Dependence of galaxy clustering on UV luminosity and stellar mass at $z \sim 4–7$

Yisheng Qiu,1,2* J. Stuart B. Wyithe,1,2 Pascal A. Oesch,3 Simon J. Mutch,1,2 Yuxiang Qin1,2 Ivo Labbé,4 Rychard J. Bouwens,5 Mauro Stefanon5 and Garth D. Illingworth6

1School of Physics, University of Melbourne, Parkville, VIC 3010, Australia
2ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia
3Department of Astronomy, University of Geneva, Chemin des Maillettes 51, CH-1290 Versoix, Switzerland
4Centre for Astrophysics & Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3112, Australia
5Leiden Observatory, Leiden University, NL-2300 RA Leiden, the Netherlands
6UCO/Lick Observatory, University of California, Santa Cruz, CA 95064, USA

Accepted 2018 September 22. Received 2018 August 23; in original form 2018 June 26

ABSTRACT

We investigate the dependence of galaxy clustering at $z \sim 4–7$ on UV luminosity and stellar mass. Our sample consists of $\sim 10,000$ Lyman-break galaxies in the eXtreme Deep Field (XDF) and Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) fields. As part of our analysis, the $M_\star - M_{UV}$ relation is estimated for the sample, which is found to have a nearly linear slope of $d\log_{10} M_\star / d M_{UV} \sim 0.44$. We subsequently measure the angular correlation function and bias in different stellar mass and luminosity bins. We focus on comparing the clustering dependence on these two properties. While UV luminosity is only related to recent starbursts of a galaxy, stellar mass reflects the integrated build-up of the whole star formation history, which should make it more tightly correlated with halo mass. Hence, the clustering segregation with stellar mass is expected to be larger than with luminosity. However, our measurements suggest that the segregation with luminosity is larger with $\simeq 90$ per cent confidence (neglecting contributions from systematic errors). We compare this unexpected result with predictions from the MERAXES semi-analytic galaxy formation model. Interestingly, the model reproduces the observed angular correlation functions and also suggests stronger clustering segregation with luminosity. The comparison between our observations and the model provides evidence of multiple halo occupation in the small-scale clustering.

Key words: galaxies: evolution – galaxies: haloes – galaxies: high-redshift.

1 INTRODUCTION

Galaxy clustering provides a probe of the host halo mass of galaxies. The clustering strength is commonly described by the two-point correlation function, which measures the probability of finding galaxy pairs at given spatial separations. Mo & White (1996) used the extended Press–Schechter formalism (Bond et al. 1991) to show that the ratio between the correlation functions of haloes and the underlying matter depends on halo mass. This ratio is known as bias. Since galaxies reside in haloes, the bias links galaxy clustering to the mass of their host haloes (see Cooray & Sheth 2002 for a review).

* E-mail: yishengq@student.unimelb.edu.au

The dependence of clustering strength on galaxy properties is known as clustering segregation and reveals the correlation between galaxy properties and halo mass. At high redshifts, clustering segregation is observed for Lyman-break galaxies (LBGs) with UV luminosity (Lee et al. 2006; Hildebrandt et al. 2009; Barone-Nugent et al. 2014; Harikane et al. 2016, 2018) and stellar mass (Ishikawa et al. 2017; Durkalec et al. 2018). One basic conclusion from these studies is that more luminous and larger stellar mass galaxies are more clustered, and therefore reside in more massive haloes.

In the context of hierarchical galaxy formation, this correlation between halo and galaxy properties is unsurprising, since halo mass is closely related to the gas reservoir available for star formation and those processes that impact on it. For instance, in the low-mass regime, supernova (SN) feedback can effectively suppress star formation (Wyithe & Loeb 2013; Duffy et al. 2014; Hopkins et al. 2014).
Therefore, it is of particular interest to explore which galaxy property is more tightly correlated with the host halo mass. While UV luminosity is directly related to the current star formation rate, stellar mass provides integrated information over the star formation history. For this reason, it is expected that, when splitting the same sample by luminosity and stellar mass, clustering segregation with stellar mass should be larger than with UV magnitude.

