0.74}_{-0.73}$	$12.52^{+0.17}_{-0.17}$
$42.83 < \log_{10} L_{\text{line}} < 43.18$	44	$6.98^{+1.12}_{-1.00}$	$12.80^{+0.19}_{-0.17}$
$9.20 < \log_{10} M < 9.70$			
$9.70 < \log_{10} M < 10.30$	56	$5.09^{+0.69}_{-0.63}$	$12.29^{+0.18}_{-0.17}$
$10.30 < \log_{10} M < 11.00$	80	$4.35^{+0.65}_{-0.55}$	$12.08^{+0.20}_{-0.17}$
$10.30 < \log_{10} M < 11.00$	29	$5.02^{+1.21}_{-1.04}$	$12.27^{+0.32}_{-0.28}$

**Table 3.** The clustering properties of [OII] as a function of line luminosity and stellar mass. Table description is the same as that of Table 2. The  $z = 3.34$  and  $4.69$  measurements are not included in this table as the sample sizes were too small to divide in line luminosity and stellar mass bins. The measurements corresponding to the full samples are shown in Table 1.

Subsample	$N_D$	$r_{0,\text{exact}}$ (Mpc $h^{-1}$ )	$\log_{10} M_{\text{min}}$ ( $M_{\odot} h^{-1}$ )
$[OII]$ $z = 1.47$ ( $\log_{10} L^* = 41.86^{+0.03}_{-0.03}$ )			
$41.05 < \log_{10} L_{\text{line}} < 41.15$	200	$1.34^{+0.27}_{-0.22}$	$10.47^{+0.61}_{-0.51}$
$41.15 < \log_{10} L_{\text{line}} < 41.25$	501	$1.41^{+0.18}_{-0.18}$	$10.62^{+0.37}_{-0.36}$
$41.25 < \log_{10} L_{\text{line}} < 41.45$	761	$1.74^{+0.20}_{-0.20}$	$11.16^{+0.27}_{-0.28}$
$41.45 < \log_{10} L_{\text{line}} < 41.65$	638	$2.47^{+0.29}_{-0.29}$	$11.89^{+0.21}_{-0.22}$
$41.65 < \log_{10} L_{\text{line}} < 41.85$	667	$2.76^{+0.32}_{-0.32}$	$12.08^{+0.20}_{-0.20}$
$41.85 < \log_{10} L_{\text{line}} < 42.00$	292	$3.34^{+0.40}_{-0.40}$	$12.39^{+0.19}_{-0.19}$
$42.00 < \log_{10} L_{\text{line}} < 42.10$	101	$3.23^{+0.46}_{-0.49}$	$12.34^{+0.23}_{-0.24}$
$42.10 < \log_{10} L_{\text{line}} < 42.20$	68	$3.32^{+0.49}_{-0.50}$	$12.38^{+0.23}_{-0.24}$
$42.20 < \log_{10} L_{\text{line}} < 42.60$	56	$4.06^{+0.88}_{-0.70}$	$12.68^{+0.31}_{-0.25}$
$8.40 < \log_{10} M < 8.80$			
$8.80 < \log_{10} M < 9.20$	217	$1.51^{+0.23}_{-0.23}$	$10.84^{+0.49}_{-0.49}$
$9.20 < \log_{10} M < 9.40$	671	$2.04^{+0.24}_{-0.24}$	$11.48^{+0.21}_{-0.21}$
$9.40 < \log_{10} M < 9.85$	429	$1.88^{+0.24}_{-0.24}$	$11.33^{+0.24}_{-0.23}$
$9.85 < \log_{10} M < 10.30$	840	$2.20^{+0.26}_{-0.25}$	$11.61^{+0.21}_{-0.21}$
$10.30 < \log_{10} M < 10.51$	492	$2.46^{+0.29}_{-0.29}$	$11.81^{+0.21}_{-0.21}$
$10.51 < \log_{10} M < 10.85$	163	$2.30^{+0.33}_{-0.33}$	$11.69^{+0.25}_{-0.26}$
$10.85 < \log_{10} M < 11.05$	203	$2.61^{+0.36}_{-0.35}$	$11.92^{+0.25}_{-0.24}$
$10.85 < \log_{10} M < 11.05$	97	$2.54^{+0.37}_{-0.39}$	$11.86^{+0.27}_{-0.28}$
$[OII]$ $z = 2.25$ ( $\log_{10} L^* = 42.34^{+0.04}_{-0.03}$ )			
$42.40 < \log_{10} L_{\text{line}} < 42.57$	102	$2.97^{+0.42}_{-0.42}$	$11.95^{+0.22}_{-0.22}$
$42.57 < \log_{10} L_{\text{line}} < 43.21$	35	$4.49^{+0.75}_{-0.73}$	$12.55^{+0.23}_{-0.22}$
$9.50 < \log_{10} M < 10.25$			
$10.25 < \log_{10} M < 11.80$	61	$3.21^{+0.66}_{-0.54}$	$11.95^{+0.36}_{-0.29}$
$10.25 < \log_{10} M < 11.80$	43	$5.01^{+0.76}_{-0.73}$	$12.56^{+0.21}_{-0.20}$

$\sim 10^{10.0} M_{\odot}$  and  $z = 2.23$   $H\beta+[OIII]$  emitters for a similar range, although the latter is within  $1\sigma$  error bars. We find that the  $z = 3.24$   $H\beta+[OIII]$  sample shows a constant halo mass of  $\sim 10^{12.3} M_{\odot}$  for the full stellar mass range ( $9.2 < \log_{10} M_{\text{stellar}}/M_{\odot} < 11.0$ ) and this is consistent with the other redshift slices for stellar masses  $> 10^{9.75} M_{\odot}$ , although with a spread in halo mass between  $12.0 < \log_{10} M_{\text{halo}}/M_{\odot} < 12.5$ . Interestingly, we find that for the stellar mass range of  $10^{9.75-11.0} M_{\odot}$ ,  $H\beta+[OIII]$  emitters between  $z = 0.84$  and  $3.24$  reside in  $\sim 10^{12.3} M_{\odot}$  halos forming a redshift-independent plateau as also seen in the line luminosity dependency (see §5.1).

The bottom panel of Figure 8 shows the dependency for [OII]-selected emitters up to  $z = 2.25$  and we include the full sample measurements (same as shown in Figure 4) for our  $z = 3.34$  and  $4.69$  samples due to small sample sizes. We find that the  $z = 1.47$  [OII] sample shows an increase in halo mass with increasing stellar mass between  $8.4 < \log_{10} M_{\text{stellar}}/M_{\odot} < 10$ . The  $z = 2.25$  sample shows an increase between  $9.5 < \log_{10} M_{\text{stellar}}/M_{\odot} < 11.8$ , although this is based only on two measurements. We find that the  $z = 3.34$  measurement for the full sample is consistent with a halo mass of  $\sim 10^{12} M_{\odot}$  which agrees with the  $z = 1.47$  and  $2.25$  measurements. We find that the  $z = 4.69$  measurements are consistent only with the most massive  $z = 2.25$  [OII] emitters with a halo mass of  $\sim 10^{12.6} M_{\odot}$ .