In order to observe the difference between clustering segregation with stellar mass and luminosity, it is essential that the stellar mass is measured from SED fitting including rest-frame optical photometry. Recently, Harikane et al. (2016) carried out an analysis of clustering segregation with stellar mass in similar fields to our work. However, the stellar mass used in that study was obtained from a simple conversion of the UV luminosity to mass using the $M_\ast - M_{\text{UV}}$ relation. Thus, the analysis could not self-consistently infer any difference of the clustering segregation between stellar mass and UV luminosity. In this paper, we measure stellar masses from spectral energy distribution (SED) fitting including rest-frame optical Spitzer/IRAC data, which allows us to measure the clustering segregation in stellar mass at these redshifts for the first time.

Semi-analytic models (SAMs) of galaxy formation are based on halo merger trees provided by N-body simulations and evolve galaxy properties within these haloes using analytic or empirical prescriptions of baryonic physics and feedback processes. Since the theory that links galaxy clustering to dark matter haloes also relies on N-body simulations (e.g. Navarro, Frenk & White 1996; Tinker et al. 2010), a comparison between observed clustering and that predicted by an SAM tests the link between galaxy properties and halo mass. In this study, we compare our clustering measurements with results from the MERAXES SAM. This model has been shown to be successful in reproducing UV-luminosity functions and stellar mass functions over a wide range of redshifts (Liu et al. 2016; Mutch et al. 2016; Qin et al. 2017).

This paper is organized as follows. We describe the observational catalogue used in the analysis and measure the $M_\ast - M_{\text{UV}}$ relation in Section 2. Methods to measure the angular correlation function (ACF) and galaxy bias are introduced in Section 3, and results are demonstrated in Section 4. We perform the comparison between the observations and predictions from MERAXES in Section 5. Finally, the work is summarized in Section 6. Throughout the paper, unless specified, the cosmology of ($h, \Omega_m, \Omega_b, \Omega_{\Lambda}, \sigma_8$) = (0.678, 0.308, 0.0484, 0.692, 0.815) (Planck Collaboration I 2016) is assumed. Magnitudes are in the AB system.

# 2 DATA

The galaxy sample used for our measurements is based on the photometric catalogue from Bouwens et al. (2015), who selected LBGs at $z \sim 4$–7 based on the Hubble Space Telescope (HST) data in all the CANDELS fields, as well as the very deep XDF and HUDF09 parallel fields. In particular, our sample is drawn from the XDF (Illingworth et al. 2013), HUDF-091 and HUDF-092 (Bouwens et al. 2011), CANDELS-GR and CANDELS-GS (Grogin et al. 2011; Koekemoer et al. 2011), ERS (Windhorst et al. 2011), and CANDELS-UDS, CANDELS-COSMOS and CANDELS-EGS (Grogin et al. 2011; Koekemoer et al. 2011). These survey regions span an aggregate of $\sim 700$ arcmin$^2$ in the sky, and $\sim 10,000$ LBGs are identified. Photometric redshifts of these sources are estimated using the EAZY code (Brammer, van Dokkum & Coppi 2008). For more information on the LBG selection and the photometric redshifts, see Bouwens et al. (2015).

We combine the HST photometry with the large archive of Spitzer/IRAC legacy data available in the CANDELS fields (Ashby et al. 2013, 2015), which includes the ultra-deep I GOODS/UDF and GREATS surveys (Labbé et al. 2015; Labbé et al. in preparation) in the GOODS fields. IRAC photometry is measured in circular apertures after subtracting the contaminating flux of neighbouring galaxies using the code MOPHONGO (Labbé et al. 2006; Labbé et al. in preparation), which is similar to the code tPHOT (Merlin et al. 2016).