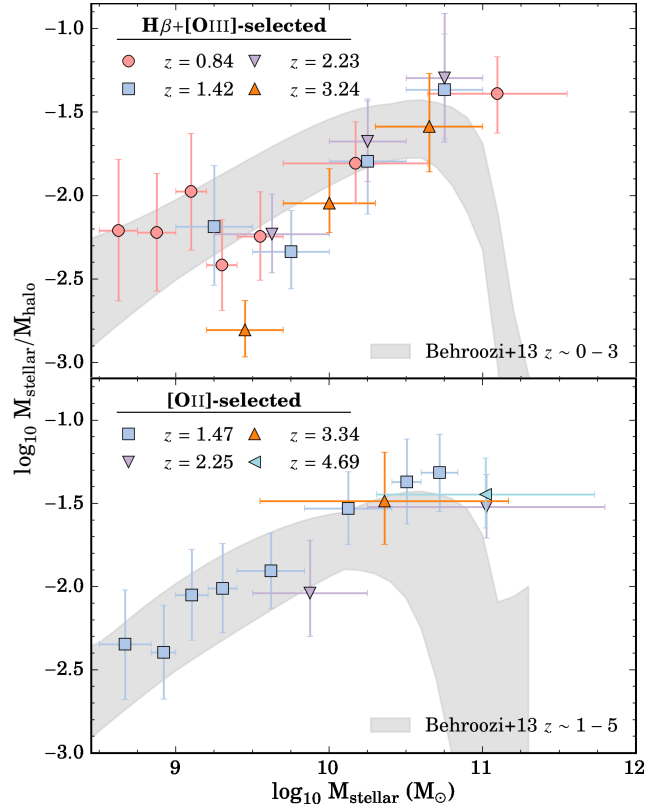


**Figure 8.** The minimum halo mass dependency with stellar mass. We find a strong relationship at all redshift slices for our  $H\beta+[OIII]$  samples and for the  $z = 1.47$   $[OII]$  sample. The  $z = 2.25$   $[OII]$  sample also shows an increasing trend, but is limited only to two stellar mass bins. The other  $[OII]$  samples are limited due to sample size and could not be separated in stellar mass bins. We find no redshift evolution in the relationship. Interestingly at stellar masses  $> 10^{9.75} M_{\odot}$  the halo mass is found to be constant at  $\sim 10^{12.3} M_{\odot}$  for  $H\beta+[OIII]$  emitters and  $\sim 10^{12} M_{\odot}$  for  $[OII]$  emitters.

### 5.2.1 Stellar-Halo Mass Ratio

A byproduct of the stellar mass dependency is the stellar-halo mass (SHM,  $M_{\text{stellar}}/M_{\text{halo}}$ ) ratio, which is defined as the stellar mass divided by the halo mass. This is a useful tracer of the star formation efficiency since the SHM ratio can be interpreted as the ratio of baryons that formed stars to dark matter (assuming a universal baryon fraction). Theoretical and observational studies have found that the maximum star formation efficiency in galaxies occurs in  $\sim 10^{12} M_{\odot}$  halos (e.g., Moster et al. 2010, 2013; Behroozi et al. 2013a). In this section, we explore the SHM ratio as a function of stellar mass for our  $H\beta+[OIII]$  and  $[OII]$  samples.

Figure 9 shows the SHM ratio where we find it to be redshift-independent for  $H\beta+[OIII]$  emitters for all stellar masses. We find the SHM ratio for  $z = 0.84$  and  $1.42$   $H\beta+[OIII]$  emitters as constant between  $8.5 < \log_{10} M_{\text{stellar}}/M_{\odot} < 9.75$  and increases for all redshift slices from  $-2.3$  to  $-1.3$  dex for  $M_{\text{stellar}} > 10^{9.75} M_{\odot}$ . The bottom panel of Figure 9 shows the SHM ratio as a function of stellar mass for  $[OII]$  emitters. We find that the SHM ratio increases with stellar mass at  $z = 1.47$  for the full stellar mass range



**Figure 9.** The stellar-halo mass (SHM) ratio as a function of stellar mass. We find our  $[OII]$  measurements show a continuous, redshift-independent increase in the SHM ratio for the full stellar mass range. The  $H\beta+[OIII]$  measurements show a constant ratio up to  $\sim 10^{9.75} M_{\odot}$  followed by a continuous, redshift-independent increase in the ratio. We compare our measurements with the abundance matching measurements of Behroozi et al. (2013b) overlaid in grey. We find that our  $[OII]$  and  $H\beta+[OIII]$  measurements are in agreement within  $1\sigma$  except for our  $H\beta+[OIII]$   $10^{9.75-10.00} M_{\odot}$  measurements.

probed. The  $z = 2.25$  sample also shows an increase with stellar mass and is consistent with the  $z = 1.47$  measurements. We find the same redshift-independent trend as for the  $H\beta+[OIII]$  emitters.

We overlay in Figure 9 the measurements of Behroozi et al. (2013b), which used the abundance matching technique and the constraints set by observational measurements of the global stellar mass functions to calculate the SHM ratio up to  $z \sim 8$ . Behroozi et al. (2013b) found that the ratio is redshift-independent and we therefore only highlight in Figure 9 the  $1\sigma$  confidence region of their measurements that correspond to the redshifts of our sample. We find all four redshift slices for the  $[OII]$  samples are in agreement with the Behroozi et al. (2013b) measurements. Our  $H\beta+[OIII]$  measurements are also in agreement for  $M_{\text{stellar}} < 10^{9.5} M_{\odot}$  and  $> 10^{10} M_{\odot}$ . Note that the Behroozi et al. (2013b) measurements are based on ‘global’ (passive+active galaxy) stellar mass functions, while our samples are comprised of ‘active’ galaxies (see Figure 3 of Khostovan et al. (2016) for the  $UVJ$  diagram) which could

explain the discrepancy at  $\sim 10^{9.75} M_{\odot}$  shown in Figure 9 for the  $H\beta+[OIII]$  samples.

### 5.2.2 Minimum or Effective Halo Mass?

The comparison with Behroozi et al. (2013b) is not exactly a like-to-like comparison as their measurements are constrained using global stellar mass functions. Our samples are emission line-selected, such that they are selecting the active population of galaxies and are not stellar mass complete. Furthermore, the halo masses reported in Behroozi et al. (2013b) are defined as the mass of a host halo similar to an effective halo mass and not a minimum halo mass, as used in our work. Their models also take into account satellite galaxies, while our model assumes one central galaxy per host dark matter halo. Our measurements shown in Figure 9 then have two caveats: (1) stellar mass incompleteness and (2) minimum halo mass.

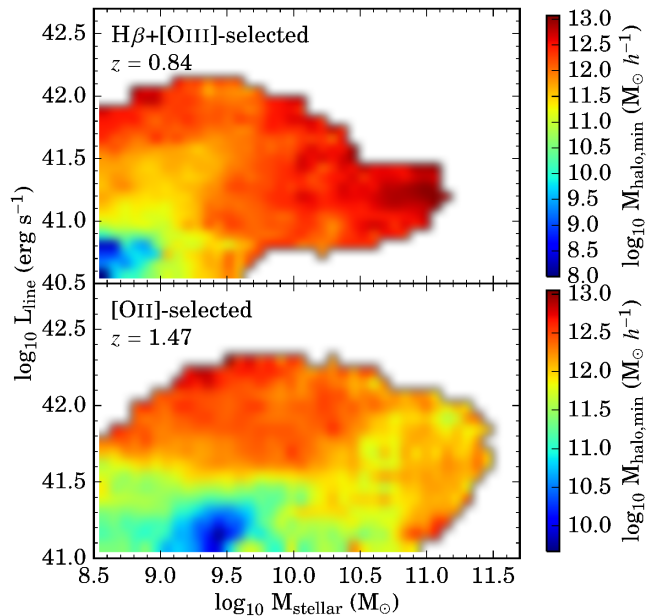
Despite these differences in assumptions and caveats, it is interesting that our measurements of the SHM ratio are consistent with those of Behroozi et al. (2013b). A possible reason for the agreement is that the stellar mass incompleteness and minimum halo mass effects are canceling each other. The stellar mass incompleteness could be underestimating the clustering signal and, as a consequence, underestimating the minimum halo mass. The strong agreement with our [OII] SHM ratio measurements shown in Figure 9 could also suggest that our [OII] samples are more representative of a stellar mass-complete sample in comparison to our  $H\beta+[OIII]$  samples.

In regards to the different definition of halo mass, the agreement with the measurements of Behroozi et al. (2013b) could suggest that our minimum halo mass measurements are more representative of the effective halo mass in HOD models due to our simplified assumption of only pure central galaxy occupation. For example, we show the  $H\alpha$  measurements of Cochrane et al. (2017) in Figure 6 using their  $r_0$  measurements and converting it to minimum halo mass using our model as described in §3.6. We find that their effective halo masses are roughly consistent with our assessment of the minimum halo masses using their  $r_0$  measurements.

Since our dark matter halo model assumes only one galaxy per host halo and given the steepness of the halo mass functions, it is then likely that our minimum halo masses are similar to effective (average) halo masses. We test this by integrating the halo mass functions to calculate the effective halo mass down to a given  $M_{\min}$ . We find that the maximum offset between the effective and minimum halo mass is 0.25 dex at  $M_{\min} \sim 10^{13.5} M_{\odot}$  and 0.07 dex at  $M_{\min} \sim 10^{12} M_{\odot}$ . Our results show that even for our simplified model, the difference between minimum and effective halo mass is negligible. Although we continue to refer to our halo mass measurements as ‘minimum’ halo mass as defined in Equation 6, we strongly caution the reader that our measurements may be more representative of the effective halo mass in comparison to HOD models.

### 5.3 Observed Line Luminosity – Stellar Mass Dependency

Observations have found a correlation between the star-formation rate and stellar mass in the local Universe (e.g.,



**Figure 10.** Shown is the halo mass dependency on line luminosity and stellar mass. Only the NB921 samples are used ( $z = 0.84$   $H\beta+[OIII]$  and  $z = 1.47$  [OII]) as these are the most populated ( $\sim 2500 - 3000$  sources each). All the measurements were done by randomly sampling the grid 10000 times and going through the clustering analysis to measure the halo mass. Overall, we find that the halo mass correlation with line luminosity is stronger than with stellar mass.

Salim et al. 2007; Lee et al. 2011), around cosmic noon (e.g., Daddi et al. 2007; Noeske et al. 2007; Rodighiero et al. 2011; Whitaker et al. 2014; Shivaei et al. 2015), and at higher redshifts (e.g., Schreiber et al. 2015; Tasca et al. 2015; Tomczak et al. 2016). Line luminosities trace star-formation activity (e.g., [OIII]: Suzuki et al. 2016; [OII]: Kennicutt 1998; Kewley et al. 2004) and we find a dependence between halo mass, line luminosity, and stellar mass. The question that arises is how much does the dependency of line luminosity affect the dependency measured with stellar mass or vice versa?

We test this by redoing our clustering analysis in 10000 randomly selected parts of the line luminosity-stellar mass grid and calculate the halo mass following the same methodology highlighted in §3. Each realization is a rectangular box randomly placed in the grid and must have  $> 50$  sources. The results are shown in Figure 10 for only the NB921 samples ( $H\beta+[OIII]$   $z = 0.84$  and [OII]  $z = 1.47$ ) as these are the most populated samples and are much easier to investigate the dual dependency of line luminosity and stellar mass with the halo mass. We find that for increasing line luminosity and stellar mass, the halo mass is increasing from as low as  $10^8$  to  $10^{13} M_{\odot}$ , although there is a significant scatter such that to assess which property dominates the dependency with halo mass requires a look at how stellar mass (line luminosity) is dependent on halo mass for a fixed line luminosity (stellar mass).

### 5.3.1 Fixed Stellar Mass

In this section, we investigate if there is a line luminosity dependency for a fixed stellar mass. We find a strong dependency between halo mass and line luminosity in  $H\beta+[OIII]$  emitters with fixed stellar masses of  $10^{8.5-9.5} M_{\odot}$  where the halo mass is found to increase from  $\sim 10^{8.0} M_{\odot}$  to  $\sim 10^{12.5-13.0} M_{\odot}$ . Beyond  $> 10^{9.5} M_{\odot}$  the halo mass is consistent with  $10^{12.5-13} M_{\odot}$  for all observed line luminosities, although this is primarily due to a small sample size ( $\sim 300$  sources, see Table 2) and a limiting range of line luminosities, especially at higher stellar masses.

For the  $z = 1.47$   $[OII]$  emitters, we find that for fixed stellar masses of  $10^{8.5-11} M_{\odot}$ , there is a strong dependency with line luminosity such that the halo mass increases from  $\sim 10^{9.5} M_{\odot}$  to  $10^{13} M_{\odot}$  with increasing line luminosity. Interestingly, the dependency is found for a wider range of fixed stellar masses in comparison to the  $H\beta+[OIII]$  sample and this could be due to the  $[OII]$  sample selecting more higher mass galaxies with low SFRs and ionization parameters compared to  $H\beta+[OIII]$ .