We measure stellar masses of galaxies based on SED fitting to the HST+Spitzer photometry using zEBRA+ (Oesch et al. 2010). The synthetic template set used here is based on Bruzual & Charlot (2003) with a constant star-formation history, subsolar metallicities (0.2Z$_\odot$) and a Chabrier (2003) initial mass function (IMF). Nebular continuum and emission lines are added self-consistently based on the number of ionizing photons emitted by each and used assuming line ratios relative to H β as tabulated by Anders & Fritze-v. Alvensleben (2003). Dust extinction is applied using the attenuation curve by Calzetti et al. (2000).

Following Bouwens et al. (2015), the absolute magnitudes, $M_{\text{UV}}$, are computed based on the fluxes in the photometric band that is closest to rest-frame 1600 Å. We first fit the $M_\ast - M_{\text{UV}}$ relation for the LBG sample, which will be used in our clustering analysis to compare stellar mass and luminosity segregation. The form of the relation is assumed to be

\[
\log_{10} M_\ast^\text{obs} = \frac{d\log_{10} M_\ast}{dM_{\text{UV}}} (M_{\text{UV}} + 19.5) + \log_{10} M_\ast^* (M_{\text{UV}} - 19.5),
\]

where mass is in units of $M_\odot$. The log-likelihood is then constructed as

\[
\ln \mathcal{L} = -\frac{1}{2} \sum \frac{[\log_{10}(M_\ast^\text{obs} / M_\ast^*)^2]}{\Delta^2} + \ln(2\pi \Delta^2),
\]

where the sum is over all LBGs, and $\Delta$ is a mass-independent free parameter representing scatter in the $M_\ast - M_{\text{UV}}$ relation. We adopt a Bayesian approach to perform the fit, assume constant priors for all parameters, and apply the EMCEE MCMC sampler developed by Foreman-Mackey et al. (2013). The resulting $M_\ast - M_{\text{UV}}$ relations are shown in Fig. 1. Best-fitting parameters are given in Table 1. We find that the $M_\ast - M_{\text{UV}}$ relations are close to linear ($M_\ast \propto \alpha L$) and that the scatter in stellar mass at fixed luminosity is $\sim 0.5$ dex. Even though our best-fitting slopes are slightly shallower, our measurements are consistent with the recent study from Song et al. (2016).

In this work, every galaxy in our HST sample has an estimate of stellar mass irrespective of the quality of Spitzer data. Low signal-to-noise ratios (S/Ns) of Spitzer bands could make the stellar masses less precise. To investigate this, in Fig. 1, galaxies that have at least one Spitzer band (3.6 and 4.5 μm) with $\text{S/N} > 3$ are shown as yellow empty circles in Fig. 1, while the others are shown as small blue dots. No systematic offset is found between them. Since the sample is large enough at $z \sim 4$, we use multiple stellar mass and luminosity bins for the clustering measurements, and avoid using bins that have no lower bound to reduce possible effects due to low S/N ratios of Spitzer data. The lower bound for the least massive stellar mass bin is shown as red dashed line in the corresponding panel of Fig. 1. At all other redshifts, limited by the sample size, we include all galaxies and use two bins to examine clustering segregation with stellar mass and luminosity. The fraction of LBGs that are included and have at least one Spitzer band with $\text{S/N} > 3$ is 74 per cent, 63 per cent, 53 per cent, and 47 per cent at $z = 4, 5, 6,$ and 7, respectively. The advantage of this approach is that the completeness of sample LBGs, which is defined by the selection, is not affected by Spitzer...
errors by bootstrap resampling (Ling, Barrow & Frenk 1986). We measure the ACFs in logarithmic bins and estimate of uniformly distributed random points. This is generated by a random Poisson process. For each field, the random catalogue contains galaxy–galaxy, galaxy–random, and random–random pairs, respectively. These probabilities are calculated by counting all pairs at separations between \( \theta \) and \( \theta + \delta \theta \), and normalizing by the total number of pairs. Estimates of DR \((\theta)\) and RR \((\theta)\) require a catalogue of uniformly distributed random points. This is generated by a random Poisson process. For each field, the random catalogue contains 10 000 points uniformly placed within the corresponding survey regions. We measure the ACFs in logarithmic bins and estimate errors by bootstrap resampling (Ling, Barrow & Frenk 1986). We construct bootstrap subsamples by replacing individual galaxies and perform the resampling for \( \sim 500 \) times. This approach is also used in Barone-Nugent et al. (2014) and Harikane et al. (2016). In order to investigate the clustering dependence on both stellar mass and UV magnitude, the total sample is split into subsamples. We choose bins such that they satisfy the \( M_\ast - M_{UV} \) relation at each redshift. The bin cuts are listed in Table 2.