### 5.3.2 Fixed Line Luminosity

In the case of a fixed line luminosity, we find that there is only a stellar mass dependency with halo mass for  $H\beta+[OIII]$  emitters with  $L \lesssim 10^{41.5} \text{ erg s}^{-1}$  and it becomes more prevalent at  $L \lesssim 10^{41.0} \text{ erg s}^{-1}$ . The stellar mass dependency in the  $L_{H\beta+[OIII]} \sim 10^{41.0-41.5} \text{ erg s}^{-1}$  regime is probably due to contaminants, such as high-mass AGNs, that reside in halos of  $\sim 10^{13} M_{\odot}$ . If we disregard this sub-population of high mass sources, then the dependency breaks down. At  $L_{H\beta+[OIII]} \lesssim 10^{41.0} \text{ erg s}^{-1}$ , we find the dependency is the strongest where emitters with stellar masses  $> 10^{8.6} M_{\odot}$  reside in increasingly higher mass halos.

Figure 10 shows no significant stellar mass dependency for  $z = 1.47$   $[OII]$  emitters at a given line luminosity  $> 10^{41.6} \text{ erg s}^{-1}$ . We only find a stellar mass dependency in the case that  $L_{[OII]} \lesssim 10^{41.3} \text{ erg s}^{-1}$  where the halo mass is between  $10^{11-11.5} M_{\odot}$  for  $8.5 < \log_{10} M_{\text{stellar}}/M_{\odot} < 9$ , drops to halo masses of  $10^{9.5-11} M_{\odot}$  for  $9 < \log_{10} M_{\text{stellar}}/M_{\odot} < 9.5$ , and then increases to halo mass of  $10^{12} M_{\odot}$  with increasing stellar mass.

### 5.3.3 Which one: Line Luminosity or Stellar Mass?

We find that for both  $H\beta+[OIII]$  and  $[OII]$  emitters, a stellar mass dependency appears for the case of faint line luminosities as opposed to the line luminosity dependency which appears for the full stellar mass range. This could suggest that the trend between halo mass and line luminosity are more significant than with stellar masses, such that the correlations we observed in stellar mass could be a result of the halo mass correlation with line luminosity for our samples. Sobral et al. (2010) came to a similar conclusion using a sample of  $z = 0.84$   $H\alpha$  emitters and the rest-frame  $K$ -band luminosity as a proxy for stellar mass. Cochrane et al. (2017) also came to a similar conclusion using samples of  $z = 0.84$ , 1.47, and 2.23  $H\alpha$  emitters. We do caution the reader that our results are for line luminosity-selected samples.

## 6 DISCUSSION

In the previous sections, we found that there is a strong, redshift-independent relationship between line luminosity and minimum halo mass (relatively independent of stellar mass for  $z = 0.84$  and 1.47  $H\beta+[OIII]$  and  $[OII]$  emitters, respectively) up to  $L \sim L^*$  for  $H\alpha$ ,  $H\beta+[OIII]$ , and  $[OII]$  emitters. For the  $L > L^*$  regime, we find that the dependency becomes shallower and is consistent with minimum halo masses between  $10^{12.5} M_{\odot}$  and  $10^{13} M_{\odot}$ . In this section, we discuss potential physical reasons for the flattening/shallower slope of this relationship for the brightest emitters with the understanding that the emission lines observed trace the underlying star formation activity.

### 6.1 Transitional Halo Mass

Current models of galaxy formation suggest that the star formation efficiency is tied to the host halo mass with the peak in the efficiency found in  $\sim 10^{12} M_{\odot}$  halos (e.g., Behroozi et al. 2013a). For  $> 10^{12} M_{\odot}$  halos, models predict that the star formation activity in galaxies diminishes as external quenching mechanisms (e.g., shock heating of infalling gas; Dekel & Birnboim 2006) become stronger and are accompanied by internal quenching mechanisms (e.g., AGN feedback; Best et al. 2006). This is referred to as ‘halo quenching’, where a specific global halo mass is related to the quenching of galaxies. We note that this is still debatable where, observationally, some studies have found that external quenching is mainly a local phenomenon (e.g., Darvish et al. 2016) and does not depend significantly on the global halo mass (e.g., Peng et al. 2012; Carollo et al. 2013). Other observational studies find that galaxy quenching does depend on halo mass (e.g., Prescott et al. 2011; also see references in Darvish et al. 2017).

A consequence of the halo quenching predictions is a possible characteristic halo mass scale for which the fraction of star-forming galaxies drops and the fraction of passive galaxies increases sharply. Current predictions are that this occurs around a few  $\times 10^{12} M_{\odot}$  to  $10^{13} M_{\odot}$  and is also redshift independent (Croton et al. 2006; Dekel & Birnboim 2006; Cen 2011; Bower et al. 2017). Observations have reported such a transitional halo mass. For example, Dolley et al. (2014) used a sample of  $\sim 23,000$   $24\mu\text{m}$ -selected sources between  $0.2 < z < 1.0$  with an areal coverage of  $8.42 \text{ deg}^2$  and find evidence for a transitional halo mass  $\sim 8 \times 10^{12} M_{\odot}$ . Hartley et al. (2013) came to a similar conclusion using the deep  $0.77 \text{ deg}^2$  UKIDSS UDS data up to  $z \sim 3$  and measure a transitional halo mass of  $5 \times 10^{12} M_{\odot}$ .

We show in Figure 6 that  $L > L^*$  emitters have a flat/shallower line luminosity dependency consistent with minimum halo masses between  $3 \times 10^{12} M_{\odot}$  and  $10^{13} M_{\odot}$ . Note that based on our short discussion in §5.2.2, our minimum halo masses may be more representative of the effective halo mass due to our assumptions made in §3.6. With this caveat taken into account, our results are then consistent up to  $z \sim 5$  with the predictions of a transitional halo mass for which the number of star-forming galaxies (traced by our sample) diminishes and the fraction of passive galaxies increases. We note that this can also be a sample selection effect since the number densities of  $> 10^{13} M_{\odot}$  halos de-



creases significantly and requires large comoving volumes to detect their residing galaxies.