Since the area of each survey region is finite, the observed ACFs are affected by border effects. This is corrected by an additive correction, which is known as the intergal constrain (IC). Following Roche & Eales (1999), we have

\[
\omega_{\text{true}}(\theta) = \omega_{\text{obs}}(\theta) + \text{IC},
\]

with

\[
\text{IC} = \frac{1}{\Omega^2} \int_{\theta_{\text{min}}}^{\theta_{\text{max}}} \omega_{\text{true}}(\theta) \, d\Omega_1 \, d\Omega_2 = \frac{1}{\Omega} \sum_i \text{RR}(\theta_i) \omega_{\text{true}}(\theta_i) / \sum_i \text{RR}(\theta_i),
\]

where \( \omega_{\text{true}}(\theta) \) is the fitting model of the ACF.

We assume the ACFs to be a power law

\[
\omega_{\text{true}}(\theta) = A_w \left( \frac{\theta}{1 \text{ arcsec}} \right)^{-\beta},
\]

and construct the log-likelihood using

\[
\ln L = -\frac{1}{2} \sum_i \sum_{\text{fields}} \left[ \frac{\omega_{\text{obs}}(\theta) - A_w (\theta^{-\beta} - \text{IC} / A_w)}{\sigma(\theta)} \right]^2,
\]

where sums are over all bins \( i \) and over all survey fields. This is equivalent to measuring the average ACF of all fields using inverse-variance weighting. Since this approach requires an estimate of the ACF in each individual field, we only include fields that are deep enough such that the mean separation of galaxy pairs is smaller than 100 arcsec. The number of LBGs that enter into the analysis is nearly 6000 in each field. The dependence of the IC on the fitting model results in some degeneracy between \( A_w \) and \( \beta \) (Lee et al. 2006), we therefore fix \( \beta = 0.6 \) following Lee et al. (2006) and Barone-Nugent et al. (2014). In addition, since the area of XDF, HUDF-091, and HUDF-092 is only 4.7 arcmin \(^2\) (i.e. one WFC3/IR pointing), the counted number of galaxy pairs in these fields decreases when the angular separation is greater than \( \sim 140 \) arcsec. We therefore only include separations smaller than that in the likelihood function.

The amplitude of the ACF \( A_w \) could be weakened by contamination of lower redshift sources. We reduce this effect by removing...
all LBGs whose best-fitting photometric redshift indicates that it might be a low-redshift contaminant ($z_{\text{phot}} < 2$). It is also noted that Harikane et al. (2016) used the contamination fraction estimated by Bouwens et al. (2015) to correct for this effect. They found that the difference is insignificant compared with the statistical error. Thus, no further treatment is employed to correct the effect of contamination.

### 3.2 Estimating the correlation length and bias

Real space parameters are obtained by applying the Limber transform to the ACFs. The real-space correlation $\xi(r)$ provides three-dimensional information on galaxy clustering, which is also approximated by a power law,

$$\xi(r) = \left( \frac{r}{r_0} \right)^{-\gamma},$$

where $r_0$ is called the correlation length. In this case, the Limber transform takes the form (Peebles 1980)

$$\beta = \gamma - 1,$$

and

$$A_\omega = r_0^2 B \left( \frac{1}{2} \right)^{-\gamma} \left( \frac{1}{2} \right)^{-\gamma} \int_0^\infty \int_0^\infty dN(z) d\Omega \left| A \right|^{1+\gamma} (1 + z)^{1-\gamma},$$

with

$$d_H(z) = \frac{c}{H(z)}, \quad d_a = \frac{1}{1+z} \int_0^z dH d\zeta,$$

where $B(x, y)$ is the beta function, $N(z)$ is the redshift distribution function of sample galaxies, and $H(z)$ is the Hubble parameter as a function of redshift. The above equations link $A_\omega$ and $\beta$ to the power-law parameters in the real space. $N(z)$ is estimated using the photometric redshifts of each LBG.

We derive the bias using the ratio between the variance of the galaxy and the matter correlation functions smoothed by a top-hat with radius $8h^{-1}$ Mpc:

$$b = \frac{\sigma_{8,s}^2}{\sigma_8},$$

where (Peebles 1980)

$$\sigma_{8,s}^2 = \frac{72(\bar{r}_0/8h^{-1}\text{Mpc})^3}{(3 - \gamma)(4 - \gamma)(6 - \gamma)\gamma^2}.$$
Clustering segregation at $z \sim 4-7$.

Figure 2. Measured ACFs and their best-fitting power laws $A_\omega(\theta/1$ arcsec)$^{-0.6}$. Top left, top right, bottom left, and bottom right-hand panels show the measurements at $z = 3.8, 5.0, 5.9, \text{and } 6.8$, respectively. For each panel, the first and second columns illustrate the stellar mass and luminosity subsamples, respectively. In all plots, black squares with error bars are measured ACFs, which are averaged over all fields using inverse-variance weighting, and dashed lines are best-fitting power laws.
samples at \( z \sim 7 \), we find that the bias increases from \( b = 7.4^{+2.0}_{-1.4} \) to \( b = 11.7^{+2.2}_{-1.6} \) and from \( b = 5.6^{+1.3}_{-1.1} \) to \( b = 8.6^{+2.3}_{-1.8} \) for stellar mass and luminosity bins, respectively. As a comparison, we also plot the bias estimated by Harikane et al. (2016) from their power-law fits as a function of mean UV magnitude. We find that the trends of clustering dependence on luminosity are consistent, while the offsets on biases themselves could be due to different methods of computing them. To summarize, we find that both more massive and more luminous galaxies are more highly clustered, implying that they are hosted by more massive dark matter haloes.

On the other hand, the comparison of segregation between stellar mass and luminosity disagrees with our prior expectations, especially at \( \bar{z} = 3.8 \) and 5.0. In particular, we find that the clustering dependence is larger for luminosity than for stellar mass. In order to quantify this trend, we fit the measured biases of the lightest (faintest) and heaviest (brightest) by straight lines, i.e.

\[
b = \alpha_{\text{SM}} \log_{10}(M_*/M_\odot) + \text{const},
\]

and

\[
b = -1.1 \alpha_{\text{UV}} \log_{10}(L_{\text{UV}}/L_\odot) + \text{const}.
\]

If clustering segregation with both properties were the same, one would expect that the ratio of the slopes should recover the slope of the \( M_*/M_{\text{UV}} \) relation. Therefore, we include a correction factor of \(-1.1\) to equation (14) in order to make \( \alpha_{\text{SM}} \) and \( \alpha_{\text{UV}} \) comparable. The value \(-1.1\) is based on the results in Table 1. If \( M_\star \propto L \), this factor would become unity. We combine the measured biases over all redshifts, assume the same slope but different intercepts for this factor would become unity. We combine the measured biases of the lightest (faintest) and heaviest (brightest) by straight lines, i.e.