### 6.1.1 Potential Causes

Although our results show evidence for this transitional halo mass, it raises the question of how the brightest emitters reside in  $10^{13} M_{\odot}$  halos. Since the line luminosity traces the star-formation activity, it then seems puzzling that systems with such high SFRs are found in massive halos when the peak SF efficiency is found in  $\sim 10^{12} M_{\odot}$  halos. One possibility is that  $L > L^*$  emitters have their emission lines powered by AGN activity rather than SF activity. Sobral et al. (2016) spectroscopically followed up 59 bright  $L > L^*$  H $\alpha$  emitters and found that the AGN fraction increases with observed line luminosity such that the fraction of AGNs is  $\sim 50\%$  by  $\sim 4L^*$ . Although this is only measured for H $\alpha$  emitters up to  $z = 2.23$  and may not be true for H $\beta$ + [OIII] and [OII] emitters, studies up to  $z \sim 1.6$  have shown that X-ray and radio-selected AGN tend to reside in halos of  $\sim 10^{13} M_{\odot}$  (Hickox et al. 2009; Koutoulidis et al. 2013; Mendez et al. 2016), which is consistent with the constant halo mass for  $L > L^*$  H $\beta$ + [OIII] and [OII] emitters shown in Figure 6. Therefore, it is quite possible that these sources are AGN, although we require spectroscopic follow-up to measure AGN fractions in this line luminosity range for the H $\beta$ + [OIII] and [OII] samples.

Another possibility is that a fraction of the brightest emitters can have their emission lines powered by major merging events, such that these systems are currently undergoing a starburst phase. Simulations of major mergers predict elevated levels of star-formation activity (e.g., Mihos & Hernquist 1996; Di Matteo et al. 2008; Bournaud et al. 2011) and observations have thus far found evidence to support this (e.g., Hung et al. 2013). Semi-analytical models have also predicted that the stellar mass assembly in high-mass halos is due to mergers (e.g., Zehavi et al. 2012). A detailed morphological study of the fraction of mergers as a function of line luminosity would help in addressing this issue and we plan to explore this in the future.

It could also be possible that environmental effects could allow for the presence of  $> L^*$  emitters in massive halos. Dekel & Birnboim (2006) used simulations and predict that cold filamentary streams can penetrate the shock heated halo gas in  $> 3 \times 10^{12} M_{\odot}$  halos and fuel star-formation activity in  $L > L^*$  galaxies above  $z > 2$ . To support this level of star-formation activity requires large cold gas accretion rates and a recent ALMA study by Scoville et al. (2017) estimated the rate to be  $> 100 M_{\odot} \text{ yr}^{-1}$  for  $z > 2$  to maintain galaxies along the main-sequence.

Overall, we find evidence for a possible transitional halo mass for which star-forming galaxies become less common and halos are increasingly populated by passive galaxies. It stands to reason that the  $L > L^*$  emitters are a mixture of AGN- and star-formation-dominated systems. This is also suggested by Kauffmann et al. (2003) in the local Universe (up to  $z \sim 0.3$ ) where they find that galaxies with AGN and bright [OIII] lines also include young stellar populations due to a recent phase of star-formation activity. Future spectroscopic and morphological studies can shed light on the physical processes involved that are powering nebular emis-

sion lines in such massive halos and provide us with valuable insight on the quenching mechanisms that are occurring at this transitional halo mass.

## 6.2 Clustering more dependent on line luminosity than stellar mass?

In §5.2 and §5.3 we found that the dependency of clustering on line luminosity was more significant than on stellar mass. We also concluded, based on the results of our  $z = 0.84$  H $\beta$ + [OIII] and  $z = 1.47$  [OII] samples in §5.3, the stellar mass dependency may be a result of the line luminosity dependency. This is a similar conclusion made by Sobral et al. (2010) where they used a  $z = 0.84$  H $\alpha$ -selected sample and found that the line luminosity dependency was more significant than the dependency with stellar mass. Coil et al. (2017) came to a similar conclusion where they found that the clustering amplitude was a stronger function of the specific star formation rate than stellar mass and that the clustering strength for a given specific star formation rate was found to be independent of stellar mass. Cochrane et al. (2017) used H $\alpha$ -selected narrowband samples at  $z = 0.84, 1.47,$  and  $2.23$  and found that the line luminosity dependency was not driven/independent of stellar mass.

We note that the lack of a strong stellar mass dependency with clustering strength/dark matter halo mass could be mainly caused by sample selection. As mentioned before, our samples are selected based on line flux such that they are complete in line luminosity down to a completeness limit. In regards to stellar mass, our samples are not complete, especially for the low stellar mass range ( $< 10^9 M_{\odot}$ ; see Khostovan et al. 2016 for the stellar mass functions of our samples). We can only conclude that for narrowband-selected samples, the clustering strength dependency with stellar mass seems to be less significant than the dependency with line luminosity and may also be a result of it as well.

## 7 CONCLUSIONS

We have presented our H $\beta$ + [OIII] and [OII] clustering measurements up to  $z \sim 3.3$  and  $\sim 4.7$ , respectively. The main results of this study are:

- (i) We find that the power law slopes of the angular correlation functions are consistent with  $\beta \sim -0.80$ . Using the exact Limber equation, we find typical  $r_0$  between  $1.45 - 4.01 h^{-1} \text{ Mpc}$  and  $1.99 - 8.25 h^{-1} \text{ Mpc}$  for H $\beta$ + [OIII] and [OII] emitters, respectively. These correspond to minimum halo masses between  $10^{10.70-12.08} M_{\odot}$  and  $10^{11.46-12.62} M_{\odot}$ , respectively.
- (ii) A  $r_0$ -line luminosity dependency is found where the brightest emitters are more clustered compared to the faintest emitters. This dependency is found to be redshift-dependent but is biased due to evolution in the line luminosity function. When rescaling based on  $L^*(z)$  and using model predictions of halo mass given  $r_0$ , we find a strong increasing dependency between minimum halo mass and line luminosity that is independent of redshift with the faintest H $\beta$ + [OIII] ([OII]) emitters found in  $10^{9.5} M_{\odot}$  ( $10^{10.5} M_{\odot}$ ) halos and the brightest H $\beta$ + [OIII] ([OII]) emitters in  $10^{13} M_{\odot}$  ( $10^{12.5} M_{\odot}$ ) halos.