\[
b = \alpha_{\text{SM}} \log_{10}(M_*/M_\odot) + \text{const},
\]

and

\[
b = -1.1 \alpha_{\text{UV}} \log_{10}(L_{\text{UV}}/L_\odot) + \text{const}.
\]

If clustering segregation with both properties were the same, one would expect that the ratio of the slopes should recover the slope of the \( M_*/M_{\text{UV}} \) relation. Therefore, we include a correction factor of \(-1.1\) to equation (14) in order to make \( \alpha_{\text{SM}} \) and \( \alpha_{\text{UV}} \) comparable. The value \(-1.1\) is based on the results in Table 1. If \( M_\star \propto L \), this factor would become unity. We combine the measured biases over all redshifts, assume the same slope but different intercepts for this factor would become unity. We combine the measured biases of the lightest (faintest) and heaviest (brightest) by straight lines, i.e.

\[
b = \alpha_{\text{SM}} \log_{10}(M_*/M_\odot) + \text{const},
\]

and

\[
b = -1.1 \alpha_{\text{UV}} \log_{10}(L_{\text{UV}}/L_\odot) + \text{const}.
\]

If clustering segregation with both properties were the same, one would expect that the ratio of the slopes should recover the slope of the \( M_*/M_{\text{UV}} \) relation. Therefore, we include a correction factor of \(-1.1\) to equation (14) in order to make \( \alpha_{\text{SM}} \) and \( \alpha_{\text{UV}} \) comparable. The value \(-1.1\) is based on the results in Table 1. If \( M_\star \propto L \), this factor would become unity. We combine the measured biases over all redshifts, assume the same slope but different intercepts for this factor would become unity. We combine the measured biases of the lightest (faintest) and heaviest (brightest) by straight lines, i.e.

\[
b = \alpha_{\text{SM}} \log_{10}(M_*/M_\odot) + \text{const},
\]

and

\[
b = -1.1 \alpha_{\text{UV}} \log_{10}(L_{\text{UV}}/L_\odot) + \text{const}.
\]

If clustering segregation with both properties were the same, one would expect that the ratio of the slopes should recover the slope of the \( M_*/M_{\text{UV}} \) relation. Therefore, we include a correction factor of \(-1.1\) to equation (14) in order to make \( \alpha_{\text{SM}} \) and \( \alpha_{\text{UV}} \) comparable. The value \(-1.1\) is based on the results in Table 1. If \( M_\star \propto L \), this factor would become unity. We combine the measured biases over all redshifts, assume the same slope but different intercepts for this factor would become unity. We combine the measured biases of the lightest (faintest) and heaviest (brightest) by straight lines, i.e.

\[
b = \alpha_{\text{SM}} \log_{10}(M_*/M_\odot) + \text{const},
\]

and

\[
b = -1.1 \alpha_{\text{UV}} \log_{10}(L_{\text{UV}}/L_\odot) + \text{const}.
\]

If clustering segregation with both properties were the same, one would expect that the ratio of the slopes should recover the slope of the \( M_*/M_{\text{UV}} \) relation. Therefore, we include a correction factor of \(-1.1\) to equation (14) in order to make \( \alpha_{\text{SM}} \) and \( \alpha_{\text{UV}} \) comparable. The value \(-1.1\) is based on the results in Table 1. If \( M_\star \propto L \), this factor would become unity. We combine the measured biases over all redshifts, assume the same slope but different intercepts for this factor would become unity. We combine the measured biases of the lightest (faintest) and heaviest (brightest) by straight lines, i.e.

\[
b = \alpha_{\text{SM}} \log_{10}(M_*/M_\odot) + \text{const},
\]

and

\[
b = -1.1 \alpha_{\text{UV}} \log_{10}(L_{\text{UV}}/L_\odot) + \text{const}.
\]
Clustering segregation at $z \sim 4-7$

Figure 4. Results of a straight line fit of measured biases over all redshifts (assuming the same slope). In top panels, dashed lines are the best-fitting results. Squares and triangles with error bars are the data that is used for the fits. Empty stars and pentagons are predicted biases from the SAM MERAXES. When computing those pentagons, 0.5 dex Gaussian scatters are added to the stellar mass of each model galaxy. In the bottom left-hand panel, solid lines depict the marginalized distributions of the slope. We show the medians and $1-\sigma$ percentiles of the distributions in the top right corner. Dashed vertical lines show the slopes derived from the model, and the dot dashed line gives the result in the case where the model stellar mass has scatter added. The values of these vertical lines are also shown in the top right corner. In this panel, all red and blue lines correspond to stellar mass and luminosity split samples, respectively. Bottom right-hand panel shows the distribution of $\alpha_{\text{UV}} - \alpha_{\text{SM}}$ obtained by subtracting the samples of $\alpha_{\text{UV}}$ and $\alpha_{\text{SM}}$. The area of the shaded region is $\simeq 90$ per cent, which is the probability that $\alpha_{\text{UV}} > \alpha_{\text{SM}}$.

We follow Park et al. (2016, 2017) to select model LBGs and calculate the ACFs. This approach mimics the incompleteness of the LBG sample by adding photometric scatter to the magnitudes of each LBG selection band. The level of the scatter is given by the $1\sigma$ field detection limits. In the present work, we assemble model LBGs according to the flux limits of the deepest field in our observations, i.e., the XDF. The resulting redshift distribution of model selected LBGs is shown in Fig. 5, which agrees with the observed one estimated by photometric redshifts. In terms of the determination of the ACF, we compute the real-space correlation function across a sequence of snapshots directly from the spatial coordinates of each galaxy and convert to an ACF by the Limber transform. The readers are referred to Park et al. (2016, 2017) for a more detailed description.

5.2 Results

We focus on the comparison between our model and observations at $z \sim 4$, where the measurements have the smallest errors. We plot predicted ACFs together with measured ACFs in Fig. 6. The model ACFs agree very well with observations and reproduce the observed clustering dependence on both stellar mass and luminosity. In order to check whether the model predicts larger clustering segregation with luminosity as indicated by our observations, we therefore calculate the bias from the model by $b^2 = \xi(r) / \xi_{\text{DM}}(r, z)$, where $\xi(r)$ and $\xi_{\text{DM}}(r, z)$ are the real-space correlation functions of galaxies and dark matter. An average value is taken in the range $5 \text{ Mpc} \leq r \leq 10 \text{ Mpc}$. The estimated biases for the lightest (faintest) and heaviest
Figure 5. Example redshift distribution of observed and selected model galaxies in the $z \sim 4$ LBG sample. The observed distribution is estimated by photometric redshifts. Our colour selection on the model galaxies results in a very well-matched redshift distribution. The field depth of the XDF is used to select model galaxies.

(brightest) samples are shown using star symbols in the left (middle) panel in Fig. 4. Subsequently, we derive the slope for these two points for comparison with observational data in Fig. 4. We conclude that the measured variation of bias with stellar mass and luminosity is consistent with predictions and that the clustering segregation with luminosity is larger than stellar mass in the model. In other words, the model predictions also contradict our expectation that stellar mass should be more tightly correlated with halo mass. The physical interpretation behind this could be complex, and we defer it to a subsequent paper.