- (iii) A stellar mass dependency trend is found with  $r_0$  and, when converted to minimum halo mass, is found to be redshift independent. We find that  $H\beta+[OIII]$  emitters with stellar masses  $> 10^{9.75} M_{\odot}$  reside in  $10^{12.3} M_{\odot}$  halos between  $z = 0.84$  and  $3.24$ . The  $[OII]$  samples also show a stellar mass dependency for the full stellar mass range.
- (iv) We investigate how the interrelation between observed line luminosity and stellar mass can affect the individual dependencies we see on minimum halo mass. By creating subsamples in a line luminosity-stellar mass grid space for the most populated samples ( $H\beta+[OIII]$   $z = 0.84$  and  $[OII]$   $z = 1.47$ ), we find that the main dependency on minimum halo mass arises from the observed line luminosity such that the stellar mass dependency is weaker and could be a result of the line luminosity dependency. This then suggests a simple connection between the nebular emission line properties of galaxies and their host halo mass.
- (v) The line luminosity-halo mass dependency shows an increase from the faintest emitters observed to  $L \sim L^*(z)$ . For emitters brighter than  $L^*$ , we find that the trend is consistent with halos between  $10^{12.5-13} M_{\odot}$ . To understand what is powering such bright emission lines, we consider three possibilities: AGN-driven, merger-driven, and/or gas inflow. There is evidence from related studies to support this hypothesis although spectroscopic and morphological studies of our samples are required to properly investigate these sources. In comparison to predictions from models, we find that the shallower trend that we observe for  $L > L^*(z)$  emitters is consistent with the transitional halo mass for which the fraction of star-forming galaxies decreases and the fraction of passive galaxies increases due to internal and external quenching mechanisms.

Our results suggest a simple connection between the clustering/dark matter halo properties and nebular emission line properties of star-forming/‘active’ galaxies up to  $z \sim 5$ . This has implications for future theoretical studies that model this connection since previous constraints were up to  $z \sim 2$  for only  $H\alpha$  emitters. On the observational side, future spectroscopic studies of bright, emission line-selected galaxies can allow us to investigate the dependency between the ISM properties (internal mechanisms) of galaxies and massive halos (external mechanisms). Morphological studies of our samples can also test to see if the shape of galaxies is connected with the host halo properties. Future space-based (e.g., *JWST*, *WFIRST*) and ground-based observatories (e.g., European Extremely Large Telescope, Thirty Meter Telescope), can also allow us to study the clustering properties of emission line-selected galaxies at higher redshifts and larger comoving volumes. This would allow us to see when the following redshift-independent trends that seem to have been in place since  $z \sim 5$  were first established, which would present a new scaling relation for galaxy formation and evolution models.

## ACKNOWLEDGMENTS

AAK thanks Anahita Alavi and Irene Shivaie for useful discussion in the making of this paper.

AAK acknowledges that this work was supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program - Grant NNX16AO92H. DS acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO) through a Veni fellowship and from Lancaster University through an Early Career Internal Grant A100679. PNB is grateful for support from STFC via grant STM001229/1. IRS acknowledges support from STFC (ST/L00075X/1), the ERC Advanced Grant DUSTYGAL (321334), and a Royal Society/Wolfson Merit award. JM acknowledges the support of a Huygens PhD fellowship from Leiden University. BD acknowledges financial support from NASA through the Astrophysics Data Analysis Program (ADAP), grant number NNX12AE20G.

## REFERENCES

- Baugh C. M., 2006, *Reports on Progress in Physics*, **69**, 3101
- Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, **717**, 379
- Behroozi P. S., Wechsler R. H., Conroy C., 2013a, *ApJ*, **762**, L31
- Behroozi P. S., Wechsler R. H., Conroy C., 2013b, *ApJ*, **770**, 57
- Benson A. J., 2010, *Phys. Rep.*, **495**, 33
- Best P. N., Kaiser C. R., Heckman T. M., Kauffmann G., 2006, *MNRAS*, **368**, L67
- Bournaud F., et al., 2011, *ApJ*, **730**, 4
- Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, *MNRAS*, **465**, 32
- Bradshaw E. J., et al., 2013, *MNRAS*, **433**, 194
- Carollo C. M., et al., 2013, *ApJ*, **776**, 71
- Cen R., 2011, *ApJ*, **741**, 99
- Cochrane R. K., Best P. N., Sobral D., Smail I., Wake D. A., Stott J. P., Geach J. E., 2017, preprint, ([arXiv:1704.05472](https://arxiv.org/abs/1704.05472))
- Coil A. L., et al., 2008, *ApJ*, **672**, 153
- Coil A. L., et al., 2011, *ApJ*, **741**, 8
- Coil A. L., Mendez A. J., Eisenstein D. J., Moustakas J., 2017, *ApJ*, **838**, 87
- Colless M., et al., 2001, *MNRAS*, **328**, 1039
- Contreras S., Baugh C. M., Norberg P., Padilla N., 2015, *MNRAS*, **452**, 1861
- Cooray A., Sheth R., 2002, *Phys. Rep.*, **372**, 1
- Croton D. J., et al., 2006, *MNRAS*, **365**, 11
- Daddi E., et al., 2007, *ApJ*, **670**, 156
- Darvish B., Sobral D., Mobasher B., Scoville N. Z., Best P., Sales L. V., Smail I., 2014, *ApJ*, **796**, 51
- Darvish B., Mobasher B., Sobral D., Rettura A., Scoville N., Faisst A., Capak P., 2016, *ApJ*, **825**, 113
- Darvish B., Mobasher B., Martin D. C., Sobral D., Scoville N., Stroe A., Hemmati S., Kartaltepe J., 2017, *ApJ*, **837**, 16
- Davis M., et al., 2003, in Guhathakurta P., ed., *Proc. SPIE Vol. 4834, Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II*. pp 161–172 ([arXiv:astro-ph/0209419](https://arxiv.org/abs/astro-ph/0209419)), [doi:10.1117/12.457897](https://doi.org/10.1117/12.457897)
- Dekel A., Birnboim Y., 2006, *MNRAS*, **368**, 2
- Di Matteo P., Bournaud F., Martig M., Combes F., Melchior A.-L., Semelin B., 2008, *A&A*, **492**, 31
- Diemer B., Kravtsov A. V., 2015, *ApJ*, **799**, 108
- Dolley T., et al., 2014, *ApJ*, **797**, 125
- Driver S. P., et al., 2011, *MNRAS*, **413**, 971
- Eisenstein D. J., Hu W., 1998, *ApJ*, **496**, 605
- Garilli B., et al., 2014, *A&A*, **562**, A23
- 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
- Guo Q., White S., Li C., Boylan-Kolchin M., 2010, *MNRAS*, **404**, 1111
- Guo H., et al., 2013, *ApJ*, **767**, 122
- Guo H., et al., 2014, *MNRAS*, **441**, 2398
- Harikane Y., et al., 2016, *ApJ*, **821**, 123
- Hartley W. G., et al., 2013, *MNRAS*, **431**, 3045
- Hickox R. C., et al., 2009, *ApJ*, **696**, 891
- Hung C.-L., et al., 2013, *ApJ*, **778**, 129
- Jose C., Lacey C. G., Baugh C. M., 2016, *MNRAS*, **463**, 270
- Kashino D., et al., 2017, preprint, ([arXiv:1703.08326](https://arxiv.org/abs/1703.08326))
- Kauffmann G., et al., 2003, *MNRAS*, **346**, 1055
- Kennicutt Jr. R. C., 1998, *ARA&A*, **36**, 189
- Kerscher M., Szapudi I., Szalay A. S., 2000, *ApJ*, **535**, L13
- Kewley L. J., Geller M. J., Jansen R. A., 2004, *AJ*, **127**, 2002
- 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., Sobral D., Mobasher B., Smail I., Darvish B., Nayyeri H., Hemmati S., Stott J. P., 2016, *MNRAS*, **463**, 2363
- Koutoulidis L., Plionis M., Georgantopoulos I., Fanidakis N., 2013, *MNRAS*, **428**, 1382
- Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, *ApJ*, **609**, 35
- Landy S. D., Szalay A. S., 1993, *ApJ*, **412**, 64
- Lawrence A., et al., 2007, *MNRAS*, **379**, 1599
- Lee J. C., et al., 2011, *ApJS*, **192**, 6
- Limber D. N., 1953, *ApJ*, **117**, 134
- Lin L., et al., 2012, *ApJ*, **756**, 71
- Marulli F., et al., 2013, *A&A*, **557**, A17
- Matarrese S., Coles P., Lucchin F., Moscardini L., 1997, *MNRAS*, **286**, 115
- Matthee J., Sobral D., Oteo I., Best P., Smail I., Röttgering H., Paulino-Afonso A., 2016, *MNRAS*, **458**, 449
- Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2017, *MNRAS*, **465**, 2381
- McCracken H. J., et al., 2015, *MNRAS*, **449**, 901
- McLure R. J., et al., 2013, *MNRAS*, **428**, 1088
- Mendez A. J., et al., 2016, *ApJ*, **821**, 55
- Meneux B., et al., 2008, *A&A*, **478**, 299
- Meneux B., et al., 2009, *A&A*, **505**, 463
- Mihos J. C., Hernquist L., 1996, *ApJ*, **464**, 641
- Mo H. J., White S. D. M., 1996, *MNRAS*, **282**, 347
- Moscardini L., Coles P., Lucchin F., Matarrese S., 1998, *MNRAS*, **299**, 95
- Mostek N., Coil A. L., Cooper M., Davis M., Newman J. A., Weiner B. J., 2013, *ApJ*, **767**, 89
- Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, **710**, 903
- Moster B. P., Naab T., White S. D. M., 2013, *MNRAS*, **428**, 3121
- Noeske K. G., et al., 2007, *ApJ*, **660**, L43
- Norberg P., et al., 2002, *MNRAS*, **332**, 827
- Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, *MNRAS*, **396**, 19
- Orsi Á., Padilla N., Groves B., Cora S., Tecce T., Gargiulo I., Ruiz A., 2014, *MNRAS*, **443**, 799
- Peacock J. A., Smith R. E., 2000, *MNRAS*, **318**, 1144
- Peng Y.-j., Lilly S. J., Renzini A., Carollo M., 2012, *ApJ*, **757**, 4
- Prescott M., et al., 2011, *MNRAS*, **417**, 1374
- Press W. H., Schechter P., 1974, *ApJ*, **187**, 425
- Roche N. D., Almaini O., Dunlop J., Ivison R. J., Willott C. J., 2002, *MNRAS*, **337**, 1282
- Rodighiero G., et al., 2011, *ApJ*, **739**, L40
- Salim S., et al., 2007, *ApJS*, **173**, 267
- Schreiber C., et al., 2015, *A&A*, **575**, A74
- Scoville N., et al., 2007, *ApJS*, **172**, 1
- Scoville N., et al., 2017, *ApJ*, **837**, 150
- Seljak U., 2000, *MNRAS*, **318**, 203
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, **323**, 1
- Shioya Y., et al., 2008, *ApJS*, **175**, 128
- Shivaei I., et al., 2015, *ApJ*, **815**, 98
- Simon P., 2007, *A&A*, **473**, 711
- Skibba R. A., et al., 2015, *ApJ*, **807**, 152
- Sobral D., et al., 2009, *MNRAS*, **398**, 75
- 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., Best P. N., Smail I., Geach J. E., Cirasuolo M., Garn T., Dalton G. B., 2011, *MNRAS*, **411**, 675
- Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, *MNRAS*, **420**, 1926
- 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., Kohn S. A., Best P. N., Smail I., Harrison C. M., Stott J., Calhau J., Matthee J., 2016, *MNRAS*, **457**, 1739
- Sobral D., et al., 2017, *MNRAS*, **466**, 1242
- Somerville R. S., Davé R., 2015, *ARA&A*, **53**, 51
- Stott J. P., et al., 2013, *MNRAS*, **436**, 1130
- Stroe A., Sobral D., 2015, *MNRAS*, **453**, 242
- Suzuki T. L., et al., 2016, *MNRAS*, **462**, 181
- Takahashi M. I., et al., 2007, *ApJS*, **172**, 456
- Tasca L. A. M., et al., 2015, *A&A*, **581**, A54
- 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
- Tomczak A. R., et al., 2016, *ApJ*, **817**, 118
- Vale A., Ostriker J. P., 2004, *MNRAS*, **353**, 189
- Wake D. A., et al., 2011, *ApJ*, **728**, 46
- Whitaker K. E., et al., 2014, *ApJ*, **795**, 104
- York D. G., et al., 2000, *AJ*, **120**, 1579
- Zehavi I., et al., 2005, *ApJ*, **630**, 1
- Zehavi I., et al., 2011, *ApJ*, **736**, 59
- Zehavi I., Patiri S., Zheng Z., 2012, *ApJ*, **746**, 145
- Zheng Z., et al., 2005, *ApJ*, **633**, 791
- da Cunha E., Charlot S., Elbaz D., 2008, *MNRAS*, **388**, 1595

## APPENDIX A: WHERE DOES LIMBER'S APPROXIMATION FAIL?

As discussed in §3.5, Limber's approximation works well up to a certain angular separation. The question that arises is to what angular scales can the Limber approximation be used? As Simon (2007) showed, this depends on the filter profile. Here, we briefly describe the method to calculate the angular scales for which Limber's approximation fails. The equations derived in Simon (2007) assumed a simple top-hat filter.