The observed clustering dependence on stellar mass could be weakened due to observational uncertainties on the estimations of stellar mass. The uncertainties, for instance, can be due to the low S/N of Spitzer data. We demonstrate this effect by adding 0.5 dex of Gaussian scatter to the stellar mass of selected model LBGs and remeasure the clustering in stellar mass bins. This level of scatter is also used in abundance matching studies (e.g. Behroozi, Wechsler & Conroy 2013; Moster, Naab & White 2013). The recalculated ACFs are shown as green lines in Fig. 6. For the most massive bin, the ACF decreases at all scales, while for the other two bins, scatter in stellar mass only affects the small-scale correlation functions. We also recalculate the bias and the corresponding slope. The results are demonstrated in Fig. 4. In the right-hand panel of Fig. 4, the dot–dashed line represents the slope in the case where scatter is added to the stellar mass, and shows that scatter reduces the slope relative to that of the original model, from $\alpha_{SM} = 1.3$ to $\alpha_{SM} = 0.7$. This effect is most significant for the most massive bins, since the massive end has fewer galaxies. We can use this systematic error of model $\alpha_{SM}$ introduced by adding scatter to the stellar mass to estimate the effect on our measured $\alpha_{SM}$. This indicates that accounting for uncertainties in stellar mass leads to clustering segregation that is similar for both mass and UV luminosity, but which is not larger with stellar mass. Hence, although uncertainties in stellar mass can weaken the clustering dependence, the unexpected trend could still result from physical reasons.

5.2.1 Satellite galaxies

Another interesting finding from the comparison between observations and the model is the deviation of the ACFs from a power law at small scales ($\theta < 10$ arcsec). In the model, this is due to the one-halo term, arising from multiple halo occupation, where more than one galaxy resides in the same halo. To provide evidence of this multiple halo occupation, we calculate the one-halo and two-halo terms explicitly from MERAXES by counting galaxy pairs in the same and different friends-of-friends groups. These two terms are shown as dashed and dot–dashed lines in Fig. 6, respectively, demonstrating that the steep increase of the ACFs at small scales is due to multiple halo occupation. Consistency is also found between observations and the model. It can also be seen that the transition of the model ACFs between the one-halo and two-halo terms becomes more rapid with decreasing stellar mass. However, there is no such trend in luminosity. This finding may suggest an additional feature of clustering segregation with stellar mass and luminosity, imply-
ing different satellite properties for the two cases. However, we caution that the satellite properties of MERAXES have yet to be fully explored and compared with observations and that achieving realistic recent satellite star formation histories has traditionally been a challenging task for SAMs. This difference can also be seen from the observed ACFs but with large uncertainties. At the very bright end, the smooth transition between one-halo and two-halo terms is observed in Harikane et al. (2018) and explained using non-linear bias (Jose, Lacey & Baugh 2016). Larger surveys with more complete samples might be used to investigate this phenomenon in more details for fainter and less massive galaxies.

6 SUMMARY
We have carried out a clustering analysis of LBGs over the range \(z \sim 4–7\), with emphasis on the comparison between clustering segregation with stellar mass and luminosity. We also compare our measurements with predictions from the MERAXES SAM. Our findings can be summarized as follows:

(i) The observed ACF amplitude and bias generally increase with stellar mass and luminosity over \(z \sim 4–7\). The ACFs obtained from the model are consistent with observations and reproduce clustering segregation with both stellar mass and luminosity. This suggests that more luminous and massive galaxies are more clustered, and hence hosted by more massive dark matter haloes.

(ii) By combining measurements over all redshifts, a systematic difference is found between clustering segregation with stellar mass and luminosity. In particular, it is observed that clustering strength is more tightly correlated with luminosity. This is in contrast to the expectation that stellar mass should be more tightly correlated with halo mass since stellar mass reflects the whole star formation history, while UV magnitude only corresponds to recent star formation. We find that the model also predicts this surprising result of larger clustering segregation with luminosity.

(iii) At \(z \sim 4\), the model predicts that the transition between the one-halo and two-halo terms of the ACFs is smoother for larger stellar mass measurements substantially reducing the systematic errors in clustering studies.

ACKNOWLEDGEMENTS
Parts of this work were performed on the gSTAR national facility at Swinburne University of Technology. gSTAR is funded by Swinburne and the Australian Government’s Education Investment Fund.

This research was also supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

REFERENCES
Harikane Y. et al., 2018, PASJ, 70, S11