The Limber equation is generally defined as:

$$w(\theta) \approx \int_0^\infty dr_1 \int_0^\infty dr_2 p_1(r)p_2(r)\xi\left(R, \frac{r_1+r_2}{2}\right)$$

$$R = \sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos \theta} \quad (\text{A1})$$

where  $p(r)$  is the filter profile defined by a center,  $r_c$ , and width,  $\Delta r$ , in comoving distance units. The comoving distance between the observer and source 1 and 2 are defined as  $r_1$  and  $r_2$  with  $R$  being the radial separation using the law of cosines. Simon (2007) showed that in the case that  $r_c \gg \Delta r$  the ACF becomes a rescaled version of the spatial correlation function with a slope of  $\gamma$  instead of  $1 - \gamma$ :

$$w(\theta_b) \approx \xi(r_c \theta_b, r_c) \quad (\text{A2})$$

Filter	True & Gaussian Filter Profile Parameters					
	$\lambda_{\text{obs}}$ ( $\mu\text{m}$ )	FWHM ( $\text{\AA}$ )	$z_{\text{True}}^{H\beta+[OIII]}$	$z_{\text{Gaussian}}^{H\beta+[OIII]}$	$z_{\text{True}}^{[OII]}$	$z_{\text{Gaussian}}^{[OII]}$
NB921	0.9196	132	$0.84 \pm 0.01$	$0.83 \pm 0.01$	$1.47 \pm 0.02$	$1.46 \pm 0.01$
NBJ	1.211	150	$1.42 \pm 0.02$	$1.42 \pm 0.01$	$2.25 \pm 0.02$	$2.25 \pm 0.02$
NBH	1.617	211	$2.23 \pm 0.02$	$2.23 \pm 0.02$	$3.34 \pm 0.03$	$3.34 \pm 0.02$
NBK	2.121	210	$3.24 \pm 0.02$	$3.24 \pm 0.02$	$4.69 \pm 0.03$	$4.69 \pm 0.02$

**Table A1.** The True filter parameters and the corresponding Gaussian-assumed filter parameters in terms of the  $H\beta+[OIII]$  and  $[OII]$  redshifts.

where we define  $\theta_b$  as the angular separation for which the Limber approximation fails. Note that the  $r_c\theta_b$  term arises from doing a small-angle approximation in the definition of  $R$  (see Equation A1 with  $r_1 = r_2 = r_c$ ).

To calculate  $\theta_b$ , we rewrite Equation A2 using the definition of  $A_w$  in the Limber approximation to get:

$$\theta_b = \frac{1}{A_w} \left( \frac{r_0}{r_c} \right)^\gamma \quad (\text{A3})$$

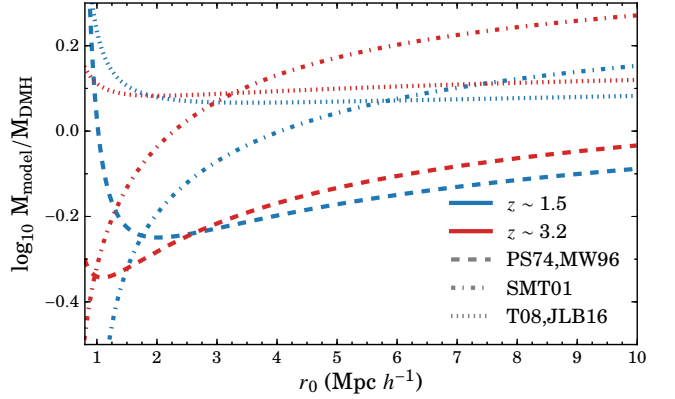
$$= \frac{1}{r_c^\gamma \sqrt{\pi}} \frac{\Gamma(\gamma/2)}{\Gamma(\frac{\gamma-1}{2})} \left[ \int_0^\infty d\bar{r} p_1(\bar{r}) p_2(\bar{r}) \bar{r}^{1-\gamma} \right]^{-1}. \quad (\text{A4})$$

where  $\bar{r}$  is defined as  $(r_1+r_2)/2$  and  $\Gamma$  is the gamma function. The above equation is generalized that for any filter  $p(\bar{r})$ , the expected angular separation for which the Limber approximation departs from the true ACF can be measured. In this study, we treated all four narrowband filters as Gaussians and show the true and Gaussian filter profile parameters in Table A1. We find that all our narrowband filters can be well-treated as Gaussian filters.

## APPENDIX B: UNCERTAINTIES IN CLUSTERING LENGTH - DARK MATTER HALO MASS PREDICTIONS

We presented our  $r_0$ -halo mass predictions in §3.6 where we assumed a Tinker et al. (2010) halo bias function and Tinker et al. (2008) halo mass function. Other prescriptions for the mass and bias functions do exist and in this section we explore how our predicted minimum halo masses change when assuming different assumptions. We consider three cases. The *first* case is assuming the Press & Schechter (1974) mass function and Mo & White (1996) bias function. This is a ‘classical’ assumption and was also used in Matarrese et al. (1997) and Moscardini et al. (1998), from which we use their methodology to make our predictions as highlighted in §3.6. The *second* case assumes the Sheth et al. (2001) halo mass and bias functions. Lastly, the *third* case assumes the Tinker et al. (2008) halo mass function and Jose et al. (2016) bias function.

Figure B1 shows a comparison between our predictions of halo mass ( $M_{\text{DMH}}$ ) and the predictions from the three cases highlighted above ( $M_{\text{model}}$ ) for a given  $r_0$  measurement at  $z \sim 1.5$  and  $\sim 3.2$ . We find that offsets can be quite large, especially towards lower  $r_0$  values where above  $3 \text{ Mpc } h^{-1}$  we find offsets of  $\pm 0.2$  dex and below  $3 \text{ Mpc } h^{-1}$  the offsets can be as high as 0.4 dex. The only case that best matches our predictions is the third case, which is not surprising as it uses the Tinker et al. (2008) halo mass function (same as the one we assumed) and the Jose et al. (2016)



**Figure B1.** Comparison of the predicted minimum halo masses for a given  $r_0$  between our assumed halo bias (Tinker et al. 2010) and mass (Tinker et al. 2008) functions against other assumptions. The cases are as follows: (1) Press & Schechter (1974) mass function and Mo & White (1996) bias function, (2) Sheth et al. (2001) mass and bias functions, and (3) Tinker et al. (2008) mass function and Jose et al. (2016) bias function. We show the difference for  $z \sim 1.5$  and  $\sim 3.2$  with  $M_{\text{model}}/M_{\text{DMH}}$  being the ratio of one of the cases highlighted above ( $M_{\text{model}}$ ) and our model ( $M_{\text{DMH}}$ ) as described in §3.6. We find that assuming different prescriptions for halo properties can introduce offsets of  $\sim \pm 0.2$  dex for  $r_0 > 3 \text{ Mpc } h^{-1}$  and become significantly worse for  $r_0 < 3 \text{ Mpc } h^{-1}$  such that at  $r_0 \sim 1 \text{ Mpc } h^{-1}$  the offset range between  $-0.4$  to  $0.2$  dex.

bias function is an update of the Tinker et al. (2010) bias function. Based on Figure B1, we caution the reader that minimum halo mass measurements, be it from our model in §3.6 or any HOD/abundance matching model, are sensitive to the assumed halo prescriptions.

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